Manuscript version with marked-up changes

Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part2: the European terrestrial CO_2 fluxes

Panagiotis Kountouris[1], Christoph Gerbig[1], Christian Rödenbeck[1], Ute Karstens[1,*], Thomas F. Koch[2], Martin Heimann[1]

[1]Max Planck Institute for Biogeochemistry, Jena, Germany

[2]Meteorological Observatory Hohenpeissenberg, Deutscher Wetterdienst, Germany

[*]now at ICOS Carbon Portal, Lund University, Lund, Sweden

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

Abstract

Optimized biogenic carbon fluxes for Europe were estimated from high resolution regional scale inversions, utilizing atmospheric CO_2 measurements at 16 stations for the year 2007. Additional sensitivity tests with different data-driven error structures were performed. As the atmospheric network is rather sparse and consequently contains large spatial gaps, we use a priori biospheric fluxes to further constrain the inversions. The biospheric fluxes were simulated by the Vegetation Photosynthesis and Respiration Model (VPRM) at a resolution of 0.1º and optimized against Eddy covariance data. Overall we estimate an a priori uncertainty of 0.54 GtC y^{-1} related to the poor spatial representation between the biospheric model and the ecosystem sites. The sink estimated from the atmospheric inversions for the area of Europe (as represented in the model domain) ranges between 0.23 and 0.38 GtC y^{-1} (0.39~~0~~ and 0.71~~49~~ GtC y^{-1} up-scaled to geographical Europe). This is within the range of posterior flux uncertainty estimates of previous studies using ground based observations.

1 Introduction

Global and regional atmospheric inversions ~~have been applied using~~assimilate atmospheric CO_2 measurements made by a global network ~~since~~ for two decades, to infer terrestrial carbon fluxes using surface in situ or flask measurements of CO_2 dry mole fractions (Tans et al., 1989; Enting and Mansbridge, 1989, Conway et al., 1994, Fan et al., 1998; Rödenbeck et al., 2003). The optimization of CO_2 biospheric fluxes for the European domain has been ~~the focus~~ of high interest in previous studies either using pseudo or real data (~~Gurney et al., 2004;~~ Peters et al., 2010; Carouge et al., 2010a; Carouge et al., 2010b; Rivier et al., 2010; Broquet et al., 2011; Broquet et al., 2013; Peylin et al., 2013). Retrieved fluxes from most of the inversions are obtained from global systems at coarse resolution; hence, ~~which makes difficult to retrieve~~ the spatial and temporal flux variability at finer scales can not be resolved. Large uncertainties in the flux retrievals are introduced due to the coarse resolution of the transport models used and due to the network sparseness (Peters et al., 2010). ~~For example the prevailing westerly winds and the fact that all atmospheric sites are mainly located in central Europe, introduce large flux uncertainties at eastern European regions.~~

Apart from ground based observations, satellite measurements have also been recently used in atmospheric inversions to infer terrestrial fluxes (Basu et al., 2013; Deng et al., 2014; Chevallier et al., 2014). The advantage of using space-borne measurements lies on the high density of the observations providing the opportunity to constrain regions ~~which the~~not seen by the ground network ~~does not see~~. However satellite based inversions significantly differ from ground based inversions, reporting a larger annual uptake for Europe. A characteristic example is the estimated European uptake in the study by Reuter et al. (2014). They calculated an uptake of 1.02 $GtCy^{-1}$ which triggered an ongoing debate on whether those estimates are data driven or they lack robustness due to deficiencies in the satellite observations and in the inverse modeling (Feng et al., 2016).

One of the largest sources of uncertainty in inversions is the atmospheric transport uncertainty. Modeled ~~tracer~~ dry mole fractions are biased particularly due to uncertainties ~~of~~ in vertical mixing near the surface (Gurney et al., 2003; Gerbig et al., 2008; Houweling et al., 2010). As a consequence, posterior flux estimates are also biased because biases in concentrations due to transport model errors are translated into biases in fluxes through the optimization procedure. Propagation of uncertainties in winds (Lin and Gerbig, 2005) and in mixing heights (Gerbig et al., 2008) for summer months with active vegetation resulted in uncertainties in simulated dry air mole fractions of 5.9 ppm and 3.5 ppm respectively.

The current study is a continuation of the Kountouris et al. (2016) study (hereafter referred to as Ko16) in which the inversion system and its set-up were assessed based on pseudo data. As a next step we apply the modeling framework to real CO_2 atmospheric observations. Our main objectives in the second part of this work are to investigate the potential to infer flux estimates for Europe with reduced uncertainties, and to estimate biospheric fluxes at high spatial resolution and for a full year. We use a spatial flux resolution of 0.25^o x 0.25^o to couple fluxes with the atmospheric transport model, and the state space allows optimizing 3-hourly NEE corrections to the prior NEE fluxes at a nominal spatial resolution of 0.5^o x 0.5^o. A data driven error structure is implemented consistent with model-data flux mismatches (Kountouris et al., 2015) as tested in part 1 (Ko16) of this study. Further, different error structures are used and assessed including also a spatial error structure with a hyperbolic correlation shape as suggested by Chevallier et al. (2012). Since spatial autocorrelations have been found to be very short, the annual aggregated uncertainty over the European domain is smaller than traditionally assumed (see also Ko16). The error inflation necessity and implementation was addressed in Ko16 either by inflating the error covariance or, more formally, by introducing a bias term. However the hyperbolic correlation shape suggested by Chevallier et al. (2012) has a stronger impact from larger distances compared to the exponential shape, leading to an aggregated uncertainty which does not require to be inflated. We perform also a number of sensitivity tests to account for misrepresentation of the fossil fuel signal and also for transport uncertainties due to vertical mixing.

This paper is structured as follows: Section 2 describes the inversion system, the network and station data which are used and details the assumed error structure. Section 3 shows the results of the goodness of fit, and the retrieved fluxes. The data fitting and the reliability of the posterior fluxes are extensively discussed in section 4.

2 Methods

2.1 Two-step inversion

Real-data inversions require a nested inversion scheme, since observations contain also contributions from regions outside of the Domain of Interest (DoI). As in part 1 of this study (Ko16), the Jena Inversion System (Rödenbeck 2005) including the two-step nesting scheme (Rödenbeck et al., 2009; Trusilova et al., 2010) was used. This scheme allows for combining regional and global inversions within a consistent system. Here we only provide a brief description as details are given in Rödenbeck et al. (2009) and Trusilova et al. (2010). The atmospheric transport models TM3 (step 1) (Heimann and Körner, 2003) and STILT (step 2) (Lin et al., 2003) were used for transport at the global and regional domain, respectively. For the global runs, TM3 was used at a spatial resolution of 4° latitude x 5° longitude, driven by meteorological fields from the ERA-Interim reanalysis produced by ECMWF (Dee et al., 2011). The transport matrix for the regional inversions was identical to the one used for the synthetic data study in part 1.

In the first step, a global inversion is performed using the global transport model. The outcome is an optimized flux field, at coarser scale for the full period (FP) and the global domain. Then two forward runs are performed. The first run uses the global transport model over the FP, computing the modeled mixing ratios Δc_{mod1}. The second run initializes again the global transport model but only within the regional DoI. This can be regarded as a regional simulation, but with coarse resolution, yielding modeled mixing ratios Δc_{mod2}. Then the "remaining mixing ratio" is calculated for all the observing sites inside the DoI:

$$\Delta c_{remain} = c_{meas} - (\Delta c_{\mathrm{mod}\,1} - \Delta c_{\mathrm{mod}\,2} + c_{ini}) \tag{1}$$

were c_{ini} the initial condition which corresponds to a well mixed atmosphere with a given initial tracer mixing ratio.

In step two, the high-resolution transport model is used for the regional inversion within the DoI, where all fluxes are represented at fine resolution. For this inversion the vector containing the measured mixing ratios c_{meas} are replaced by the "remaining mixing ratios" Δc_{remain}. The optimized fluxes from this step are the high-resolution fluxes of interest.

2.2 Atmospheric network and data

For step 1 we used the same station network as in version s04_v3.6 of the Jena Carboscope CO_2 inversion (http://www.bgc-jena.mpg.de/CarboScope/?ID=s04_v3.6), with 64 stations globally. For step 2 (regional inversion) continuous and flask measurements from 16 stations within Europe were used as described in Ko16 (see also Table 1). Of those 16 stations 7 are already included in the step 1 inversion. All provided valid values were used, except those paired flask measurements that differ more than 0.34 ppm which were omitted. Measurements from the continuous stations were aggregated to hourly values where needed~~, and neighboring values differing by more than 1 ppm were omitted~~. Night and day time observations were selected depending on the type of station (Ko16). As all institutions report mixing ratio values traceable to WMO (World Meteorological Organization) calibration scale, we expect compatibility between the different datasets (also see Rödenbeck et al., 2006).

In this study we use the site HEI (Heidelberg) which is traditionally not used for European CO_2 flux inversions as being considered too local (Broquet et al., 2013; Rödenbeck et al., 2009; Rivier et al., 2010). The Heidelberg region is considered to be one of the most polluted regions in Germany (Fiedler et al. 2005) and therefore could bias the flux estimates. Moreover the WES (Westerland) site contains long periods with no data. This could potentially affect posterior flux estimates since extended data gaps can lead to jumps in the presence of biases. Thus we evaluate the performance and the sensitivity of the European flux estimates to the network configuration, by performing also an inversion (referred to as ~~S2~~nBV14, see Table 2) excluding HEI and WES.

2.3 A-priori information and uncertainties

A set of inversion cases differing in the prior information, the error structure and the station configuration was realized (see overview in Table 2). Prior information derived from both biosphere models (VPRM and GBIOME-BGCv1) is used to investigate the impact of the prior fields to the posterior flux estimates. Furthermore an ensemble of inversions using different error structures is used to investigate the impact on the posterior flux estimates and uncertainties.

Similarly to the synthetic inversion (Ko16) the model-data mismatch uncertainties are the same as in the Ko16 study (see also fig. 2 therein). Further, we use the base case ~~B1~~nBV (No Bias VPRM as prior, B1 in Ko16) which inflates the prior uncertainty by up-scaling the error covariance matrix, and case ~~S1~~BVR (Bias VPRM as prior Respiration as shape, S1 in Ko16) which includes a bias term. In the base case the VPRM model provides the prior flux field, and exponentially decaying correlations are assumed. The bias component in the ~~S1~~BVR scenario will always have a correction with the same sign for all grid-cells as it just scales a predefined flux field. In the ~~S1~~BVR case it follows the shape of the annually averaged respiration flux, in the ~~S1a~~BVN case that of the a priori ~~net~~ Net biogenic flux, and in the ~~S1b~~BVRT case again that of the annually averaged respiration flux, but with monthly temporal resolution of the bias term to allow for some temporal flexibility. The ~~B2~~ nBB inversion refers to the scenario where GBIOME-BGCv1 was used as a priori information instead of VPRM, and the error structure does not contain a bias term. With this case we can evaluate how sensitive the posterior flux estimates are with respect to the prior information which has been used. We also examine a spatial error structure based on a hyperbolic (instead of an exponential) spatial correlation shape as suggested in Chevallier et al. (2012) which we will refer to as ~~S3~~nBVH scenario.

Note that in most of the inversions performed, VPRM fluxes were used as prior information. Those fluxes are already optimized using EC measurements, therefore evaluation of the posterior flux estimates against EC data at the local scale could result in posterior fluxes that are limited or even not further constrained (since they are already optimized). In contrast, posterior fluxes produced with BIOME-BGC used as prior are expected to show significantly larger corrections compared to the prior estimations, and are therefore used for evaluation against EC data.

Nevertheless in most cases we use VPRM as prior in order to keep our estimates as data-driven as possible through the overall optimization procedure; at local scale by using EC data, and at regional scale using the atmospheric dry mole fractions.

As in the synthetic experiment (Ko16) the temporal decorrelation time was set to 31 days. In Kountouris et al. (2015), model-data comparisons representative at site scale (around 1 km) showed spatial correlation lengths of 40 km whilst model-model comparisons representative at 50 km resolution identified a correlation scale of 370 km. Considering also that the state space has a resolution of 50 km, the spatial decorrelation length was chosen to be approximately 100 km (66 km in meridional, and 130 km in zonal direction). In the prior error covariance, diagonal elements of 2.27 μmol$m^{-2}s^{-1}$ were assumed, consistent with the model-data flux mismatches as calculated in Kountouris et al. (2015). Propagating this spatiotemporal error structure yields a domain-integrated uncertainty (E_{st}) of 0.15 GtC y^{-1}. Note that this is substantially smaller than for the synthetic experiment due to the much shorter spatial correlation length scales. A total annual, domain integrated uncertainty E_{tot} of 0.3 GtC y^{-1} was assumed, which corresponds to twice the standard deviation of annual terrestrial flux estimates for 2007 between terrestrial biosphere models taken from the global carbon atlas (http://www.globalcarbonatlas.org). This is also consistent with the prior uncertainty (for Europe) assumed for the global inversions performed by the Jena inversion system. For those inversions in which the additional bias term was considered (~~S1~~BVR, ~~S1a~~BVN, and ~~S1b~~BVRT scenarios), its error E_{BT} was calculated using

$$E_{tot}^{2} = E_{ST}^{2} + E_{BT}^{2} \qquad (2)$$

For the ~~S3~~nBVH scenario using hyperbolic correlations similar to Chevallier et al. (2012) ($\frac{1}{1+d}$), a characteristic value d (lag distance) was used such that the correlation drops after around 60 km to 1/e of its initial value, consistent with the hyperbolic fit to the model-data flux residual autocorrelation in Kountouris et al. (2015). For this case no additional bias term was needed, as the spatially and temporally aggregated uncertainty was found to be 0.32 GtC y^{-1}, which is very close to the uncertainty assumed for the inversions (0.3 GtC y^{-1}).

Furthermore, we include ocean fluxes from Mikaloff-Fletcher et al. (2007), and anthropogenic emissions from the EDGAR v4.1 inventory scaled at national level for individual years according to the BP (British Petroleum) statistical review of world energy (BP, 2012) following Steinbach et al. (2011). Anthropogenic emissions are considered to be perfectly known (with no prior uncertainty), as one typically assumes that there is more a-priori knowledge regarding the anthropogenic emissions as compared to biogenic fluxes. As the inversion cannot distinguish between biogenic and anthropogenic signals, any errors in the a-priori anthropogenic emissions will be included as corrections to the NEE flux.

2.4 Diagnostics and aggregation of fluxes

Similar to Ko16 we use the χ_c^2 metric to evaluate the goodness of fit for each station (Eq. 3)

$$\chi_c^2 = \frac{\sum_t \frac{(\Delta c_t)^2}{\sigma_t^2}}{n} \qquad (3)$$

where Δc_t is the model-data mismatch in dry mole fractions for a given observation time t, n the number of observations and σ_t the assumed uncertainty. Further we make use also of the reduced χ_r^2 (Eq. 4) where J_{min} is the cost function at its minimum

$$\chi_r^2 = 2\frac{J_{\min}}{n} \qquad (4)$$

For more details about the chi-square metric the reader is referred to Ko16 study.

The optimized fluxes are derived at 0.25º spatial and daily temporal resolution from the inversion system. We post-process the fluxes by aggregating them spatially at country/domain-wide scales and temporally at monthly/annual scales.

Flux comparisons with other studies require that both fluxes refer to the same geographical region. Typically studies refer to TransCom regions with a European domain that expands more into the Eurasian region. To scale our results to the TransCom EU region, we calculated the ~~flux~~ area ratio between the TransCom EU region and our European domain.~~ from the first step of the two-step inversion (global inversion), and used t~~This ratio (about 1.69~~3~~) was used to scale our posterior estimates and the corresponding uncertainties assuming linearity in the variances (presented in Fig. 8~~Fig. 7~~).

3 Results

3.1 Simulated CO_2 and goodness of fit

Figure 1~~Figure 1~~presents a comparison of observed and modeled daily averages of the nighttime (hours 23, 00, 1, 2, 3, 4 UTC) CO_2 dry air mole fractions for the Schauinsland station (SCH), a mountain station, for the year 2007. The prior estimates (gray line) as derived from a forward model run using VPRM flux fields are systematically lower than the observations (black line) with the most divergent values occurring during the growing season. A similar pattern was found for the other atmospheric stations. Posterior CO_2 timeseries from all the inversions are in much closer agreement with the observations.

