ASSIMILATION OF MERIS FAPAR INTO A TERRESTRIAL VEGETATION MODEL AND MISSION DESIGN

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

[1]*FastOpt GmbH, Schanzenstraße 36, 20357 Hamburg, Germany, Email: Thomas.Kaminski@FastOpt.com*
[2]*QUEST – Dep. of Earth Sciences, Wills Memorial Building, University of Bristol, Bristol, BS8 1RJ, U.K., Email: wolfgang.knorr@bristol.ac.uk*
[3]*Dep. of Earth Sciences, Wills Memorial Building, University of Bristol, Bristol, BS8 1RJ, U.K., Email: marko.scholze@bristol.ac.uk*
[4]*European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, Global Environment Monitoring Unit, TP 272, via E. Fermi, 21020 Ispra (VA), Italy, Email: nadine.gobron@jrc.ec.europa.eu*
[5]*European Commission, DG Joint Research Centre, Institute for Environment and Sustainability, Global Environment Monitoring Unit, TP 272, via E. Fermi, 21020 Ispra (VA), Italy, Seconded to the Earth Observation Directorate, ESA-ESRIN, via Galileo Galilei, 00044 Frascati, Italy, Email: bernard.pinty@jrc.ec.europa.eu*
[6]*FastOpt GmbH, Schanzenstraße 36, 20357 Hamburg, Germany, Email: Ralf.Giering@FastOpt.com*
[7]*European Space Agency, Earth Observation Science & Applications, Via Galileo Galilei, Casella Postale 64, 00044 Frascati (Rm), Italy, Email: Pierre.Philippe.Mathieu@esa.int*

ABSTRACT

The current and future strength of the terrestrial carbon sink has a crucial influence on the expected degree of climate warming humanity is going to face. Usually, Earth Observation (EO) by its very nature focuses on diagnosing the current state of the planet. However, it is possible to use EO products in data assimilation systems to improve not only the diagnostics of the current state, but also the accuracy of future predictions.

This contribution reports from an on-going ESA funded study (see http://rs.ccdas.org) in which the MERIS FAPAR product is assimilated into a terrestrial biosphere model within the global Carbon Cycle Data Assimilation System (see http://CCDAS.org). Using methods of variational data assimilation, CCDAS relies on first and second derivatives of the underlying model for estimating process parameters with uncertainty ranges. In a subsequent step these parameter uncertainties are mapped forward onto uncertainty ranges for predicted carbon fluxes.

In this contribution, we quantify how MERIS data improve the accuracy of the current and future (net and gross) carbon flux estimates for a range of sites spanning the major biomes of the globe. We further present first assimilation experiments of MERIS FAPAR at the global scale together with in situ observations of atmospheric CO_2 in a coarse-resolution setup of CCDAS and address the systematic application of CCDAS for the design of future space missions.

Key words: Data Assimilation, Carbon Cycle, MERIS, mission design.

1. INTRODUCTION

The Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC) has shown that, while the terrestrial biosphere appears to have been a significant sink for atmospheric CO_2 during the past, its further ability to take up atmospheric CO_2 could be reduced by the effect of climate change (*Denman et al.*, 2007). Current projections of the terrestrial carbon sink exhibit large uncertainties, which to a considerable extent can be attributed to uncertainties in the values of the parameters in the process representations of the underlying models. Systematic calibration of these models against a range of observations of the carbon cycle can narrow down these uncertainties.

A mathematically rigorous assessment of the uncertainty reduction in decadal-scale predictions of carbon fluxes via current observations was first demonstrated by *Scholze et al.* (2007). They constrained the prognostic Biosphere Energy Transfer HYdrology scheme (BETHY, *Knorr* (2000); *Knorr and Heimann* (2001)) by 20 years of atmospheric carbon dioxide observations provided by the GLOBALVIEW flask sampling network (*GLOBALVIEW-CO_2*, 2004). The study uses the Carbon Cycle Data Assimilation System (CCDAS, http://CCDAS.org) and shows, among other things, that their observing system can constrain only part of BETHY's parameter space.

