Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design

T. Kaminski[1], P. J. Rayner[2], M. Voßbeck[1], M. Scholze[3], and E. Koffi[4]

[1]FastOpt, Lerchenstraße 28a, 22767 Hamburg, Germany
[2]School of Earth Sciences, University of Melbourne, Melbourne, Australia
[3]Dep. of Earth Sciences, University of Bristol, Bristol, UK
[4]LSCE/IPSL, Gif-sur-Yvette Cedex, France

Correspondence to: T. Kaminski (thomas.kaminski@fastopt.com)

Abstract. This paper investigates the relationship between the heterogeneity of the terrestrial carbon cycle and the optimal design of observing networks to constrain it. We combine the methods of quantitative network design and carbon-cycle data assimilation to a hierarchy of increasingly heterogeneous descriptions of the European terrestrial biosphere as indicated by increasing diversity of plant functional types. We employ three types of observations, flask measurements of CO_2 concentrations, continuous measurements of CO_2 and pointwise measurements of CO_2 flux. We show that flux measurements are extremely efficient for relatively homogeneous situations but not robust against increasing or unknown complexity. Here a hybrid approach is necessary, and we recommend its use in the development of integrated carbon observing systems.

1 Introduction

CO_2 and methane are the most important anthropogenic greenhouse gases. Their increasing concentration is the major reason for global warming (Solomon et al., 2007). It is thus of paramount interest to quantify and ultimately predict the exchanges of these gases between the terrestrial biosphere and the atmosphere. At a number of points on the globe, carbon and water fluxes are sampled directly (see, e.g, http://www.fluxnet.ornl.gov). The interpolation of these measurements to the globe (upscaling) requires external information about the uncertain spatio-temporal flux structure. The same type of information is required by atmospheric transport inversions (see, e.g., Rayner et al., 1999; Gurney et al., 2002; Enting, 2002) which infer surface fluxes from atmospheric concentration measurements. The most sophisticated tools for quantifying the structure and variability of carbon fluxes

20 are process models of the terrestrial carbon cycle like those used for the assessments of the IPCC (Solomon et al., 2007). Underlying these models is the assumption of fundamental equations that govern the processes controlling the terrestrial carbon fluxes. Uncertainty in the simulation of these fluxes arises from four sources: first, there is uncertainty in the forcing data (such as precipitation or temperature) driving the terrestrial processes. Second, there is uncertainty regarding the formulation

25 of individual processes and their numerical implementation (structural uncertainty). Third, there are uncertain constants (process parameters) in the formulation of these processes (parametric uncertainty). Forth, there is uncertainty about the state of the terrestrial biosphere at the beginning of the simulation.

Observational information helps to reduce these uncertainties. Currently there are several initia-

30 tives to extend the observational network of the carbon cycle. Europe's Integrated Carbon Observing System (ICOS, see http://www.icos-infrastructure.eu/), for example, aims at setting up an integrated sampling network for land, ocean, and atmosphere. Ideally, all data streams are interpreted simultaneously with the process information provided by the model to yield a consistent picture of the carbon cycle that balances all the observational constraints, thereby taking the respective uncertainty

35 ranges into account. Data assimilation systems around prognostic models of the carbon cycle are the ideal tools for this integration allowing us to assimilate a wide range of observations. They can, for example, be applied to systematically reduce parametric uncertainty (Kaminski et al., 2002, e.g.) or to expose structural errors (Rayner, 2010). In a first step, they use the observations to constrain the uncertain process parameters (calibration), and in a second step they use the calibrated model for

40 analysis and prediction. Ideally, both steps include the propagation of uncertainties. This allows us to derive uncertainty ranges on simulated *target quantities* that are consistent with the uncertainties in the observations and the model. Examples of such target quantities are fluxes of carbon on regional, continental, or global scale, integrated over part of the assimilation period (*diagnostic target quantity*) or some period before or after (*prognostic target quantity*). With regard to a specified target

45 quantity, such an assimilation system can assess the performance of a given observational network. This performance is typically quantified by the uncertainty range.

Quantitative Network Design (QND) aims at constructing an observational network with optimal performance. The approach is based on Hardt and Scherbaum (1994) who optimised the station locations for a seismographic network. It was first applied to observational networks of the global

50 carbon cycle by Rayner et al. (1996), who optimised the locations of atmospheric CO_2 and $\delta^{13}C$ measurements. A pioneering study for sensor design has been performed by Rayner and O'Brien (2001) who established the required precision for observations of the column-integrated atmospheric CO_2 concentration from space.

The latter two studies investigated purely atmospheric networks. To assist the design of an inte-

55 grated carbon observing system, we need the capability of evaluating the complementarity of various observational data streams including those of the terrestrial biosphere. As outlined by Kaminski and

2

Rayner (2008) assimilation systems are the ideal tool for this task. The Carbon Cycle Data Assimilation System (CCDAS, see http://ccdas.org) can assimilate several observational data streams and infers uncertainty ranges on diagnosed (Rayner et al., 2005) or prognosed carbon (Scholze et al., 2007; Rayner et al., 2011) and water (Kaminski et al., 2011) fluxes. The first QND applications investigated the utility of space borne observations of atmospheric CO_2 (Kaminski et al., 2010) or vegetation activity (Kaminski et al., 2011) in constraining various surface fluxes. Another study explores the atmospheric in situ network and its ability to constrain the productivity of the terrestrial biosphere (*Koffi et al.*, 2012).

Kaminski and Rayner (2008) noted two general aspects of QND studies. The first is the dependence on the target quantity; clearly different networks are optimal for constraining different things (Rayner et al., 1996). The second is the dependence on prior knowledge brought to the problem. For traditional inversions of fluxes this information takes the form of the covariance of prior uncertainty. For CCDAS it is determined by the process resolution of the underlying dynamical model (how many processes are modelled) and the spatial detail at which these processes are allowed to vary independently. The level of heterogeneity of the biosphere is a fundamental question which goes beyond CCDAS; it determines how much any understanding of processes gained locally can be more widely applied. However it is clear that observing networks presupposing a given heterogeneity are at some risk. Current earth system models map this spatial heterogeneity by dividing the global vegetation into a small number of plant functional types (PFTs). *Groenendijk et al.* (2011) demonstrate through calibration of a terrestrial model against direct flux measurements the limit of this approximation and the difficulty in deriving a realistic PFT classification.

This paper uses the network designer, a CCDAS-based interactive QND tool, to investigate the performance of several networks composed of direct flux observations and flask or continuous samples of the atmospheric carbon dioxide concentration. In particular we investigate the robustness of network performance to various choices of target quantities and levels of heterogeneity. The outline of the paper is as follows. Section 2 describes our QND methodology and Sect. 3 the networks we consider. Then Sect. 4 will present and discuss the evaluations. Finally, in Sect. 5 we summarise our conclusions.

2 Methods

CCDAS is built around the Biosphere Energy Transfer HYdrology scheme (BETHY, Knorr, 2000; Knorr and Heimann, 2001), a global model of the terrestrial vegetation. The version used here is described in Rayner et al. (2005). This section gives brief descriptions of BETHY, the observational data types, CCDAS and of the QND approach.

