**Biogeosciences
Discussions**

Consistent assimilation of MERIS FAPAR and atmospheric CO$_2$ into a terrestrial vegetation model and interactive mission benefit analysis

T. Kaminski[1], W. Knorr[2,3,4], M. Scholze[2,5], N. Gobron[6], B. Pinty[6], R. Giering[1], and P.-P. Mathieu[7]

[1]FastOpt, Lerchenstraße 28a, 22767 Hamburg, Germany
[2]Dep. of Earth Sciences, University of Bristol, Bristol, BS8 1RJ, UK
[3]Dep. of Meteorology and Climatology, Aristotle University of Thessaloniki, Greece
[4]Dep. of Earth and Ecosystem Sciences, Sölvegatan 12, 223 62 Lund, Sweden
[5]University of Hamburg, Grindelberg 5, 20144 Hamburg, Germany
[6]European Commission, DG Joint Research Centre, Inst. for Environment and Sustainability, Global Environment Monitoring Unit, TP 272, via E. Fermi, 21020 Ispra (VA), Italy
[7]European Space Agency, Earth Observation Science & Applications, Via Galileo Galilei, Casella Postale 64, 00044 Frascati (Rm), Italy

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Assimilation of MERIS FAPAR into a terrestrial model

T. Kaminski et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

The terrestrial biosphere is currently a strong sink for anthropogenic CO_2 emissions. Through the radiative properties of CO_2 the strength of this sink has a direct influence on the radiative budget of the global climate system. The accurate assessment of this sink and its evolution under a changing climate is, hence, paramount for any efficient management strategies of the terrestrial carbon sink to avoid dangerous climate change. Unfortunately, simulations of carbon and water fluxes with terrestrial biosphere models exhibit large uncertainties. A considerable fraction of this uncertainty is reflecting uncertainty in the parameter values of the process formulations within the models.

This paper describes the systematic calibration of the process parameters of a terrestrial biosphere model against two observational data streams: remotely sensed FAPAR provided by the MERIS sensor and in situ measurements of atmospheric CO_2 provided by the GLOBALVIEW flask sampling network. We use the Carbon Cycle Data Assimilation System (CCDAS) to systematically calibrate some 70 parameters of the terrestrial biosphere model BETHY. The simultaneous assimilation of all observations provides parameter estimates and uncertainty ranges that are consistent with the observational information. In a subsequent step these parameter uncertainties are propagated through the model to uncertainty ranges for predicted carbon fluxes.

We demonstrate the consistent assimilation for two different set-ups: first at site-scale, where MERIS FAPAR observations at a range of sites are used as simultaneous constraints, and second at global scale, where the global MERIS FAPAR product and atmospheric CO_2 are used simultaneously. On both scales the assimilation improves the match to independent observations. We quantify how MERIS data improve the accuracy of the current and future (net and gross) carbon flux estimates (within and beyond the assimilation period).

We further demonstrate the use of an interactive mission benefit analysis tool built around CCDAS to support the design of future space missions. We find that, for long-term averages, the benefit of FAPAR data is most pronounced for hydrological

Assimilation of MERIS FAPAR into a terrestrial model

T. Kaminski et al.

quantities, and moderate for quantities related to carbon fluxes from ecosystems. The benefit for hydrological quantities is highest for semi-arid tropical or sub-tropical regions. Length of mission or sensor resolution is of minor importance.

1 Introduction

5 The terrestrial biosphere is a significant sink for atmospheric CO_2 and thus plays a key role in the radiative budget of the global climate system (Denman et al., 2007). Prognostic terrestrial vegetation models are used to simulate the strength and distribution of this sink and its response to climate change. These prognostic models solve the equations governing the evolution of the carbon, water, and energy balance. In their

10 formulation, these equations rely on a set of constants, which we call process parameters. There is uncertainty in both the correct formulation of the equations and then the correct values of the process parameters. This uncertainty yields significant uncertainties in the simulated terrestrial carbon sinks on decadal and longer time scales (Denman et al., 2007). On shorter time scales parameter uncertainty is reflected in

15 large uncertainties in the hydrological cycle on all spatial scales.

The use of observational information is required to reduce this uncertainty. Systematic model calibration through inversion procedures can infer parameter ranges that are consistent with the observations and exclude parameter ranges that are inconsistent with observations (Tarantola, 1987). Remaining inconsistencies can be attributed to

20 weaknesses in the formulation of the model equations or errors in the observational data. For such calibration procedures it is desirable to use multiple data streams and sample at multiple locations and points in time. To assure consistency, it is then essential to impose all observational constraints simultaneously, an approach we call consistent assimilation. In a non-linear model, any step-wise inclusion of the observa-

25 tional information typically yields a suboptimal estimate of final parameter values, i.e. consistency with the observational information used in early steps is not assured.

T. Kaminski et al.

The first mathematically rigorous calibration of a prognostic terrestrial biosphere model was performed within the Carbon Cycle Data Assimilation System (CCDAS, http://CCDAS.org) built around the Biosphere Energy Transfer HYdrology scheme (BETHY, Knorr, 2000; Knorr and Heimann, 2001). CCDAS estimates the values of BETHY's process parameters including their uncertainty ranges and maps them onto simulated carbon and water fluxes. The system was first used with 20 yr of atmospheric carbon dioxide observations provided by the GLOBALVIEW flask sampling network (GLOBALVIEW-CO_2, 2008). The system evaluated the effect of this observational constraint on the 0 net and gross fluxes of CO_2 over the assimilation period (Rayner et al., 2005), and also on their predictions from years (Scholze et al., 2007) to decades (Rayner et al., 2011).

The above studies showed that the flask sampling data can only constrain part of BETHY's parameter space. Fortunately there is an ever-increasing set of observational constraints on the terrestrial biosphere becoming available. One of the requirements for assimilation of a given data stream is the capability to simulate (by a so-called observation operator) its counterpart in the model. For the assimilation of atmospheric carbon dioxide the role of the observation operator was taken by an atmospheric transport model (TM2, Heimann, 1995) that was coupled to BETHY.

A further observational constraint on the terrestrial biosphere is provided by "Fraction of Absorbed Photosynthetically Active Radiation" (FAPAR) (Gobron et al., 2008) products. FAPAR is an indicator of healthy vegetation, which exhibits a strong contrast in reflectance between the visible and the near-infrared domains of the solar spectrum (Verstraete et al., 1996). FAPAR products can thus be derived from observations provided by space-borne instruments, e.g. by ESA's Medium Resolution Imaging Spectrometer (MERIS). The extensions of CCDAS for assimilation of FAPAR are detailed by Knorr et al. (2010), who also demonstrate the consistent assimilation of FAPAR at multiple sites. These extensions include modules for simulation of hydrology and leaf phenology and, as observational operator, a two flux scheme of the radiative balance within the canopy.