Table 3 summarizes the statistics between the modeled and the observed CO_2 dry mole fractions for all stations based on daily averages using the respective sampling times (see also Ko16) for mountain (nighttime) and other stations (daytime). Of note is that the real data inversions include errors due to the modeling of transport, which is not the case in the synthetic experiment in Ko16 as the same transport model was used for forward and inversion runs. Standard deviations of the posterior residuals (observed – modeled) show an average decrease for all inversion setups and for all stations of 59% compared to the prior residuals. Correlations between prior and observed as well as posterior and observed mole fractions (also Table 3) were likewise increased on average from 0.48 to 0.93. Of note is that ~~B1~~nBV and ~~B2~~nBB, which use an inflated prior error covariance for the spatiotemporal component, show larger improvement relative to the prior in RMSD and some limited improvement in correlation coefficient, compared to those inversions where a bias component was included (~~S1~~BVR, ~~S1a~~BVN, ~~S1b~~BVRT). Figure 2 visually summarizes the goodness of fit in a Taylor diagram for cases ~~B1~~nBV and ~~S1~~BVR, presenting prior and posterior estimates of the correlation and the normalized standard deviation between modeled and observed CO_2 dry mole fraction time-series. It is obvious that the additional flexibility of ~~B1~~nBV in the spatiotemporal flux distribution results in a better reproduction of the concentration variability. The same picture emerges when comparing the ~~B1~~nBV and ~~B2~~nBB inversions to ~~S3~~nBVH (see Table 3). Although all these cases assume no explicit bias term in the error structure, the larger correlations from areas farther away for the ~~S3~~nBVH case with a hyperbolic correlation causes a reduced number of effective degrees of freedom, which results in larger residuals in posterior-observed mole fractions (Table 3) comparable to those of the ~~S1~~BVR case.

Calculating the goodness of fit using the station-specific χ_c^2 values from Eq. (3), most of the sites (Table 3) show values around 1, indicating that the misfits are inside the 1 sigma site specific uncertainty~~ range~~. For the CBW, HEI, JFJ, KAS sites, values above 1 regardless the error structure were found, with the most extreme value of 5.17 for the HEI site in the ~~S3~~nBVH inversion. This could suggest that for a polluted site as HEI larger uncertainties should be considered.

The reduced χ_r^2 values regarding the overall model performance (Eq. 4) for all inversion set ups is found to be close to 1 with χ^2 values of 1.08 (~~B1~~nBV), 1.16 (~~B2~~nBB), 1.17 (~~S1~~BVR), 1.17 (~~S1a~~BVN), 1.19 (~~S1b~~BVRT), 0.89 (~~S2~~nBV14) and 1.25 (~~S3~~nBVH), suggesting that the assumed prior uncertainty describes well the actual uncertainties.

3.2 Posterior flux estimates at different scales

The annually integrated spatial flux distribution is presented in Fig. 3~~Fig. 3 ~~for all the different inversion settings. Differences between the results based on the two general error structures (with and without the bias term) were observed mainly in central and Western Europe (longitudes less than 20º E), where the network provides a strong constraint. This difference is characterized by stronger spatial flux variability for the general ~~B1~~nBV case, with multiple transitions between carbon sources and sinks at regional scales. The same picture emerges for the western part of Europe. In contrast, all the inversions including a bias component (~~S1~~BVR, ~~S1a~~BVN, ~~S1b~~BVRT) yield a more homogeneous flux distribution with somewhat finer structure in the flux retrievals (e.g France and north-east part of Europe). Comparisons between ~~S1~~BVR, ~~S1a~~BVN, ~~S1b~~BVRT flux distributions do not show any significant difference. Almost the same picture emerges when comparing ~~B1~~nBV and ~~S2~~nBV14 cases, indicating that excluding the 2 stations does not have a very strong influence on our annual flux estimates. However spatial differences were observed for the areas close to the two sites. The most important one applies for the area near the HEI station where we observed a transition from source to net carbon sink when excluding the corresponding site. The choice of the prior does only have a small impact on the mean flux as can be seen by comparing posterior fluxes from ~~B1~~nBV and ~~B2~~nBB despite the significant differences in the flux innovations (Fig. 3~~Fig. 3~~). All innovations show that positive fluxes were added mainly in central Europe and more intensively for the cases were no bias term was used. The positive flux corrections is something to be expected since prior fluxes from VPRM show a strong European sink of 0.96 GtC y^{-1} which is most likely to be unrealistic. Overall the results suggest that the general error structure matters, i.e. whether or not to include a bias term, but how the bias is implemented is of less importance for the retrieved flux patterns. One would expect that the flux distribution from the ~~S3~~nBVH case would follow the general flux structure from the inversions without the bias term. Interestingly the distribution is similar to the one obtained from the inversions with the bias term (cases ~~S1~~BVR, ~~S1a~~BVN, and ~~S1b~~BVRT). This shows that inversions assuming correlations with a strong contribution from the far field have similar characteristics as inversions that assume a flat bias term.

Figure 3~~Figure 4~~ shows the spatially aggregated posterior flux estimates for the full domain with the corresponding uncertainties integrated at monthly and at annual temporal scales. The same prior uncertainty was used for cases ~~B1~~nBV and ~~B2~~nBB although they differ in prior flux field. Posterior estimates from ~~B1~~nBV (blue line/shading) and ~~B2~~nBB (green line/shading) inversions do not show any significant difference at monthly and annual scales despite the large difference in prior fluxes. We observe that the maximum uptake occurs slightly earlier for the ~~B2~~nBB case. Monthly fluxes from the ~~S3~~nBVH inversion also show the same temporal evolution. We do not observe any significant difference in monthly fluxes for the ~~S1~~BVR (red line/shading) and ~~S1a~~BVN (violet line/shading) inversions. Both cases are comparable to the ~~B1~~nBV and ~~B2~~nBB cases at monthly and annual scales. A slightly different picture emerges from the ~~S1b~~BVRT inversion, where the bias term allowed for more degrees of freedom for monthly corrections. The resulting seasonal cycle is somewhat smaller, with reduced summer carbon uptake. Inversions that included the bias term yielded smaller posterior uncertainties at both temporal scales, which is expected as the spatiotemporal component of the uncertainty was not inflated as was the case for the ~~B1~~nBV scenario. Flux retrievals from the reduced network (sensitivity case ~~S2~~nBV14) show a slightly deeper sink, but the differences to the base case ~~B1~~nBV are insignificant (i.e. clearly within the posterior uncertainties).

All of the inversions suggest Europe to be a carbon sink, with a range of -0.23 ± 0.13 GtC y^{-1} to -0.38 ± 0.17 GtC y^{-1} for the ~~S1b~~BVRT and ~~B1~~nBV inversions respectively. The mean annual posterior flux estimate for Europe averaged over different inversions amounts to - 0.32 GtC y^{-1}.

Posterior monthly flux estimates at smaller spatial scales (country level) are shown in Fig. 6~~Fig. 5~~. Areas that are not well constrained by the current network show some divergence in the posterior flux estimates although not significant considering the uncertainty range. For example Germany, which is better constraint, shows a limited spread of the posterior fluxes with an annually averaged standard deviation between the different posterior flux estimates being 0.~~000~~9 M~~G~~tC y^{-1}, while United Kingdom (which is less well constrained) shows a slightly larger spread of the posterior estimates with an annually averaged standard deviation of ~~0.00~~2 M~~G~~tC y^{-1}. Note that the posterior uncertainties are smaller by about 36% for the ~~S1~~BVR case, which is related to the smaller prior uncertainties at monthly time scales (see also section 3.2 in Ko16).

3.3 Validation against eddy covariance measurements

As shown in Broquet et al. (2013) and in Ko16, eddy covariance measurements in principle have the potential for quantitative evaluation of the retrieved fluxes from the inversions. Here we used posterior flux estimates from the ~~B2~~nBB inversion for evaluation against eddy covariance measurements, as the prior flux fields in ~~B2~~nBB (GBIOME-BGCv1) were not optimized using eddy covariance measurements. Gap-filled data were downloaded from the European Fluxes Database Cluster (http://www.europe-fluxdata.eu). A modified flux-site network compared to the one reported in Kountouris et al. (2015) was used. Specifically we omitted sites that they have not been used for the VPRM optimization (*CH-Fru, CH-Lae, CH-Oe1, ES-LMa, FR-Avi, FR-Mau, IT-Cas, IT-LMa, IT-Ro2, NL-Dij, NL-Lut, SE-Sk1, SK-Tat*) as well as sites that were not available as gap-filled data (*CH-Dav, ES-Agu, FR-Aur*). Further some more sites were added both for the VPRM optimization and for the flux comparisons (*CZ-wet, DK-Sor, HU-Bug, IT-Non, NL-Ca1, PL-wet, RU-Fyo, UK-PL3*). Monthly averaged fluxes were extracted, with weights for each vegetation class that compensate for the asymmetry between number of flux towers per vegetation type and the fraction of land area covered by the specific vegetation type, similar to Ko16.

The analysis of the monthly prior biospheric fluxes in ~~Fig. 6~~Fig. 7 reveals significant differences between observed and prior fluxes from the inversion. The GBIOME-BGCv1 model systematically overestimates the observed fluxes throughout the year. The retrieved fluxes from the inversion (dark green line) are closer to the observed fluxes, with a stronger uptake compared to the prior during spring and summer time. The timing of the peak uptake is shifted to one month earlier in comparison to the observations. The mean absolute bias (averaged absolute differences between prior/posterior and observed fluxes) is significantly reduced by 52% from 0.84 to 0.40 $gCm^{-2}day^{-1}$. The standard deviation of the residuals is reduced by around 24%, from 0.68 for the prior to 0.40 $gCm^{-2}day^{-1}$ for the posterior residuals. Splitting the sites into two main categories, the first only with crops, and the second with non crop sites, revealed differences on how well those sites can be represented. Clearly best matches were found for the non crop sites with a reduction in the mean absolute bias of 51% whilst for the crop sites it is limited to 38%.

4 Discussion

We performed a series of atmospheric CO_2 inversions based on atmospheric data taken from 16 European stations for 2007. Different data-driven error structures in the prior error covariance were assessed, and optimized biospheric fluxes were retrieved and post-processed at various temporal and spatial scales for further evaluation. In this part we discuss the fitting performance of the inversion system, and we detail the comparisons between our flux estimates at grid, national and continental scales against eddy covariance data and reported flux estimates from previous studies. Finally we discuss how sensitive flux retrievals are in the presence of erroneous representation of the fossil fuel fluxes, and the site selection.

4.1Goodness of fit

Site-specific misfits show a reasonable fit to the atmospheric data. Nevertheless in 4 cases (CBW, HEI, JFJ, and KAS) site-specific χ_c^2 values were found to be larger than 1 (see also Table 3), indicating that either the model-data mismatch errors were chosen too small, or the spatiotemporal resolution of the flux model is too coarse compared to the biosphere fluxes and therefore small scale variations are not resolved (Rödenbeck et al., 2003). In fact this seems to be the case for the JFJ and KAS sites as those are high altitude sites with steep cliffs. In such a complex terrain the atmospheric circulation is hard to be simulated from the transport models. Regarding CBW and in particular HEI, those are polluted sites and it would be reasonable to assume larger model-data mismatch uncertainty since the model is too coarse to resolve the fossil fuel emission patterns. One could argue that using higher spatial resolution to couple fossil fuel fluxes with transport models might reduce the model-data mismatch uncertainties, and hence improve posterior fluxes. To investigate that, we performed a forward run at coarser (0.25º) and higher (1/12º lat. X 1/8º lon.) spatial resolution using only the fossil fuel emissions. As we use a Lagrangian transport model, fluxes at higher resolution than that of the meteorological fields can be used such that the simulated fossil fuel signals contain more spatially detailed information (Lin et al., 2003). The derived concentration signal was subtracted from the observations and subsequently an atmospheric inversion was performed. We report no significant differences between the retrieved fluxes indicating that simply increasing the spatial resolution to about 10 km is not enough to correctly represent the fossil fuel distribution.

~~We note though that~~ A common approach in atmospheric inversion studies to evaluate the defined uncertainties is to ~~ex~~examin~~ing~~e the reduced χ_r^2 values. However, this ~~is~~might not always be a sufficient metric ~~to evaluate the defined uncertainties~~ (Michalak et al., 2005; Chevallier, 2007). The reduced χ_r^2 values in our study (between 1.08 and 1.25) are larger than those found by Tolk et al. (2011) where values between 0.34 and 0.78 were found for their pixel based inversion, indicating a more conservative choice for their model-data mismatch errors. Even lower values were reported in the study by Peylin et al. (2005) with values ranging from 0.01 up to 0.6 depending on the assumed correlations. χ^2 values from Zhang et al. (2015) were within a range of 1 to 4, but were modified by inflating the error covariances through an iterative procedure, resulting in χ_r^2 values comparable to ours. Concluding, the χ_r^2 values give confidence that the assumed prior uncertainties are well defined. ~~We note though that examining the~~ $\cancel{\chi_r^2}$

~~values is not always a sufficient metric to evaluate the defined uncertainties (Michalak et al., 2005; Chevallier, 2007).~~

4.2 Validation against eddy flux measurements

At the local scale the inversion shows ability to capture the observed flux variability at monthly scale, as shown for the ~~B2~~nBB case (see Fig. ~~6~~7). The residuals between posterior model and eddy covariance flux-data for monthly and site averaged fluxes show a range of misfits not exceeding 1.04 $gCm^{-2}day^{-1}$ which is comparable with Broquet et al. (2013), where misfits up to 1.5 $gCm^{-2}day^{-1}$ were found using 6 years of data (2002-2007). Of note is that the estimated carbon uptake agrees well with the estimated uptake for 2007 in Broquet et al. (2013) (within the uncertainty range). However, in contrast to the synthetic inversion of Ko16, the real data inversion showed a larger monthly averaged posterior bias equal to 0.40 $gCm^{-2}day^{-1}$ compared to the -0.04 $gCm^{-2}day^{-1}$ for the synthetic case~~ in Ko16~~. The poorer performance in terms of bias compared to the synthetic case is presumably mainly caused by the representation error. In the synthetic inversion we created a true flux field at the same spatial resolution as the posterior flux estimates, and sampled this true flux distribution at the specific eddy covariance measurement location. This does not include any spatial representation error of the EC measurements (footprint about 1 km) with respect to the spatial resolution of 25 km at which the fluxes are used within the inversion. A further cause for this poorer performance is related to the transport error, as in the synthetic case the same transport was used to create the synthetic observations and to perform the inversion, while in the real data inversions the observed atmospheric mole fraction are a result of real transport which can only be approximated with the transport model used for the inversion.

Differences between posterior flux retrievals and observed NEE fluxes at the eddy covariance stations are clearly driven ~~from~~ by the crop sites. The good agreement between posterior inverse flux estimates and fluxes measured with the eddy covariance technique at non-crop sites can be attributed to the relatively stable, within the year, land condition. Contrastingly, crop areas are subject to human activities throughout the year. Soil enrichment with organic fertilizers, irrigation and harvesting, can severely influence the carbon balance of the local ecosystem. Thus the poor performance between inverse estimates and eddy covariance flux measurements at crop sites can be linked to the extensive anthropogenic influence on those areas. Further, another difficulty which is common for all the ecosystems, ~~it is worth to mention~~ is the fact that atmospheric concentrations implicitly contain more components than just the NEE signal e.g. fire emissions. Such emissions are captured in the atmospheric observations (representative scale of hundreds of km) but might not be captured from the eddy covariance flux measurements which they have a very short representative scale of around 1 km.

Posterior fluxes showed a shift by one month earlier (in May), for the maximum carbon uptake (see also fig. 7). An initial hypothesis that this might be driven from sites which are difficult to simulate, such as those located in mountain regions, can not be justified. In specific, mountain sites were excluded in an additional sensitivity analyses, but the temporal shift remains. However, looking into the error of the difference between two months suggests that the flux difference between May and June is not significant. The error of the difference was calculated using a Monte Carlo experiment. Fluxes were averaged over the stations and the monthly differences were calculated. Then we used the standard deviation of the differences over the ensemble members to describe the month-to-month uncertainty.

4.3 Reliability of European flux estimates

4.3.1 Mismatch in bottom-up and top-down methods

Of note is the strong flux correction when using a-priori fluxes from VPRM with an uptake of 0.96 GtC y^{-1} compared to the 0.3 GtC y^{-1} after the inversion. The large correction of about 0.66 GtC y^{-1} corresponds to roughly twice the prior uncertainty. We note that VPRM is a diagnostic model which uses simple light use efficiency and respiration equations and MODIS indices, with parameters optimized to match hourly observations of NEE fluxes (Mahadevan et al., 2008). It does not account for land management and land use changes (i.e. crop harvest, deforestation), thus it will estimate a strong sink even for lands that have been harvested, with the respiration fluxes resulting from the use of the harvest (e.g. as food) not included. Those so-called lateral carbon fluxes, that are seen by the atmospheric inversion, account for approximately 0.165 GtC y^{-1}) of the prior-posterior flux difference (Ciais et al., 2008). The rest of the difference of about 0.5 GtC y^{-1} might be related to local characteristics of eddy covariance sites, which VPRM is not able to represent. Spatial variations of NEE from VPRM are driven by those of EVI (Enhanced Vegetation Index), which is used at a spatial resolution of 1 km. For example, a crop field with typical dimensions of 100 m – 200 m surrounded by other fields with different crop rotation (and differing phenology) are hard to represent with 1 km resolution EVI (even with the highest possible resolution of 250 m for MODIS reflectances). To quantitatively assess the impact of this representation error in combination with the selection of sites used for the VPRM optimization, the annual domain wide C-budget from VPRM was recalculated after omitting one site per vegetation type at a time and optimizing the VPRM parameters (Jackknife delete-1 method). Detailed results are shown in Table 4. The derived Jackknife standard error amounted to 0.54 GtC y^{-1}, with a dominant contribution from the cropland vegetation class (0.50 GtC y^{-1}). This uncertainty can fully explain the mismatch between the a priori and the posterior fluxes, and it emphasizes the importance of site selection and site representativeness in up-scaling local eddy covariance measurements to larger regions.