2.1 BETHY

Following Wilson and Henderson-Sellers (1985) BETHY decomposes the global terrestrial vegetation into 13 PFTs as listed in Table 1. Each grid cell can be covered by up to three PFTs. Figure 1 shows the distribution of the dominant PFT. As in Scholze et al. (2007) we integrate the model over 21 yr from 1979 to 1999 on a global 2 by 2 degree grid and use observed meteorological driving data (Nijssen et al., 2001).

The process formulations within BETHY are controlled by a set of process parameters (see Table 2). For this study we use the model version of Scholze et al. (2007) with the extension of simulating hourly Net Ecosystem Productivity (NEP). This is done by dividing the daily calculated heterotrophic respiration flux into 24 equal-sized hourly fluxes and subtracting these fluxes from the hourly simulated Net Primary Productivity (NPP). BETHY simulates 13 PFTs including 21 different parameters. Three of these parameters are PFT-specific and 18 are applied globally, i.e. they refer to all PFTs. We thus have $18 + 3 \times 13 = 57$ parameters. The role of the individual parameters is described elsewhere (Rayner et al., 2005; Scholze et al., 2007). In our context of network design it is important to know to which parameters our respective target quantities are sensitive. We will use regional integrals of the NPP and the NEP as target quantities. The latter is net CO_2 flux between the atmosphere and the biosphere and defined as the difference of NPP and heterotrophic soil respiration. Except for one atmospheric parameter c_0, all parameters impact NEP. NPP is sensitive to all parameters, except c_0 and the soil and carbon balance parameters.

2.2 Observational data types

In this study we use three types of observational data: direct (NEP) flux measurements, flask and continuous samples of the atmospheric CO_2 concentration. Within the model, a flux measurement is represented by a time series of hourly NEP samples of the grid cell the site is located in. The atmospheric data types require, as a so-called observation operator, an atmospheric transport model to transform the global NEP field into atmospheric concentrations. Flask samples are represented by a time series of monthly mean concentrations at the sampling location as simulated by the atmospheric transport model TM2 (Heimann, 1995), which is run at 8 by 10 degree horizontal resolution and with nine vertical levels. As in Carouge et al. (2010a,b) continuous samples are represented by a time series of daily mean concentrations at the sampling location as simulated by the atmospheric transport model LMDZ (Hauglustaine et al., 2004), which is run at 3.75 by 2.5 degree resolution over most of the globe but a zoomed 0.5 degree resolution over Europe. For each data type the observational time series covers the 20 yr period from 1980 to 1999. By representing flask samples in the model as monthly means, much of the synoptic signal is averaged out. Likewise by representing continuous measurements by daily means the diurnal signal is averaged out. This averaging reduces the information content of the observations but is also less demanding of the models' performance,

125   i.e. a conservative choice.

2.3   CCDAS

CCDAS uses a gradient method to adjust BETHY's process parameters in order to minimise a cost function. This cost function quantifies the fit to all observations plus the deviation from prior knowledge on the process parameters:

130
$$J(\tilde{\mathbf{x}}) = \frac{1}{2} \left[(M(\tilde{\mathbf{x}}) - \mathbf{d})^T \mathbf{C}(d)^{-1} (M(\tilde{\mathbf{x}}) - \mathbf{d}) + (\tilde{\mathbf{x}} - \mathbf{x_0})^T \mathbf{C}(x_0)^{-1} (\tilde{\mathbf{x}} - \mathbf{x_0}) \right] \tag{1}$$

where M denotes the model considered as a mapping from parameters to observations, d the observations with data uncertainty $\mathbf{C}(d)$, x_0 the prior parameter values with uncertainty $\mathbf{C}(x_0)$, and the superscript T the transpose.

The second derivative (Hessian) of the cost function at the optimum x is used to approximate the

135   inverse of the covariance matrix $\mathbf{C}(x)$ that quantifies the uncertainty ranges on the parameters that are consistent with uncertainties in the observations and the model. In a second step, the linearisation $\boldsymbol{N'}$ (Jacobian) of the model N used as a mapping from parameters to target quantities is used to propagate the parameter uncertainties forward to the uncertainty in a target quantity $\sigma(y)$:

$$\sigma(y)^2 = \boldsymbol{N'} \mathbf{C}(x) \boldsymbol{N'}^T + \sigma(y_{mod})^2. \tag{2}$$

140    $\sigma(y_{mod})$ quantifies all uncertainty in the simulation of the target quantity except the uncertainty in x (which we resolve explicitly). If the terrestrial model was perfect, $\sigma(y_{mod})$ would be zero. In contrast, if the parameters were perfectly known, the first term on the right hand side would be zero. Likewise the data uncertainty $\mathbf{C}(d)$ is the sum of the observational uncertainty and all uncertainty in the simulation of the observations except the uncertainty in the parameter vector.

145   All derivative information is provided with the same numerical accuracy as the original model in an efficient form via automatic differentiation of the model code by the automatic differentiation tool TAF (Giering and Kaminski, 1998).

2.4   QND

In network design mode, CCDAS is restricted to the uncertainty propagation for candidate networks.

150   It builds on the optimal parameter set estimated from data of the available network for the evaluation of the required first and second derivatives. In our case the optimal parameter vector is taken from the study of Scholze et al. (2007). For the evaluation of potential networks, the Hessian is evaluated for $d = M(x)$. In this case the posterior target uncertainty solely depends on the prior and data uncertainties and linearised model responses at observational locations and for target quantities. The

155   approach does not require real observations, and can thus evaluate hypothetical candidate networks (see Kaminski and Rayner, 2008; Kaminski et al., 2010). Candidate networks are defined by a set of observations characterised by observational data type, location, and data uncertainty. In practise for

pre-defined target quantities and observational types and locations, model sensitivities can be pre-computed and stored. A network composed of these pre-defined observations, can then be evaluated in terms of the pre-defined target quantities without further model evaluations. Only matrix algebra is required to combine the pre-computed sensitivities with the data uncertainties.

This is the approach implemented in the network designer (see http://imecc.ccdas.org), an interactive software tool that evaluates networks composed of flask and continuous samples of atmospheric CO_2 and direct flux measurements. Available target quantities are NPP and NEP over three regions: Europe, Brazil, and Russia (see Fig. 2). They are provided in the form of annual mean values averaged over the 20 yr assimilation period. Model sensitivities have been pre-computed for a list of atmospheric sampling sites (see Fig. 3 and Table 4). For flux measurements, model sensitivities have been pre-computed for every terrestrial grid cell and all PFTs that are available in the grid cell. When defining the site, the user can specify a mix among these PFTs. Uncertainties for data sampled at different sites and times are assumed to be uncorrelated. The uncertainty for each site is quantified by a standard deviation $\sigma(d)$, that reflects the combined effect of observational $\sigma(d_{obs})$ and model error $\sigma(d_{mod})$:

$$\sigma(d)^2 = \sigma(d_{obs})^2 + \sigma(d_{mod})^2 \tag{3}$$

The unit of the data uncertainties depends on the data type. For flask and continuous samples of atmospheric CO_2 it is ppm, for eddy flux measurements it is $\mathrm{gCm^{-2}day^{-1}}$ (where gC stands for grams of carbon). The output of the network designer is the list of posterior uncertainties $\sigma(y)$ of the target quantities according to Eq. (2). $\sigma(y_{mod})$ can be specified by the user as a percentage of the 20 yr average of annual mean NPP.