Fortunately an ever increasing number of data streams suitable for observing the carbon cycle is now becoming available. For instance, the presence of healthy vegetation can be captured well from space, because it exhibits a strong contrast in reflectance between the visible and the near-infrared part of the solar spectrum (*Verstraete et al.*, 1996). Measurements from ESA's Medium Resolution Imaging Spectrometer (MERIS) can be used to compute the Fraction of (vegetation) Absorbed Photosynthetically Active Radiation (FAPAR) (*Gobron et al.*, 2008). The assimilation of FAPAR into CCDAS requires an extension of the system as detailed by *Knorr et al.* (2010), who also present the assimilation of the MERIS FAPAR product at a set of sites spanning the major biomes of the globe.

The obvious next step, and the focus of the present contribution, is the simultaneous assimilation of the MERIS FAPAR product and atmospheric CO_2 data from the flask sampling network at the global scale. This is the first time such an exercise has ever been carried out and the present study therefore explores unchartered territory associated with a set of technological and scientific challenges. We hence opted for a fast, coarse spatial resolution, which greatly facilitated development, testing, and debugging.

The layout of this contribution is as follows: We first describe CCDAS (Section 2) followed by the observational data (Section 3) and the results of the two sets of data assimilation experiments (Section 4). This latter section, after presenting a summary for the site scale, also reports on the results of the first global-scale assimilation, and subsequently discusses the use of CCDAS as a generic tool for mission design. Finally, Section 5 draws various conclusions related to carbon cycle observation and options for related mission design.

2. DESCRIPTION OF CCDAS

The set-up, data and models used in CCDAS have been described by *Scholze* (2003), *Rayner et al.* (2005), *Scholze et al.* (2007), and *Knorr et al.* (2010) to which we refer for details. In brief, BETHY, the core CCDAS model, is a process-based model of the terrestrial biosphere (*Knorr*, 2000). It simulates carbon uptake and plant and soil respiration embedded within a full energy and water balance and phenology scheme. BETHY is a fully prognostic model, and is thus able to predict the future evolution of the terrestrial carbon cycle under a prescribed climate scenario. Global vegetation is mapped onto 13 plant functional types (PFT) based on *Wilson and Henderson-Sellers* (1985). Each grid cell of arbitrary size can contain up to three different PFTs, with the amount specified by their fractional coverage. The model is run with daily precipitation, minimum and maximum temperatures and incoming solar radiation. The data were generated through a combination of available monthly gridded and daily station data [*R. Schnur, pers. comm.*] by a method by *Nijssen et al.* (2001), using gridded data from the Summary of the Day Observations (Global CEAS), National Climatic Data Center and the latest updates of gridded data by *Jones et al.* (2001) and *Chen et al.* (2002) and using the available data nearest to the site.

Assimilation of atmospheric CO_2 requires an atmospheric transport model (TM2, *Heimann* (1995)) coupled to BETHY, as well as additional background 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. We use the same background fluxes as *Scholze et al.* (2007).

The assimilation of FAPAR data in CCDAS is based on minimisation of the difference between satellite and model-derived FAPAR (*Knorr et al.*, 2010). Within BETHY, FAPAR is calculated as the vertical integral of absorption of photosynthetically active radiation by healthy green leaves divided by the difference between the incoming and outgoing radiation flux at the top and bottom of the canopy. This integration is carried out by a two-flux scheme, which takes into account soil reflectance, solar angle and amount of 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 only with respect to the absorption by photosynthesising plant parts (*Pinty et al.*, 2009), which is consistent with the definition used for deriving the MERIS FAPAR product.

Previous CCDAS implementations used a two-stage inversion procedure, where BETHY's sub-models for soil moisture and phenology where split off the core CCDAS model and the assimilation of FAPAR derived from satellite data was carried out in pre-step. Here, the full BETHY model is included in the CCDAS framework as detailed by (*Knorr et al.*, 2010). CCDAS then allows the rigorous propagation of uncertainties as described by *Kaminski et al.* (2002, 2003); *Rayner et al.* (2005) and demonstrated by *Kaminski et al.* (2002); *Rayner et al.* (2005); *Scholze et al.* (2007). It uses a probabilistic framework, described in detail by *Tarantola* (1987) or *Enting* (2002), who also gives an exhaustive overview on applications to biogeochemistry.