The estimated uncertainty for VPRM fluxes based on jackknifing is larger than the prior uncertainties assumed for the atmospheric inversions. Hence, one could argue that the prior fluxes using VPRM (which indicate a too strong sink) combined with a too small prior uncertainty in the inversion leads to erroneous posterior flux estimates. However the optimized biogenic fluxes from all inversions converge at the annual and domain-integrated scale. A particular example is that of the ~~B2~~nBB inversion. Even though the GBIOME-BGCv1 fluxes differ greatly from those produced by VPRM, this inversion is fully in line with the results from the rest of the inversions, indicating that the optimized flux estimates are not biased by the a priori flux fields but instead are driven by the atmospheric data.

4.3.2 Sensitivity to anthropogenic emissions

Another source of biospheric flux misrepresentation is the fossil fuel inventories. As mentioned in section 2.3 we do not allow for corrections in anthropogenic emissions, as they are assumed to be better known than the terrestrial fluxes. An overestimation/underestimation in anthropogenic emissions will thus lead to a stronger/weaker biospheric sink in atmospheric inversions. The anthropogenic emissions we use are 0.32 $GtCy^{-1}$ (27%) lower for the EU-12 countries compared to those used by Rivier et al. (2010) (1.2 $GtCy^{-1}$). Peylin et al. (2011) estimates the difference between national totals for the different emission inventories to be around 10%. In a study by Ciais et al. (2009) uncertainties of total fossil-fuel CO_2 emissions in the European Union 25 member states were estimated to 19%, based on four different emission inventories. For the EU-25 countries, EDGAR emissions were found to be 12% larger than the mean of the GAINS (Greenhouse Gas and Air Pollution Interactions and Synergies), UNFCC (United Nations Framework Convention on Climate Change) and CDIAC (Carbon Dioxide Information Analysis Center) inventories (Ciais et al. 2009, table 2). Sensitivity tests with increased prior fossil fuel emissions showed that the added fossil fuel increases the estimated uptake by almost 50% relative to the added anthropogenic emissions. Taking an extreme scenario where the fossil fuel emissions are increased by 17% or 0.3 GtC y^{-1} (resulting in 1.77 GtC y^{-1} compared to 1.47 GtC y^{-1} total emissions for EU-domain), we estimate a European carbon sink for the ~~B1~~nBV set up of -0.51 ± 0.17 $GtCy^{-1}$ compared to -0.38 ± 0.17 GtC y^{-1} for the standard ~~B1~~nBV case. Thus the additional assumed fossil fuel emissions increased the estimated uptake by 0.13 $GtCy^{-1}$, which is about 44% of the added anthropogenic emissions. The fact that the resulting increase in the biospheric sink does not fully correspond to the increase in assumed emissions is likely a result of the sparse network, where emissions from regions further away from the measurement sites are not fully registered in the simulated mole fractions.

In this study we assumed that anthropogenic emissions are perfectly known (which is a traditional assumption in atmospheric inversions), although this is not the case. As a result of not allowing for a correction in the fossil fuel component, this correction will be added to the correction of the biogenic signal. In this paragraph we already discussed how uncertain fossil fuel emissions may be. Further, we estimated how the uncertainty in the fossil fuel component impacts, the carbon flux estimates; the magnitude but also spatial and temporal flux distributions may be significantly erroneous. For better future carbon flux estimations, fossil fuel optimization seems to be necessary. However, that would require 14C tracer measurements which are currently not available.

4.3.3 Sensitivity to site selection

Uncertainties in vertical mixing and especially in the nocturnal boundary layer (Gerbig et al., 2008) should be carefully addressed as they might lead to erroneous estimations of the carbon uptake. Typically, in atmospheric inversions the model-data mismatch error (measurement error covariance) accounts also for uncertainties due to the transport (i.e. wrong representation of the nocturnal boundary layer). The set of network stations includes 7~~6~~ mountain stations, for which we use night-time observations (day-time for non mountain stations) as these measurements are considered to be representative for the free troposphere. Errors can be introduced if the measurement height assumed in the transport model is within the modeled nocturnal stable boundary layer while in the real world it is not, which would lead to an overestimation in the simulated CO_2 signals from respiration or vice versa. In the inversion this would be compensated by introducing stronger uptake fluxes to match the observed CO_2 time series. In order to investigate whether our results are influenced by the use of mountain stations, we performed an additional inversion using the ~~B1~~nBV error structure, but excluding all these stations. The resulting sink in Europe was found to be -0.41 ± 0.17 $GtCy^{-1}$ which is fully in line with ~~B1~~nBV inversion using all sites, suggesting that our estimates are not biased due to misrepresentation of the mountain stations at least at annual and domain wide aggregation scales.

However, the spatial flux distribution seems to depend on the site selection and in particular on the mountain sites used in a given inversion. Ambiguous carbon fluxes e.g. carbon sinks over non productive areas like Alps, England, and west Chech Republic, as well as carbon sources over cultivated lands like western France, Poland and Ukraine were derived from the inversions (fig. 3). Figure 4 presents the annual spatial flux distribution by using a network of stations with no mountain sites (MS0 case) and using an error structure which does not contain a bias term. This sensitivity test is equivalent to the nBB case where we used also the GBIOME-BGC model as prior. Subsequently we plot the flux distribution by adding one mountain site at a time (cases 1:7 where the number denotes how many mountain sites are being used). The add-one mountain site sequence is as follows: CMN, OXK, PTR, JFJ, KAS, SIL, PUY. For the MS0 case, we observe that in the region around the Alps, and the neighboring countries, the sink is smaller compared to the rest inversions. The OXK and the KAS sites seem to be responsible for the sink over the Czech Republic. The KAS site seems also to be the driver for the high carbon flux sources around Poland, Ukraine and the Black Sea coasts.

Using an error structure which allows for a bias term as the one in BVR case, seems to moderate the spatial flux misrepresentation. Comparing in fig. 3 the subplots nBV: without bias term, BVR: with bias term, we see that the abovementioned highly productive regions (according to the simulation), show somewhat weaker sinks for the BVR case compared to the nBV (indicated by the less bluish contours). Subsequently, regions that appear to be strong carbon sources (in nBV case), show weaker flux signal when the bias term is used (BVR).

Although this study uses as much information as possible, in terms of the available atmospheric observations still, large areas are poorly or not constrained at all from the atmospheric network e.g. West France, the whole East European part. Hence, the spatial flux distribution at those areas, is prone to large uncertainties.

4.3.4 Retrieved fluxes and comparison to previous inverse estimates

The retrieved spatially resolved fluxes showed a sensitivity in their spatial patterns to the a priori error structure, specifically to the inclusion of a bias component, as indicated by differences between the ~~B1~~nBV and ~~S1~~BVR cases. Such differences were not identified in the synthetic experiment in Ko16, however there a much larger spatial correlation length scale was assumed. In the synthetic inversions the long correlation length (766 km at the zonal and 411 km at the meridional direction) drastically reduces the effective number of degrees of freedom, forcing the fluxes to be smoothly corrected, regardless of the use of the bias component. In the real data inversions the shorter correlation length (around 100 km), combined with the required larger error inflation (compared to the synthetic inversions) for the ~~B1~~nBV and ~~B2~~nBB cases, increases the effective number of degrees of freedom. By using a bias component (~~S1~~BVR, ~~S1a~~BVN, ~~S1b~~BVRT cases) or by using the hyperbolic correlation shape (~~S3~~nBVH) with stronger large-scale correlation, instead of inflating the spatiotemporal error component, fluxes remain less flexible at gridscale.

Our knowledge regarding annual CO_2 flux estimates for Europe is still highly uncertain, in part due to the limited number of regional inversions focusing on this domain. Flux estimates from previous studies, mainly global inversions, show a wide range (Fig. 8~~Fig. 7.~~). We estimated an annual European carbon sink (ranging between -0.23 ± 0.13 and -0.38 ± 0.17 GtC y^{-1} for the different inversion scenarios, Fig. 5 d)), which is however representative for a smaller European region compared to the TransCom European region typically used in other studies. The up-scaled flux estimates (see also section 2.4) for the TransCom EU region have a range of -0.39~~0~~ to -0.71~~49~~ GtC y^{-1}. Ciais et al. (2000) estimated a European sink of -0.3 ± 0.8 GtC y^{-1} for the target period 1985-1995, however in contrast to our study they used a global system and a gap filling algorithm since 42% of the observational data were missing. A recent study from Peylin et al. (2013) computed the mean European sink for the period 1998-2001 to be -0.44 ± 0.45 GtC y^{-1} by utilizing eleven different global inversion systems. Gurney et al. (2004) performed also global inversions and found the mean European annual fluxes for 1992 – 1996 period to be -0.98 ± 0.4 GtC y^{-1} which is larger compared to our estimations. Moreover, our results for the mean net monthly fluxes over Europe agreed very well with Rivier et al. (2010) who estimated for the 1998-2001 time frame using five different transport models in their inversion that the maximum seasonal uptake occurs in July and lies between -10 and -80 $gCm^{-2}month^{-1}$, while our results show maximum uptake in June with a range of -33 to -37 $gCm^{-2}month^{-1}$ for the different inversion cases. We note that the annual flux differences between our flux estimates and those from other studies may be also caused due to the interannual flux variability. Nevertheless this should not be expected to critically drive those differences since posterior uncertainties found to be larger than interannual variations (Broquet et al., 2013) making the significance of the variations questionable.

A recent study from Reuter et al. (2014) based on inversions using satellite observations estimated the carbon budget for the TransCom European region. For the year 2007 the sink was found to be -1.1 ± 0.30 GtC y^{-1}, much larger compared to most of other inversion estimates using ground observations. However Feng et al. (2016) tried to investigate why atmospheric inversions using satellite observations, show an elevated European uptake, through a series of sensitivity tests. They linked the increased uptake when using satellite measurements to potential observation biases and to the emission spatial patterns. Further Feng et al. (2016) highlighted that the large European uptake is related up to 60-90 % from systematically higher modeled CO_2 fluxes transported into Europe from regions outside of the domain. As this looks to be a problem related with column measurements this is not the case in our study since ground observations were used. In addition we use the two step inversion scheme which limits the influence from the far field as we calculate the concentration signal from outside the domain and subtract that from the observations. Whilst the flux uncertainties outside the domain are not propagated, still they can be expressed as uncertainties in the observation space. However if biases introduced from the global inversion to the fluxes outside of the domain, then regional flux estimations may differ.

At national scale we can compare our results to those obtained by Meesters et al. (2012) for the Netherlands, who estimated the annual national carbon sink to about -0.017 ± 0.004 $GtCy^{-1}$. Our estimations are very close, with a range of – 0.012 ± 0.004 $GtCy^{-1}$ (~~S1~~BVR inversion) to – 0.014 ± 0.005 for the ~~B2~~nBB inversion. Of note is that the carbon budget estimates for Netherlands agree remarkably well despite the substantial differences between the two studies: Meesters et al. (2012) used an inversion scheme that solves for scaling factors of the gross prior fluxes. Spatial correlations of 100 km were assumed but only for photosynthetic fluxes within the same land use class. In addition the domain of interest (Netherlands) has a stronger constraint as four stations located within the domain were used, while our inversion only uses one station (CBW), with the rest of the stations being at least 360 km away (WES). Both studies assume approximately the same fossil fuel emissions (0.051 GtC y^{-1} vs. 0.053 GtC y^{-1} in Meesters et al. (2012)).

5 Conclusions

This study is a follow up work from Kountouris et al. (2016). In this second part, the inverse modeling framework was deployed using real atmospheric data from 16 stations in Europe, to infer biospheric carbon fluxes. Different prior error structures were assumed to investigate how sensitive posterior fluxes are. The results are validated and compared at different temporal and spatial scales. Satisfactory agreement was found when posterior inverse flux estimates were compared against eddy covariance observations at local scale, as well as against previous studies at national and continental scales, which gives us confidence for our carbon flux estimations. We calculated a sink for the European continent which amounts of -0.23 ± 0.13 GtC y^{-1} to -0.38 ± 0.17 GtC y^{-1} depending on the assumed prior error structure.

A special effort was also made to avoid potential biased flux estimations due to site selection (i.e. heavily polluted sites, or sites that are within the nocturnal boundary layer e.g. mountain stations) by performing inversions using different network configurations. We did not observe any significant impact for domain-wide aggregated fluxes at least for monthly and annual scales. However changes in spatial flux patterns at the pixel scale should be expected, when then network configuration is changed. Further we studied also how sensitive biospheric carbon fluxes are, when wrong fossil fuel emissions are assumed. We found that due to the network sparseness the fossil fuel emissions are not fully captured in the simulated mole fractions which may bias the flux estimates.

What do we learn or should we expect then from the top down approach? The current analysis part one and two, suggests that aggregated fluxes at monthly (temporally) and country (spatially) scales can be successfully retrieved from the inversion system. However, retrieving spatially resolved fluxes at finer scales is still rather challenging. Lack of observations for extended European regions, complexity of the terrain especially in mountainous regions as well as the absence of fossil fuel measurements which would otherwise, allow the separation of fossil fuel signals from biospheric signals in observed CO_2 time-series, complete the mosaic of the current problems that regional inversions are facing. Whilst ICOS (Integrated Carbon Observing System) will introduce more stations in the European continent still, inversions should use all the available information; that could be achieved by assimilating multiple data streams like continuous and flask measurements in combination with satellite derived information, aiming to constrain as tight as possible the European continent. Further, new stations should also aim in measuring combustion tracers. That would be of a great help in future inversion systems to be able to update the anthropogenic emission maps and subsequently to compute more accurately the biogenic signal.

Acknowledgments

This work contributed to the European Community's Seventh Framework Program (FP7) project ICOS-INWIRE, funded under grant agreement no. 313169. The authors would also like to thank the Deutsches Klimarechenzentrum (DKRZ) for using the high performance computing facilities. The authors are also grateful to Max – Planck – Gesellschaft (MPG) for funding the article fees and the high performance computing. The authors would like to thank all station PI's for providing the data and in specific: A.T. Vermeulen, E. Dlugokencky, M. Galli, J.A. Morgui, J. Necki, J.V. Lavric, M. Leuenberger, M. Ramonet, L. Haszpra, F. Meinahardt, M. Schumacher and S. Hammer. This work used atmospheric CO2 dataacquired by the following sites, sorted by project/funding agency: JFJ, HUN were funded by CarboEurope IP (GOCE-CT-2003-505572) and IMECC (026188-I3), PRS is funded under Contract Agreement between RSE and the Ministry of Economic Development-General Directorate for Nuclear Energy, Renewable Energy and Energy Efficiency, CBW is funded by CarboEurope IP, CHIOTTO, and Min. of Environment NLand BSIK ME02, OXK was funded by MPG, BIK was funded by MPG, CHIOTTO (EVK2-CT-2002-00163) and CarboEurope IP (GOCE-CT-2003-505572), HEI was funded by CarboEurope IP (GOCE-CT-2003-505572), MHD and PUY were coordinated by LSCE (CEA/CNRS/UVSQ) as part of the SNO-RAMCES/ICOS monitoring network. This publication is an outcome of the International Space Science Institute (ISSI) Working Group on "Carbon Cycle Data Assimilation: How to consistently assimilate multiple data streams.

References

Alemanno, M., Di Diodato, A., Lauria, L. and Santobuono, N.: Environmental measurements at Monte Cimone GAW station, Int. J. Global Warming, Vol. 6, No. 4, 424–454, doi: 10.1504/IJGW.2014.066048 2014.

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO_2 fluxes estimated from GOSAT retrievals of total column CO_2, Atmos. Chem. Phys., 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.

BP (British Petroleum): Statistical Review of World Energy 2012: http: //www.bp.com/statisticalreview, last access: December 2013, 2012.

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

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

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

Carouge, C., Peylin, P., Rayner, P. J., Bousquet, P., Chevallier, F., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 2: Sensitivity of flux accuracy to inverse setup, Atmos. Chem. Phys., 10, 3119-3129, doi:10.5194/acp-10-3119-2010, 2010b.

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

Chevallier, F., Palmer, P. I., Feng, L., Bösch, H., O'Dell, C., and Bousquet, P.: Towards robust and consistent regional CO_2 flux estimates from in situ and space-borne measurements of atmospheric CO_2, Geophys. Res. Lett., 41, 1065–1070, doi:10.1002/2013GL058772, 2014.