3   Experimental setup

We will be evaluating several networks. To define these networks we have to select the sampling locations and the respective data uncertainties. Data uncertainty is generally difficult to estimate, especially in advance of actual measurements. In the following we give some motivation for our choices and for some cases we will test the effect of an alternative choice in Sect. 4.

The thrust of our study is the interaction between the spatial density of various classes of measurements and assumed heterogeneity of the spatial biosphere. It is important therefore that our choice of data uncertainty does not overly influence the results. We therefore make the most neutral possible choice of a uniform data uncertainty for each class of measurement. We also assume uncorrelated uncertainties in space and time. This is partly justified by the reduction in the underlying datasets to either daily or monthly means and, more importantly, by the focus of our study. We note that, in principle, systematic errors (biases) in the observations or the model (which would give rise to uncertainty correlations along the entire time series) can be removed or at least reduced by bias correction schemes. For example, *Pillai et al.* (2010) assess biases for the atmospheric data types and

6

derive a recipe for their reduction.

For the flux measurements we use an uncertainty of $10\,\mathrm{gCm^{-2}day^{-1}}$. With respect to the minimum uncertainty of $3\times10^{-6}\,\mathrm{molm^{-2}s^{-1}}\approx3.11\mathrm{gCm^{-2}day^{-1}}$ chosen by Knorr and Kattge (2005) this is a factor of about $\sqrt{10}$ larger. This effective sample size of 10 corresponds to ignoring half of the data because of nighttime sampling and allowing another factor of 5 to account of correlated uncertainties.

For the atmospheric data types we assume the combined error in the terrestrial and transport models to be the dominant contribution to data uncertainty. For flask samples (represented by monthly mean values) we use a data uncertainty of $1.0\,\mathrm{ppm}$, above the average assigned by *Rödenbeck et al.* (2003) for the combined observational and transport model error. For continuous observations, which are more difficult to simulate, we use an uncertainty of $1.5\,\mathrm{ppm}$. We can regard the factor of 1.5 compared to flask samples as an inflation of the data uncertainty, to achieve an effective sample size that is reduced by a factor of 2. With roughly 30 times as many measurements this still gives continuous observations greater weight than flask measurements but this is reasonable given their greater ability to represent a monthly mean.

Next we have to define the sampling locations. For each observational data type we define a base network:

- The atmospheric flask sampling network *flask*, which consists of the 41 monitoring stations listed in Table 2 of Kaminski et al. (2002) and shown in Fig. 3.

- The atmospheric continuous sampling network *cont*, which consists of the 15 sites listed in Table 4 and indicated with symbol "X" in Fig. 4.

- The eddy flux network *flux*, which consists of a dedicated site for each of the ten PFTs that are available to the model over Europe (PFT numbers 3–5 and 7–13 of Table 1). Each site is defined such that it is covered to $100\,\%$ by the respective PFT. Table 3 lists the sites and Fig. 4 indicates their locations with the symbols "+".

We evaluate the networks in terms of the uncertainty reduction (Kaminski et al., 1999) in six target quantities:

$$1-\frac{\sigma(y)}{\sigma(y_{prior})}, \tag{4}$$

where $\sigma(y_{prior})$ denotes the uncertainty in the target quantity without any observational constraint and $\sigma(y)$ is taken from Eq. (2). The prior uncertainties for our target quantities are computed by propagating the prior parameter uncertainties of Scholze et al. (2007) via the Jacobian N' (Eq. 2). They are $0.45\,\mathrm{GtC}$, $1.45\,\mathrm{GtC}$, and $1.13\,\mathrm{GtC}$ for NEP over Europe, Russia, and Brazil, respectively, and $0.66\,\mathrm{GtC}$, $1.08\,\mathrm{GtC}$, and $4.86\,\mathrm{GtC}$ for NPP. $\sigma(y_{mod})$ is an offset in Eq. (2). If the term was very high it would dominate the posterior uncertainty. To render the contrasts between the networks

more drastic, we use a value of zero, i.e. we only analyse the effect of the networks on the parametric uncertainty in the target quantities.

In the above-described default set up BETHY runs with 13 PFTs. To investigate the robustness of the network performance with respect to model complexity in terms of the number of available PFTs, we extend the default set up as follows: We split the global vegetation into several equal fractions. Each fraction has its own set of 57 independent parameters with uncorrelated prior uncertainty. All fractions of a PFT share the location of the original PFT. In other words, a grid cell that in the default setup is populated by a single PFT is now composed of equal subgrid patches, each with their own PFT; the corresponding surface fluxes add up to one grid cell flux to be used for the atmospheric networks (hence the patches can be said to have the same location) but they are separately monitored by the flux network. In the following we will call the number of fractions *multiplicity*. With multiplicity 4, for example, we have $4 \times 13 = 52$ PFTs and $57 \times 4 = 228$ parameters. A parameter that was global in the default configuration is no longer global, because its validity is restricted to one of the fractions of the global vegetation. A change of multiplicity also affects the prior uncertainty in the target quantities. Introducing the multiplicity m means that m copies have to share the same area. Hence, compared to the original flux y the flux y_i from each copy (i counting the copies) is reduced by a factor of m. And with it the original flux uncertainty $\sigma(y_{prior})$ is also reduced by a factor of m for each copy $\sigma(y_{i,prior})$. Since there is no correlation of the prior uncertainty among the copies, the total flux uncertainty $\sigma(y_{prior,m})$ is the square root of the sum of squares:

$$\sigma(y_{prior,m}) = \sqrt{\sum_{i=1,m} \sigma(y_{i,prior})^2} = \sqrt{\sum_{i=1,m} (\frac{\sigma(y_{prior})}{m})^2} = \frac{\sigma(y_{prior})}{\sqrt{m}}. \tag{5}$$

This means, for example, quadrupling the multiplicity halves the prior uncertainty.

4   Results and discussion

We start this section with evaluations of simple networks composed of one or two flux sites. Then we move on to the base networks defined in Sect. 3 and, finally, study the effect of increasing the number of PFTs that are available to the model.

4.1   Simple configurations of flux sites

The selection of a site location for sampling a particular PFT defines the Jacobian matrix that provides the link from the model parameters (required for simulating that PFT) to the simulated flux. To understand the effects which we will later see in larger networks, it is instructive to evaluate first a series of small networks consisting of one or two flux sites. We start with the separate evaluation of two sites which both observe PFT 9 (C3 grass) to 100 %, namely "site1731-9" in Southern Spain and "site143-9" in Northern Scandinavia. Note that we can populate any given location with up to three PFTs. For the current experiment we take the location of "site143-5" from Table 3 but populate

260   it to 100 % with PFT 9. For convenience, for the remainder of this subsection, we will refer to the sites just as "143" and "1731". The respective uncertainty reductions are displayed by blue (site "143") and orange (site "1731") bars in Fig. 5. First we note that flux measurements over Europe can reduce the uncertainty of target quantities over Russia and Brazil. This reflects our assumption of fundamental processes with a combination of universal and PFT-specific parameters: An obser-

265   vation provides information beyond its sampling time and location helping to reduce uncertainty everywhere. Figure 6 shows for site "1731" the uncertainty reduction in NEP per grid cell. This quantifies how the observational information of the site is spread around the globe. Comparing with Fig. 1 we note high uncertainty reduction where the dominant PFT is C3 grass.