The state of information on a specific physical quantity is conveniently formulated in terms of a probability density function (PDF). The prior information is quantified by a PDF in the space of control variables (here: process parameters of BETHY and the initial atmospheric CO_2 concentration), 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 $\mathbf{C_d}$. Note that $\mathbf{C_d}$ has to account for uncertainties in the observations and uncertainties from errors in simulating their counterpart. We approximate the posterior PDF by a Gaussian with mean x_{post} and covariance matrix $\mathbf{C_{post}}$. The mean is the minimum of the following cost function:

$$J(x) = \frac{1}{2}[(M(x)-d)^T \mathbf{C_d}^{-1}(M(x)-d) + (x-x_0)^T \mathbf{C_0}^{-1}(x-x_0)] \ , \qquad (1)$$

where $M(x)$ denotes the model operated as a mapping of the control variables onto simulated counterparts of the observations. In practice, the minimisation of J is performed iteratively by a gradient algorithm, where the search direction is determined via the gradient of J, evaluated by adjoint code. The use of adjoint model code greatly enhances computational performance of the non-linear optimisation.

We approximate the covariance matrix of the model parameters as

$$\mathbf{C_{post}}^{-1} = \mathbf{H}(x_{post}) \ , \qquad (2)$$

where $\mathbf{H}(x_{post})$ denotes the Hessian matrix of J, i.e. the matrix composed of its second partial derivatives $\frac{\partial^2 J}{\partial x_i \partial x_j}$. Since the dimension of x_{post} never exceeds a few hundred, it is computationally feasible to evaluate the full Hessian by running efficient second derivative code.

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

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

and covariance

$$\mathbf{C_y} = \mathbf{N}'(x_{post})\mathbf{C_{post}}\mathbf{N}'(x_{post})^T + \mathbf{C_{y,mod}} \ , \qquad (4)$$

where $N(x)$ is the model operated as a mapping of the control variables 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 rate of photosynthesis at some desired time step. $\mathbf{N}'(x_{post})$, the Jacobian matrix of N, is its linearisation around x_{post}, and $\mathbf{C_{y,mod}}$ is the uncertainty in the simulation of y resulting from errors in the model. In the hypothetical case of a perfect model, only the first term would contribute to $\mathbf{C_y}$. On the other hand, if the control variables were known to perfect accuracy, only the second term would contribute to $\mathbf{C_y}$.

The minimisation of Equ. (1) and the propagation of uncertainties are implemented in a normalised parameter space with Gaussian prior. 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. We further assume that Hessian Eigenvalues less than 1 reflect small-scale noise. To remove this noise Hessian Eigenvalues around 1 are set to 1 as in the procedure detailed by *Rayner et al.* (2005).

All CCDAS derivative code is generated from the model code by the automatic differentiation tool Transformation of Algorithms in Fortran (TAF) (*Giering and Kaminski*, 1998).

3. OBSERVATIONAL DATA

3.1. MERIS FAPAR

We assimilate daily data from the Level 2 FAPAR land product derived from the Medium Resolution Imaging Spectrometer (MERIS) of the European Space Agency (ESA). For the site-scale assimilation we used the operational resolution of 1.2 km for the period June 2002 to September 2003, from which square 15 by 15 pixel scenes have been processed. Each site consists of a rectangular study area over one to several satellite pixels as described in Table 1. The last site shown in the table has been included for validation purposes and is therefore excluded from the data assimilation exercise. The areas were chosen in such a way that they constitute homogeneous land cover as identified through Google Earth images. BETHY represents the vegetation of each site by two to three PFTs and a corresponding surface cover fraction, where the remainder corresponds to bare ground. Out of the total of 13 PFTs of the global version of CCDAS, seven distributed over eight sites are included in the simultaneous assimilation. For details we refer to *Knorr et al.* (2010).

The global-scale assimilation is based on the Level 2 product provided by 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 averages for the period June 2002 to September 2003. These data have been interpolated to the a 10 by 8 degree global grid corresponding to the model's coarse resolution setup.