Here we report on the first consistent assimilation of flask samples of atmospheric CO_2 and FAPAR at global scale, i.e. the simultaneous assimilation of both data streams. To limit the computation time in development, testing, and debugging, this challenging exploration of unchartered territory was conducted in BETHY's fast, coarse spatial resolution.

A further application of advanced assimilation systems that can propagate uncertainties from the observations to target quantities of interest is quantitative network design (QND). QND is particularly appealing because it can evaluate the benefit of hypothetical data streams based on their assumed uncertainty. Kaminski and Rayner (2008) describe the methodological framework and present a set of examples related to the global carbon cycle. Within CCDAS, the QND concept was applied to support the design of an active LIDAR mission sampling atmospheric CO_2 from space (Kaminski et al., 2010). For FAPAR assimilation at site-scale the concept was applied to evaluate the effect of modifications of sensor characteristics on uncertainties in current and future carbon fluxes (Knorr et al., 2008). This paper describes the development of an interactive mission benefit analysis (MBA) software tool based on the global version of CCDAS. The MBA tool quantifies the benefit of space missions in terms of their constraint on various carbon and water fluxes, and we demonstrate the effect of design aspects such as mission length and sensor resolution.

The remainder of the paper is organised as follows. Section 2 describes first CCDAS (Sect. 2.1) and then the MBA tool (Sect. 2.2). The observational data are presented in Sect. 3. Next, Sect. 4 describes the consistent assimilation of the MERIS FAPAR product at multiple sites (Sect. 4.1) and the consistent global-scale assimilation of MERIS FAPAR and atmospheric CO_2 (Sect. 3.2), and Sect. 5 presents our simulations for mission design. Finally, we draw conclusions and give perspectives in Sect. 6.

Assimilation of MERIS FAPAR into a terrestrial model

T. Kaminski et al.

2 Methods

2.1 CCDAS

The Carbon Cycle Data Assimilation System (CCDAS) is a variational assimilation system built around the Biosphere Energy Transfer HYdrology scheme. The system is described in full detail elsewhere (Scholze, 2003; Kaminski et al., 2003; Rayner et al., 2005; Scholze et al., 2007; Knorr et al., 2010).

In brief, BETHY, simulates carbon uptake and plant and soil respiration embedded within a full energy and water balance and phenology scheme (Knorr, 2000). The model is fully prognostic and is thus able to predict the future evolution of the terrestrial carbon cycle under a prescribed climate scenario. The process formulation distinguishes 13 plant functional types (PFTs) based on the classification by Wilson and Henderson-Sellers (1985). Each model grid cell can be populated by up to three different PFTs. Driving data (precipitation, minimum and maximum temperatures, and incoming solar radiation) were derived from a combination of available monthly gridded and daily station data (R. Schnur, personal communication, 2008) by a method by Nijssen et al. (2001).

As mentioned above, assimilation of atmospheric CO_2 requires, as an observation operator, an atmospheric transport model (TM2, Heimann, 1995) coupled to BETHY. CO_2 fluxes from processes not represented in BETHY, i.e. fossil fuel emissions, exchange fluxes with the ocean and emissions from land use change, were prescribed as in Scholze et al. (2007). The observation operator for FAPAR calculates the vertical integral of the absorbed photosynthetically active radiation (PAR) by healthy green leaves between the canopy top and the canopy bottom, divided by the incoming radiation. FAPAR thus equals the net PAR flux entering the canopy at the top (incoming minus outgoing) minus the net PAR flux leaving the canopy at the bottom (outgoing minus incoming, i.e. reflected from soil background), divided by the incoming PAR flux at the top of the canopy. The PAR flux is calculated by a two-flux scheme (Sellers, 1985), which takes into account soil reflectance, solar angle and incoming amount of

Title Page

Abstract   Introduction

Conclusions   References

Tables   Figures

◄   ►

◄   ►

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

diffuse radiation.

Equating satellite and model FAPAR means that given the same illumination conditions, the same number of photons enter the photosynthetic mechanism of the vegetation, even if some of the assumptions differ between BETHY and the model used to derive FAPAR (Gobron et al., 2000). It also means that FAPAR in the model is defined based on the assumption that the canopy consists only of photosynthesising plant parts (Pinty et al., 2009), which is consistent with the definition used for deriving the MERIS FAPAR product.

Assimilation of FAPAR required the extension of CCDAS by components included in BETHY for simulating hydrology and leaf phenology. In the previous CCDAS setup, these components were used in a preliminary assimilation step that provided input to CCDAS. This extension was necessary to allow consistent assimilation of FAPAR and atmospheric CO_2.

CCDAS implements a probabilistic inversion concept (see Tarantola, 1987) that describes the state of information on a specific physical quantity by a probability density function (PDF). The prior information on the process parameters is quantified by a PDF in parameter space and the observational information by a PDF in the space of observations. Their respective means are denoted by x_0 and d and their respective covariance matrices by $\mathbf{C}_0$ and $\mathbf{C}_d$, where $\mathbf{C}_d$ accounts for uncertainties in the observations as well as uncertainties from errors in simulating their counterpart (model error). If the prior and observational PDFs were Gaussian and the model linear, the posterior PDF would be Gaussian, too, and completely characterised by its mean x_{post} and covariance matrix $\mathbf{C}_{post}$. Further x_{post} would be the minimum of the following cost function:

$$J(x) = \frac{1}{2}[(M(x) - d)^T \mathbf{C}_d^{-1}(M(x) - d) + (x - x_0)^T \mathbf{C}_0^{-1}(x - x_0)] \; , \tag{1}$$

where $M(x)$ denotes the model operated as a mapping of the parameters onto simulated counterparts of the observations. Further $\mathbf{C}_{post}$ would be given by:

$$\mathbf{C}_{post}^{-1} = \mathbf{J}''(x_{post}) \; , \tag{2}$$

where $\mathbf{J}''(x_{\text{post}})$ denotes the Hessian matrix of J, i.e. the matrix composed of its second partial derivatives $\frac{\partial^2 J}{\partial x_i \partial x_j}$.

Our model is non-linear and we approximate the posterior PDF by a Gaussian with x_{post} as the minimum of Eq. (1) and $\mathbf{C}_{\text{post}}$ from Eq. (2).