    Among the two sites in terms of NEP site "1731" performs only marginally better, but in terms

270   of NPP it performs about 10 percentage points better.[1] Next we investigate the complementarity between the two sites, i.e. we use a network that consists of both sites and note a slight improvement for NPP over Europe and Russia (yellow bar in Fig. 5) compared to the better site "1731" alone. For these two target quantities the weaker site "143" is not *redundant* in this two site network, because it brings at least a little bit of extra information. In other words there is at least a slight *complementarity*

275   between the two sites with respect to the two target quantities.

    For the analysis of the above effects, recall that each scalar target quantity is (through the vector $\boldsymbol{N'}$ of Eq. 2) influenced by its own one-dimensional sub-space of the parameter space, i.e. a *target direction* in parameter space. Likewise each scalar observation constrains a direction in parameter space (*observed direction*). We can use the analogy of a *perspective* under which the target direction

280   is observed. If the target and observed directions are orthogonal, the observation can not reduce the uncertainty in the target quantity. If both directions are collinear, i.e. in the same subspace of the parameter space, the observation can most efficiently reduce the uncertainty in the target quantity. This means, for example, that even a hypothetically perfect measurement that removed all uncertainty for all parameters pertinent to one PFT would not completely constrain any of our

285   target quantities (which are all influenced by several PFTs). In other words a one-site flux network is *incomplete* with respect to our target quantities. The strength of an observational constraint on a target quantity depends 1) on the sensitivity of the observed quantity to a parameter change in the observed direction (*signal size*), 2) on how well the observed direction projects onto the target direction (perspective), and 3) on the data uncertainty. We use the same data uncertainty for both

290   sites and the same target directions. The observed direction and signal size depend 1) on the PFT, 2) on the sampling time, and on 3) the meteorological driving data. Our two flux sites provide measurements at the same times (hourly for 20 yr) and of the same PFT. The only different factors are the meteorological driving data. Indeed the meteorology in Southern Spain is quite different from Scandinavia.

---

[1]The unit *percentage point* quantifies an absolute change in the percentage value. For example, for a value of 40 % an increase by 50 percentage points yields 90 %. By contrast, an increase by 50 % corresponds to 20 percentage points and yields 60 %.

295   To isolate the effects of the perspective and the signal size on performance of the individual sites we reduce their respective data uncertainties by a factor of 100 (green and brown bars in Fig. 5). This can compensate for a weaker signal but does not change the perspective. Now both sites show exactly the same performance, i.e the Scandinavian site has just a smaller signal. In other words, we find the relevant information at both sites, but at sites with a larger signal we can afford a larger data

300   uncertainty or, probably, a shorter observational period.

     A common property between all networks evaluated in Fig. 5 is the larger uncertainty reduction for NPP compared to NEP. This happens although we sample hourly NEP, i.e. we should match the perspective for long-term NEP quite well. On the other hand, the target space for NPP has fewer dimensions, because it depends on fewer parameters. The extra parameters in NEP play an

305   important role. This effect would probably be even more pronounced if NEP was compared with the Gross Primary Productivity (GPP) which is influenced by even fewer parameters (*Koffi et al.*, 2012). Another point to note is that for Brazil the prior uncertainty in NPP is about four times higher than for NEP, and thus easier to reduce.

4.2   Base networks and their combinations

310   The performance of the three base networks *flask* (blue bars), *cont* (orange bars), and *flux* (yellow bars) is shown in Fig. 7. Over Europe, the flux network achieves an uncertainty reduction of about 99 % for both NEP and NPP and outperforms both atmospheric networks. The reason for the strong performance of *flux* over Europe is its completeness with respect to the European target quantities, i.e. the fact that for each PFT over Europe it contains a dedicated site. With respect to the Brazilian

315   target quantities, in turn, the network *flux* is incomplete because it does not cover the tropical PFTs. This is why *flux* is weaker than the global network *flask*, in particular for NEP where the performance difference between both networks is over 50 percentage points.

     The above suggests we would always attempt complete flux networks. In reality this will be hard to achieve, because we do not know how many PFTs are required to simulate the terrestrial carbon

320   cycle, nor do we know their spatial distribution (*Groenendijk et al.*, 2011). Hence, it may happen that we accidently miss a PFT in our flux network. We can test the effect of this by removing from network *flux* the site "1731-9" (network *flux-C3*). The performance over Europe drops by about 69 percentage points for NEP and 58 for NPP (green bars). Over Brazil the effect of missing the C3 grass site is only marginal (performance drop of less than four percentage points for NEP and less

325   than two for NPP).

     For the atmospheric networks, *flask* outperforms *cont* over Europe by 2 and 10 percentage points for NEP and NPP despite the European focus of *cont*. Obviously, for the atmosphere, the large-scale information matters. For Brazil or Russia it is not surprising that the global network *flask* is more powerful than the network *cont*. The most important aspect is that the atmospheric networks

330   outperform the incomplete flux network *flux-C3*. The only exception is NPP over Brazil, where the

loss of C3 had only a marginal effect on the performance of the flux network and *flux-C3* is stronger than *cont* but not than *flask*. We note that the relatively coarse resolution of TM2 may yield a slight overestimation in the integrative capacity of *flask*. For any given monthly mean sample, the higher resolution of LMDZ would resolve a finer influence structure (footprint) within the TM2 grid cells.

335 On the other hand, our sampling period of 20 yr would probably average out much of this time-dependent fine-scale structure, a mechanism that tends to increase the footprint. Increasing the data uncertainty of *cont* yields only small performance reductions of 3 percentage points over Europe, 6-7 percentage points over Russia and below one percentage point over Brazil (not shown).

To assess the *complementarity* of atmospheric and flux networks, we combine the networks *flux-*

340 *C3* and *flask*. Over Europe the resulting network *flux-C3 + flask* performs almost as well as the *complete* flux network *flux*, and over Brazil and Russia even better. Both networks (*flux-C3* and *flask*) complement each other. Given the experience from the two grass sites we evaluated initially (Sect. 4.1), we can think of the atmospheric network as an observer of averages over multiple sites. We can regard its addition to the flux network as an insurance against the *incompleteness* of the flux

345 network.