Ciais, P., Peylin, P. and Bousquet, P.:, Regional biospheric carbon fluxes as inferred from atmospheric CO_2 measurements, Ecol Appl, 10, 1574-1589, doi: 10.2307/2641225, 2000.

Ciais, P., Borges, A. V., Abril, G., Meybeck, M., Folberth, G., Hausglustaine, D., and Janssens, I. A.: The impact of lateral carbon fluxes on the European carbon balance, Biogeosciences, 5, 1259-1271, doi: 10.5194/bg-5-1259-2008, 2008.

Ciais, P., Paris, J. D., Marland, G., Peylin, P., Piao, S. L., Levins, I., Pregger, T., Scholz, Y., Friedrich, R., Rivier, L., Houwelling, S., Schulze, E. D., and members of the CARBOEUROPE SYNTHESIS TEAM: The European carbon balance. Part1: fossil fuel emissions, Glob Change Biol, 16, 1395-1408, doi: 10.1111/j.1365-2486.2009.02098.x, 2009.

Conway, T. J., Tans, P. P., Waterman, L. S., Thoning, K. W., Kitzis, D. R., Masarie, K. A., and Zhang, N.: Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, J. Geophys. Res., 99(D11), 22831–22855, doi: 10.1029/94JD01951, 1994.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N. and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc., 137: 553–597. doi: 10.1002/qj.828, 2011.

Deng, F., Jones, D. B. A., Henze, D. K., Bousserez, N., Bowman, K. W., Fisher, J. B., Nassar, R., O'Dell, C., Wunch, D., Wennberg, P. O., Kort, E. A., Wofsy, S. C., Blumenstock, T., Deutscher, N. M., Griffith, D. W. T., Hase, F., Heikkinen, P., Sherlock, V., Strong, K., Sussmann, R., and Warneke, T.: Inferring regional sources and sinks of atmospheric CO_2 from GOSAT XCO_2 data, Atmos. Chem. Phys., 14, 3703–3727, doi:10.5194/acp-14-3703-2014, 2014.

Dlugokencky, E.J., Lang, P.M., Masarie, K.A., Crotwell, A.M., and Crotwell, M. J.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2014, Version: 2015-08-03, Path: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/.

Enting, I. G. and Mansbridge, J. V.: Seasonal sources and sinks of atmospheric CO2: Direct inversion of filtered data, Tellus 41B, 111–126, doi: 10.3402/tellusb.v41i2.15056, 1989.

Fan, S., Gloor, M., Mahlman, J., Pacala, S., Sarmiento, J., Takahashi, T., and Tans, P.: A Large Terrestrial Carbon Sink in North America Implied by Atmospheric and Oceanic Carbon Dioxide Data and Models, Science, 282, 442–446, doi : 10.1126/science.282.5388.442, 1998.

Feng, L., Palmer, P. I., Parker, R. J., Deutscher, N. M., Feist, D. G., Kivi, R., Morino, I., and Sussmann, R.: Estimates of European uptake of CO_2 inferred from GOSAT XCO_2 retrievals: sensitivity to measurement bias inside and outside Europe, Atmos. Chem. Phys., 16, 1289-1302, doi: 10.5194/acp-16-1289-2016, 2016.

Ferrarese S., Apadula F., Bertiglia F., Cassardo C., Ferrero A., Fialdini L., Francone C., Heltai D., Lanza A., Longhetto A., Manfrin M., Richiardone R., Vannini C.:. Inspection of high–concentration CO_2 events at the Plateau Rosa Alpine station, Atmos. Pollution Res., 6, 3, 415-427, doi:10.5094/APR.2015.046, 2015.

Fiedler, V., Dal Maso, M., Boy, M., Aufmhoff, H., Hoffmann, J., Schuck, T., Birmili, W., Hanke, M., Uecker, J., Arnold, F., and Kulmala, M.: The contribution of sulphuric acid to atmospheric particle formation and growth: a comparison between boundary layers in Northern and Central Europe, Atmos. Chem. Phys., 5, 1773-1785, doi:10.5194/acp-5-1773-2005, 2005.

Gerbig, C., Körner. S. and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmos. Chem. Phys., 8, 591-602, doi:10.5194/acp-8-591-2008, 2008.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Kowalczyk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: TransCom and CO_2 inversion intercomparison 1. Annual and mean control results and sensitivity to transport and prior flux information, Tellus 55B , 555-579, doi: 10.1034/j.1600-0889.2003.00049.x, 2003.

Gurney, K. R., Rachel M. L., Denning, A. S., Rayner, P. J., Bernard C. P., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyotov, S., Peylin, P., Prather, M. and Taguchi, S.: Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, Global Biogeochem. Cy., 18, GB1010, doi:10.1029/2003GB002111, 2004.

Hammer, S., Glatzel-Mattheier, H., Müller, L., Sabasch, M., Schmidt, M., Schmitt, S., Schönherr, C., Vogel, F., Worthy, D. E., and Levin, I.: A gas chromatographic system for high-precision quasi-continuous atmospheric measurements of CO2, CH4, N2O, SF6, CO and H2, available at: "http://www.iup.uni-heidelberg.de/institut/forschung/groups/kk/GC_html" (last access: 25 January 2016), 2008.

Haszpra, L., Barcza, Z., Bakwin, P.S., Berger, B.W., Davis, K.J., Weidinger, T.: Measuring system for the long-term monitoring of biosphere/atmosphere exchange of carbon dioxide. J. Geophys Res, 106D, 3057-3069, DOI: 10.1029/2000JD900600, 2001.

Heimann, M. and Körner, S.: The global atmospheric tracer model TM3, Tech. Rep. 5, MPI BGC, Jena (Germany), online available at: http://www.bgc-jena.mpg.de/mpg/websiteBiogeochemie/Publikationen/Technical Reports/tech report5.pdf, 2003.

Houweling, S., Aben, I., Breon, F.-M., Chevallier, F., Deutscher, N., Engelen, R., Gerbig, C., Griffith, D., Hungershoefer, K., Macatangay, R., Marshall, J., Notholt, J., Peters, W., and Serrar, S.: The importance of transport model uncertainties for the estimation of CO2 sources and sinks using satellite measurements, Atmos. Chem. Phys., 10, 9981-9992, doi:10.5194/acp-10-9981-2010, 2010.

Kountouris, P., Gerbig, C., Totsche, K. U., Dolman, A. J., Meesters, A. G. C. A., Broquet, G., Maignan, F., Gioli, B., Montagnani, L., Helfter, C.: An objective prior error quantification for regional atmospheric inverse applications, Biogeosciences, 12, 7403-7421, doi: 10.5194/bg-12-7403-2015, 2015.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, F. Th., Heimann, M.: Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part1: Methodology and system evaluation, submitted in Atmos. Chem. Phys., 2016.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res., 108, 4493, doi: 10.1029/2002JD003161, 2003.

Lin, J. C., and C. Gerbig: Accounting for the effect of transport errors on tracer inversions, Geophys. Res. Lett., 32, L01802, doi:10.1029/2004GL021127, 2005.

Lopez, M., Schmidt, M., Ramonet, M., Bonne, J. L., Colomb, A., Kazan, V., Laj, P., and Pichon, J. M.: Three years of semicontinuous greenhouse gas measurements at the Puy de Dôme station (central France), Atmos. Meas. Tech., 8, 3941-3958, doi:10.5194/amt-8-3941-2015, 2015.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y. and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Glob. Biogeochem. Cy. 22, GB2005, doi: 10.1029/2006GB002735, 2008.

Meesters, A. G. C. A., Tolk, L. F., Peters, W., Hutjes, R. W. A., Vellinga, O. S., Elbers, J. A., Vermeulen, A. T., van der Laan, S., Neubert, R. E. M., Meijer, H. A. J., Dolman, A. J.: Inverse carbon dioxide flux estimates for the Netherlands, J. Geophys. Res.-Atmos. 117, D20306, 1984-2012 , doi: 10.1029/2012jd017797, 2012.

Michalak, A., Hirsch, A., Bruhwiller, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 100, D24107, doi:10.1029/2005JD005970, 2005.

Mikaloff F., S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., Gloor, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J. L.: Inverse estimates of the oceanic sources and sinks of natural CO_2 and the implied oceanic transport, Glob. Biogeochem. Cy., 21, GB1010, doi:10.1029/2006GB002751, 2007.

Necki, J.M.,Chmura, L., Zimnoch, M., Rozanski, K.: Impact of emissions on atmospheric composition at Kasprowy Wierch based on results of carbon monoxide and carbon dioxide monitoring, Polish J. Environm. Studies, 22, 4, 1119-1127, 2013.

Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., DI Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Stroem, J., Haszpra, L., Meijer, H. A. J., van der Laan, S., Neubert, R. E. M., Jordan, Rodo. X., Morgui, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Glob. Change Biol., 16, 1317-1337, doi: 10.1111/j.1365-2486.2009.02078.x, 2010.

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

Peylin, P., Houweling, S., Krol, M. C., Karstens, U., Rödenbeck, C., Geels, C., Vermeulen, A., Badawy, B., Aulagnier, C., Pregger, T., Delage, F., Pieterse, G., Ciais, P., and Heimann, M.: Importance of fossil fuel emission uncertainties over Europe for CO_2 modeling: model intercomparison, Atmos. Chem. Phys., 11, 6607-6622, doi:10.5194/acp-11-6607-2011, 2011.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences 10 , 6699-6720 , doi: 10.5194/bg-10-6699-2013, 2013.

Popa, M. E., Gloor, M., Manning, A. C., Jordan, A., Schultz, U., Haensel, F., Seifert, T., and Heimann, M.: Measurements of greenhouse gases and related tracers at Bialystok tall tower station in Poland, Atmos. Meas. Tech., 3, 407-427, doi:10.5194/amt-3-407-2010, 2010.

Ramonet, M., Ciais, P., Aalto, T., Aulagnier, C., Chevallier, F., Cipriano, D., Conway, T. J., Haszpra, L., Kazan, V., Meinardt, F., Paris, J. D., Schmidt, M., Simmonds, P., Xueref-Remy, I. and Necki, J. N.: A recent build-up of atmospheric CO2 over Europe. Part 1: observed signals and possible explanations. Tellus B, 62: 1–13. doi: 10.1111/j.1600-0889.2009.00442.x, 2010

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J. P., Bösch, H., Parker, R., Butz, A., Hasekamp, O., O'Dell, C. W., Yoshida, Y., Gerbig, C., Nehrkorn, T., Deutscher, N. M., Warneke, T., Notholt, J., Hase, F., Kivi, R., Sussmann, R., Machida, T., Matsueda, H., and Sawa, Y.: Satellite-inferred European carbon sink larger than expected, Atmos. Chem. Phys., 14, 13739-13753, doi:10.5194/acp-14-13739-2014, 2014.

Rivier, L., Peylin, P., Ciais, P., Gloor, M., Roedenbeck, C., Geels, C., Karstens, U., Bousquet, P., Brandt, J. and Heimann, M.: European CO_2 fluxes from atmospheric inversions using regional and global transport models, Climatic Change, 103, 93-115, doi: 10.1007/s10584-010-9908-4, 2010.

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

Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Technical Report 6, Max Planck Institute for Biogeochemistry, Jena, http://www.bgc-jena.mpg.de/mpg/websiteBiogeochemie/ Publikationen/Technical Reports/tech report6.pdf, 2005.

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

Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F. and Heimann, M.: Interannual sea–air CO2 flux variability from an observation-driven ocean mixed-layer scheme, Biogeosciences, 11(17), 4599–4613, doi:10.5194/bg-11-4599-2014-supplement, 2014.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C. and Mukai, H.: The CO2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, Atmos. Chem. Phys., 11(14), 6855–6870, doi:10.5194/acp-11-6855-2011, 2011.

Tans, P. P., Conway, T. J., and Nakazawa, T: Latitudinal distribution of the sources and sinks of atmospheric carbon dioxide derived from surface observations and an atmospheric transport model, J. Geophys. Res., 94, 5151–5172, 1989.

Thompson, R. L., Manning, A. C., Gloor, E., Schultz, U., Seifert, T., Hänsel, F., Jordan, A., and Heimann, M.: In-situ measurements of oxygen, carbon monoxide and greenhouse gases from Ochsenkopf tall tower in Germany, Atmos. Meas. Tech., 2, 573-591, doi:10.5194/amt-2-573-2009, 2009.

Tolk, L. F., Dolman, A. J., Meesters, A. G. C. A. and Peters, W.: A comparison of different inverse carbon flux estimation approaches for application on a regional domain, Atmos. Chem. Phys., 11, 10349-10365, doi: 10.5194/acp-11-10349-2011, 2011.

Trusilova, K., Rödenbeck, C., Gerbig, C., and Heinmann, M.: Technical Note: A new coupled system for global to regional downscaling of CO_2 concentration estimation, Atmos. Chem. Phys. 10, 3205-3213, doi:10.5194/acp-10-3205-2010, 2010.

Vermeulen, A. T., Hensen, A., Popa, M. E., Bulk, W. C. M., Jongejan, P. A. C.: Greenhouse gas observations from Cabauw Tall Tower (1992-2010), Atmos. Meas. Tech., 4, 617-644, doi:10.5194/amt-4-617-2011, 2011.

Zhang, S., Zheng, X., Chen, J. M., Chen, Z., Dan, B., Yi, X., Wang, L. and Wu, G.: A global carbon assimilation system using a modified ensemble Kalman filter, Geoscientific. Model Development., 8, 805-816, doi: 10.5194/gmd-8-805-2015, 2015.

1

Table 1. Information on the stations used for the regional inversions. Same network applied for the synthetic, and the real data inversions in Kountouris et al. (2016). In first column the term “type” stands for continuous (C) or flask (F) data. Under “Data origin” WDCGG means “World Data Centre for Greenhouse Gases”.

Site Code / type	Name	Latitude (°)	Longitude (°)	Height (m.a.s.l.) (m)	Measurement height (above ground) (m)	Model height	Data provider	Data origin	Citation
BAL/F	Baltic Sea, Poland	55.50	16.67	8	57	28	NOAA	Direct contact	Dlugokencky et al. 2015
BIK/C	Bialystok, Poland	53.23	23.03	183	90	90	MPI-BGC	Direct access	Popa et al. (2010)
CBW/C	Cabauw, Netherlands	51.58	4.55	-2	200	200	ECN	Direct contact	Vermeulen et al. (2011)
CMN/C	Monte Cimone, Italy	44.18	10.7	2165	12	670	IAFMS	WDCGG	Alemanno et al. (2014)
HEI/C	Heidelberg, Germany	49.42	8.67	116	30	30	University of Heidelberg	CarboEurope	Hammer et al. (2008)
HPB/F	Hohenpeissenberg, Germany	47.80	11.01	934	50	10	NOAA	Direct contact	-
HUN/C	Hegyhatsal, Hungary	46.95	16.65	248	115	96	HMS	WDCGG	Haszpra et al. (2001)
JFJ/C	Jungfraujoch, Switzerland	46.55	7.98	3572	10	720	University of Bern	CarboEurope	-
KAS/	Kasprowy	49.23	19.93	1987	5	480	UKRAK	CarboEur	Necki et

C	Wierch						, AGH	ope	al. (2013)
LMU/C	La Muela, Spain	41.36	-1.6	570	79	80	University of Barcelona	CarboEurope	-
MHD/C	Mace Head, Ireland	53.33	-9.90	25	10	15	LSCE	WDCGG	Ramonet et al. (2010)
OXK/C	Ochsenkopf, Germany	50.03	11.81	1022	163	163	MPI-BGC	CarboEurope	Thompson et al. (2009)
PRS/C	Plateau Rosa, Italy	45.93	7.71	3480	-	500	RSE	WDCGG	Ferrarese et al. (2015)
PUY/C	Puy De Dome, France	45.77	2.97	1465	10	400	LSCE	CarboEurope	Lopez et al. (2015)
SCH/C	Schauinsland, Germany	47.92	7.92	1205	-	230	UBA	WDCGG	-
WES/C	Westerland, Germany	54.93	8.32	12	-	15	UBA	WDCGG	-

Glossary for the data providers: AGH: University of science and Technology Polland, ECN: Energy research Centre of the Netherlands, HMS: Hungarian Meteorological Service, IAFMS: Italian Air Force Meteorological Service, LSCE: Le Laboratoire des Sciences du Climat et de l'Environnement, MPI-BGC: Max Planck Institute for BioGeoChemistry, NOAA: National Oceanic and Atmospheric Administration, RSE: Ricerca sul Sistema Energetico, UBA: Umweltbundesamt, UKRAK: Department of Environmental Physics Polland

Table 2. Overview of the inversion scenarios. “Shape” describes the internal structure of the bias component (proportional to respiration R or to Net Ecosystem Exchange NEE), and “Time vary” indicates whether the bias component also has temporal variations or not. The fifth column “Prior” represents the terrestrial model used as prior, and “Correlation shape” describes the functional form used for the spatial prior uncertainty correlation, either exponential (E) or hyperbolic (H). The last column indicates whether the full or the reduced station network was assumed.

