Geoscientific
Model Development

Optimization of a prognostic biosphere model for terrestrial biomass and atmospheric CO_2 variability

M. Saito[1],*, **A. Ito[1], and S. Maksyutov[1]**

[1]Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan
*previously at: Laboratoire des Sciences du Climat et de l'Environnement, UMR8212, CEA-CNRS-UVSQ, 91191 Gif sur Yvette, France

Correspondence to: M. Saito (saito.makoto@nies.go.jp)

Abstract. This study investigates the capacity of a prognostic biosphere model to simulate global variability in atmospheric CO_2 concentrations and vegetation carbon dynamics under current environmental conditions. Global data sets of atmospheric CO_2 concentrations, above-ground biomass (AGB), and net primary productivity (NPP) in terrestrial vegetation were assimilated into the biosphere model using an inverse modeling method combined with an atmospheric transport model. In this process, the optimal physiological parameters of the biosphere model were estimated by minimizing the misfit between observed and modeled values, and parameters were generated to characterize various biome types. Results obtained using the model with the optimized parameters correspond to the observed seasonal variations in CO_2 concentration and their annual amplitudes in both the Northern and Southern Hemispheres. In simulating the mean annual AGB and NPP, the model shows improvements in estimating the mean magnitudes and probability distributions for each biome, as compared with results obtained using prior simulation parameters. However, the model is less efficient in its simulation of AGB for forest type biomes. This misfit suggests that more accurate values of input parameters, specifically, grid mean AGB values and seasonal variabilities in physiological parameters, are required to improve the performance of the simulation model.

1   Introduction

The terrestrial biosphere generally absorbs CO_2 from the atmosphere, and its annual global carbon uptake rate is considered to be similar to that of the ocean (Tans et al., 1990). Many studies have attempted to accurately quantify the total carbon exchange rate between the terrestrial biosphere and the atmosphere and to determine the role of the terrestrial biosphere in the global carbon cycle (e.g., Schimel, 1995; Field et al., 1998). Modeling of the terrestrial biosphere is one of the key strategies used in these studies (e.g., Potter et al., 1993; Running and Hunt, 1993; Ito and Oikawa, 2002).

Biosphere models use a variety of approaches and representations to predict the processes of ecosystem carbon dynamics. Therefore, quantities in the carbon budgets derived from these models differ. Cramer et al. (1999) compared 17 biosphere models and found that global terrestrial net primary productivity (NPP) calculated by the models ranged from approximately 40 to $65\,Pg\,C\,yr^{-1}$ and that the models yielded different NPP spatial distributions. Even at regional scales, Ichii et al. (2010) reported large differences in the magnitudes of annual gross primary productivity (GPP) and ecosystem respiration generated by nine simulation models. The discrepancies revealed by the systematic comparisons between different biosphere models indicate that the current biosphere models are still in need of improvement (Friedlingstein et al., 2006; Jung et al., 2007; Sitch et al., 2008).

Model–data synthesis is one approach to addressing the problem of discrepancies between models because such a

synthesis reduces uncertainties and optimizes controlling processes in a model to improve its fit to observed data. This method is widely applied in terrestrial biosphere modeling, as well as in atmospheric and oceanic modeling, based on various a priori information. Examples include modeling approaches to the estimation of residence times of carbon in soils and plants in regional areas (Barrett, 2002; Zhou and Luo, 2008), the fit of eddy covariance flux variables at point scales (Braswell et al., 2005; Sacks et al., 2006), the impact of seasonal water stress on ecosystem gas exchange (Reichstein et al., 2003), and atmospheric inversion schemes (Ciais et al., 1995; Enting et al., 1995; Bousquet et al., 2000; Peylin et al., 2005). An overview of model–data synthesis methods in terrestrial carbon studies and the improvements required in such methods have been discussed in several previous studies (Raupach et al., 2005; Wang et al., 2009; Luo et al., 2011).

Using a model–data synthesis approach, Kaminski et al. (2002) introduced a method to systematically infer optimal model parameters and applied the method to the assimilation of observed atmospheric CO_2 concentrations into an integrated terrestrial biosphere–atmospheric transport model. Atmospheric CO_2 concentrations reflect the distributions of CO_2 sources and sinks at various spatial and temporal scales, rather than those distributions restricted to small-scale footprints. In this way, these inputs are more likely to yield reliable results when applied to an optimized global biosphere model. In these studies, two parameters, light use efficiency and the temperature sensitivity of respiration (Q_{10}) were individually assessed for 12 biomes, which resulted in good simulations of the phases and amplitudes of seasonal atmospheric CO_2 at the observation sites. Rayner et al. (2005) further developed the method of Kaminski et al. (2002) by replacing the biosphere model with a more mechanistic approach and optimizing more of the model parameters, in what is known as the "Carbon Cycle Data Assimilation System" (CCDAS).

To date, studies have focused predominantly on CO_2 fluxes associated with atmospheric CO_2 variability, and less effort has been directed toward the effects of parameter optimization in simulations of ecosystem variability in carbon pools or in plant materials. However, the importance of accurate estimates of the forest carbon pools, as well as atmospheric CO_2 variability, is increasingly recognized in scientific and political circles (e.g., Carvalhais et al., 2010; Saatchi et al., 2011), as such estimates are required to quantify the CO_2 emissions attributable to deforestation and forest degradation, which account for about 12 % or more of total anthropogenic CO_2 emissions (van der Werf et al., 2009). Improvements in the estimates of forest carbon pools and atmospheric CO_2 variability are necessary for assessing the global carbon balance and the role of the terrestrial biosphere in the state of the current climate.