For both scales, global and site-scale, we use an uncorrelated data uncertainty of 0.1 irrespective of how many pixels where used in the spacial averaging of the FAPAR pixels (*Gobron et al.*, 2008). Thus, $\mathbf{C_d}$ in Equ. (1) contains only diagonal elements with values of either 0.1^2, or ∞ if no data are available for the day and site concerned (in practice set to a very large value).

3.2. Atmospheric CO_2

In the case of the global-scale application, we assimilate the FAPAR product jointly with monthly mean values of the atmospheric CO_2 concentration provided by the GLOBALVIEW flask sampling network (*GLOBALVIEW-CO_2*, 2008).

4. ASSIMILATION EXPERIMENTS

4.1. Site-Scale Assimilation

The site-scale assimilation is discussed in full detail by *Knorr et al.* (2010). Here we give a brief summary. Of the global model's 72 parameters, only 38 affect FAPAR

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

when BETHY is run at the eight selected sites of this study. For example, parameters controlling the carbon balance in the soil have no impact on simulated leaf area and FAPAR.

The assimilation led to an improvement of the fit at all sites, including the one where FAPAR data were not assimilated. The smallest improvement is found (in this order) for Loobos, Sodankylä and Zotino. In these cases, simulated FAPAR changes only slightly from prior to posterior case. Here prior agreement with the data is already good. It is interesting to note that of the two grass sites included, the fit using optimised parameters is better at the one not included in the assimilation than at the one included.

To quantify the observational constraint of the MERIS data on the model, we computed the posterior parameter uncertainty as a percentage of the prior parameter uncertainty. Out of the 14 parameters in the phenology model, five had a posterior uncertainty reduced by more than 50% compared to the prior, and nine had a posterior uncertainty reduced by more than 33% compared to the prior. The remaining 24 model parameters showed only marginal uncertainty reductions, with only one of them exceeding 10%.

In order to assess to what extent the MERIS FAPAR data helped to constrain simulations of vegetation-atmosphere carbon fluxes, we select annual mean net primary production (NPP) at each site as the set of target quantities (i.e. as y in Equ. (4)), including the site for which no FAPAR data were assimilated. The period chosen for those prognostic simulations is January 2001 to December 2003, which is almost twice as long as the period for which FAPAR data are available. Inferring information for outside the "diagnostic" period is a major strength of the process-based data assimilation technique, as demonstrated before by *Knorr and Kattge* (2005) and *Scholze et al.* (2007).

The computed prior and posterior means and uncertainties of annual NPP are shown in Table 2. Relative change in NPP is again shown as a fraction 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, owing to the low value of V_{max}^{25} for PFT 5, 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 ones.

The only site where there is a large relative change (absolute value greater than two) 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 error reduction 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 Hainich grass, the latter not included in the data assimilation.

4.2. Global-Scale Assimilation

For the global case, we assimilate the MERIS FAPAR product in conjunction with the atmospheric CO_2 product provided by the GLOBALVIEW flask sampling network (*GLOBALVIEW-CO$_2$*, 2008). As mentioned before this first set of experiments uses the fast coarse resolution of CCDAS (8 by 10 degree). We also restrict the simulation period to five years. For the final setup used for mission design we plan to use the standard resolution of CCDAS of 2 by 2 degrees, with a number of grid cells increased by about a factor of 20. Since we assimilate the average FAPAR value on the model grid, the number of FAPAR data points entering the cost function is by about a factor 20 smaller for the coarse-grid compared to the standard setup. In order to have about the same ratio of FAPAR to CO_2 observations than in the standard setup, we also reduce the number of CO_2 data points entering the cost

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

Site	prior NPP	post. NPP	rel. change [%]	prior unc.	post. unc.	unc. reduction [%]
Sodankylä	137	151	68	112	98	5
Zotino	201	216	54	28	28	0
Aardhuis	853	842	-7	164	101	38
Loobos	449	424	-40	62	59	5
Hainich forest	689	657	-29	112	98	13
Manaus	1465	964	-196	255	168	34
Maun	350	346	-10	50	46	8
Hainich grass	619	786	0.97	172	89	48

function by about a factor of 20 by selecting only Mauna Loa (MLO) and South Pole (SPO) out of the 41 observational sites used in the standard setup. This means we assimilate simultaneously over 3000 FAPAR observations (18 months in more than 170 grid cells) and 120 CO_2 observations (60 months at two sites). We further adjust the default PFT fractions by a scaling factor per grid cell such as to match the observed long-term mean FAPAR.