Inversion code	Bias component	Shape	Time vary	Prior	Correlation shape	No. of Stations
~~B1~~nBV	-	-	-	VPRM	E	16
~~B2~~nBB	-	-	-	GBIOME	E	16
~~S1~~BVR	Yes	R	Flat	VPRM	E	16
~~S1a~~BVN	Yes	NEE	Flat	VPRM	E	16
~~S1b~~BVRT	Yes	R	Vary	VPRM	E	16
~~S2~~nBV14	-	-	-	VPRM	E	14
~~S3~~nBVH	-	-	-	VPRM	H	16

Table 3. RMSD (first column in ppm) and correlation coefficients (second column) between observations and prior/posterior CO_2 dry mole fractions for daily "daytime" or "nighttime" averaged values and for each station. The third column shows χ_c^2, the normalized dry mole fraction mismatch per degree of freedom for 7-day averaged residuals, as a measure of how well the data were fitted. The format for each station is as follows: RMSD | r^2 | χ^2.

<table>
<tr><th></th><th>Prior</th><th>~~B1~~nBV</th><th>~~B2~~nBB</th><th>~~S1~~BVR</th><th>~~S1a~~BVN</th><th>~~S1b~~BVRT</th><th>~~S2~~nBV14</th><th>~~S3~~nBVH</th></tr>
<tr><td>BAL</td><td>7.12 | 0.20 | 69.35</td><td>1.48 | 0.97 | 0.89</td><td>1.53 | 0.97 | 0.93</td><td>2.26 | 0.93 | 2.04</td><td>2.26 | 0.93 | 2.03</td><td>2.25 | 0.93 | 2.02</td><td>1.41 | 0.97 | 0.83</td><td>2.37 | 0.92 | 2.07</td></tr>
<tr><td>BIK</td><td>8.20 | 0.52 | 60.10</td><td>2.93 | 0.93 | 0.88</td><td>3.17 | 0.92 | 0.99</td><td>3.52 | 0.90 | 1.51</td><td>3.52 | 0.90 | 1.53</td><td>3.51 | 0.90 | 1.53</td><td>2.93 | 0.93 | 0.88</td><td>3.78 | 0.88 | 1.70</td></tr>
<tr><td>CBW</td><td>8.71 | 0.23 | 83.98</td><td>3.43 | 0.88 | 2.05</td><td>3.49 | 0.88 | 2.18</td><td>4.09 | 0.83 | 2.47</td><td>4.09 | 0.83 | 2.48</td><td>4.09 | 0.83 | 2.49</td><td>3.42 | 0.88 | 1.99</td><td>4.33 | 0.81 | 2.61</td></tr>
<tr><td>CMN</td><td>4.20 | 0.40 | 31.73</td><td>1.26 | 0.96 | 0.16</td><td>1.35 | 0.95 | 0.19</td><td>1.45 | 0.94 | 0.19</td><td>1.44 | 0.95 | 0.19</td><td>1.46 | 0.94 | 0.21</td><td>1.25 | 0.92 | 0.15</td><td>1.57 | 0.94 | 0.26</td></tr>
<tr><td>HEI</td><td>14.04 | 0.37 | 31.28</td><td>6.93 | 0.84 | 3.05</td><td>7.07 | 0.83 | 3.07</td><td>7.92 | 0.79 | 4.22</td><td>7.91 | 0.79 | 4.23</td><td>7.92 | 0.79 | 4.23</td><td>-</td><td>8.34 | 0.77 | 5.17</td></tr>
<tr><td>HPB</td><td>5.06 | 0.43 | 15.61</td><td>1.41 | 0.91 | 0.34</td><td>1.70 | 0.94 | 0.50</td><td>2.00 | 0.96 | 0.65</td><td>2.01 | 0.91 | 0.66</td><td>2.00 | 0.91 | 0.65</td><td>1.41 | 0.96 | 0.33</td><td>2.03 | 0.91 | 0.67</td></tr>
<tr><td>HUN</td><td>7.44 | 0.55 | 66.36</td><td>2.58 | 0.94 | 0.84</td><td>2.74 | 0.93 | 0.88</td><td>3.07 | 0.92 | 1.32</td><td>3.08 | 0.92 | 1.34</td><td>3.08 | 0.92 | 1.33</td><td>2.58 | 0.94 | 0.87</td><td>3.43 | 0.90 | 1.98</td></tr>
<tr><td>JFJ</td><td>4.52 | 0.03 | 21.39</td><td>1.96 | 0.77 | 1.59</td><td>2.23 | 0.72 | 1.53</td><td>2.07 | 0.75 | 1.83</td><td>2.07 | 0.75 | 1.82</td><td>2.07 | 0.75 | 1.84</td><td>1.95 | 0.78 | 1.58</td><td>2.10 | 0.74 | 1.98</td></tr>
<tr><td>KAS</td><td>6.35 | 0.39 | 52.58</td><td>3.41 | 0.87 | 2.90</td><td>3.43 | 0.87 | 2.89</td><td>3.88 | 0.82 | 3.96</td><td>3.88 | 0.82 | 3.99</td><td>3.87 | 0.83 |3.93</td><td>3.29 | 0.77 | 2.77</td><td>4.01 | 0.81 | 4.67</td></tr>
<tr><td>LMU</td><td>6.01 | 0.05 | 29.00</td><td>1.45 | 0.94 | 0.29</td><td>1.51 | 0.94 | 0.28</td><td>1.74 | 0.92 | 0.59</td><td>1.74 | 0.92 | 0.58</td><td>1.76 | 0.92 | 0.60</td><td>1.44 | 0.95 | 0.29</td><td>1.84 | 0.91 | 0.68</td></tr>
<tr><td>MHD</td><td>4.50 | 0.21 | 22.24</td><td>1.23 | 0.94 | 0.24</td><td>1.20 | 0.94 | 0.21</td><td>1.29 | 0.92 | 0.31</td><td>1.74 | 0.93 | 0.31</td><td>1.76 | 0.94 | 0.31</td><td>1.23 | 0.94 |</td><td>1.26 | 0.94 | 0.27</td></tr>
</table>

							0.24	
OXK	5.39 | 0.28 | 38.95	2.45 | 0.85 | 0.79	2.52 | 0.84 | 0.85	2.78 | 0.81 | 1.19	2.78 | 0.81 | 1.20	2.79 | 0.81 | 1.20	2.41 | 0.86 | 0.70	2.98 | 0.78 | 1.59
PRS	2.98 | 0.07 | 20.75	1.06 | 0.89 | 0.46	1.10 | 0.88 | 0.49	1.16 | 0.87 | 0.52	1.16 | 0.87 | 0.52	1.17 | 0.87 | 0.52	1.07 | 0.89 | 0.45	1.22 | 0.86 | 0.53
PUY	4.86 | 0.29 | 39.48	2.05 | 0.87 | 0.67	2.16 | 0.86 | 0.75	2.40 | 0.82 | 0.97	2.40 | 0.82 | 0.97	2.40 | 0.82 | 0.95	2.02 | 0.88 | 0.71	2.48 | 0.81 | 1.27
SCH	5.18 | 0.24 | 41.77	1.90 | 0.89 | 0.27	2.00 | 0.88 | 0.28	2.23 | 0.85 | 0.51	2.23 | 0.85 | 0.51	2.23 | 0.85 | 0.51	1.84 | 0.90 | 0.24	2.38 | 0.84 | 0.70
WES	8.06 | 0.23 | 41.77	2.21 | 0.94 | 0.27	2.00 | 0.94 | 0.28	2.23 | 0.91 | 0.51	2.23 | 0.91 | 0.51	2.23 | 0.91 | 0.51	-	2.38 | 0.90 | 0.70

Table 4. Results from Jackknife delete-1 statistics for VPRM estimated domain-wide NEE for different vegetation classes and for all of the land area. The uncertainty in NEE from all land area was derived assuming independence in the vegetation class specific uncertainties. Note the strong asymmetry between the fraction of land area covered by the different vegetation classes and the number of eddy covariance sites used, indicating over/under representation: for example 8 crop sites represent 51% of the land area, while 15 grassland sites represent 5.6% of the land area of Europe.

	NEE [GtC/y]	NEE uncertainty [GtC/y]	Number of sites	Fraction of land area [%]
Evergreen forest	-0.165	0.039	16	16.5
Deciduous forest	-0.174	0.020	5	4.4
Mixed forest	-0.025	0.176	2	8.4
Open shrub[a]	-0.201	-	1	13.8
Savanna[a]	-0.012	-	0	0.3
Crop	-0.443	0.502	8	51.0
Grass	0.059	0.026	15	5.6
Total	0.960	0.536	47	100

[a]Uncertainties for open shrubland and savanna could not be derived due to the lack of representative eddy covariance sites

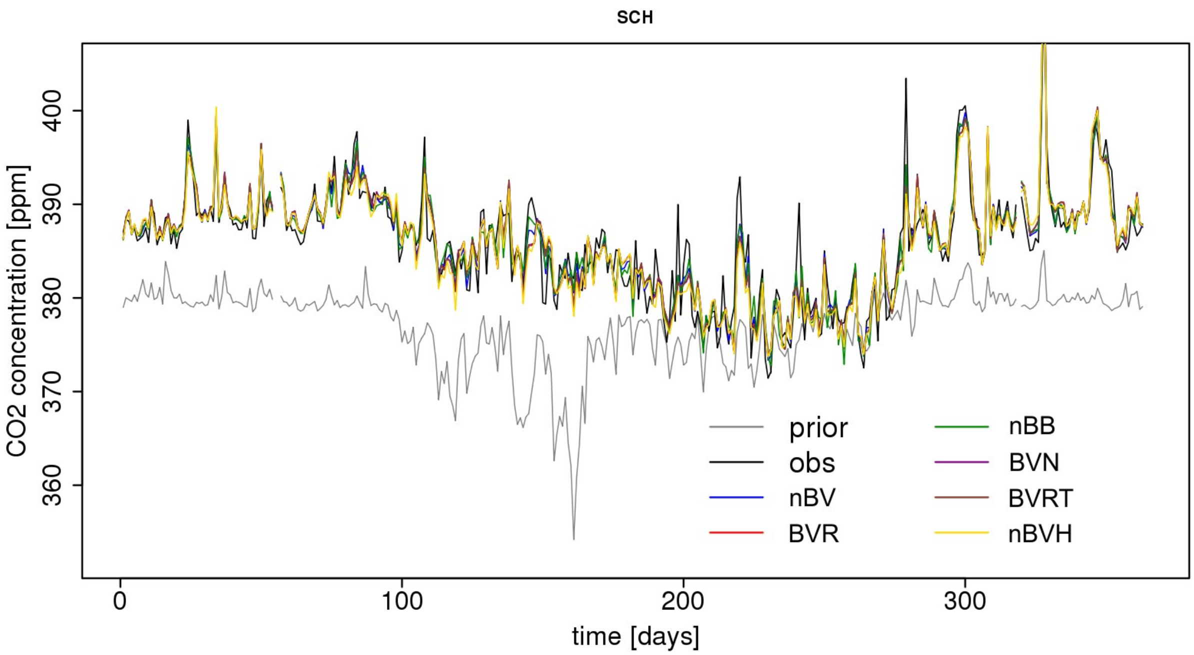

Figure 1. Daily nighttime (23:00-4:00 UTC) averages for prior, true, and posterior CO_2 dry mole fraction time series for the Schauinsland site for the real data inversion. Time starts at 1st January 2007.

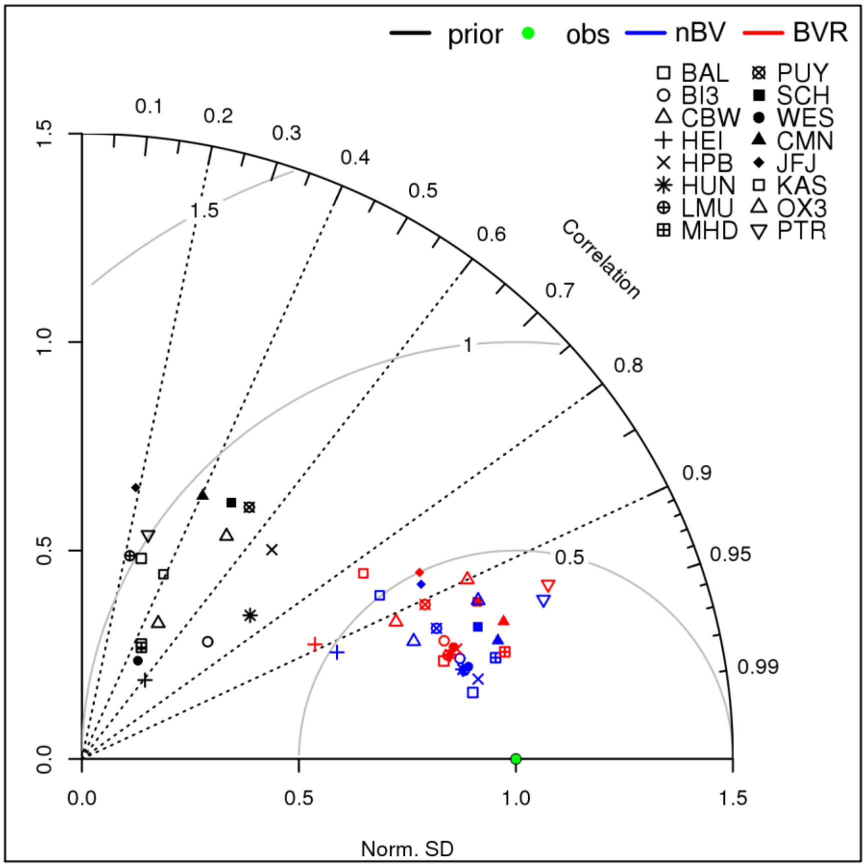

Figure 2 Taylor diagram for modeled and observed time-series of CO_2 dry mole fractions. Prior (black), observed (green, the perfect match of modeled and observed time-series) and the different inversion cases (~~B1~~nBV blue; ~~S1~~BVR red) are displayed. Different symbols denote different atmospheric stations. The normalized SD was calculated as the ration of the SD of the modeled time-series to the SD of observations. Gray semi-circles show contours of the standard deviation of the model error.

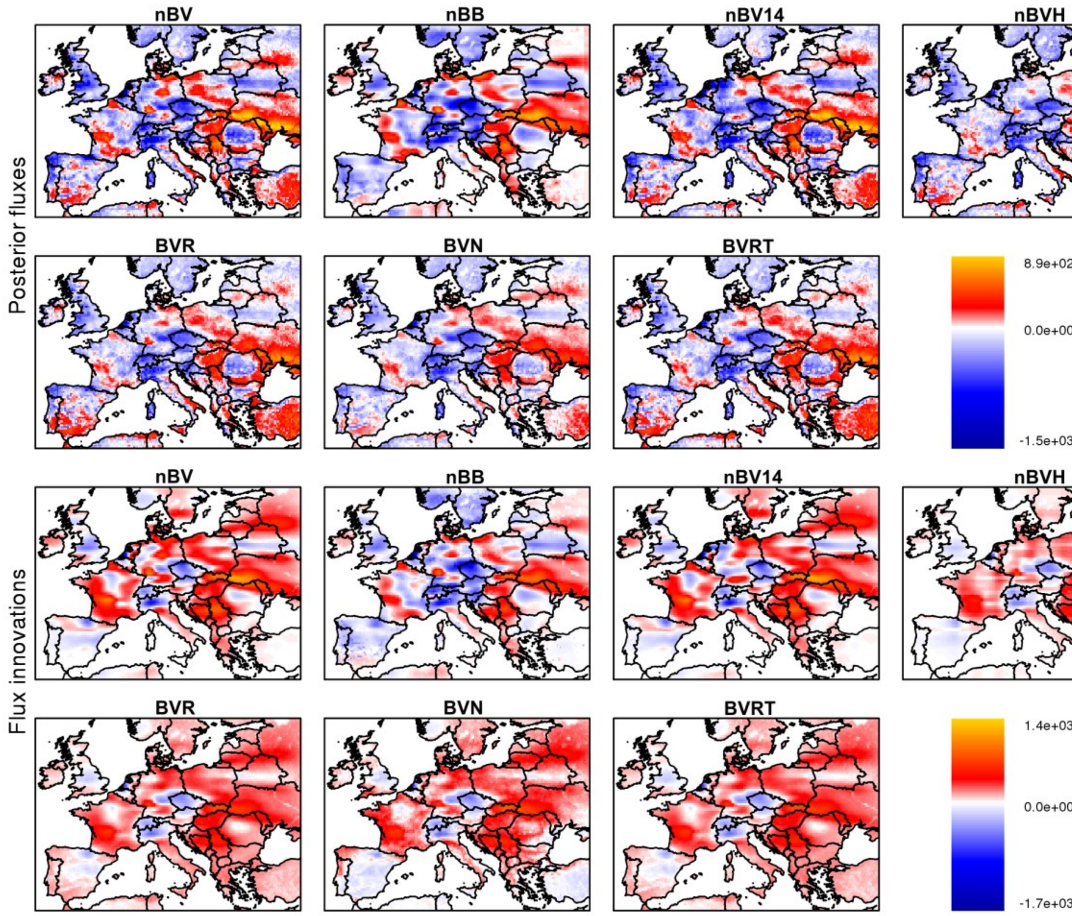

Figure 3. Annual biogenic flux spatial distribution (top two rows) and flux innovations (posterior - prior) (bottom two rows) as estimated from the different inversions for the real data case. Units are in $gCy^{-1}m^{-2}$.