Here, we develop an optimization scheme for a global biosphere model for atmospheric CO_2 variability, above-ground biomass (AGB), and NPP data, combined with an atmospheric transport model and a Bayesian inversion method. Our primary goal is to improve current assessments of the terrestrial carbon cycle under current climatic conditions and atmosphere–biosphere interactions. For this purpose, we focus on constructing a model that is capable of comprehensively simulating the carbon dynamics of natural vegetation, vegetation structure, and atmospheric CO_2 variability on a global scale.

2   Method and data

2.1   Terrestrial biosphere model

A prognostic biosphere model, the Vegetation Integrative SImulator for Trace gases (VISIT; Ito, 2010) is used in this study to simulate the global terrestrial carbon cycle. All variables in VISIT are calculated at a 2.5° × 2.5° spatial resolution and with a time step of one day. The global vegetation types in the model are classified into 14 biomes (Fig. 1), and a map of their distributions is produced using the synergetic land cover product (SYNMAP) (Jung et al., 2006). The reanalysis/assimilation data sets released by the Japan Meteorological Agency (JMA) and the Japan 25-year reanalysis (JRA-25)/JMA Climate Data Assimilation System (JCDAS) (Onogi et al., 2007) are used to force the VISIT simulation. The JRA-25 reanalysis covers the period from 1979 to 2004, and JCDAS covers the period from 2005 to 2009. The meteorological data used to operate VISIT are downward shortwave radiation, total cloudiness, air temperature, ground surface temperature, soil temperatures, specific humidity, precipitation, and wind velocity. The precipitation bias in JRA-25/JCDAS is corrected based on the study of Saito et al. (2011). We use the climate data for the 31-year period from 1979 to 2009 for both spin-up and model simulations. In the spin-up simulation, the 31-year climate data are repeated over 4000 years until carbon pools reach an equilibrium state.

In the model, plant physiology and the carbon budget are simulated in each grid cell. Net ecosystem productivity (NEP) is defined as follows:

$$NEP = NPP - HR = -NEE, \tag{1}$$
$$NPP = GPP - AR, \tag{2}$$

where HR is the heterotrophic respiration, AR is the autotrophic respiration, and NEE is net ecosystem CO_2 exchange rate. The daily GPP rate is expressed using the dry-matter production theory of Monsi and Saeki (1953):
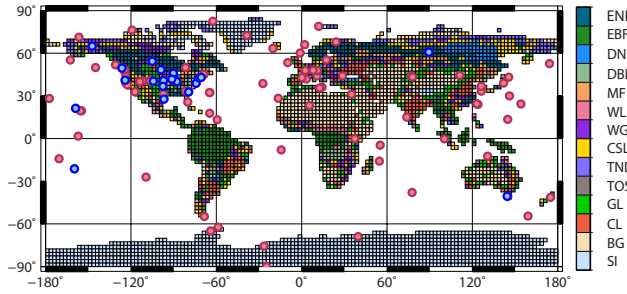
Figure 1. Distributions of the dominant biomes in each grid and the locations of atmospheric CO_2 observation sites. ENF: evergreen needle-leaf forest; EBF: evergreen broad-leaf forest; DNF: deciduous needle-leaf forest; DBF: deciduous broad-leaf forest; MF: mixed forest; WL: woodland; WGL: wooded grassland; CSL: closed shrubland; TND: tundra; TOS: temperate open shrubland; GL: grassland; CL: cropland; BG: bare ground; SI: snow and ice. The blue circles are partial-column CO_2 concentrations and the red circles are surface CO_2 concentrations.

$$GPP = \epsilon \int_0^{DL} \int_0^{LAI} P_c \, \mathrm{d}LAI \, \mathrm{d}t \qquad (3)$$

$$= \frac{2\epsilon P_{sat} DL}{K_a} \left[\ln \left\{ 1 + \sqrt{1 + \frac{K_a \cdot Q_e \cdot PPFD_{mid}}{P_{sat}}} \right\} \right.$$
$$\left. - \ln \left\{ 1 + \sqrt{1 + \frac{K_a \cdot Q_e \cdot PPFD_{mid} \cdot \exp(-K_a \cdot LAI)}{P_{sat}}} \right\} \right],$$

where $\epsilon (= 4.32 \times 10^{-4})$ is a unit converter from $\mu\mathrm{mol}\,CO_2\,\mathrm{m}^{-2}\,\mathrm{s}^{-1}$ to $\mathrm{Mg\,C\,ha}^{-1}\,\mathrm{day}^{-1}$; DL is day length (h), LAI is a leaf-area index estimated as a function of the given specific leaf area, SLA ($\mathrm{cm}^2\,\mathrm{g}^{-1}$); P_c is the single-leaf photosynthetic rate ($\mu\mathrm{mol}\,CO_2\,\mathrm{m}^{-2}$); P_{sat} is the single-leaf photosynthetic rate under light saturation ($\mu\mathrm{mol}\,CO_2\,\mathrm{m}^{-2}$); K_a is a dimensionless function of solar height; Q_e is light use efficiency ($\mathrm{mol}\,CO_2\,\mathrm{mol}^{-1}$ photon); and $PPFD_{mid}$ is the photosynthetic photon flux density (PPFD) at the canopy top at midday ($\mu\mathrm{mol}\,\mathrm{photon}\,\mathrm{m}^{-2}\,\mathrm{s}^{-1}$). The parameter $PPFD_{mid}$ is estimated from the equation of Kuroiwa (1966) with inclusion of the cloudy decline effect. Seasonal variations in P_{sat} are assumed to be dependent on the ground surface temperature (T_g; °C), the intercellular CO_2 concentration (C_i; ppmv) (which is estimated from the ambient CO_2 concentration), and soil moisture (Φ; mm), according to

$$P_{sat} = P_{max} \cdot F_{tmp}(T_g) \cdot F_{stl}(C_i) \cdot F_{nstl}(\Phi), \qquad (4)$$

where P_{max} is the potential maximum value of P_{sat}; F is the coefficient function used to calculate seasonal variations in P_{sat}; F_{tmp} is formulated using the maximum, minimum, and optimum temperatures for photosynthesis (T_{max}, T_{min}, and