What can we do in the case where we can not afford enough sites to sample all PFTs over our target region? Is it useful to have a flux site which observes two PFTs? We test this by removing the site "site1731-9" from the network *flux* and modify the PFT fractions at site "site143-9" to 50 % each for PFTs 5 and 9. This network has the same number of sites as *flux-C3* but much better performance

350 (not shown). Uncertainty reduction for NEP over Europe is 76 %, and for the other target quantities the performance is only marginally (less than one percentage point) inferior to *flux*. This performance enhancement is based on the same principle as atmospheric sampling, the integration of a multi-PFT signal. This result seems surprising. It arises from the ability of a long time series to observe the different dynamics of the two underlying PFTs.

355 We can also investigate the complementarity of the base networks. Since the uncertainty reduction for the network *flux* is above 99 % already, in this set of assessments we rather quantify the performance gain by the reduction in posterior uncertainty relative to the posterior uncertainty of *flux*. Adding both atmospheric networks to network *flux* reduces NEP uncertainty over Europe by over 30 % and NPP by 20 %. For the other regions the effect is much larger (up to 99 % reduction

360 for NEP over Brazil).

4.3 Increased model complexity

The above network evaluations are based on the default model setup with 13 PFTs. In the following we investigate the robustness of the results with respect to model complexity in terms of the number of available PFTs. To increase the number of PFTs we use the procedure described in Sect. 3.

365 For multiplicity 4, Fig. 8 shows the performance of the three base networks. Among the atmospheric networks *flask* is superior to *cont* for all target quantities, except for NEP over Europe where

11

flask is slightly inferior. The network *flux*, in turn, is superior to *flask* except for NEP over Brazil and Russia. We define the network *M4-1 flux*, which is incomplete over Europe, by excluding one parameter copy out of the 4 from the network *flux*. This means *M4-1 flux* samples 30 out of the
370 40 PFTs that are available over Europe. As in the case of multiplicity 1, the incompleteness is reflected in a strong drop in uncertainty reduction in particular over Europe, where the performance is roughly halved (green bars in Fig. 8). Combining *M4-1 flux* with *flask* is only marginally superior to *flask* alone. Apparently *M4-1 flux* is too incomplete to bring extra information. Put another way, the unobserved parts of the domain dominate the final uncertainty.

375   For multiplicity 25 (not shown), *flux* achieves uncertainty reductions close to 100 % over Europe and above 85 % elsewhere. We define two incomplete flux networks over Europe, one with one parameter copy out of 25 removed (over Europe 240 of 250 PFTs sampled) and the other one with two parameter copies out of 25 removed (230 PFTs sampled). Over Europe, compared to *flux*, the first network suffers a performance drop of about 20 percentage points, and the other one of almost
380 30 percentage points. Over Europe the *flask* performance (78 % for NEP and 74 % for NPP) lies in-between both incomplete flux networks, which also holds for NPP over Russia. Elsewhere *flask* is better than the two networks. Even with a highly increased number of PFTs, an incomplete flux network that misses only a small fraction of the total PFTs is outperformed by *flask*. Combining *flask* with either one of the incomplete networks increases the *flask* performance over Europe by about ten
385 percentage points for the smaller flux network and by another three percentage points for the larger network. This means both incomplete networks exhibit enough complementarity to *flask* to achieve a significant performance gain. Note that with multiplicity 25 the prior uncertainty is reduced by a factor of five (see Eq. (5)). For example, an uncertainty reduction of 80 % corresponds to the same posterior uncertainty as an uncertainty reduction of 96 % (100-(100-80)/5 %) in BETHY's default
390 setup (i.e., multiplicity 1). This means that the posterior uncertainty in NEP over Europe of *flask* (uncertainty reduction of 78 %) is similar to that in the default setup (uncertainty reduction of 94 %).

5   Conclusions

QND is well-suited to explore the performance of observational networks of the carbon cycle. The network designer is a fast and easy-to-use QND implementation, that enables interactive network
395 evaluations, e.g. within a meeting. Its current focus is on the continental-scale carbon balance. Figures 2, 3 and 5 in this paper are directly obtained from the network designer.

    As mentioned above the particular performance values are consequences of specific choices such as the complexity of the underlying terrestrial model. There are, however, a set of general findings that follow from the above-mentioned assumption of fundamental equations that govern the pro-
400 cesses controlling the terrestrial carbon fluxes. First, for direct flux observations, it is important to cover the full range of different PFTs and not the range of climates to which a given PFT is exposed.

An incomplete flux network, i.e. one that misses a fraction of the PFTs risks a considerable performance loss. Atmospheric measurements are less prone to this problem, thus we can say that flux networks are more powerful while concentration networks are more robust. The combination can provide both qualities, i.e. atmospheric and flux networks complement each other.

The implications for the design of integrated observing strategies for the continental carbon balance seem clear. The baseline requirement is an atmospheric sampling network. That way if we underestimate the heterogeneity we will not find ourselves suddenly terribly undersampled. The strongest constraint, however, will come by overlaying this with a flux network which is as comprehensive as possible. Oversampling important PFTs will also give a diagnostic of heterogeneity. If parameters retrieved from one flux site enable us to predict the fluxes at a second then these are properly considered the same PFT for CCDAS, otherwise we need to increase the multiplicity.

This study addressed parametric and, to a certain extent, initial value uncertainty. To resolve structural uncertainty, it is important to build into the network the flexibility to detect features that are not or badly included in the model, i.e. the capability to discover surprises. Here, we have focused on carbon dioxide fluxes, however, observational networks for other trace gases, e.g. methane, can be evaluated with the same approach. Also, it is possible to evaluate networks that combine observations from space with in situ measurements as shown by Kaminski et al. (2010) and Kaminski et al. (2011). Similarly the column integrated CO_2 measurements collected by the Total Carbon Column Observing Network (TCCON, http://www.tccon.caltech.edu/) can be included, as an extra data type, in the network designer. The approach can also be extended to oceanic networks.

13

References

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, http://dx.doi.org/10.5194/acp-10-3107-2010doi:10.5194/acp-10-3107-2010, 2010a.

Carouge, C., Rayner, P. J., Peylin, P., 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, http://dx.doi.org/10.5194/acp-10-3119-2010doi:10.5194/acp-10-3119-2010, 2010b.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, Cambridge, 2002.

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM T. Math. Software, 24, 437–474, 1998.

Groenendijk, M., et al., Assessing parameter variability in a photosynthesis model within and between plant functional types using global fluxnet eddy covariance data, *Agricultural and Forest Meteorology*, *151*, 22–38, 2011.

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

Hardt, M. and Scherbaum, F.: The design of optimum networks for aftershock recordings, Geophys. J. Int., 117, 716–726, 1994.

Hauglustaine, D. A., Hourdin, F., Jourdain, L., Filiberti, M. A., Walters, S., Lamarque, J. F., and Holland, E. A.: Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: description and background tropospheric chemistry evaluation, J. Geophys. Res., 109, 44 pp., http://dx.doi.org/10.1029/2003JD003957doi:10.1029/2003JD003957, 2004.

Heimann, M.: The Global Atmospheric Tracer Model TM2, Technical Report No. 10, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 1995.

Kaminski, T. and Rayner, P. J.: Assimilation and network design, in: Observing the Continental Scale Greenhouse Gas Balance of Europe, edited by: Dolman, H., Freibauer, A., and Valentini, R., Ecological Studies, chap. 3, Springer-Verlag, New York, 2008.