The inverse step is followed by a second step, the estimation of a diagnostic or prognostic target quantity y. The corresponding PDF is approximated by a Gaussian with mean

$$y = N(x_{\text{post}}) \tag{3}$$

and covariance

$$\mathbf{C}_y = \mathbf{N}'(x_{\text{post}})\mathbf{C}_{\text{post}}\mathbf{N}'(x_{\text{post}})^T + \mathbf{C}_{y,\text{mod}} \, , \tag{4}$$

where $N(x)$ is the model operated as a mapping of the parameters onto the target quantity. In other words, the model is expressed as a function of the vector of its parameters x and returns a vector of quantities of interest, for example the uptake of carbon integrated over a region and time interval. The linearisation (derivative) of N around x_{post} is denoted by $\mathbf{N}'(x_{\text{post}})$ and also called Jacobian matrix. $\mathbf{C}_{y,\text{mod}}$ denotes the uncertainty in the simulation of y resulting from errors in N. If the model was perfect (a hypothetical case), $\mathbf{C}_{y,\text{mod}}$ would be zero, and only the first term would contribute to $\mathbf{C}_y$. Conversely, if all parameters were known to perfect accuracy (an equally hypothetical case), $\mathbf{C}_{\text{post}}$ would be zero and only the second term would contribute to $\mathbf{C}_y$.

The minimisation of Eq. (1) and the propagation of uncertainties are implemented in a normalised parameter space with Gaussian priors. The normalisation is such that parameter values are specified in multiples of their standard deviation, i.e. $\mathbf{C}_0$ is the identity matrix (for details see Kaminski et al., 1999; Rayner et al., 2005). In addition, for some bounded parameters a suitable variable transformation is included.

Technically, the minimisation of J is performed by a powerful iterative gradient algorithm, where, in each iteration, the gradient of J is used to define a new the search direction. The gradient (plus J itself) are efficiently evaluated by a single run of the

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

so-called adjoint code of J. The associated computational cost is independent of the number of parameters and is in the current case comparable to 3–4 evaluations of J. Likewise $\mathbf{J}''(x_{post})$ is evaluated by a single run of the derivative code of the adjoint code (Hessian code). Here the associated computational cost grows roughly linearly with the number of parameters (more precisely an affine function of the number of parameters). In the present case of 71 parameters the cost is comparable to about 60 evaluations of J. These numbers are only a rough indication of performance as they vary with platform, compiler, and even compiler flags. For performance numbers of the previous CCDAS implementation we refer to Kaminski et al. (2003). All CCDAS derivative code (adjoint, Hessian, Jacobian) is generated from the model code by the automatic differentiation tool Transformation of Algorithms in Fortran (TAF) (Giering and Kaminski, 1998). The Hessian code is generated by reapplying TAF to the adjoint code.

2.2 Mission benefit analysis

Our mission benefit analysis is based on the Quantitative Network Design (QND) methodology presented by Kaminski and Rayner (2008). The approach exploits the fact that the uncertainty propagation from the observations to the parameters (via Eq. 2) and then further to the target quantities (Eq. 4) can be performed independently from the parameter estimation. The requirements for the evaluation of $\mathbf{J}''$ in Eq. (2) are the data uncertainty ($\mathbf{C}_d$), the capability to simulate (expressed by $M(x)$) a counterpart of the data stream via an appropriate observational operator, and a reasonable parameter vector. We can then evaluate the benefit of hypothetical or planned observational data streams on the uncertainty reduction in relevant target quantities.

A QND system for mission benefit analysis was built around the extended CCDAS framework for global scale assimilation described in Sect. 4.2. The tool can combine prior information, flask samples of atmospheric carbon dioxide, and global coverage FAPAR within a single cost function (see Fig. 1). For the tool, the sensitivity of each data item (each observation of FAPAR and atmospheric CO_2) with respect to the process parameters was precomputed and stored for a run of 14 yr. These sensitivities are the

derivatives of $M(x)$ (see Eq. 1), which are evaluated for the optimal parameter vector x determined by the global scale assimilation run (see Sect. 4.2). To approximate the posterior parameter uncertainty (Eq. 2) resulting from a user-defined data uncertainty ($\mathbf{C}_d$ of Eq. 1), requires just matrix multiplications and a matrix inversion. In this inversion step, the user can choose the length of the mission. This will determine how many of the 14 yr of data for which sensitivities were precomputed are actually used in the assessment. Further, the user can choose to include or exclude the information from the atmospheric CO_2 observations.

Evaluation of Eq. (4) yields posterior uncertainties in a set of target quantities. The target quantities we offer are annual mean values of net ecosystem production (NEP), net primary production (NPP), evapotranspiration, and plant available soil moisture averaged over five years. Each of these quantities is available for six regions of the globe. The Jacobian matrix $\mathbf{N}'$ (of Eq. 4) representing the derivative of the target quantities with respect to the model parameters was also precomputed and stored. For this step, just as for the previous step, the tool only requires matrix multiplications.

In summary, all steps to assess a mission configuration from the precomputed CC-DAS output only involve matrix algebra. On a standard notebook these operations take only a few seconds, which enables the tool to run in interactive mode. The options for the configuration comprise the uncertainty in the FAPAR product, the length of the mission, and whether atmospheric CO_2 observations are included or excluded. A similar tool (including a web interface) was set up for the in situ network of the carbon cycle (see http://imecc.ccdas.org).

3 Observational data

3.1 MERIS FAPAR

We use FAPAR products derived from the Medium Resolution Imaging Spectrometer (MERIS) of the European Space Agency (ESA). At site scale we use the Level 2 FAPAR

land product for the period June 2002 to September 2003. We use daily data at the operational 1.2 km resolution, from which scenes consisting of 15 by 15 pixels have been processed. For each site, we selected a rectangular subset of the respective scene in such a way that it constitutes visually homogeneous land cover as identified through Google Earth images. For each site, the number of pixels within the respective rectangular study areas is indicated in Table 1, along with further site-properties. Note that data for the last site were used for validation purposes and therefore withheld from the assimilation procedure. As mentioned in Sect. 2 BETHY allows up to three PFTs (out of a total of 13) per site. Each of the selected sites contains two to three PFTs with a corresponding surface cover fraction, where the remainder corresponds to bare ground. In total, the above sites cover seven PFTs. For further details we refer to Knorr et al. (2010).

At global scale we use the Level 3 product for the period June 2002 to September 2003. The data were processed at ESA's Grid Processing on Demand (GPoD, http://gpod.eo.esa.int) facility on a global 0.5 degree grid in the form of monthly composites and then interpolated to the model's coarse resolution 10 by 8 degree grid.

For both scales, global and site-scale, we use an uncorrelated data uncertainty of 0.1 for the definition of $\mathbf{C}_d$ in Eq. (1) irrespective of how many pixels where used in the spatial averaging of the FAPAR pixels (Gobron et al., 2008).