T_{opt}, respectively); and F_{stl} and F_{nstl} are formulated using the photosynthesis-limiting factors for intercellular CO_2 concentration (kmci) and soil moisture (km_nstl), according to

$$F_{tmp} = \qquad (5)$$
$$\max \left[\min \left[\frac{(T_g - T_{max})(T_g - T_{min})}{(T_g - T_{max})(T_g - T_{min}) - (T_g - T_{opt})^2}, 1 \right], 0 \right],$$

$$F_{stl} = \qquad (6)$$
$$\max \left[\min \left[(1 - C_{stl}) + \frac{C_{stl} \cdot (C_i - CD_{cmp})}{kmci + C_i}, 1 \right], 0 \right],$$

$$F_{nstl} = \qquad (7)$$
$$\max \left[\min \left[(1 - C_{sntl}) + \frac{C_{nstl} \cdot \Phi}{\Phi + FP_{sw} \cdot km\_nstl}, 1 \right], 0 \right],$$

where CD_{cmp} is the CO_2 compensation point (Brooks and Farquhar, 1985), FP_{sw} is the soil water holding capacity (mm) (given by soil property data), and C_{stl} and C_{nstl} are constant coefficients.

Parameter AR represents both the growth and maintenance respiration from foliage, stem and branch, and root components. The growth respiration (AR_G) of each component is the cost to produce new biomass, given as

$$AR_{GX} = rg_X \cdot TP_X, \qquad (8)$$

where rg_X is the specific growth respiration rate, TP_X is the carbon translocation rate, and the subscript X represents the different components: foliage (FL), stem and branch (SB), and roots (RT). In contrast, the maintenance respiration (AR_M) is represented as a function of T_g:

$$AR_{MX} = rm_X \cdot \exp \left[\frac{\ln Q_{10}}{10}(T_g - 15) \right] \cdot M_X, \qquad (9)$$

where rm_X is the specific maintenance respiration rate for each component and M_X is the carbon mass for each component.

The parameter HR represents respiration rates in two layers: the litter (HR_L) and humus (HR_H) layers. Both HR_L and HR_H are calculated as functions of the soil temperature at two depths, namely 10 cm (T_{s10}) and 200 cm (T_{s200}), and Φ in the upper and lower soil layers:

$$HR_L = shl \cdot M_L \cdot F_{HR}(T_{s10}, \Phi_u), \qquad (10)$$
$$HR_H = shm \cdot M_H \cdot F_{HR}(T_{s200}, \Phi_l), \qquad (11)$$

where shl and shm are the specific heterotrophic respiration rates and M_L and M_H are the carbon masses of the organic matter in the litter and humus layers, respectively. The annual mean value of T_{s10} is used as a substitute for T_{s200} because JRA-25/JCDAS provides the soil temperature at depth.

In VISIT, the amount of litter fall (LF) is assumed to be proportional to the carbon mass of each component (foliage, stems and branches, and roots), according to

$$\text{LF}_X = \text{lf}_X \cdot M_X, \tag{12}$$

where lf_X is the specific litter fall rate.

Parameter AGB is the sum of the carbon mass in foliage M_{FL} and stems and branches M_{SB}. Changes in M_X for a given period are given by

$$\Delta M_X = \text{TP}_X - \text{AR}_{GX} - \text{LF}_X. \tag{13}$$

The value of AGB is estimated by accumulating ΔM_{FL} and ΔM_{SB} on a given initial AGB for both spin-up and model simulations.

In this study, the prior values of the parameters are the same for all biome types (Table 1). The posterior parameters are estimated by Bayesian inversion. An average sensitivity of each parameter to three variables (NEE, AGB, and NPP of whole biomes) was estimated using 10 % change in the prior parameters and a 2 °C variation in temperature parameters, using first-order sensitivity analysis. From this analysis the 12 parameters with the highest sensitivities (T_{min}, SLA, lf_{FL}, P_{max}, shl, T_{opt}, Q_e, TP_{FL}, km_nstl, shm, Q_{10}, and lf_{SB}) were selected to optimize the fit of the model simulations to the observations (Table 1). These parameters explain 90 % of the variations in the three variables for whole biomes. Uncertainties associated with these parameters are unknown; thus, they are defined using the results of the sensitivity analysis. The sensitivity of lf_{SB}, which shows the lowest sensitivity (1.9 %) of the selected 12 parameters to changes in the three variables, is defined as a reference value for parameter uncertainty; i.e., it is assumed that all parameters potentially have uniform sensitivity to the three variables. A scale factor is calculated for each parameter using the reference sensitivity of 1.9 % divided by the sensitivity of the parameter; then, an uncertainty of the parameter is defined by multiplying by 10 % of the prior parameter, or 2 °C, using a scale factor (Table 1). Here we assume a linear relationship between parameter sensitivity and changes in the parameters.