The minimisation of Equ. (1) is carried out five times from different starting points, of which four runs find the same minimum. The minimisation starting from the prior value takes 153 iterations to reduce the cost function J from from 4574 to 2830 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 of the prior term, the CO_2 observations, and the MERIS observations are 124, 61, and 2644.

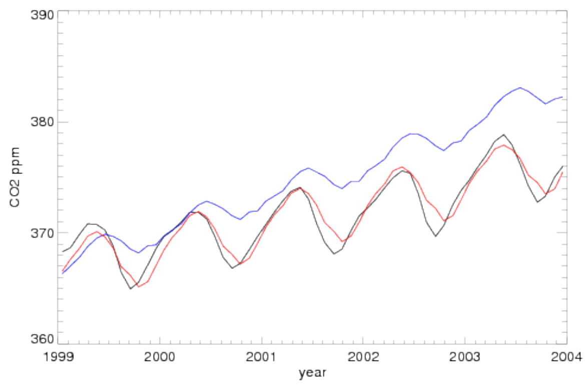

Figure 1. Atmospheric CO_2 at Mauna Loa in ppm: Observations (black), prior (blue), and posterior (red).

The fit to atmospheric CO_2 at MLO (Fig. 1) and at SPO (Fig. 2) has improved considerably. The change in simulated FAPAR through the assimilation (posterior – prior) is displayed in Fig. 3 for four months of 2003. We note a reduction over the Amazon forest, an increase over Australia, and an increased seasonal cycle over East Asia and

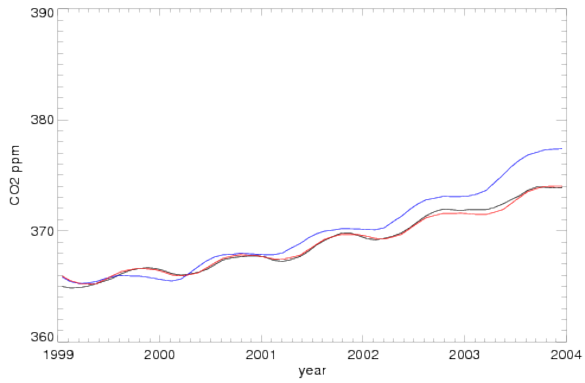

Figure 2. Atmospheric CO_2 at South Pole in ppm: Observations (black), prior (blue), and posterior (red).

the North American high latitudes.

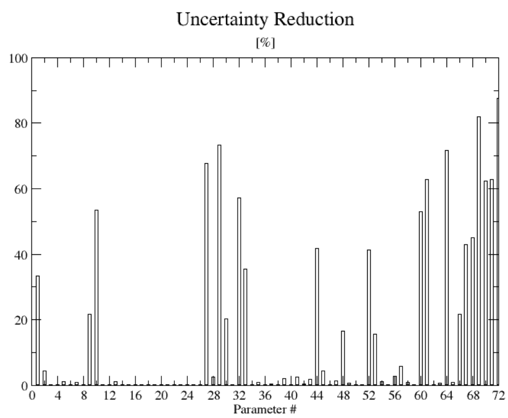

Figure 4. Uncertainty reduction in process parameters.