3.2 Atmospheric CO_2

We use monthly mean values of atmospheric CO_2 concentrations provided by the GLOBALVIEW flask sampling network (GLOBALVIEW-CO_2, 2008). We use data for the period from 1999 to 2004 at two sites, Mauna Loa (MLO) and South Pole (SPO). We use the time series of residual standard deviations (RSD) from the compilation to assign a data uncertainty to the observations. We only use data from years when sufficient measurements are made to assign values without the gap-filling procedures in the GLOBALVIEW compilation.

4 Assimilation experiments

4.1 Site-scale assimilation

The consistent assimilation at site scale uses the MERIS FAPAR product at all sites (except the validation site) as simultaneous constraints. We give a summary of the main findings and refer for details to Knorr et al. (2010). The first point to mention is that only 38 of the global model's 72 parameters affect FAPAR at our eight study sites. This is due to two reasons: first, as mentioned, the sites cover only a subset of BETHY's PFTs. Second, FAPAR simulation is insensitive to some of BETHY's processes (e.g. the carbon balance in the soil), and, hence, to their parameters.

Table 2 summarises the assimilation results. At all sites, including the validation site where FAPAR data were not assimilated, the posterior fit to the observations (RMS post) is improved compared to the prior (RMS prior). The improvement is small for Loobos, Sodankylä and Zotino, where the prior agreement with the data is already good. Interestingly, of the two grass sites included, the posterior fit at the validation site (Hainich grass) is better than at the site included in the assimilation (Aardhuis). This validation confirms the quality of both the assimilation approach and the process formulations in the model.

In order to assess to what extent the MERIS FAPAR data helped to constrain simulations of vegetation-atmosphere carbon fluxes, we select as target quantities (i.e. as y in Eq. 4) the vector composed of annual mean net primary production (NPP) at each site, including the validation site. For these prognostic simulations we deliberately choose a period beyond the assimilation interval, namely from January 2001 to December 2003, which is almost twice as long as the assimilation interval. We did so in order to demonstrate a major strength of the process-based variational data assimilation technique: as previously shown by Scholze et al. (2007) it can propagate information to periods before and after the assimilation period.

The computed prior and posterior means and uncertainties of annual NPP are shown in Table 2. Relative change in NPP is shown as a percentage of the prior uncertainty,

which is computed at the optimal parameter point. The lowest NPP is found at the far northern sites, a rather low value also for Loobos and for the semi-arid Maun site, intermediate values for the temperate sites at Hainich and Aardhuis, and high values for the evergreen tropical site at Manaus. Prior uncertainties are considerable for Sodankylä, and moderate for the remaining sites.

The only site where there is a large relative change (over 100 % of prior uncertainty) in the simulated NPP is Manaus. We suspect that with either larger uncertainties for FAPAR or a more conservative screening algorithm to account for remaining effects by clouds or cloud shadows, the posterior NPP would be closer to the prior value. This would also mean much less uncertainty reduction (also quantified as a percentage of prior uncertainty) for Manaus, which here is shown as 34 %. The other sites where we find a considerable uncertainty reduction (by more than 10 %) are Aardhuis, a grass site, Hainich forest, and our validation site, Hainich grass. This demonstrates the capability of the assimilation approach to transfer observational information to an unobserved site.

4.2 Global-scale assimilation

The consistent assimilation at global scale uses both data streams, the MERIS FAPAR product and the atmospheric CO_2 observations, as simultaneous constraints. Figure 1 displays the flow of information in the forward sense, i.e. from process parameters to the cost (or misfit) function. As mentioned we use the computationally fast, 8 by 10 degree resolution with about 170 land grid cells. Our assimilation interval is the five year period from 1999 to 2004.

Several approaches to address the problem of bias in the FAPAR data product have been investigated. For the global-scale assimilation, we resolved to the following solution: we computed the average FAPAR over three years for each model grid cell and compared this value to the average observed value. We then multiplied the cover fraction of each PFT within the grid cell concerned by the ratio averaged observed FAPAR divided by average model FAPAR. If this ratio was above 1, which only occurred in very

Assimilation of MERIS FAPAR into a terrestrial model

T. Kaminski et al.

few grid cells, no correction was applied.

In order to investigate the occurrence of multiple minima, we started five minimisations of the cost function from different starting points including the prior parameter value. Out of these five minimisations, four find the same minimum. The minimisation starting from the prior parameter value takes 153 iterations to reduce the cost function J from from 4574 to 2829 and the norm of its gradient by more than eight orders of magnitude from 4×10^3 to 2×10^{-5}. At the minimum the respective contributions (see Eq. 1) of the prior term, the CO_2 observations, and the MERIS observations to the total cost function J are 124, 61, and 2644.

At both stations, MLO (left hand panel of Fig. 2) and SPO (left hand panel of Fig. 2) the fit to atmospheric CO_2 has improved considerably. Figure 3 displays the change in simulated FAPAR through the assimilation (posterior – prior) for four months of 2003. FAPAR is reduced over the Amazon forest, increased over Australia, and exhibits an increased seasonal cycle over East Asia and the North American high latitudes.

For validation of the calibrated model, i.e. the model with the posterior parameter values, we need independent information. This information is provided by flask samples of the atmospheric CO_2 concentration at extra sites withheld from our assimilation procedure. Figure 4 displays observed concentration (black) together with concentrations simulated with prior (blue) and posterior (red) parameter values for Point Barrow, a marine site in Alaska (left hand panel), and Izaña, a mountain site on the Canary Islands (right hand panel). We note that the posterior provides a considerably better fit than the prior, i.e. the validation confirms that the calibrated model performs better than the uncalibrated model.

The uncertainty reduction for the parameters is displayed in Fig. 5. Parameters 1 through 71 are control parameters of BETHY, while Parameter 72 is the initial atmospheric CO_2 concentration used by the transport model. Of the BETHY parameters, numbers 57 to 71 relate to the phenology model, which controls leaf area and thus has an immediate impact on simulated FAPAR. While the site-scale assimilation constrains the parameters outside the phenology model only marginally, in the global scale

assimilation of FAPAR or atmospheric CO_2 ten of these parameters show an uncertainty reduction of about 20 % or more.

Of more general interest are uncertainty reductions in target quantities such as predicted fluxes, because they are less specific to the model used than the process parameters. Here, we select net ecosystem production and net primary production (NEP and NPP) integrated over the period from 1999 to 2003 and six regions (Fig. 6). For all regions and both target quantities, we find a considerable degree of uncertainty reduction, where fluxes in Australia are somewhat less constrained by the data than it is the case for the other continents. It is interesting to note that, even though the observed atmospheric CO_2 is more closely related to the net atmosphere-biosphere flux (NEP) than to only one component of it (NPP), the impact of the data is to constrain NPP more than NEP compared to the prior case.