2.2 Atmospheric tracer transport model

Seasonal variability in atmospheric CO_2 concentrations at the observation sites is simulated using the National Institute for Environmental Studies, Frontier Research Center for Global Change (NIES/–FRCGC) off-line global atmospheric tracer transport model (NIES-TM; Maksyutov et al., 2008). The NIES-TM is one of the transport models evaluated in the Atmospheric Transport Model Intercomparison Project (TransCom; Law et al., 1996; Gurney et al., 2002). The vertical turbulent diffusion in the boundary layer of NIES-TM is parameterized using the monthly mean planetary boundary layer (PBL) heights derived from a 3-hourly PBL height data set (Schubert et al., 1993).

We operate the NIES-TM with a $2.5° \times 2.5°$ grid resolution and nine vertical levels, with forcing data from JRA-25/JCDAS. The oceanic CO_2 flux, fossil fuel emission inventory, and NEE in the terrestrial biosphere are required a priori information to run the NIES-TM. In this study, the ocean flux is derived from the Offline ocean Tracer Transport Model (OTTM; Valsala and Maksyutov, 2010), and the fossil fuel emission data are from the Open-source Data Inventory of Anthropogenic CO_2 emissions (ODIAC; Oda and Maksyutov, 2011). We run the NIES-TM for a 6-year period (2000–2005). The data for the first year are used for the spin-up simulation, and the data for 2001–2005 are used in the analysis. Trends in simulated atmospheric CO_2 concentrations with periods longer than around 1.5 years and shorter than about 15 days are removed using an orthogonal wavelet decomposition (Howell and Mahrt, 1994). Then, the mean monthly atmospheric CO_2 is calculated.

2.3 Bayesian inversion

We optimize a set of biophysical parameters in VISIT against mean monthly atmospheric CO_2 and mean annual AGB and NPP. To derive an optimal parameter set m, the deviations of the model estimates from the observed data are minimized using a Bayesian inversion scheme. When the probability distributions of all observed and a priori biophysical parameters are assumed to be Gaussian, and their means are d_{obs} and m_p, respectively, the misfit between d_{obs} and the modeled values $\mathbf{G}(m)$ for atmospheric CO_2, AGB, and NPP is defined by a cost function (Tarantola, 2005):

$$S(m) = \frac{1}{2} \Big[(\mathbf{G}(m) - d_{\text{obs}})^{\text{T}} \mathbf{C}_D^{-1} (\mathbf{G}(m) - d_{\text{obs}}) \tag{14}$$
$$+ (m - m_p)^{\text{T}} \mathbf{C}_M^{-1} (m - m_p) \Big],$$

where the superscript T denotes the transpose, and $\mathbf{C}_D$ and $\mathbf{C}_M$ are the covariance matrices defining the uncertainties in d_{obs} and m_p, respectively. The values of $\mathbf{C}_D$ are described in the following section. The model result $\mathbf{G}$ consists of a linear atmospheric transport model A and a nonlinear terrestrial biosphere model B. In the minimization of $S(m)$, in the case of atmospheric CO_2, $\mathbf{G}$ is determined by the first derivatives of $B(m_p)$, which, using finite differences Δm, is

$$\mathbf{G} = A \left[B(m_p) + \frac{B(m_p + \Delta m) - B(m_p)}{\Delta m} \right], \tag{15}$$

where a linear relationship is assumed between the first derivative of B and m_p for Δm. In the cases of NPP and AGB, the operator $\mathbf{G}$ is determined by Eq. (15) without multiplication by A.

Table 1. List of variables in prior parameters (Prior) and their uncertainties (Unc). The values given in parentheses are for grassland.

Symbol	Parameter description (unit)	Prior	Scale	Unc	Scale
P_{max}	Potential maximum photosynthetic rate (μmol CO_2 m^{-2} s^{-1})	1.75	$\times 10$	2.8	$\times 10^{-1}$
Q_{10}	Temperature dependence of respiration rate (dimensionless)	2.0	–	1.8	$\times 10^{-1}$
Q_e	Light use efficiency (mol CO_2 m^{-2})	5.2	$\times 10^{-2}$	2.1	$\times 10^{-3}$
SLA	Specific leaf area (cm^2 g^{-1})	1.4 (1.65)	$\times 10^2$	2.0 (2.4)	–
T_{min}	Minimum temperature for plant activities (°C)	1.0	–	2.1	$\times 10^{-1}$
T_{opt}	Optimum temperature for plant activities (°C)	2.4	$\times 10$	4.5	$\times 10^{-1}$
TP_{FL}	Carbon translocation rate to foliage (dimensionless)	5.5	$\times 10^{-1}$	3.3	$\times 10^{-2}$
km_nstl	Photosynthesis limitation factor in terms of soil moisture	0.3	–	2.2	$\times 10^{-2}$
lf_{FL}	Specific litter fall of leaf	4.3	$\times 10^{-4}$	6.4	$\times 10^{-6}$
lf_{SB}	Specific litter fall of stem and branch	2.0	$\times 10^{-4}$	2.0	$\times 10^{-5}$
shl	Specific heterotrophic respiration rate at upper soil	0.6	–	1.3	$\times 10^{-2}$
shm	Specific heterotrophic respiration rate at lower soil	0.3	–	2.6	$\times 10^{-2}$

The center of the posterior Gaussian distribution, which shows the optimized values for the model parameter m, can be expressed following Tarantola (2005) as:

$$m = \tag{16}$$
$$m_p + \left(\mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G} + \mathbf{C}_M^{-1} \right)^{-1} \mathbf{G}^T \mathbf{C}_D^{-1} \left(d_{obs} - \mathbf{G}(m_p) \right).$$

The expression $\left(\mathbf{G}^T \mathbf{C}_D^{-1} \mathbf{G} + \mathbf{C}_M^{-1} \right)^{-1}$ in Eq. (16) is inverted using single-value decomposition.