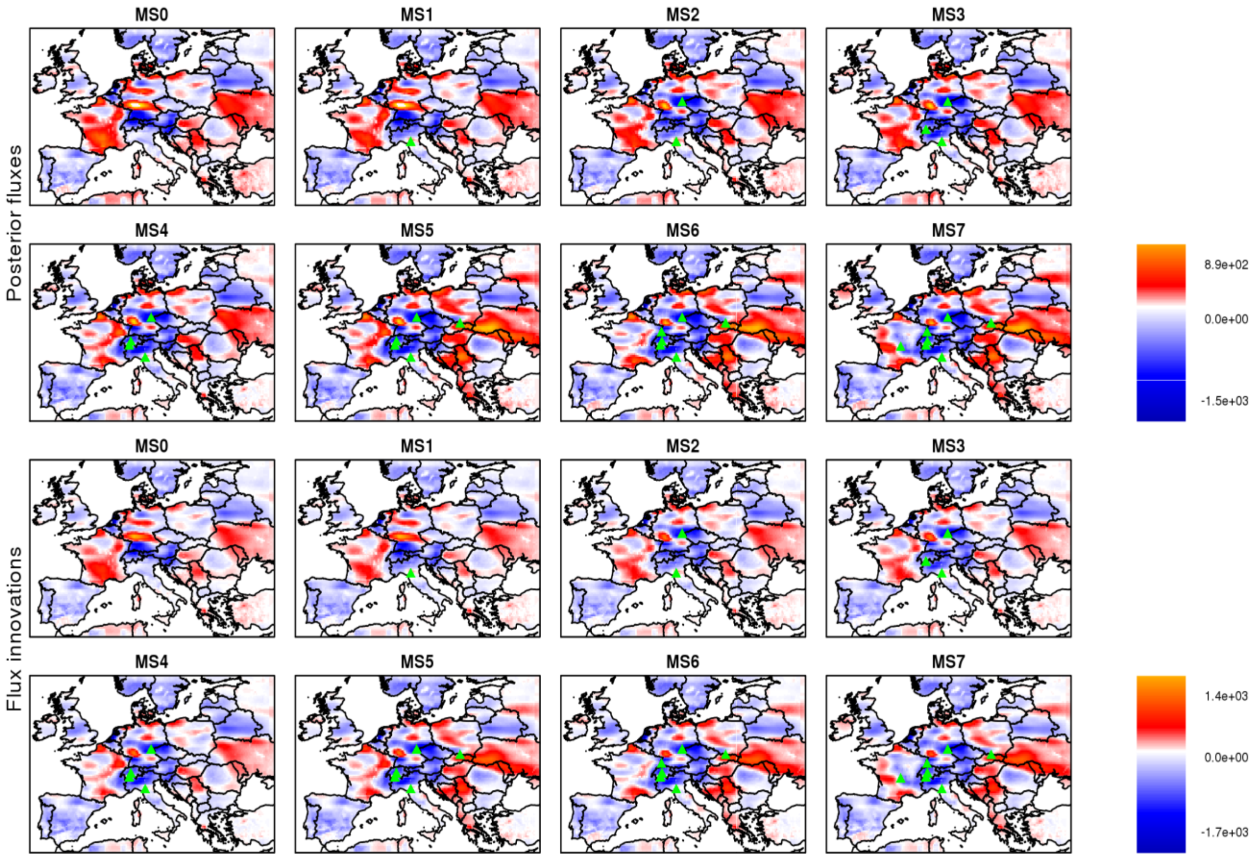

Figure 4. As figure 3, but only for the nBB inversion case. The numbers denote the number of mountain sites used in the inversions e.g. MS0: no mountain site. Units are in $gCy^{-1}m^{-2}$.

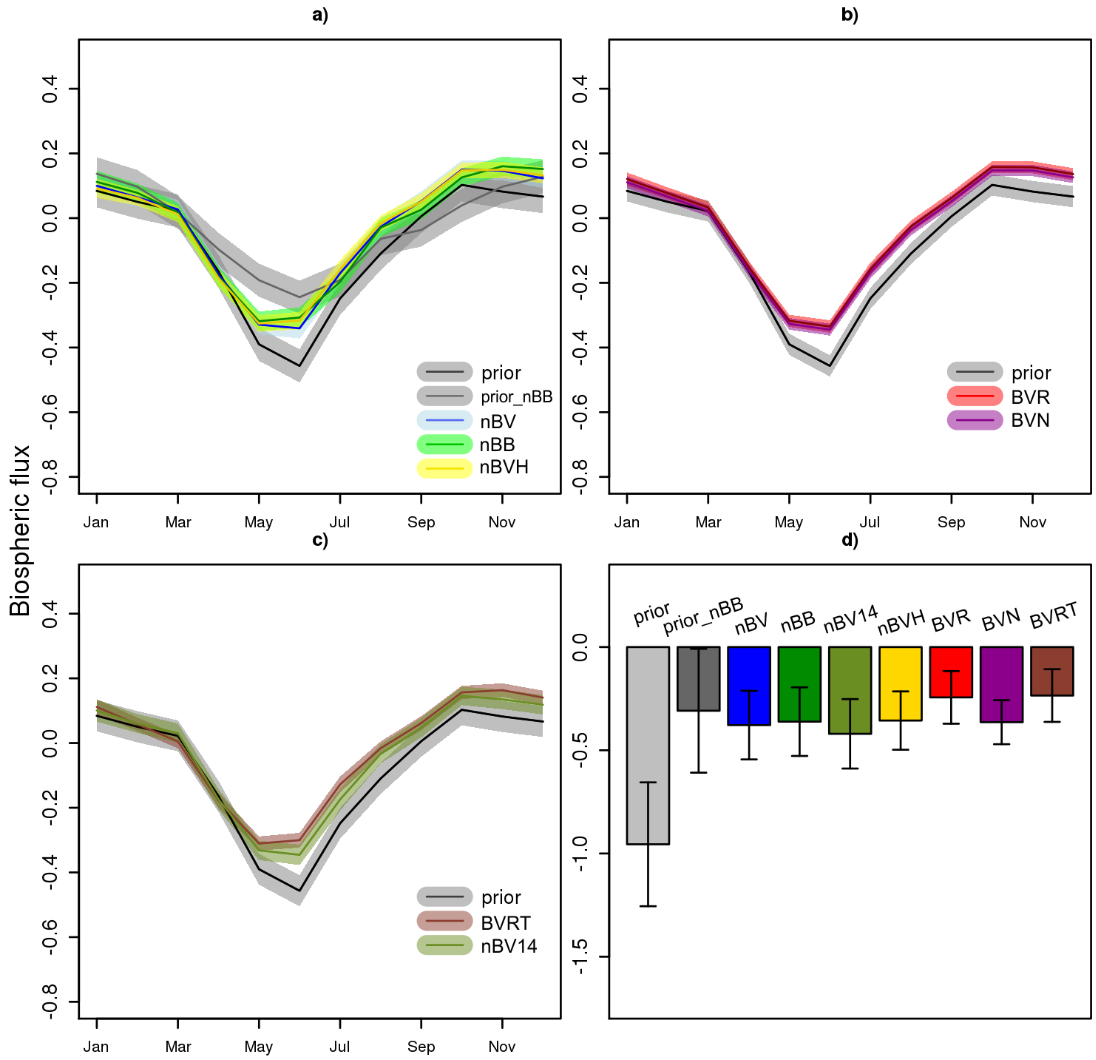

Figure ~~4~~5. Monthly and annual (panel d) biosphere fluxes integrated over the domain. Panel a) shows ~~B1~~nBV, ~~B2~~nBB and ~~S3~~nBVH cases, b) ~~S1~~BVR and ~~S1a~~BVN and the c) panel shows ~~S1b~~BVRT and ~~S2~~nBV14 cases. Note that all inversions share the same annual prior uncertainty but monthly prior uncertainties differ. Units are in GtC month^{-1} and GtC y^{-1} for monthly and annual fluxes, respectively

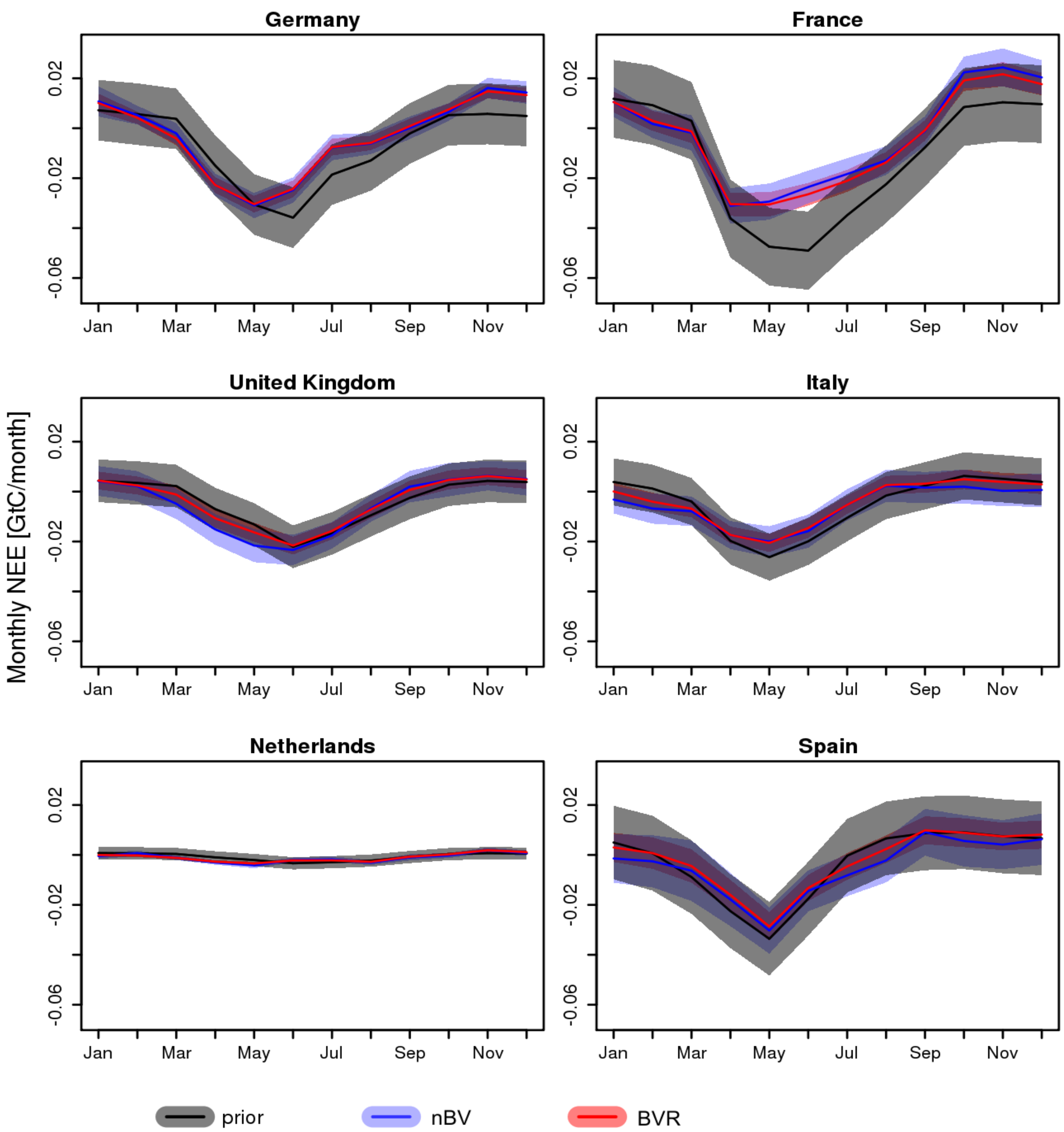

Figure ~~5~~6. Temporal evolution of prior and posterior monthly NEE for selected European countries.

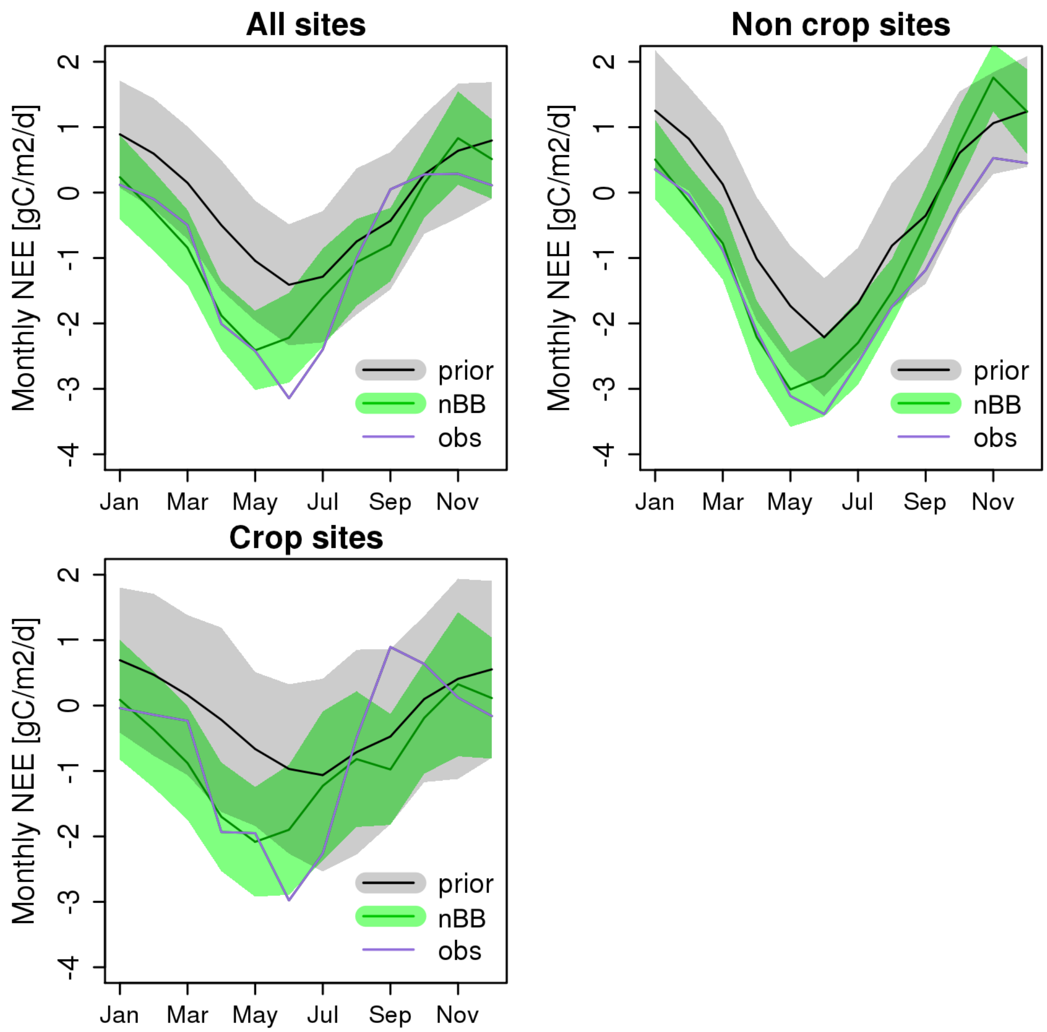

Figure ~~6~~7. Temporal evolution of monthly NEE (gCm^{-2}day^{-1}) averaged over all EC sites (top left), excluding crop (top right), and using only crop sites (bottom). Uncertainties (error of the mean monthly NEE) are indicated by the shaded areas.