5 Mission benefit analysis

As a first example we analyse the individual information content in our two data streams (Fig. 7). We assume a long mission of 14 yr. For simulation of regional NEP (left hand panel) we note that the FAPAR constraint is marginal, and that most of the uncertainty reduction can be attributed to the atmospheric CO_2 observations. The same holds for NPP (right hand panel).

Interestingly the picture is reversed for hydrological target quantities, i.e. evapotranspiration (left hand panel) and plant available soil moisture (right hand panel). It appears that FAPAR is a powerful constraint for those parameters with a strong effect on hydrological fluxes while atmospheric CO_2 is powerful in constraining parameters with a strong effect on the carbon fluxes for the case of long-term averages.

The weak constraint of FAPAR on carbon fluxes is, mathematically, reflected by a sub-space within the overall parameter space that is at the same time crucial to simulate long-term carbon fluxes and either not at all or only weakly constrained by the MERIS FAPAR data (Fig. 7). In the first case the model simulated FAPAR data would

have zero sensitivity to this part of the parameter space, while in the second case the sensitivity would be only small. There is an important difference between both cases: unlike the zero sensitivity, the weak sensitivity can be compensated for by a reduced data uncertainty.

Such a reduced data uncertainty would correspond to a new hypothetical mission concept. We investigate two hypothetical sensor concepts: the first sensor has higher spatial resolution than the MERIS sensor and the second is a hypothetical sensor with ideal resolution. We reproduce the characteristics of the sensor with higher resolution by reducing the data uncertainty for FAPAR from 0.1 (corresponding to our data uncertainty for the MERIS sensor, see Sect. 3.1) to 0.05. For the sensor with ideal resolution, we use a data uncertainty of 0.001. We stress that this low value is selected to explore an extreme case, not a case we can hope to achieve in reality. Even if future instruments might allow considerably higher precision, the theoretical limitations imposed by radiative transfer through heterogeneous canopy would prevent data uncertainties as low as this.

Figure 9 shows the reduction in parameter uncertainty for the MERIS sensor and both hypothetical mission concepts. We see that while for some parameters the uncertainty reduction improves with sensor resolution, a large fraction of the parameters remains unobserved. Figure 10 shows the corresponding uncertainty reductions in annual NEP and NEP averaged over the mission period of 14 yr (note change of scale on y-axis). Indeed the uncertainty reduction improves only marginally with sensor resolution, i.e. the unobserved parameters are important for constraining these carbon fluxes.

We further studied the effect of mission length. Figure 11 indicates that for the hydrological target quantities the gain in uncertainty reduction through a mission length extension from 3 to 14 yr is hardly larger than 10 %.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

6 Conclusions and perspectives

The study demonstrates the potential of consistent assimilation of multiple data streams, i.e. as a simultaneous constraint on the process parameters of a terrestrial biosphere model. This is the first study to combine in a mathematically rigorous framework observed FAPAR and atmospheric CO_2.

The most novel result of this study is that the MERIS-derived FAPAR product can indeed be highly valuable and beneficial for local to global scale ecosystem, hydrology and carbon cycle modelling when applied within a data assimilation framework. This includes prognostic studies, where data from climate simulations are used and predictions are made beyond the period of observations. Validation of the calibrated model resulting from the assimilation against independent observations shows a clear performance improvement.

The systematic application of the mathematically rigorous uncertainty propagation capability implemented by CCDAS allows to support the design of space missions with maximised benefit expressed in terms of uncertainties of carbon or water fluxes. The study has developed an interactive mission benefit analysis (MBA) tool that allows instantaneous evaluation of a range of potential mission designs. Applying the MBA tool, the study showed that the benefit of FAPAR data is most pronounced for hydrological quantities, and moderate for quantities related to carbon fluxes from ecosystems. In semi-arid regions, where vegetation is strongly water limited, the constraint delivered by FAPAR for hydrological quantities was especially large, as documented by the results for Africa and Australia. Sensor resolution is less critical for successful data assimilation, and with even relatively short time series of only a few years, significant uncertainty reduction can be achieved.

We also note that the approach used here to constrain process parameters of a global model can be considered an automated procedure for scientific investigation of the processes the parameters represent. We further note that the approach of multi-data stream assimilation presented here could easily be extended to include more than

one data stream from remotely sensed products. Obvious candidates are land surface temperature from the Advanced Along-Track Scanning Radiometer (AATSR), surface soil moisture from the Soil Moisture and Ocean Salinity (SMOS) mission, and possibly column-integrated CO_2 observations. This would allow a rigorous assessment of the consistency of multiple data streams (as done here for FAPAR and atmospheric CO_2).

The complementary nature of existing and potential future data streams could be explored by an extension of the MBA tool. A prominent candidate observation would be a column-integrated CO_2 product. The MBA tool could be extended such that observational data uncertainty and sampling strategy for the mission are assessed in terms of the uncertainty reduction in the tool's target quantities, i.e. terrestrial carbon fluxes but also hydrological quantities. The tool's concept is, however, general and thus also applicable to other sensor types, such as RADAR (e.g. BIOMASS, SMOS, or the Advanced Orbiting Satellite, ALOS) or LIDAR (e.g. the Geoscience Laser Altimeter System, GLAS, on ICEsat), individually or combined.

While the study emphasised improvement of process parameters, the highly flexible structure of the variational approach allows, as a slight modification of the existing CCDAS framework, to devise a soil moisture monitoring system that adjusts state variables through time such as soil moisture instead of static parameters. If input data for the ecosystem model can be derived from near-real time sources such as weather forecasting analyses or satellite data, this could result in an effective operational monitoring system for soil moisture.

References

Denman, K., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P., Dickinson, R., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P., Wofsy, S., and Zhang, X.: Couplings between changes in the climate system and bio-geochemistry, in: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., M.Tignor, and Miller, H., chap. 7, Cambridge University Press, Cambridge, UK and New York, NY, USA, 2007. 10764

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM T. Math. Software, 24, 437–474, doi:10.1145/293686.293695, 1998. 10770

GLOBALVIEW-CO_2: Cooperative atmospheric data integration project – carbon dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado, also available on Internet via anonymous FTP to: ftp://ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW, 2008. 10765, 10772

Gobron, N., Pinty, B., Verstraete, M. M., and Widlowski, J.: Advanced vegetation indices optimized for up coming sensors: design, performance and applications, IEEE T. Geosc. Remote S., 38, 2489–2505, 2000. 10768

Gobron, N., Pinty, B., Aussedat, O., Taberner, M., Faber, O., Mlin, F., Lavergne, T., Robustelli, M., and Snoeij, P.: Uncertainty estimates for the FAPAR operational products derived from MERIS – impact of top-of-atmosphere radiance uncertainties and validation with field data, Remote Sens. Eviron., 112, 1871–1883, 2008. 10765, 10772