The optimal parameter set m is given by minimizing $S(m)$ iteratively. In this process, the values of $\mathbf{C}_D$ and $\mathbf{C}_M$ are fixed, while the value of m_p for all parameters evolved at each iteration by taking a posterior m as a prior m_p. The minimization of S is assessed by calculating χ^2, which is the mean square mismatch between d_{obs} and $\mathbf{G}(m)$, expressed as

$$\chi^2 = \sum_{i=0}^{N_{obs}} (G_i - d_{obs,i})^2 / C_{D,i} \tag{17}$$
$$+ \sum_{j=0}^{N_p} (m_j - m_{p,j})^2 / C_{M,j}.$$

The value of χ^2 is calculated for the entire data set of atmospheric CO_2, AGB and NPP. The iteration for the optimization is aborted when the value of χ^2 between the current and previous iterations is less than 0.1 and the difference is continuously maintained or decreases in further iterations.

2.4 Data

The atmospheric CO_2 data from GLOBALVIEW-CO_2 (2010) are used as the observation data for CO_2 concentrations. The GLOBALVIEW-CO_2 product is based on atmospheric measurements and provides information about CO_2 variability at over 100 sampling locations around the globe. The CO_2 variability is stored in a file named "extended

records" (ext), which records weekly smoothed atmospheric concentrations, and which is used in the following analysis. Mean monthly CO_2 variability is calculated from the ext data in the same manner as that for simulated atmospheric CO_2 (Sect. 2.2). The standard error of CO_2 variability, recorded in a file named "statistical summary of the average seasonal pattern", is used as a measure of the uncertainty in the observed CO_2 in $\mathbf{C}_D$ in Eqs. (14) and (16). We analyze sampling locations at 102 sites, including at 19 sites with multiple vertical CO_2 observations sampled by aircraft (Fig. 1). For those sites with vertical profiles, partial-column CO_2 concentrations are estimated from the vertical profiles by assuming that the CO_2 concentration at each observation height represents the concentration at altitudes centered around the observation height. We analyze this partial-column concentration instead of individual CO_2 concentrations at multiple heights.

AGB data from the International Institute for Applied Systems Analysis (IIASA; Kindermann et al., 2008) are used in this study to optimize AGB in the biosphere model. The IIASA data provides half-degree global biomass and carbon stock data estimated from a downscaling model on the basis of a map provided by the Global Forest Resources Assessment. The data for AGB per hectare are calculated from the gridded AGB data divided by the grid area data. To reduce discrepancies attributable to differences in the dominant biome in each grid between the model and IIASA, the grid resolution for AGB is expanded to $7.5° \times 7.5°$. Therefore, the IIASA data with a grid resolution of $0.5° \times 0.5°$ and the VISIT data with a resolution of $2.5° \times 2.5°$ are aggregated to the desired resolution. Here, the mean AGB values in VISIT for the 10-year period of 2000–2009 are used for the analysis. We compare AGB in 344 grids over the global terrestrial area.

The third data set contains the NPP data from the Global Primary Production Data Initiative (GPPDI; Scurlock et al., 1999; Olson et al., 2001). The GPPDI data set consists of three classes of NPP data: NPP measurements at intensively

studied sites with site-specific information, NPP measurements at extensive sites with less site-specific information, and gridded NPP data with a grid resolution of 0.5° compiled from collections of inventory, modeling, and remote-sensing data. All of the NPP data from the three classes are aggregated into a 7.5° grid cells, resulting in 166 grid cells being analyzed in this study. As for AGB, the NPP data in VISIT are averaged over the period 2000–2009 and are aggregated into 7.5° grid cells.

Global maps of AGB and NPP include various estimating errors, including scale mismatches between measurements and models, measurement errors, and representation errors, leading to large uncertainties in $\mathbf{C}_D$ (Raupach et al., 2005). This study defines the uncertainties in the observation data of AGB and NPP as 50 % of the 7.5° grid mean values, based on the errors in relative estimates of annual NPP given in Raupach et al. (2005), with minimum uncertainties of 1.0 Mg C ha^{-1} and 0.5 Mg C ha^{-1} yr^{-1} for AGB and NPP, respectively.

3 Results

3.1 Atmospheric CO$_2$ simulation

The prior value of χ^2 that is weighted by number of data points for atmospheric CO$_2$, AGB, and NPP is 7.80. This value decreases with each iterations, and the iteration is aborted after 10 repetitions because no large changes in χ^2 are subsequently observed. The final value of weighted mean χ^2 after 10 iterations for the whole data set is 3.93; $\chi^2 = 4.83$ for atmospheric CO$_2$, 1.46 for AGB, and 1.90 for NPP. The large value of χ^2 for atmospheric CO$_2$ can be partly attributed to the combined effect of minor uncertainties in the observations.