The uncertainty reduction for the parameters is displayed in Fig. 4. Parameters 1 through 71 are control parame-

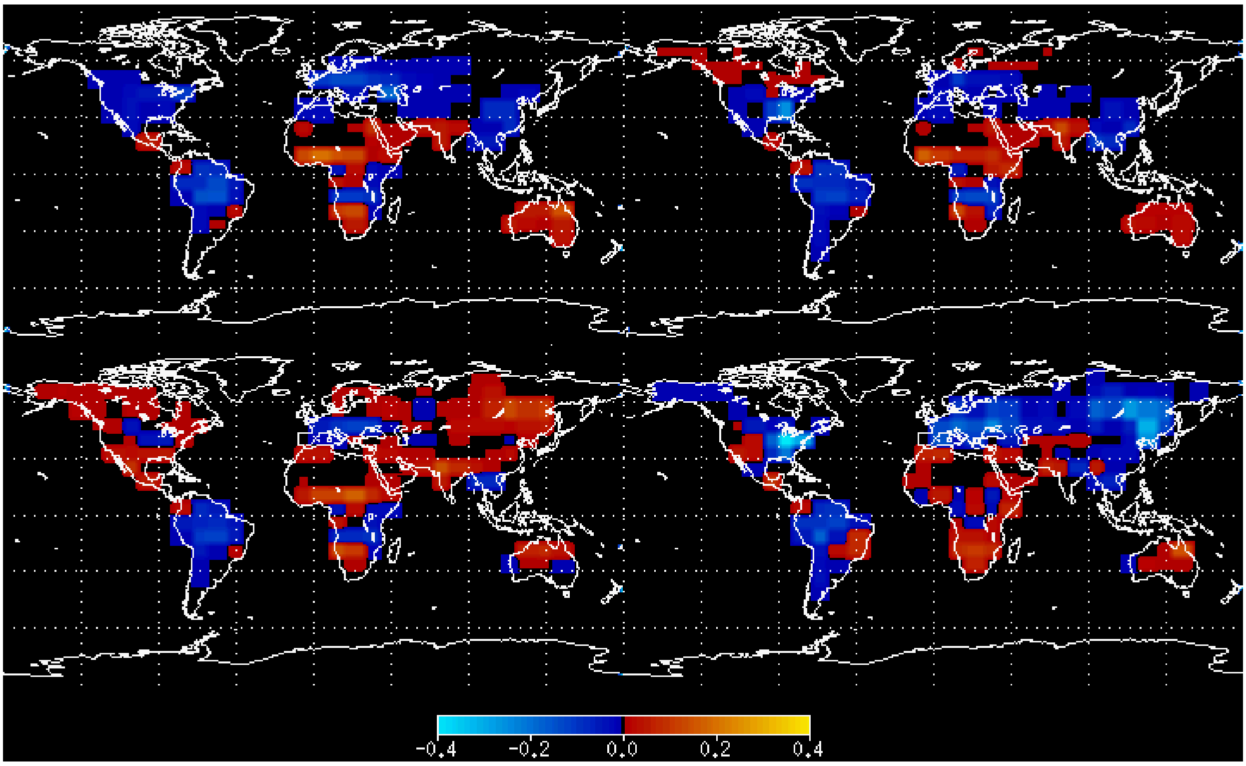

Figure 3. Posterior-prior FAPAR for 4 months in 2003: January (upper left panel), April (upper right panel), July (lower left panel), and October (lower right panel).

ters 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 case ten of these parameters show an uncertainty reduction of 20% or more.

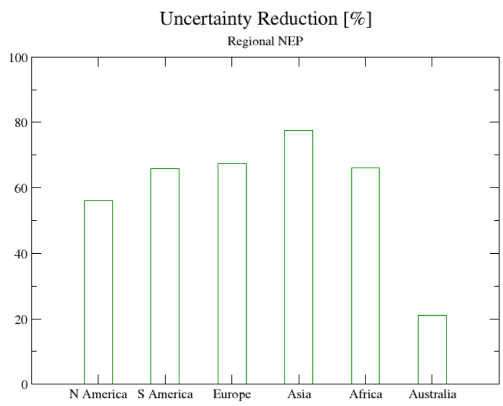

Figure 5. Uncertainty reduction in simulated NEP over six regions.

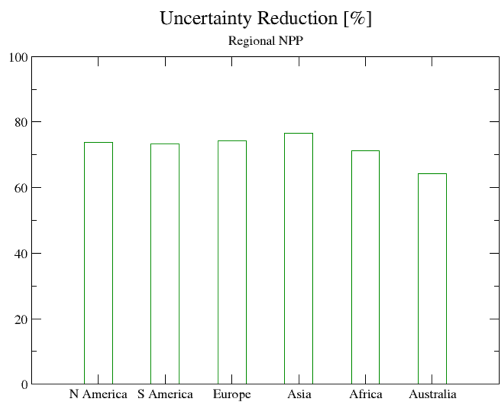

Figure 6. Uncertainty reduction in simulated NPP over six regions.