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three dimensional global inverse model of the atmospheric transport, 1, Adjoint Model and Jacobian Matrix, J. Geophys. Res., 104, 18535–18553, 1999.

Kaminski, T., Knorr, W., Rayner, P. J., and Heimann, M.: Assimilating atmospheric data into a terrestrial biosphere model: a case study of the seasonal cycle, Global Biogeochem. Cy., 16, 14-1–14-16, available online at: http://www.agu.org/journals/gb/gb0204/2001GB001463/index.html, 2002.

Kaminski, T., Scholze, M., and Houweling, S.: Quantifying the benefit of A-SCOPE data for reducing uncertainties in terrestrial carbon fluxes in CCDAS, Tellus B, 62, 784–796, http://dx.doi.org/10.1111/j.1600-0889.2010.00483.xdoi:10.1111/j.1600-0889.2010.00483.x, available at: http://journals.sfu.ca/coaction/

14

index.php/tellusb/article/view/16634, 2010.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO_2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences Discuss., 8, 10761–10795, http://dx.doi.org/10.5194/bgd-8-10761-2011doi:10.5194/bgd-8-10761-2011, 2011.

Knorr, W.: Annual and interannual CO_2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecol. Biogeogr., 9, 225–252, 2000.

Knorr, W. and Heimann, M.: Uncertainties in global terrestrial biosphere modeling: 1. A comprehensive sensitivity analysis with a new photosynthesis and energy balance scheme, Global Biogeochem. Cy., 15, 207–225, 2001.

Knorr, W. and Kattge, J.: Inversion of terrestrial biosphere model parameter values against eddy covariance measurements using Monte Carlo sampling, Glob. Change Biol., 11, 1333–1351, 2005.

Koffi, E., P. J. Rayner, M. Scholze, and C. Beer, Atmospheric constraints on gross primary productivity and net flux: Results from a carbon-cycle data assimilation system, *Glob. Biogeochem. Cyc.*, *26*, doi:10.1029/2010GB003,900, 2012.

Nijssen, B., Schnur, R., and Lettenmaier, D.: Retrospective estimation of soil moisture using the VIC land surface model, 1980–1993, J. Climate, 14, 1790–1808, 2001.

Pillai, D., C. Gerbig, J. Marshall, R. Ahmadov, R. Kretschmer, T. Koch, and U. Karstens, High resolution modeling of co2 over europe: implications for representation errors of satellite retrievals, *Atmospheric Chemistry and Physics*, *10*, 83–94, 2010.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial Carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS), Global Biogeochem. Cy., 19, 20 pp., http://dx.doi.org/10.1029/2004GB002254doi:10.1029/2004GB002254, available online at: http://www.agu.org/pubs/crossref/2005/2004GB002254.shtml, 2005.

Rayner, P. J., Koffi, E., Scholze, M., Kaminski, T., and Dufresne, J.: Constraining predictions of the carbon cycle using data, Philos. T. R. Soc. A, 369, 1955–1966, http://dx.doi.org/10.1098/rsta.2010.0378doi:10.1098/rsta.2010.0378, available online at: http://rsta.royalsocietypublishing.org/content/369/1943/1955.abstract, 2011.

Rayner, P. J.: The current state of carbon-cycle data assimilation, Current Opinion in Environmental Sustainability, 2, 289–296, http://dx.doi.org/10.1016/j.cosust.2010.05.005doi:10.1016/j.cosust.2010.05.005, available online at: http://www.sciencedirect.com/science/article/pii/S1877343510000370, 2010.

Rayner, P. J. and O'Brien, D. M.: The utility of remotely sensed CO_2 concentration data in surface source inversions, Geophys. Res. Let., 28, 175–178, 2001.

Rayner, P. J., Enting, I. G., and Trudinger, C. M.: Optimizing the CO_2 observing network for constraining sources and sinks, Tellus B, 48, 433–444, 1996.

Rayner, P. J., Enting, I. G., Francey, R. J., and Langenfelds, R. L.: Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, Tellus B, 51, 213–232, 1999.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann, Co2 flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmospheric Chemistry And Physics*, *3*, 1919–1964, 2003.

15

Scholze, M., Kaminski, T., Rayner, P. J., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic CCDAS simulations, J. Geophys. Res., 112, 13 pp., http://dx.doi.org/10.1029/2007JD008642doi:10.1029/2007JD008642, 2007.

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

Wilson, M. F. and Henderson-Sellers, A.: A global archive of land cover and soils data for use in general-515 circulation climate models, J. Climatol., 5, 119–143, 1985.

Table 1. Plant Functional Types (PFTs) defined in CCDAS and their abbreviations, taken from Rayner et al. (2005).

| PFT No.	PFT Name	Abbreviation
0	Not vegetated	
1	Trop. broadleaved evergreen tree	TrEv
2	Trop. broadleaved deciduous tree	TrDec
3	Temp. broadleaved evergreen tree	TmpEv
4	Temp. broadleaved deciduous tree	TmpDec
5	Evergreen coniferous tree	EvCn
6	Deciduous coniferous tree	DecCn
7	Evergreen shrub	EvShr
8	Deciduous shrub	DecShr
9	C3 grass	C3Gr
10	C4 grass	C4Gr
11	Tundra vegetation	Tund
12	Swamp vegetation	Wetl
13	Crops	Crop
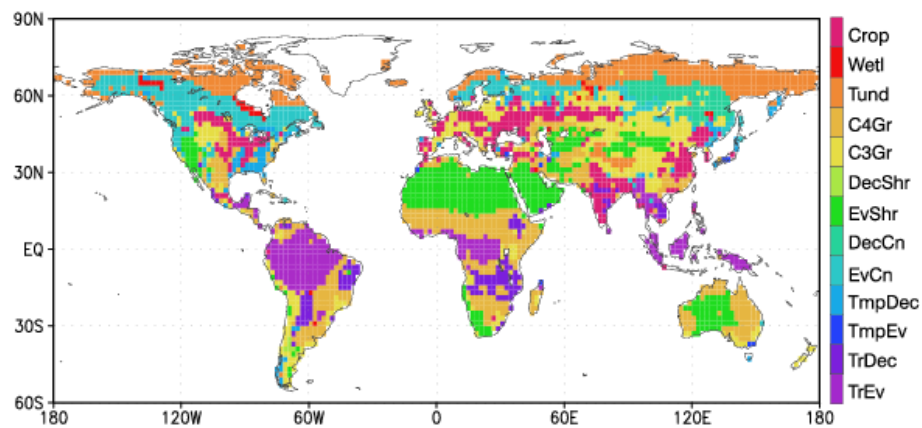
Fig. 1. Distribution of the dominant CCDAS Plant Functional Type (PFT) per grid cell, PFT labels are given in Table 1, taken from Rayner et al. (2005).

Table 2. Process parameters, their symbols (2nd column), their description (3rd column), whether their scope is PFT specific or global (4th column), whether NEP (5th column) or NPP (6th column) are sensitive to them.