Heimann, M.: The Global Atmospheric Tracer Model TM2, Tech. Rep. 10, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 1995. 10765, 10767

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, 33–52, Springer-Verlag, New York, 2008. 10766, 10770

Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three dimensional global inverse model of the atmospheric transport, 2, Inversion of the transport of CO_2 in the 1980s, J. Geophys. Res., 104, 18555–18581, 1999. 10769

Kaminski, T., Giering, R., Scholze, M., Rayner, P., and Knorr, W.: An example of an automatic differentiation-based modelling system, in: Computational Science – ICCSA 2003, Interna-

tional Conference Montreal, Canada, May 2003, Proceedings, Part II, edited by: Kumar, V., Gavrilova, L., Tan, C. J. K., and L'Ecuyer, P., vol. 2668 of Lecture Notes in Computer Science, 95–104, Springer, Berlin, 2003. 10767, 10770

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, doi:10.1111/j.1600-0889.2010.00483.x, 2010. 10766

Knorr, W.: Annual and interannual CO_2 exchanges of the terrestrial biosphere: process-based simulations and uncertainties, Global Ecol. Biogeogr., 9, 225–252, 2000. 10765, 10767

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. 10765

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., and Giering, R.: Remote sensing input for regional to global CO_2 flux modelling, in: Proceedings of 2nd MERIS/(A)ATSR User Workshop, Frascati, Italy, 22–26 September 2008, European Space Agency, 2008. 10766

Knorr, W., Kaminski, T., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Carbon cycle data assimilation with a generic phenology model, J. Geophys. Res., 115, G04017, 16 pp., doi:10.1029/2009JG001119, 2010. 10765, 10767, 10772, 10773, 10783

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. 10767

Pinty, B., Lavergne, T., Widlowski, J.-L., Gobron, N., and Verstraete, M. M.: On the need to observe vegetation canopies in the near-infrared to estimate visible light absorption, Remote Sens. Environ., 113, 10–23, doi:10.1016/j.rse.2008.08.017, 2009. 10768

Rayner, P., 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, GB2026, 20 pp., doi:10.1029/2004GB002254, 2005. 10765, 10767, 10769

Rayner, P., Koffi, E., Scholze, M., Kaminski, T., and Dufresne, J.: Constraining predictions of the carbon cycle using data, Phil. T. R. Soc. A, 369, 1955–1966, 2011. 10765

Scholze, M.: Model studies on the response of the terrestrial carbon cycle on climate change and variability, Examensarbeit, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 2003. 10767

Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic CCDAS simulations, J. Geophys. Res., 112, 13 pp.,

doi:10.1029/2007JD008642, 2007. 10765, 10767, 10773

Sellers, P. J.: Canopy reflectance, photosynthesis and transpiration, Int. J. Remote Sens., 6, 1335–1372, 1985. 10767

Tarantola, A.: Inverse Problem Theory – Methods for Data Fitting and Model Parameter Estimation, Elsevier Sci., New York, 1987. 10764, 10768

Verstraete, M. M., Pinty, B., and Myneni, R. B.: Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing, Remote Sens. Environ., 58, 201–214, 1996. 10765

Wilson, M. F. and Henderson-Sellers, A.: A global archive of land cover and soils data for use in general-circulation climate models, J. Climatol., 5, 119–143, 1985. 10767

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. List of sites for assimilation taken from Knorr et al. (2010), showing central coordinates, elevation in m, N-S and E-W extent in km of the rectangular satellite scenes, and n the number of daily data points after spatial averaging. The site in the last row has been included for validation only.

Site	Description	Country	Latitude	Longitude	Elevation	N-S	E-W	n
Sodankylä	Boreal evergreen forest	Finland	67.3619° N	26.6378° E	180	1.2	1.2	80
Zotino	Boreal mixed forest	Russia	60.8008° N	89.2657° E	116	1.2	1.2	101
Aardhuis	C3 grassland	Netherlands	52.2381° N	5.8672° E	7	1.2	1.2	91
Loobos	Temperate pine forest	Netherlands	52.1679° N	5.7440° E	25	1.2	1.2	103
Hainich forest	Temperate deciduous forest	Germany	51.0793° N	10.4520° E	430	1.2	1.2	106
Manaus	Tropical rainforest	Brazil	2.5892° S	60.1311° W	80	18.0	14.4	146
Maun	Tropical savanna	Botswana	19.9155° S	23.5605° E	940	3.6	3.6	154
Hainich grass	C3 grassland	Germany	51.0199° N	10.4348° E	302	2.4	1.2	119

Table 2. Mean annual prior and posterior NPP for the period 2000–2003 (inclusive) with change relative to prior uncertainty, prior and posterior RMS fits to observations, prior and posterior uncertainties and relative uncertainty reduction. Units are $gC\,m^{-2}\,yr^{-1}$ or percentage when stated.