More interesting in the context of mission design are uncertainty reductions in target quantities such as predicted fluxes, because they are not specific for the model used. Here, we select net ecosystem production (NEP, Fig. 5) and NPP (Fig. 6) integrated over the period from 1999 to 2003 and six regions. 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.

4.3. Mission Design

Our approach to mission design relies on the fact that the CCDAS framework decouples the uncertainty propagation (via Equ. (2) and Equ. (4)) from the parameter estimation (via the minimisation of Equ. (1)). It is thus possible to test the effect of hypothetical or planned observational data streams on the uncertainty reduction in target quantities as long as our model ($M(x)$) can simulate the data stream and we are able to define the data uncertainty ($\mathbf{C_d}$), both of which enter the evaluation of H in Equ. (2). This allows efficient and easy modification of various aspects of a mission during its planning phase, e.g. characteristics of the sensor such as its accuracy, and quantify the mission's value in terms of uncertainty reduction in user-selected target quantities.

The methodological framework just described is called quantitative network design and is presented in *Kaminski and Rayner* (2008) along with a set of examples. *Kaminski et al.* (2010) demonstrate its application for the design of an active LIDAR mission for observing column integrated atmospheric CO_2 from space. For the reported case at the site-scale, we refer to *Knorr et al.* (2008) for an evaluation of two modifications of the characteristics of the MERIS sensor in terms of diagnostic and prognostic (until 2039) NEP and NPP.

5. CONCLUSIONS

We demonstrated the successful assimilation of the MERIS FAPAR product on the site scale. With one parameter set for all sites, the model is able to reproduce the observed FAPAR spanning boreal, temperate, humid-tropical and semi-arid climates. Assimilation of FAPAR has led to a moderate reduction in NPP uncertainty.

On the global scale we demonstrated the first simultaneous assimilation of MERIS FAPAR and atmospheric CO_2 in a mathematically rigorous framework including propagation of uncertainties. Owing to the coarse global resolution, the results are only preliminary and need to be confirmed in the standard CCDAS resolution, which is currently under developing as part of the project this contribution refers to.

The systematic application of the mathematically rigorous uncertainty propagation capability implemented by CCDAS allows the design of space missions with maximised benefit expressed in terms of uncertainties of carbon fluxes. The project is currently designing an interactive uncertainty prediction tool, which will enable the Agency to instantaneously evaluate a range of potential mission designs.

ACKNOWLEDGMENTS

The authors would like to thank the European Space Agency for financing this project under contract number 20595/07/I-EC, Philippe Goryl and Olivier Colin from ESA/ESRIN, Frascati, for support with the ESA MERIS product, Monica Robustelli and Ioannis Andredakis for help with data processing, Reiner Schnur for provision of meteorological data, and Michael Voßbeck for his help with code administration.

REFERENCES

Chen, M., P. Xie, J. Janowiak, and P. Arkin (2002), Global land precipitation: A 50-yr monthly analysis based on gauge observations, *J. Hydrometeorol.*, *3*, 249–266.

Denman, K., et al. (2007), Couplings Between Changes in the Climate System and Biogeochemistry, 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 S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M.Tignor, and H. Miller, chap. 7, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Enting, I. G. (2002), *Inverse Problems in Atmospheric Constituent Transport*, 392 pp., Cambridge University Press, Cambridge, U.K.

Giering, R., and T. Kaminski (1998), Recipes for Adjoint Code Construction, *ACM Trans. Math. Software*, *24*(4), 437–474, doi: http://doi.acm.org/10.1145/293686.293695.

GLOBALVIEW-CO_2 (2004), Cooperative Atmospheric Data Integration Project - Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado, [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

GLOBALVIEW-CO_2 (2008), Cooperative Atmospheric Data Integration Project - Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado, [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW].