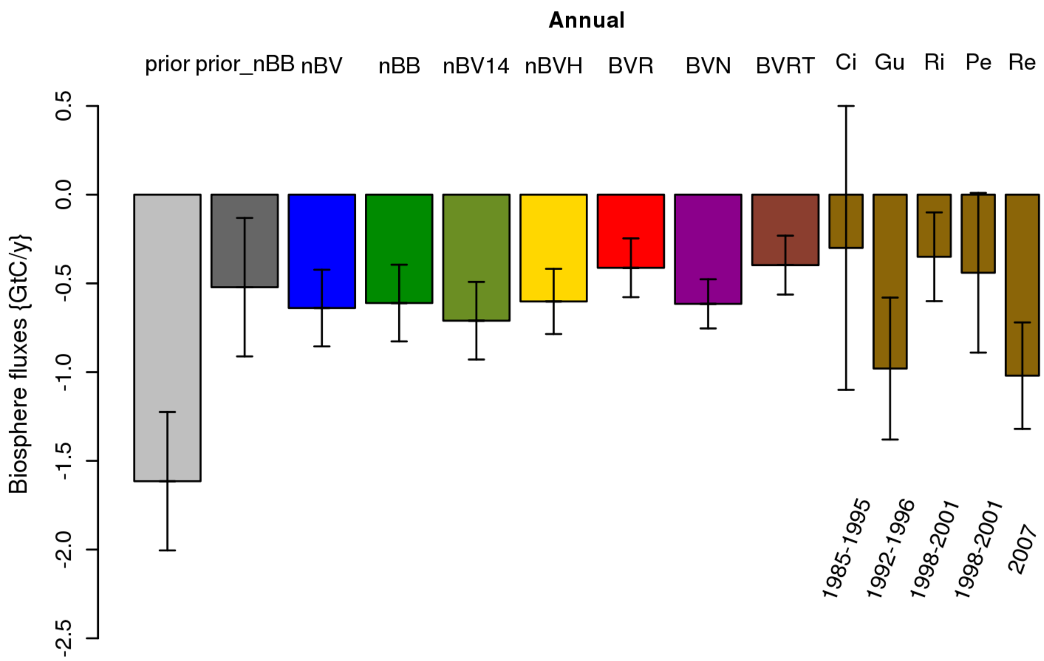

Figure ~~7~~8. Annual European biogenic CO_2 fluxes in $GtCy^{-1}$ for the different inversions and comparison to previous studies. Fluxes are upscaled to the TransCom EU domain. Labels of the references are as follows: Ci : Ciais et al. (2000); Gu : Gurney et al. (2004); Ri : Rivier et al. (2010); Pe : Peylin et al. (2013); Re : Reuter et al. (2014). Periods for the inverted fluxes are given below the flux estimates.

Point-by-point response to the reviews including relevant changes made in the manuscript:

Information on how the response is structured

We show all the comments and questions from the referees with bold text. The page and line numbers were retained, and they are linked to the first version of the manuscript. Our responses are with non-bold text whilst the page and line numbering refer to the current track-changed version of the manuscript.

Anonymous Referee #1

We sincerely thank the referee for carefully reading our manuscript and for giving such constructive comments, which helped improving our study. In the revised version, we have tried to consider all the points that were raised. We hope that our answers and the modifications are satisfactory.

The study presents an atmospheric inversion over Europe using surface tower measurements of atmospheric CO2. The methodology was described in the first part of the study, whereas the application of the method is presented in the second part. The inversion results are compared to previously published estimates for different years and to eddy-flux measurements for the same year. Overall, the inversion results show a better agreement with independent flux measurements and fall within the range of continental flux estimates. The inversion system uses an unprecedented number of observation sites and provide a fairly robust assessment of the inverse estimates with the posterior uncertainties. The one and only criticism is related to the initial objective in comparison to the final results. The regional inversion aims at improving the spatial and temporal distributions of the fluxes but I am concerned by the mismatch in the seasonal cycle of the fluxes and unexpected spatial features in the inverse fluxes. Considering the seasonality, both crop and non-crop sites start with a correct timing for the maximum uptake but are shifted by one month after inversion. Why is the inversion unable to capture the time of maximum uptake? Is there an error from the global inversion that shifts the peak? If the inversion is supporting ecosystem modeling activities, it is important to understand the potential causes of discrepancies between top-down and bottom-up approaches. A shift of one month shows that both approaches possibly disagree in the processes even if the annual and monthly estimates are closer the observed eddy-flux measurements.

The temporal shift in the maximum carbon uptake was extensively investigated. We have confirmed that the seasonal carbon flux cycle is not biased from a potential misuse of the measurement gaps. Further, we explored whether the station selection and in particular by including mountain sites in the inversion, can trigger this behavior. With respect to the former, both observed NEE and modeled fluxes share the same gaps hence, no ambiguities are expected to be introduced from this step.

Regarding the second potential issue, atmospheric transport around mountain stations is difficult to simulate. Moreover, mountain stations are prone to large errors due to misrepresentation of the nocturnal boundary layer in the model (see also response to the second comment and paragraph 4.3.3). An additional analysis was held, by performing atmospheric inversions without mountain stations, and then, by adding 1 mountain station at a time. The inversions are using GBIOME-BGC model as prior and they are equivalent to the nBB (former B2 error structure. Please see response to referee 1 regarding the new abbreviations for the different inversion cases). Fluxes at EC site locations were extracted from the spatially and temporally resolved posterior fluxes. The fluxes were averaged over all non crop sites (equivalent to right panel of fig. 7 within the paper). Their seasonal cycle seems to be insensitive to the different network configurations (see also fig.1).

An interesting part of this flux behavior is to investigate if the shift is significant. We calculated the uncertainty of the difference between two months by performing a Monte Carlo experiment. Forty ensemble members were assumed and fluxes were extracted at the EC station locations. Fluxes were averaged over all non crop sites (as we assume better representation of those sites) and we calculated the difference between month i+1 and month i. Then the error of the difference was calculated as the standard deviation over the ensemble of those differences. The uncertainty of the difference between May and June was found to be 0.35 which makes the actual difference (-2.78 for May and -3.03 for June) insignificant. Further, such a shift can also been observed in a study from Broquet et al. 2013(fig. 2 years 2003 and 2007). Later in the document, in our response to the referee 1, we highlight what we have added in the paper.

Concerning the spatial distribution, the comparison to Meesters et al. (2012) is a first step towards a more complete assessment of the inverse flux estimates at the country scale. However, looking at the maps in Figure 3, several spatial features are difficult to explain. For example, the largest (or among the largest) sinks are located over the Alps (i.e. Switzerland and northern Italy), western Czeck Republic, The Netherlands, Belgium, and England. Some of these areas, e.g. the Alps or England, are not extremely productive areas in terms of vegetation, unless recent reforestation took place there. One would expect that the most agricultural regions would represent the largest uptake of CO2. At the opposite, large sources are visible in eastern Europe (e.g. Ukraine), western Atlantic of France and Poland, where vast areas of arable land are being cultivated. Are these signals only due to the lack of observations? The argument is used for England but sites over the Alps should constrain central Europe fairly well. The comparison to existing inversions and inventories seem straightforward and would provide a better overview of the sub-continental fluxes, more than comparing annual estimates over Europe. These flux signals are the most fundamental part of the regional inversion system if one claims that higher resolution is able to improve the current inverse fluxes from global inversion systems. More thoughts and analyses should be dedicated to demonstrating the accuracy of the inversion at these scales (i.e. national scales) to confirm that the deployment of new atmospheric sites will help improve our understanding of the European carbon balance, or if regional systems are

still too uncertain to provide a clear answer to bottom-up differences. One could argue that these analyses are beyond the scope of a single inversion, which is a fair argument assuming that some errors are difficult to diagnose. Nevertheless, each inversion system should be able to produce reasonable flux distributions that provide information to bottom-up estimates. More results and comments on that matter would be beneficial to the science community and to the use of continental scale inversions in general, and possibly argue whether or not higher resolution is helpful.

We agree with the reviewer, that mountain areas like Alps and regions nearby, should be less productive compared to the optimized flux results. Flux innovations showed a high spatial flux variability which is something to be expected, since short scale correlations were assumed. Whilst such scales might seem to be ambiguous, they are fully supported by the analysis of differences between eddy flux data and a priori fluxes (Kountouris et al., 2015). The assumptions in correlations should not be unrealistic as long as each atmospheric measurement site is sufficiently well represented. The model data mismatch error should be able to account for the representation error. The model data mismatch error as implemented in the inversion system is typically 1.5 ppm for a one week period; for a full year period those temporally uncorrelated errors reduce to about 0.2 ppm, posing strong requirements on the actual representation of the individual sites in the transport model. In mountain areas were topography is quite complex, makes the atmospheric transport very difficult to be simulated within the model. Higher resolutions although they are computational expensive, would be of help and they have already proved their value (Pillai et al., 2011). Typically, in atmospheric inversions, mountain stations they assumed to measure the free troposphere. Only nighttime observations (23:00 – 04:00 local time) are considered, as this time can be better represented by the transport model. As we explain in paragraph 4.3.3., errors can be introduced if the measurement height assumed in the transport model is within the modeled nocturnal stable boundary layer while in the real world it is not. This would lead to an overestimation in the simulated CO2 signals from respiration.

In 4.3.3 paragraph we discussed the results by performing an inversion excluding the mountain stations. Domain-wide aggregated fluxes show no significant difference compared to the simulations where those sites are included. However, this might not be considered as a surprise. In fact, the inversion is constrained from the atmospheric data, which means that if fluxes from some areas are changed, then other areas should also be changed in such a way, that when running the flux field forward, this should be in agreement with the atmospheric observations. In a sense, we can consider that the aggregated fluxes will be preserved. However the spatial flux distribution seems to change and this can be justified in fig. 1 (equivalent to fig. 3 in the paper.). Figure 1 presents the annual spatial flux distribution by using a network of stations with no mountain sites (MS0 case in the subplot title). Subsequently we plot the flux distribution by adding one mountain site at a time (cases MS1:MS7 where the number denotes how many mountain sites are being used). For the MS0 case, we observe that in the region around the Alps, and the neighboring countries, the sink is smaller compared to the rest inversions i.e. Northwestern Italy, Czeck Republic. Regarding the results for Netherlands we note that we are in line with results published in Meesters et al. (2012) and we even compute a somewhat smaller sink for Netherlands. Although at smaller extent, spatial flux distribution ambiguities seem to be connected also with the prior error structure. Comparing the spatial maps in fig 3 in the paper (subplots with and without the bias term), we see that the abovementioned highly productive regions (i.e. Alps, England and west Czech Republic), show somewhat weaker sinks for the BVR case compared to the NBV (indicated by the less bluish contours), due to the reduced flexibility at the pixel scale.

The second part of the reviewer's comment, encourages a thorough analysis at the national scale, in order to demonstrate/validate the accuracy of the inversion system. We fully agree with the reviewer and this was our intention, to have more comparisons at those scales by comparing not only aggregated fluxes but the spatial flux patterns as well. Unfortunately we were not able to find studies, showing such results especially for the European domain. Nevertheless the system should be somehow evaluated. Since direct comparison seems to be very difficult (if not impossible), we performed a synthetic experiment which we present in a companion paper (Kountouris et al., (2016)). Therein, synthetic fluxes and posterior estimates were compared to each other at several temporal and spatial scales. Results are promising for annual and monthly temporal scales. Further it seems also reasonable to trust the posterior estimates up to the national level. For more information the companion paper might be of help and shed some light regarding the inversion performance at the country scale.

P20 L17 we added: "…misrepresentation of the mountain stations at least at annual and domain wide aggregation scales.

However, the spatial flux distribution seems to depend on the site selection and in particular on the mountain sites used in a given inversion. Ambiguous carbon fluxes e.g. carbon sinks over non productive areas like Alps, England, and west Chech Republic, as well as carbon sources over cultivated lands like western France, Poland and Ukraine were derived from the inversions (fig. 3). Figure 4 presents the annual spatial flux distribution by using a network of stations with no mountain sites (MS0 case) and using an error structure which does not contain a bias term. This sensitivity test is equivalent to the nBB case where we used also the GBIOME-BGC model as prior. Subsequently we plot the flux distribution by adding one mountain site at a time (cases 1:7 where the number denotes how many mountain sites are being used). The add-one mountain site sequence is as follows: CMN, OXK, PTR, JFJ, KAS, SIL, PUY. For the MS0 case, we observe that in the region around the Alps, and the neighboring countries, the sink is smaller compared to the rest inversions. The OXK and the KAS sites seem to be responsible for the sink over the Czech Republic. The KAS site seems also to be the driver for the high carbon flux sources around Poland, Ukraine and the Black Sea coasts.

Using an error structure which allows for a bias term as the one in BVR case, seems to moderate the spatial flux misrepresentation. Comparing in fig. 3 the subplots nBV: without bias term, BVR: with bias term, we see that the abovementioned highly productive regions (according to the simulation), show somewhat weaker sinks for the BVR case compared to the nBV (indicated by the less bluish contours). Subsequently, regions that appear to be strong carbon sources (in nBV case), show weaker flux signal when the bias term is used (BVR).

Although this study uses as much information as possible, in terms of the available atmospheric observations still, large areas are poorly or not constrained at all from the atmospheric network e.g. West France, the whole East European part. Hence, the spatial flux distribution at those areas, is prone to large uncertainties.

Finally, the last comment about this study is a possible conclusion from your results, at least a point that I have in mind looking at the magnitude of the flux components. As you discuss in Section 4.3.2, the magnitude of carbon sources and sinks at subcontinental scales depends on the uncertainties of the fossil fuel emissions. Because both components and their uncertainties are similar, errors in anthropogenic estimates may impact the biogenic fluxes after inversion. The magnitude of total emissions can suffer from large errors but even more singificantly the spatial and temporal distributions as well. I want to conclude here that fossil fuel emissions should not be prescribed here but instead optimized in a joint optimization framework. Comments about that statement and possibly a discussion would help. Furthermore, indications on whether or not future inversion efforts should address this issue would be welcome.

The reviewer makes a very good point regarding the necessity of optimizing the fossil fuel component. This would definitely be the right way for future systems to correct also this component. We would like to note though a practical difficulty which arise, from the lack of available tracer observations (such as 14C), which is connected with the fossil fuel combustion.

P19 L21 we added: “In this study we assumed that anthropogenic emissions are perfectly known (which is a traditional assumption in atmospheric inversions), although this is not the case. As a result of not allowing for a correction in the fossil fuel component, this correction will be added to the correction of the biogenic signal. In this paragraph we already discussed how uncertain fossil fuel emissions may be. Further, we estimated how the uncertainty in the fossil fuel component impacts, the carbon flux estimates; the magnitude but also spatial and temporal flux distributions may be significantly erroneous. For better future carbon flux estimations, fossil fuel optimization seems to be necessary. However, that would require 14C tracer measurements which are currently not available.”

Specific comments:

page 3-line 3: "...have been applied using..." Please re-phrase ("assimilate"?)

P3 L3 we corrected: “…atmospheric inversions assimilate atmospheric…”

page 3-line 5: Add references

P3 L5 we added: “…dry mole fractions (Tans et al., 1989; Enting and Mansbridge, 1989, Conway et al., 1994, Fan et al., 1998; Rödenbeck et al., 2003).”

page 3-line 6: Rephrase "focus of interest".

P3 L7 we corrected: “…has been of high interest…”

page 3-line 10: "makes difficult" Explain more clearly what you mean.

P3 L11 we clarified: “coarse resolution; hence, the spatial and temporal flux variability at finer scales can not be resolved.”

page 3-line 12-14: This example is very specific, out of place for an introduction, and not supported by a reference. Delete or move this example.

P3 L14-16 we deleted it.

page 3-line 18: ... not seen by the ground network...

P3 L20 we corrected: "…to constrain regions not seen by the ground network."

page 3-line 25: Modeled dry air mole fractions

P3 L 28 we deleted: "tracer"

page 3-line 25: in vertical mixing

P3 L28 we corrected: "in vertical mixing"

page 3-line 27: ...biases in concentrations due to transport model errors are translated...

P3 L30 we corrected: "…because biases in concentrations due to transport model errors are translated…"

page 4-line 1-2: When and where are these errors applied? The numbers have to be explained here.

Here we refer to studies from Lin and Gerbig, 2005 and Gerbig et al. (2008). Those uncertainties derived through error propagation of uncertainties in winds and in mixing heights. We do not use this approach for propagation of transport uncertainties for this study. The model-data mismatch uncertainties used for this study are fully explained in Ko16 study which we refer now in the 2.3 section.

P7 L8 we added: "Similarly to the synthetic inversion (Ko16) the model-data mismatch uncertainties are the same as in the Ko16 study (see also fig. 2 therein). Further, we…"

page 4-line 10: Does it mean that the resolution of the state space is 0.5 degree? If so, what does 0.25 degree resolution correspond to?

Indeed the state space resolution is 0.5 degrees. The 0.25 degrees correspond to the coupling between biosphere fluxes and the transport model.

P4 L10 we clarified: "…of 0.25^o x 0.25^o to couple fluxes with the atmospheric transport model, and the state space…"

page 6-line 13: For hourly concentrations, 1ppm changes seem fairly small. Why discarding these data? What fraction has been removed using this filter?

We thank the reviewer for commenting on this. We actually did not make use of the filter and the atmospheric data contain all the temporal variability.

P6 L12 we deleted: "…and neighboring values differing by more than 1 ppm were omitted."

page 7-line 8-21: The codes used for the various inversions, i.e. B1, S1, S1a,... are difficult to remember and confusing for the readers. Short but self-explanatory codes would be easier to track in the figures and the results section.

We substituted all the codes through the paper and the plots with the following abbreviations. We hope the reviewer finds that more satisfactory.