Site	prior NPP	post. NPP	rel. change (%)	RMS prior	RMS post	prior unc.	post. unc.	unc. reduction (%)
Sodankylä	137	151	68	0.085	0.080	112	98	5
Zotino	201	216	54	0.081	0.072	28	28	0
Aardhuis	853	842	−7	0.336	0.267	164	101	38
Loobos	449	424	−40	0.088	0.085	62	59	5
Hainich forest	689	657	−29	0.198	0.147	112	98	13
Manaus	1465	964	−196	0.250	0.130	255	168	34
Maun	350	346	−10	0.105	0.052	50	46	8
Hainich grass	619	786	97	0.280	0.165	172	89	48
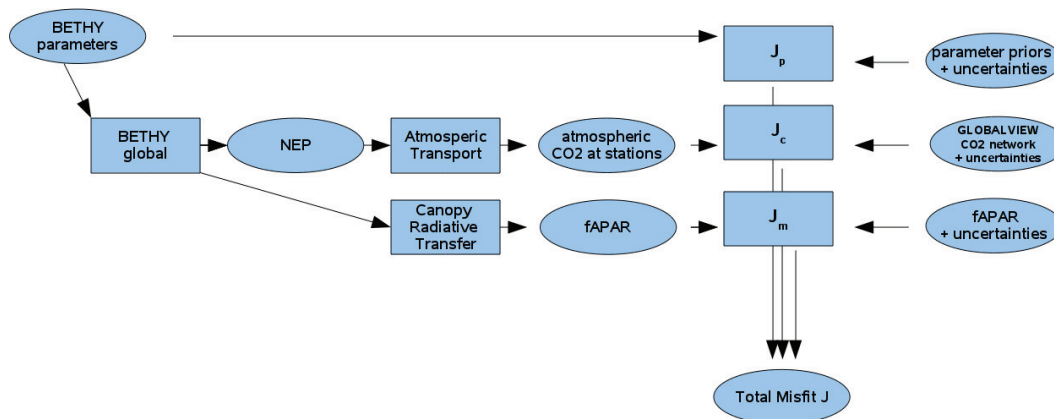
Fig. 1. Flow of information for evaluation of the cost function. J is the sum of the cost function contributions from the individual information items. Ovals denote data and rectangular boxes denote processing (i.e. code modules).
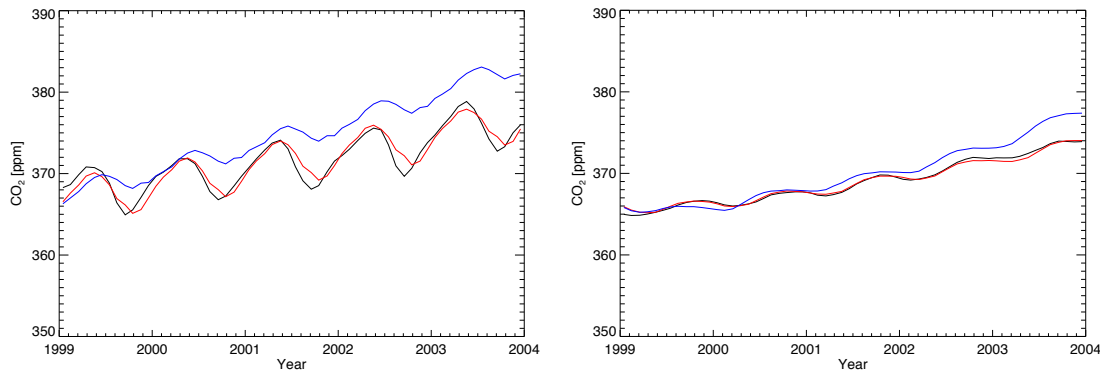
Fig. 2. Atmospheric CO_2 at Mauna Loa (left hand panel) and South Pole (right hand panel) in ppm: Observations (black), prior (blue), and posterior (red).
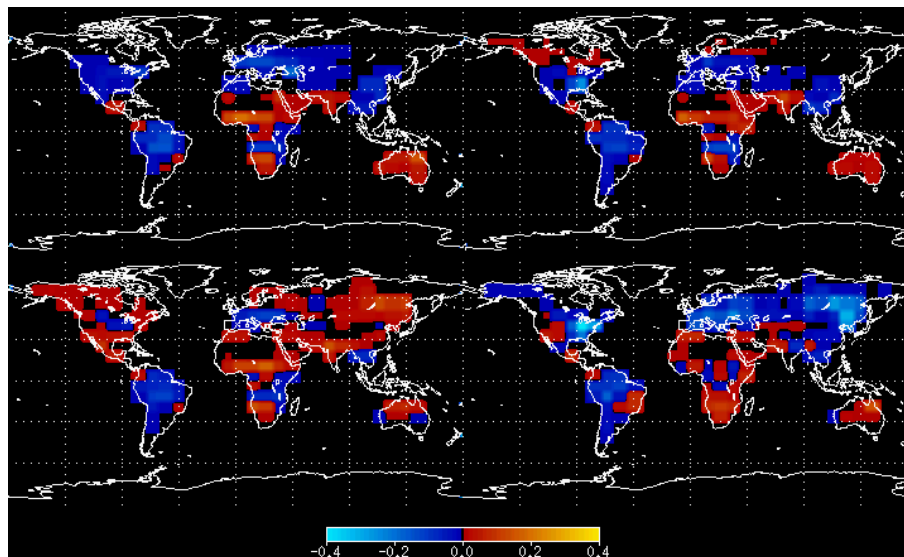
Fig. 3. Posterior-prior FAPAR for 4 months in 2003: January (upper left panel), April (upper right panel), July (lower left panel), and October (lower right panel).
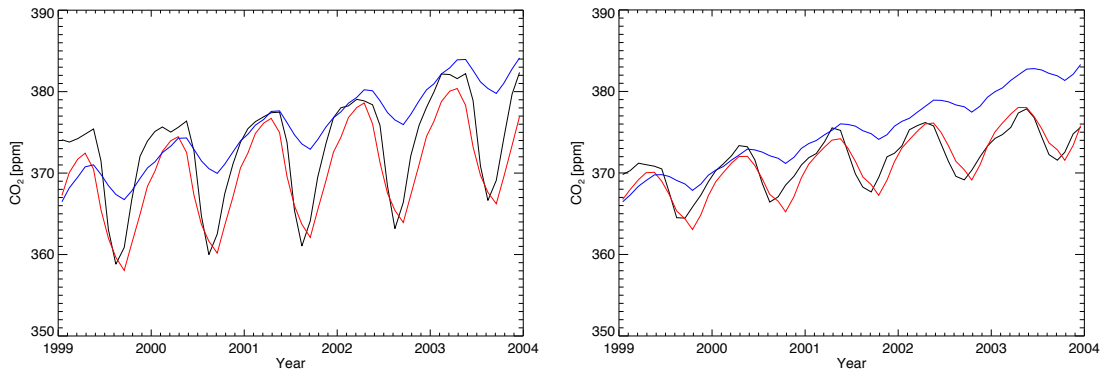
Fig. 4. Atmospheric CO_2 at Point Barrow (left hand panel) and Izaña (right hand panel) in ppm: observations (black), prior (blue), and posterior (red).
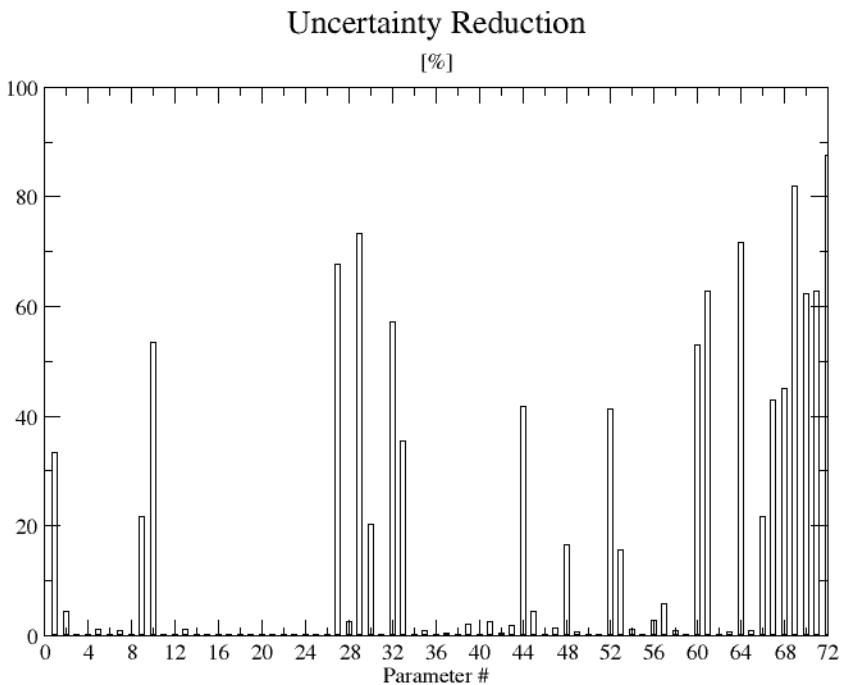
Fig. 5. Uncertainty reduction in process parameters.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
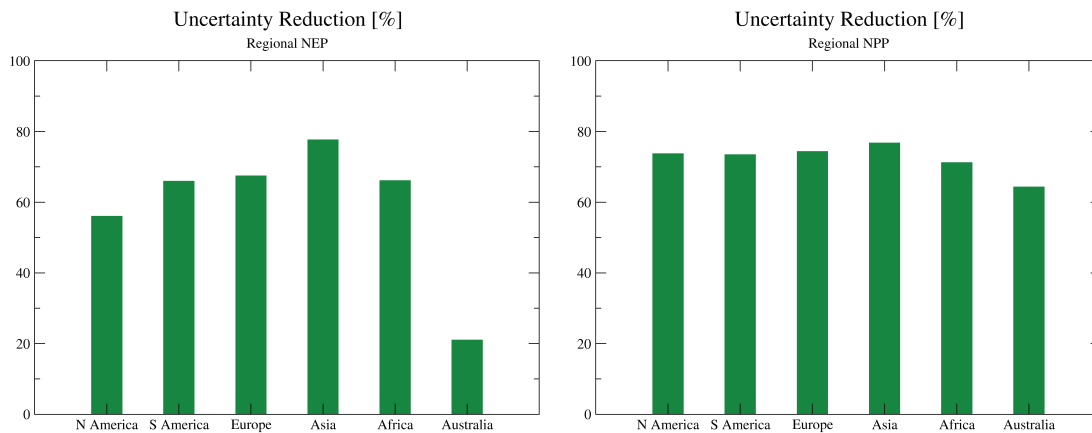
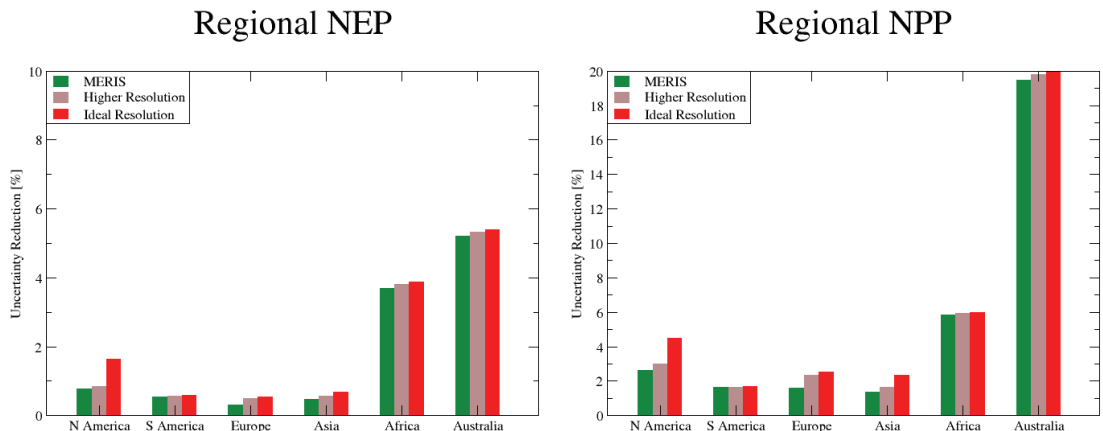**Fig. 6.** Uncertainty reduction in simulated NEP (left hand panel) and NPP (right hand panel) over six regions.