Gobron, N., B. Pinty, M. M. Verstraete, and J. Widlowski (2000), Advanced vegetation indices optimized for up coming sensors: Design, performance and applications, *IEEE Trans. Geosc. Remote Sens.*, *38*, 2489–2505.

Gobron, N., B. Pinty, O. Aussedat, M. Taberner, O. Faber, F. Mlin, T. Lavergne, M. Robustelli, and P. Snoeij (2008), 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.

Heimann, M. (1995), The global atmospheric tracer model TM2, *Tech. Rep. 10*, Max-Planck-Inst. für Meteorol., Hamburg, Germany.

Jones, P., T. Osborn, K. Briffa, C. Folland, B. Horton, L. Alexander, D. Parker, and N. Rayner (2001), Adjusting for sampling density in grid-box land and ocean surface temperature time series, *J. Geophys. Res.*, *106*, 3371–3380.

Kaminski, T., and P. J. Rayner (2008), Assimilation and network design, in *Observing the continental scale Greenhouse Gas Balance of Europe*, edited by H. Dolman, A. Freibauer, and R. Valentini, Ecological Studies, chap. 3, pp. 33–52, Springer-Verlag, New York.

Kaminski, T., M. Heimann, and R. Giering (1999), 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*(D15), 18,555–18,581.

Kaminski, T., W. Knorr, P. Rayner, and M. Heimann (2002), Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, *Global Biogeochemical Cycles*, *16*(4), 14–1–14–16.

Kaminski, T., R. Giering, M. Scholze, P. Rayner, and W. Knorr (2003), An example of an automatic differentiation-based modelling system, in *Computational Science – ICCSA 2003, International Conference Montreal, Canada, May 2003, Proceedings, Part II*, *Lecture Notes in Computer Science*, vol. 2668, edited by V. Kumar, L. Gavrilova, C. J. K. Tan, and P. L'Ecuyer, pp. 95–104, Springer, Berlin.

Kaminski, T., M. Scholze, and S. Houweling (2010), Quantifying the Benefit of A-SCOPE Data for Reducing Uncertainties in Terrestrial Carbon Fluxes in CCDAS, *in press in Tellus B*.

Knorr, W. (2000), Annual and interannual CO_2 exchanges of the terrestrial biosphere: Process-based simulations and uncertainties, *Glob. Ecol. Biogeogr.*, *9*(3), 225–252.

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

Knorr, W., and J. Kattge (2005), Inversion of terrestrial biosphere model parameter values against eddy covariance measurements using monte carlo sampling, *Global Change Biology*, *11*, 1333–1351.

Knorr, W., T. Kaminski, M. Scholze, N. Gobron, B. Pinty, and R. Giering (2008), Remote Sensing Input for regional to global CO2 flux modelling, in *Proceedings of 2nd MERIS/(A)ATSR User Workshop, Frascati, Italy*, European Space Agency.

Knorr, W., T. Kaminski, M. Scholze, N. Gobron, B. Pinty, R. Giering, and P.-P. Mathieu (2010), Carbon cycle data assimilation with a generic phenology model, *in press in J. Geophys. Res.*

Nijssen, B., R. Schnur, and D. Lettenmaier (2001), Retrospective estimation of soil moisture using the VIC land surface model, 1980-1993, *J. Climate*, *14*(8), 1790–1808.

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

Rayner, P., M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann (2005), Two decades of terrestrial Carbon fluxes from a Carbon Cycle Data Assimilation System (CCDAS), *Global Biogeochemical Cycles*, *19*, doi:10.1029/2004GB002,254.

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

Scholze, M., T. Kaminski, P. Rayner, W. Knorr, and R. Giering (2007), Propagating uncertainty through prognostic CCDAS simulations, *J. Geophys. Res.*, *112*, doi:10.1029/2007JD008,642, doi: http://dx.doi.org/10.1029/2007JD008642.

Tarantola, A. (1987), *Inverse Problem Theory - Methods for Data Fitting and Model Parameter Estimation*, Elsevier Sci., New York.

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

Wilson, M. F., and A. Henderson-Sellers (1985), A global archive of land cover and soils data for use in general-circulation climate models, *Journal of Climatology*, *5*(2), 119–143.