Number	Symbol	Description	scope	NEP	NPP
1	$V_{\max}^{25}$	maximum carboxylation rate (C3)	PFT	X	X
2	$a_{J,V}$	ratio $V_{\max}^{25}$ over max electron transport $J_{\max}^{25}$	PFT	X	X
3	α_q	photon capture efficiency (C3)	global	X	X
4	α_i	quantum efficiency (C4)	global	X	X
5	K_{C}^{25}	Michaelis-Menten constant CO_2	global	X	X
6	K_{O}^{25}	Michaelis-Menten constant O_2	global	X	X
7	$a_{\Gamma,T}$	temperature slope CO_2 compensation point	global	X	X
8	$E_{K_{\mathrm{O}}}$	activation energy, O_2	global	X	X
9	$E_{K_{\mathrm{C}}}$	activation energy, CO_2	global	X	X
10	$E_{V_{\max}}$	activation energy, carboxylation rate (C3)	global	X	X
11	E_k	activation energy, carboxylation rate (C4)	global	X	X
12	$E_{R_{\mathrm{d}}}$	activation energy, dark respiration	global	X	X
13	$f_{R,leaf}$	leaf respiration ratio	global	X	X
14	$f_{R,growth}$	growth respiration ratio	global	X	X
15	f_{S}	fraction of fast soil decomposition	global	X	
16	κ	soil moisture exponential for soil respiration	global	X	
17	$Q_{10,\mathrm{f}}$	soil respiration temperature factor, fast pool	global	X	
18	$Q_{10,\mathrm{s}}$	soil respiration temperature factor, slow pool	global	X	
19	τ_{f}	fast pool soil carbon turnover time	global	X	
20	β	net CO_2 sink factor	PFT	X	
21	c_0	atmospheric concentration offset	global		

Table 3. Network *flux*. First number in site name indicates model grid cell and second number PFT.

Name	Included	Uncertainty $[\mathrm{gCm^{-2}day^{-1}}]$	lon	lat
site1413-3	yes	10.0	9.0	45.0
site1025-4	yes	10.0	35.0	53.0
site143-5	yes	10.0	19.0	69.0
site148-7	yes	10.0	29.0	69.0
site1495-8	yes	10.0	-5.0	43.0
site1731-9	yes	10.0	-5.0	37.0
site1578-10	yes	10.0	-5.0	41.0
site377-11	yes	10.0	-21.0	65.0
site388-12	yes	10.0	29.0	65.0
site621-13	yes	10.0	13.0	61.0

Table 4. Network *cont* of continuous atmospheric sampling sites over Europe.

Name	Included	Uncertainty	lon	lat
CBW	yes	1.5	4.9	52.0
CMN	yes	1.5	10.7	44.1
DGR	yes	1.5	22.07	54.15
FDA	yes	1.5	25.3	45.47
HUN	yes	1.5	16.6	46.9
MHD	yes	1.5	−9.9	53.33
NGB	yes	1.5	13.05	53.15
PAL	yes	1.5	24.12	67.97
PRS	yes	1.5	7.7	45.9
PUY	yes	1.5	3.0	45.8
SAC	yes	1.5	2.2	48.7
SBK	yes	1.5	12.98	47.05
SCH	yes	1.5	8.0	48.0
WES	yes	1.5	8.0	55.0
WHF	yes	1.5	10.77	52.8
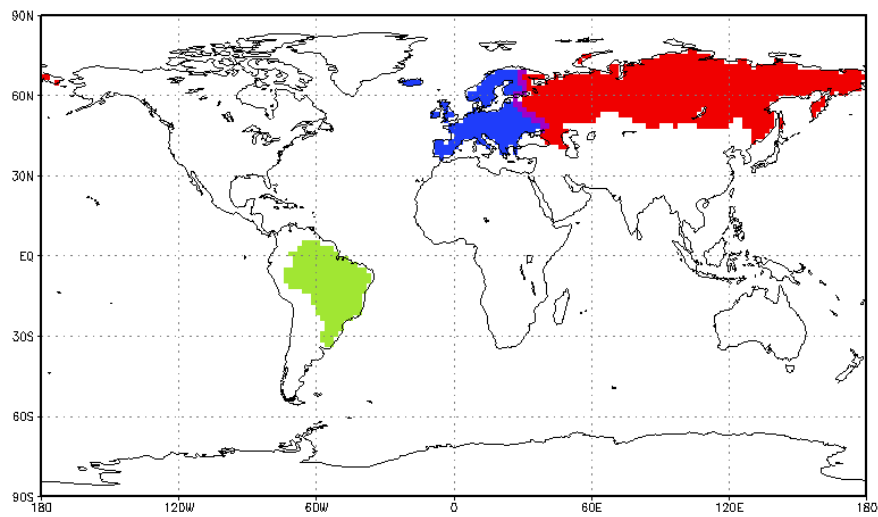
Fig. 2. Regions for computation of target quantities: Grid cells contributing to fluxes over Europe (blue), Russia (red), both (violet), and Brazil (green).
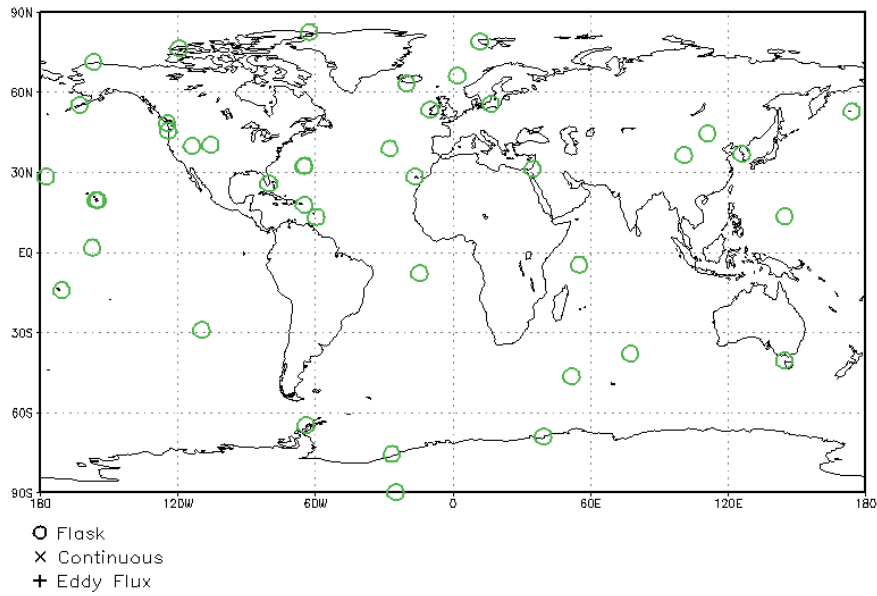
Fig. 3. Observational network *flask* providing atmospheric flask samples.
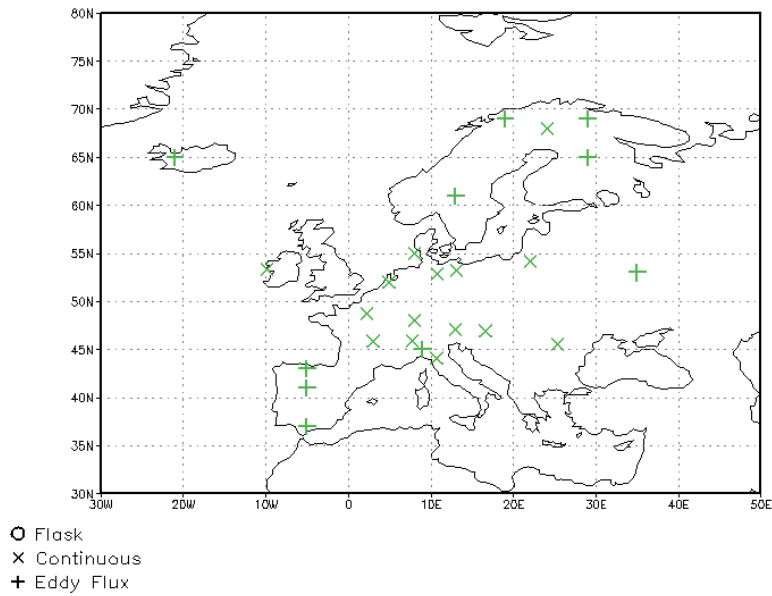
Fig. 4. Observational networks *flux* (+) providing flux measurements and *cont* (X) providing continuous atmospheric samples.
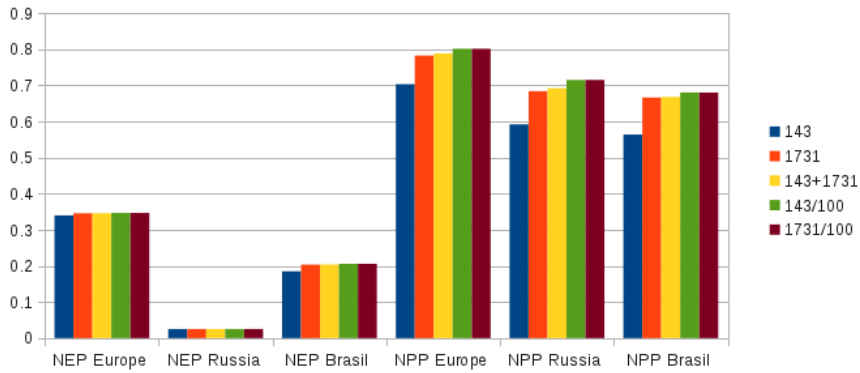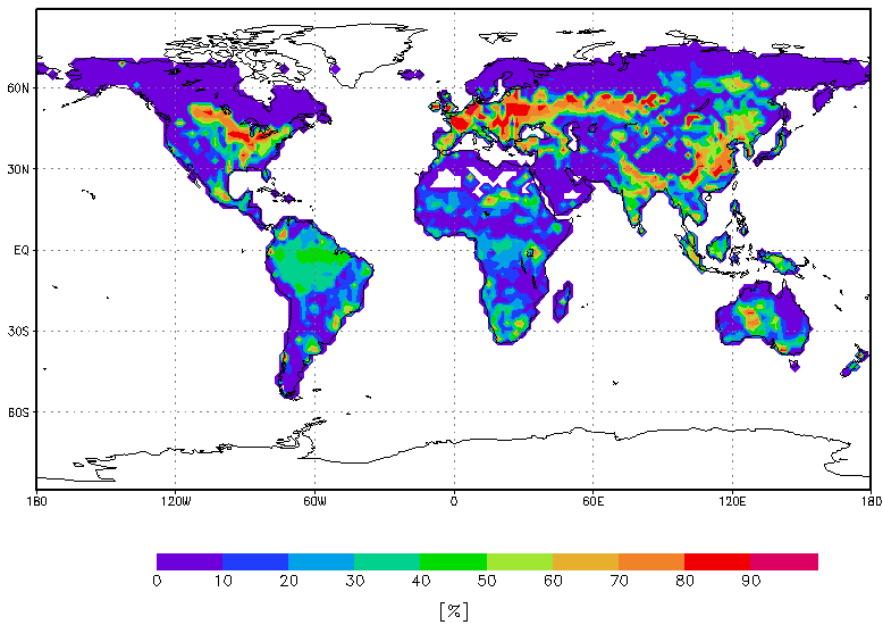
Fig. 5. Evaluation of two flux sites (blue and orange bars), of their combination (yellow bars), and of each site with data uncertainty reduced by a factor of 100 (green and brown bars): Uncertainty reduction for NEP and NPP integrated over three regions.