Atmospheric CO$_2$ variability is estimated globally, and mean monthly concentrations are compared with observed data from GLOBALVIEW-CO$_2$ at 102 stations. The result estimated from the posterior NEE is in good agreement with observations for the large domain and seasonality ($r^2 = 0.86$, $P < 0.001$; Fig. 2). The slope of the linear regression indicates that the modeled CO$_2$ variability is slightly underestimated, with an offset of 0.2 ppm. The misfit of mean monthly atmospheric CO$_2$ between observations and model values is defined at each station as $\chi_N^2 = 1/N_{\mathrm{mnt}} \sum_{i=1}^{N_{\mathrm{mnt}}} (G_i - d_{\mathrm{obs},i})^2 / C_{D,i}$, where N_{mnt} is the number of months, $N_{\mathrm{mnt}} = 12$. The value of χ_N^2 at each station varies from 0.16 to 52.75, with a mean χ_N^2 of 4.90, and the annual mean value of the root mean squared error (RMSE) varies from 0.11 to 4.71 ppm, with a mean RMSE of 1.46 ppm. The values of χ_N^2 show large misfits ($\chi_N^2 > 10$) at 11 stations; misfits of 2–3 ppm in atmospheric CO$_2$ in months with minor observation uncertainties result in the large χ_N^2 value. Such observation uncertainties occur at stations located on islands,
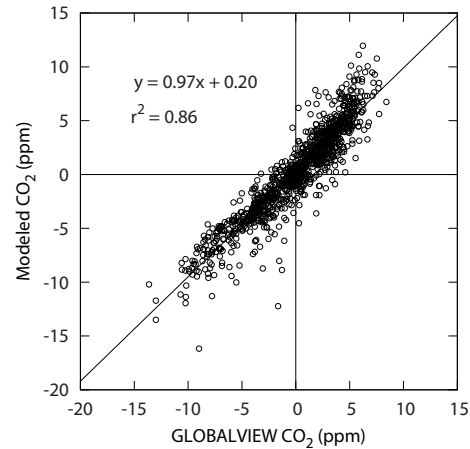

Figure 2. Comparison of the mean monthly CO$_2$ variability calculated with GLOBALVIEW and with the model presented here.

peninsulas, bare ground, complex terrains, and oceans, where the biospheric CO$_2$ flux exerts less influence on atmospheric CO$_2$ variability, or where local mesoscale motions control the variability. Improvements in the atmospheric CO$_2$ simulation at these stations is necessary to improve not only the biospheric model, but also the atmospheric transport and ocean models; thus, the 11 stations at which $\chi_N^2 > 10$ are excluded from the following analysis. Consequently, the mean values of χ_N^2 and RMSE for the whole data set become 3.02 and 1.36 ppm, respectively.

Figure 3 shows examples of the mean seasonal variations in atmospheric CO$_2$ at the University of Maine, Argyle, Maine, USA (AMT; 45.03° N, 68.68° W) and Tierra Del Fuego, Ushuaia, Argentina (TDF; 54.85° S, 68.31° W). The biome types at the AMT and TDF stations are classified as mixed forest and closed shrubland, respectively, in the model. The annual mean χ_N^2 and RMSE values for the observed and posterior data are 3.59 and 2.01 ppm, respectively, at AMT and 0.17 and 0.11 ppm, respectively, at TDF. The peak of photosynthetic uptake in August at AMT is slightly underestimated in the model, resulting in a discrepancy in the position of the peak between the observations and the model (Fig. 3a). An improvement can be seen in the results for the posterior parameter as compared with the prior parameter, although the CO$_2$ variability for some months is still outside the domain of the uncertainties in the observed data. At the TDF station, the prior concentration shows overestimates of both the photosynthetic uptake in summer and the respiration release in winter (Fig. 3b). The posterior concentration reduces these to a difference of within a few tenths of a ppm, and the discrepancy of the CO$_2$ release peak in the prior data is moderated in the posterior data.

Figure 4 shows a comparison of the mean seasonal amplitudes of atmospheric CO$_2$ based on the observed and posterior parameters at all stations along a single latitude. The seasonal amplitude at each station is defined as the difference
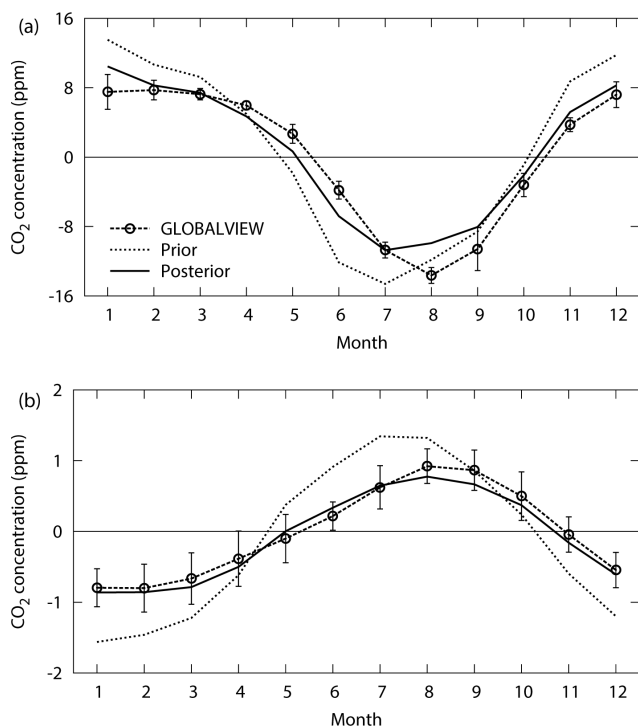
Figure 3. Seasonal variations in the observed and modeled mean monthly CO_2 concentrations (ppm) at **(a)** University of Maine, USA, and **(b)** Tierra Del Fuego, Argentina. Open circles with the dashed lines are GLOBALVIEW-CO_2 data, the dotted lines represent the prior parameters, and the solid line represents the posterior parameters. The error bars indicate the uncertainties in the observations.