Fig. 7. Reduction in uncertainty in NEP (left hand panel) and NPP (right hand panel) over six regions from MERIS sensor for a 14-yr mission. For assimilation of CO_2 (red) and FAPAR (brown) separately and jointly (green).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Fig. 8. Reduction in uncertainty in evapotranspiration (left hand panel) and plant available soil moisture (right hand panel) over six regions from MERIS sensor for a 14-yr mission. For assimilation of CO_2 (red) and FAPAR (brown) separately and jointly (green).

Fig. 9. Reduction in parameter uncertainty for a 14-yr mission for FAPAR data from the MERIS sensor (left hand panel) a hypothetical higher resolution sensor (middle panel) and from a hypothetical ideal resolution sensor (right hand panel).

Fig. 10. Reduction in uncertainty in NEP (left hand panel) and NPP (right hand panel) over six regions from three sensor concepts: the MERIS sensor (green), the higher resolution sensor (brown), and the ideal resolution sensor (red).

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

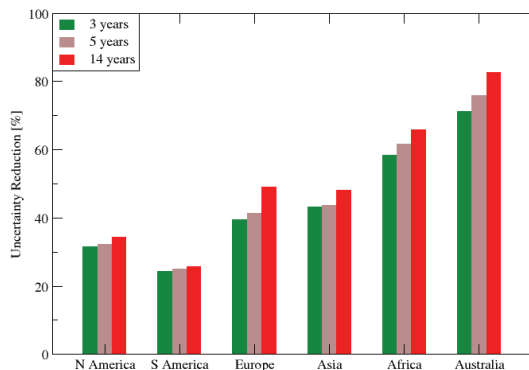Regional Evapotranspiration
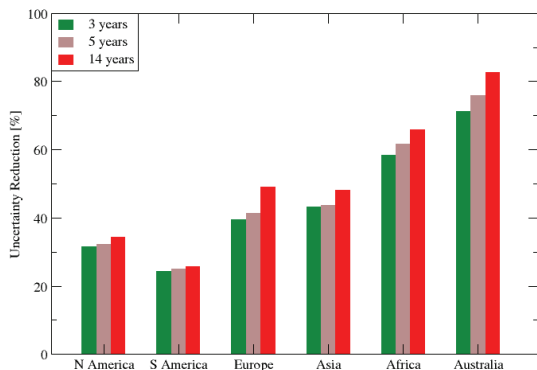
Regional Evapotranspiration

Fig. 11. Reduction in uncertainty in evapotranspiration (left hand panel) and plant available soil moisture (right hand panel) over six regions from MERIS sensor for a mission length of 3 yr (green), 5 yr (brown) and 14 yr (red).