Fig. 6. Uncertainty reduction in NEP per grid cell, for a network consisting of a single flux site (site 1731-9 in Table 4)
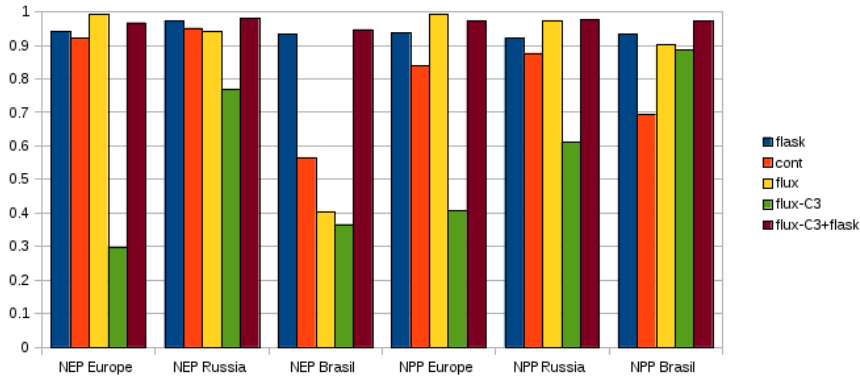
Fig. 7. Evaluation of three base networks, a flux network *flux-C3* that is incomplete over Europe and the combined *flux-C3+flask* network.
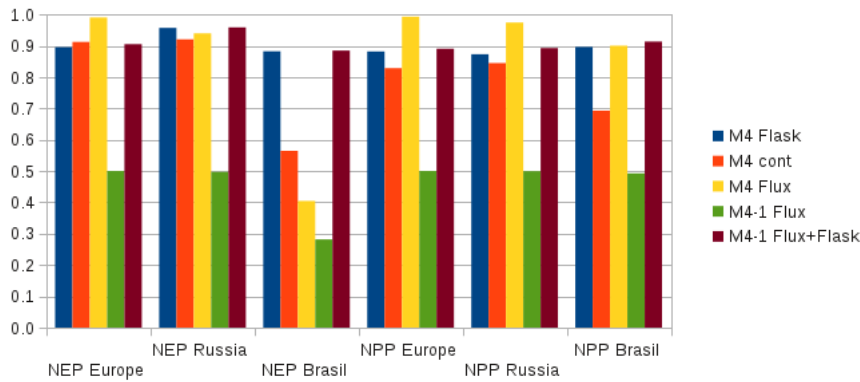

Fig. 8. Evaluation for multiplicity 4 of three base networks, an incomplete flux network with one copy of each PFT unsampled, and the combination of the incomplete flux network with the *flask* network.