between the maximum and minimum mean monthly concentrations. The posterior simulation shows a representation of the seasonal amplitude with an RMSE of 1.65 ppm for all stations, although the differences in the seasonal amplitude between the observation and the model are significant ($P < 0.001$). Erroneous overestimates (> 5 ppm) in the mean amplitude occurs at two stations in the Northern Hemisphere: Sary Taukum, Kazakhstan (KZD; 44.08° N, 76.87° W), and Schauinsland, Germany (SCH; 48.00° N, 8.00° E). The seasonal variability in the posterior parameters at both stations shows large overestimates in the winter concentration and slight overestimates of the maximum photosynthetic uptake, which occur in early summer. Unfortunately, in this study we failed to improve these mismatches and produce a good agreement with observations at other stations. This is partly because of the inability of the global biosphere model to represent NEE variability in the areas around observation points and also because unsuitable biophysical parameters are selected for adjustment in the optimization process.
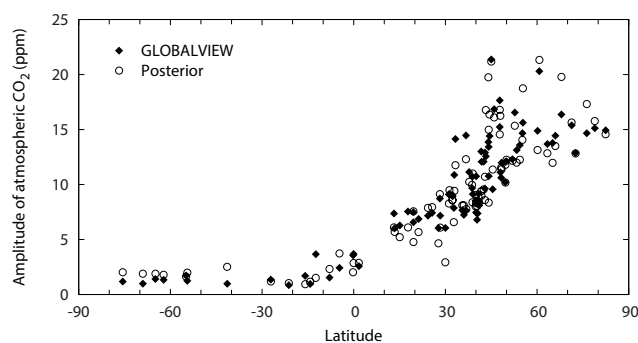

Figure 4. Annual amplitudes of the atmospheric CO_2 variables (ppm) along specific latitudes. The closed diamonds represent the GLOBALVIEW-CO_2 data and the open circles represent the posterior parameters. The latitudinal mean amplitudes and standard deviations for each 30° band from north to south are 15.1 ± 1.9 ($n = 11$), 11.3 ± 3.1 ($n = 51$), 6.7 ± 1.4 ($n = 13$), 2.2 ± 1.1 ($n = 9$), 1.3 ± 0.3 ($n = 3$), and 1.2 ± 0.2 ppm ($n = 4$) for the GLOBALVIEW-CO_2 data and 15.3 ± 2.9, 11.2 ± 3.6, 6.1 ± 1.6, 1.9 ± 0.9, 2.1 ± 0.4, and 1.9 ± 0.1 ppm for the posterior data.

3.2 Biomass simulations

The χ_N^2 values of mean annual AGB and NPP for the observed and prior data range from 0.00 to 4781 among the grid cells, with a mean of 71.50. This large misfit of AGB and NPP is improved by using posterior values, which give χ_N^2 values in the range of 0.00 to 58, with a mean of 1.74. Six of the total of 510 grid cells (at a 7.5° × 7.5° resolution) show χ_N^2 values much greater than 10 for the posterior data. In these grid cells, the posterior AGB or NPP are overestimated as compared with observations, suggesting that discrepancies exist between observations and the model in some biome types.

In the simulation using the posterior parameters, the mean annual AGB and NPP values are estimated at a spatial resolution of 2.5° × 2.5° before being compared with IIASA and GPPDI data for each biome type that had been aggregated to a comparable resolution (Figs. 5 and 6). The biome type for each grid cell is classified using the SYNMAP product (see above). The IIASA data show biome-specific patterns in the median AGB: the biomes showing AGB values of > 30 Mg C ha^{-1}, and with relatively large variances, include evergreen broad-leaf forest (EBF), deciduous broad-leaf forest (DBF), and mixed forest (MF). Those with AGB values of approximately 20–30 Mg C ha^{-1} include evergreen needle-leaf forest (ENF), deciduous needle-leaf forest (DNF), and woodland (WL); other biome types show AGB values of < 20 Mg C ha^{-1} (Fig. 5). The simulation using the prior parameters shows greater variability in AGB values for most biome types but does not overestimate AGB in any of the forest biome types relative to the IIASA data. These patterns in the AGB distributions are improved in the simulations using the posterior values. The variations in AGB with respect
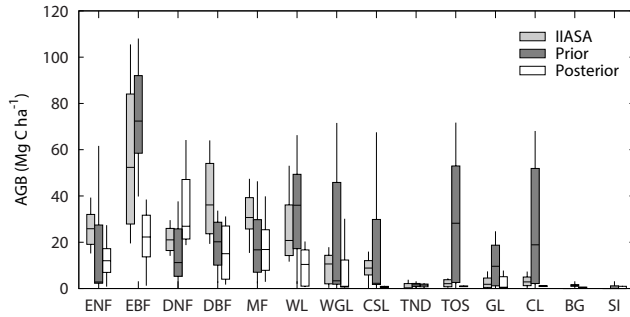
Figure 5. Comparison of AGB (Mg C ha^{-1}) values for each biome according to IIASA (light grey box) and VISIT with prior (dark grey box) and posterior (open box) parameters. The box-and-whisker plots show the median, upper, and lower quartiles, and the upper and lower 10 % of the data.
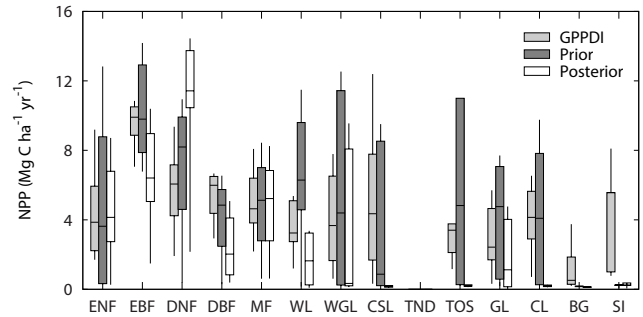
Figure 6. As for Fig. 5, except showing a comparison of NPP values (Mg C ha^{-1} yr^{-1}) for GPPDI and VISIT with prior and posterior parameters. The GPPDI does not contain NPP data in the grids corresponding to tundra (TND).

to biome and the small variations in AGB for each type are reproduced in the model. However, the model using the posterior values underestimates the magnitudes of AGB for forest and woodland biomes (except for DNF), and a significant difference is noted between observations and model results in median AGB for the whole biomes ($P < 0.001$). The largest difference between model results in terms of AGB is present in EBF; the lower and upper 10 % AGB in the IIASA are equal to 19.5 and 105.5 Mg C ha^{-1}, respectively, with a median value of 52.4 Mg C ha^{-1}, whereas those of the model range from 1.2 to 38.5 Mg C ha^{-1}, respectively, with a median value of 22.3 Mg C ha^{-1}.