B1 --> nBV : No Bias Vprm prior
B2 --> nBB : No Bias Biome prior
S1 --> BVR : Bias Vprm Respiration as shape
S1a --> BVN : Bias Vprm NEE as shape
S1b --> BVRT : Bias Vprm Respiration as shape Time varying bias
S2 --> nBV14 : No Bias Vprm 14 stations used
S3 --> nBVH : No Bias Vprm hyperbolic function for the spatial error correlations

page 8: Dependence to ecosystem types has not been considered here whereas previous ecosystem model assimilation studies often compute model parameters based on pft. Is there a reason to describe flux error correlations only based on distance?

The current inversion system does not optimize model parameters but rather optimizes directly fluxes at the grid scale. It does not differentiate or take into account different pft's and hence, pft dependent flux error correlations would be uninformative in the current system configuration. We note that the prior error structure is fully characterized and we do compute also the temporal correlation times (see also Kountouris et al. (2015)). Both metrics (spatial and temporal correlation lengths) are fully characterizing the prior error covariance matrix.

page 10-line 5: Remove "Figure 1".

It is not clear here why we should remove that. By removing it, then the sentence would not make sense.

page 10 and Table 3: the use of the goodness of fit is a simple weighted total mismatch divided by the number of unknowns. The actual Degree of Freedom of the System would be more informative as it describes also the weigths of the observations compared to the prior errors. It will indicate if the solution is over- or under-constrained by the atmospheric data.

Equation 3 describes the goodness of fit which is defined as the sum of the ratios between the observed and the expected uncertainties. In this equation we mentioned (incorrectly) that n, is the number of observations. We thank the reviewer for noticing that. Dividing with the actual Degrees of Freedom would indicate if the solution is over/under constrained by the atmospheric data. As the inversion system assumes implicitly that the atmospheric data is correlated, it is obvious that the number of observations do not represent the Degrees of Freedom. What we actually use, was the de-weighted number of observations, which are statistically independent. This number (n in equation 3) is basically the Degrees of freedom for the current station.

page 12-line 1-3: Are these very large positive corrections realistic? The discussion is very brief here.

The large positive corrections are connected with the large prior sink showed by VPRM model. This sink apparently is not explained by the atmospheric data and therefore we have a significant positive correction. On the contrary using prior fluxes from BIOME-BGC (nBB case, this model estimates a much less European sink) positive corrections are limited.

P12 L6 we added: ". The positive flux corrections is something to be expected since prior fluxes from VPRM show a strong European sink of 0.96 GtC y^{-1} which is most likely to be unrealistic."

page 12-line 16: Comments on the peak mismatch are important here (cf. general comment). Why is the inversion systematically shifting the maximum?

Regarding the peak mismatch we added explanatory text under the 4.2 paragraph, since therein we discuss the validation of the posterior estimates against EC fluxes.

P17 L9 we added: "Posterior fluxes showed a shift by one month earlier (in May), for the maximum carbon uptake (see also fig. 7). An initial hypothesis that this might be driven from sites which are difficult to simulate, such as those located in mountain regions, can not be justified. In specific, mountain sites were excluded in an additional sensitivity analyses, but the temporal shift remains. However, looking into the error of the difference between two months suggests that the flux difference between May and June is not significant. The error of the difference was calculated using a Monte Carlo experiment. Fluxes were averaged over the stations and the monthly differences were calculated. Then we used the standard deviation of the differences over the ensemble members to describe the month-to-month uncertainty.".

P24 L6 we added: "…nocturnal boundary layer e.g. mountain stations) by performing inversions using different network configurations. We did not observe any significant impact for domain-wide aggregated fluxes at least for monthly and annual scales. However changes in spatial flux patterns at the pixel scale should be expected, when then network configuration is changed."

page 16-line 23-27: the argument is valid but applies to both crop and non-crop eddy-flux sites. Clarify that this problem is common to all the ecosystems.

P17 L3 we clarified: "Further, another difficulty which is common for all the ecosystems, is the fact that atmospheric concentrations"

page 17-line 22: To study the representativity of the flux sites to constrain a pft, a Leave-One-Out cross-validation would help evaluate if the optimization applies to land classes, or if each eddy-flux site is specific to its own area.

VPRM uses PFTs (each vegetation type has a set of parameters) and the leave-one-out statistics was done on a PFT level.

Fact that large errors are found for cropland shows that the sites are not that representative (as stated in the text). It is not clear if the reviewer suggest any additional analysis here, for example to assess the uncertainty of VPRM-modeled flux at each eddy covariance site.

page 21-line 15-25: This example of country-scale flux evaluation provides a first assessment of the inverse fluxes at higher resolution. Is the agreement representative of any European country?

This is very difficult to say. Initially, this was one of the assessments we would like to perform. Unfortunately, there is no information about regional or country carbon budgets derived from regional systems, that we can compare our results with them.

Anonymous Referee #2

We sincerely thank the referee for his thorough review and highly appreciate the comments which significantly improved the quality of the publication. We hope that our answers and the modifications are satisfactory.

The authors estimate the European terrestrial carbon budget with atmospheric data at relatively high resolution. The study comes without surprises or great findings, but rather expresses a mature domain with sophisticated scientific tools. The paper is very clear and will likely serve as a useful reference both for its methodological synthesis and for its results. I recommend publication provided the following minor comments are addressed.

P. 2, l. 9 (also p. 4, l. 2): the actual quantity behind "uncertainty" should be defined.

Please see also respond to the first reviewer in the question: "page 4-line 1-2: When and where are these errors applied? The numbers have to be explained here."

P. 3, l. 4: "since" -> "for"

P3 L4 we corrected: "…for…"

P. 3, l. 7: Gurney et al. is an outlier in the list.

P3 L8 we deleted: "Gurney et al., 2004"

P. 3, l. 10: "makes" -> "makes it"

P3 L11 we corrected the whole sentence: "; hence, the spatial and…"

P. 3, l. 25: "of" -> "in"

P3 L28 we corrected: "…to uncertainties in vertical…"

P. 9, l. 24: do the authors assume linearity in the variances or in the standard deviations? I guess variances are more likely to be linear than standard deviations.

We thank the reviewer for this comment. Indeed the scaling was made such that we assumed linearity in the sd, whilst the scaling factor was taken as the flux ratio between the two regions. We have modified now the scaling of the uncertainties such that the variance now, is scaled linearly with the area (not the flux ratio any more).

P9 L21 we modified: "…we calculated the area ratio between the TransCom EU region and our European domain. This ratio (about 1.69) was used to scale our posterior estimates and the corresponding uncertainties assuming linearity in the variances (presented inFig. 8)."

Figure 8 is also updated.

P. 10, l. 13: first mention of sampling times here. Needs to be explained.

For sake of not repeating text, we added a reference regarding the sampling times.

P10 L13 we added: "…respective sampling times (see also Ko16) for mountain (nighttime) and other stations (daytime)."

P. 11, l. 4: a normal distribution extends to infinity: how is the uncertainty range defined?

The uncertainty range is defined as within 1 sigma of the respective assumed site uncertainty.

P11 L5 we clarified: "…the misfits are inside the 1 sigma site specific uncertainty."

P. 11, l. 10: a reduced chi-square only suggests something when the inversion upstream hypotheses are satisfied: are the authors confident that this is the case?

In Bayesian inversions the uncertainties are assumed to follow a Gaussian distribution. We are aware though that this might not be the case, and they might follow a different distribution e.g. Cauchy. However the Gaussian assumption is necessary for the Bayesian inversions. Nevertheless, the impact of a wrong assumption in the uncertainty distribution would be limited compared to the impact of the transport uncertainties.

P. 13, l. 3, 5: the authors should change the unit to something more appropriate.

P13 L9,10 we corrected : "flux estimates being 0.9 MtC y^{-1}, while United Kingdom (which is less well constrained) shows a slightly larger spread of the posterior estimates with an annually averaged standard deviation of 2 MtC y^{-1}"

P. 13, l. 11: "shown" is not the right word since there was at least Broquet et al. before Ko16.

P13 L16 we added: "…in Broquet et al. (2013) and in…"

P. 15, l. 21: this note of caution comes too late.

P15 L20: We moved and modified the text: "A common approach in atmospheric inversion studies to evaluate the defined uncertainties is to examine the reduced χ_r^2 values. However, this might not always be a sufficient metric (Michalak et al., 2005; Chevallier, 2007)."

P. 16, l. 3: the authors should insert "of Ko16" after "inversion" for clarity.

P16 L12 we corrected: "…synthetic inversion of Ko16, the real…"

P. 16, l. 6: why not transport errors as well?

We do refer also to the transport error in P16 L20-24.

P. 16, l. 17: the authors should change "from".

P16 L26 we corrected: "…by…"

P. 16, l. 23-24: "to mention"-> "mentioning".

P17 L3 we modified: "Further, another difficulty which is common for all the ecosystems, is the fact…"

P. 17, l. 15: EVI is still undefined.

P18 L4 we clarified: "…EVI (Enhanced Vegetation Index)…"

P. 19, l.8: the authors should insert a comma after "Typically".

P20 L5 we corrected: ". Typically, in…"

P. 22, l. 16: the authors should end their main text on a wider perspective.

P24 L14 We added : "What do we learn or should we expect then from the top down approach? The current analysis part one and two, suggests that aggregated fluxes at monthly (temporally) and country (spatially) scales can be successfully retrieved from the inversion system. However, retrieving spatially resolved fluxes at finer scales is still rather challenging. Lack of observations for extended European regions, complexity of the terrain especially in mountainous regions as well as the absence of fossil fuel measurements which would otherwise, allow the separation of fossil fuel signals from biospheric signals in observed CO_2 time-series, complete the mosaic of the current problems that regional inversions are facing. Whilst ICOS (Integrated Carbon Observing System) will introduce more stations in the European continent still, inversions should use all the available information; that could be achieved by assimilating multiple data streams like continuous and flask measurements in combination with satellite derived information, aiming to constrain as tight as possible the European continent. Further, new stations should also aim in measuring combustion tracers. That would be of a great help in future inversion systems to be able to update the anthropogenic emission maps and subsequently to compute more accurately the biogenic signal."

References

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

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, F. Th., Heimann, M.: Atmospheric CO2 inversions at the mesoscale using data driven prior uncertainties. Part1: Methodology and system evaluation, discussion Atmos. Chem. Phys. doi:10.5194/acp-2016-577, 2016.

Meesters, A. G. C. A., Tolk, L. F., Peters, W., Hutjes, R. W. A., Vellinga, O. S., Elbers, J. A., Vermeulen, A. T., van der Laan, S., Neubert, R. E. M., Meijer, H. A. J., Dolman, A. J.: Inverse carbon dioxide flux estimates for the Netherlands, J. Geophys. Res.-Atmos. 117, D20306, 1984-2012 , doi: 10.1029/2012jd017797, 2012.

Pillai, D., Gerbig, C., Ahmadov, R., Roedenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B. and Lavric, J., V. : High resolution simulation of atmospheric CO2 over complex terrain-representing the Ochsenkopf mountain tall tower, Atmos. Chem. Phys., 11, 7445–7464 doi:10.5194/acp-11-7445-2011, 2011.

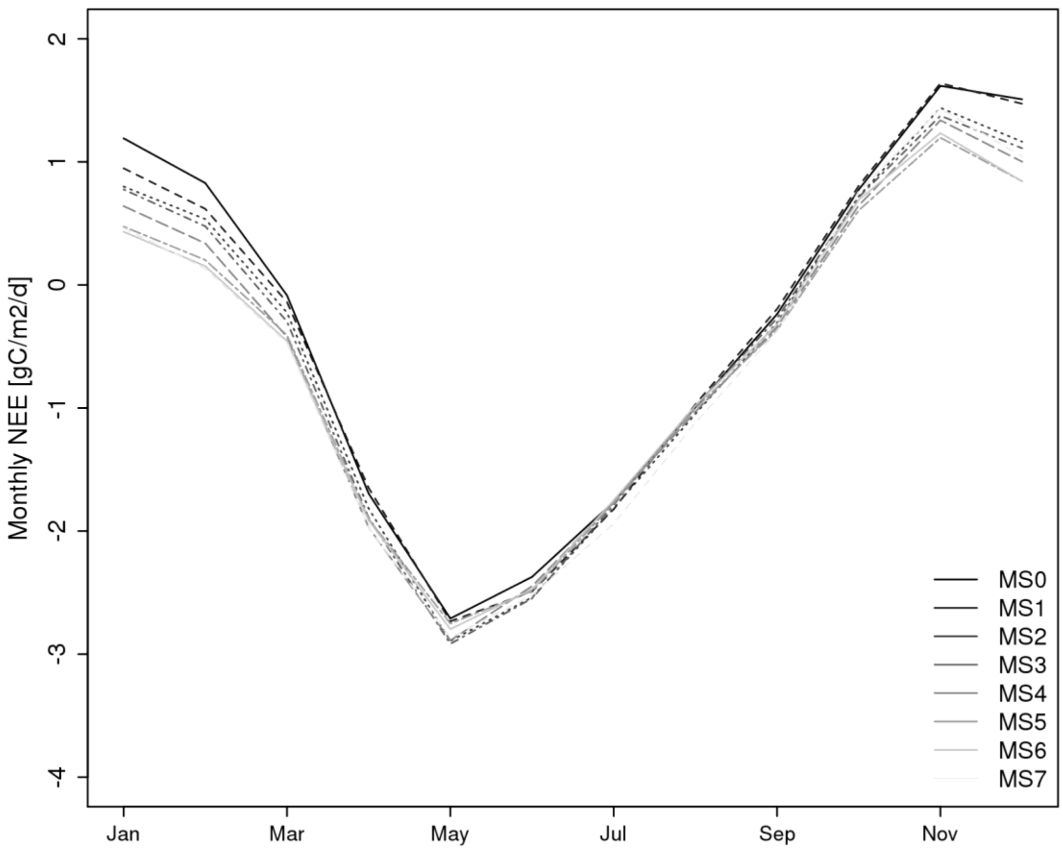

Figure 1: Temporal evolution of monthly NEE ($gCm^{-2}day^{-1}$) averaged over crop sites. MS0 case refers to a network configuration with no mountain stations and the MS7 case denotes that all (7 in total) mountain stations, were included in the inversion.

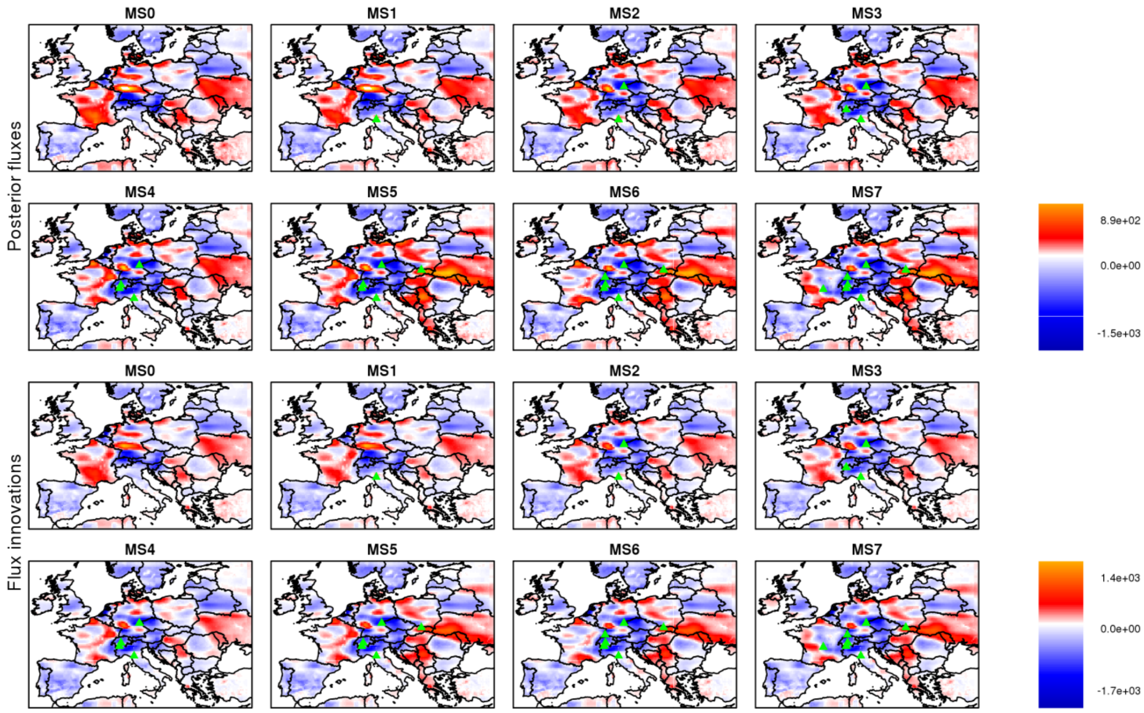

Figure 2: Annual biogenic flux spatial distribution (top two rows) and flux innovations (posterior - prior) (bottom two rows) using the nBB case. The number in the title denotes the number of the mountain stations have been used for this particular inversion (i.e. MS0: no mountain stations). Color bar and units are the same as in fig3 in the paper. Units are in $gCy^{-1}m^{-2}$.