The median NPP for each biome type in the GPPDI ranges from 2 to 6 Mg C ha^{-1} yr^{-1}, except for the much higher NPP of EBF and lower NPP of bare ground (BG) (Fig. 6). As in the AGB distributions, the simulation with the prior values shows greater variability in NPP values across biomes relative to observations. The NPP distributions using the posterior values show biome-specific patterns similar to those of AGB. As compared with AGB distributions, no large underestimations in the NPP values of forest type biomes are present; however, there is a significant difference in the median NPP values between observations and the model for whole biomes ($P < 0.05$). No forest type biomes show NPPs comparable with those of forest type biomes in the GPPDI, suggesting that there may be discrepancies in the dominant biome type in a grid cell between the GPPDI data and the model.

4 Discussion

4.1 Model parameterization

The differences between posterior and prior parameters, divided by the uncertainties of the prior parameters (Table 1), are shown in Fig. 7. The posterior parameters show various changes relative to the prior parameters. Here, we examine

some of the parameters that show clear changes. The largest increases and reductions in the magnitudes of the posterior parameters are observed for the optimum temperature for plant activities T_{opt}. This parameter specifies the optimum temperature for photosynthetic uptake, which directly controls the rate of photosynthesis and its seasonal variability in response to changes in air temperature. The temperature dependence of photosynthesis is strongly influenced by other environmental factors, such as light intensity, intercellular CO_2 pressure, and water and nutrient regimes; however, T_{opt}, which is generally between 20 and 30 °C, represents the adaptation of a plant to the temperature regimes of its habitat (Berry and Björkman, 1980; Monson et al., 1982). The posterior parameters of T_{opt} are approximately observed within 20 and 30 °C, which range from 19.8 to 28.7 °C, with high T_{opt} values for EBF, closed shrubland (CSL), and cropland, and low T_{opt} values for DNF, MF, and tundra biomes. The variability in posterior T_{opt} values among biomes is roughly consistent with the latitudinal distribution that the biomes occupy (Fig. 1): high T_{opt} value with high mean temperatures at low latitudes, and low T_{opt} values with low mean temperatures at high latitudes. The minimum temperature for plant activities, T_{min}, which mainly controls the phases of leafing and senescence and the duration of photosynthetic activity, also shows similar patterns of change in the posterior parameters to those of T_{opt}. These results suggest that variability in biome-specific parameters associated with temperature dependence can be reproduced by our optimization scheme.

The SLA describes the allocation of leaf biomass per unit of leaf area. Values of SLA generally decrease with increasing leaf life span, and the mean SLA values of forbs, broad-leaf shrubs, broad-leaf trees, and needle-leaf trees are 188, 112, 112, and 48 cm^2 g^{-1}, respectively (Reich et al., 1997, 1998). Regardless of biome, the SLA is well predicted by the posterior parameters for the corresponding biomes, of 175, 131, 131, and 195 cm^2 g^{-1} for grassland, CSL, EBF, and ENF, respectively, although the SLA for ENF is obviously a misfit relative to values obtained in previous studies. In our
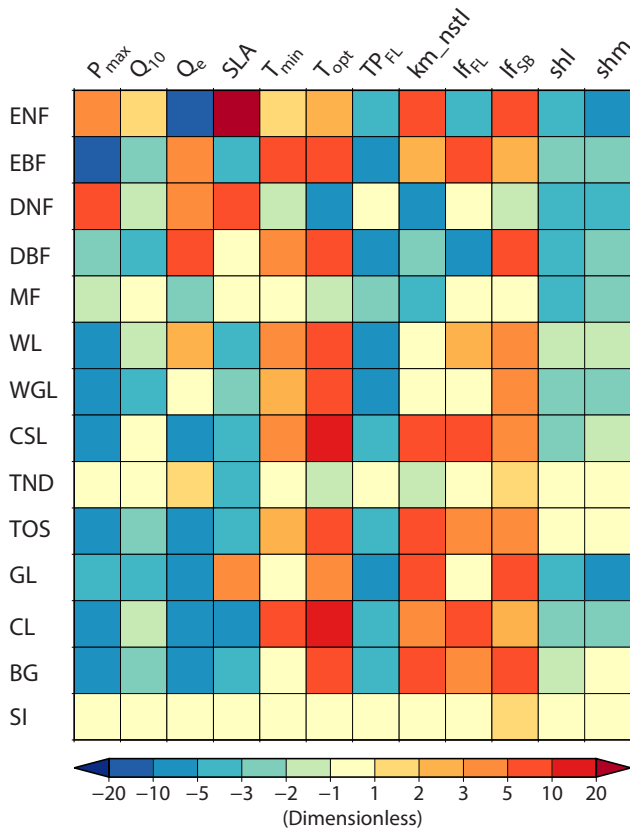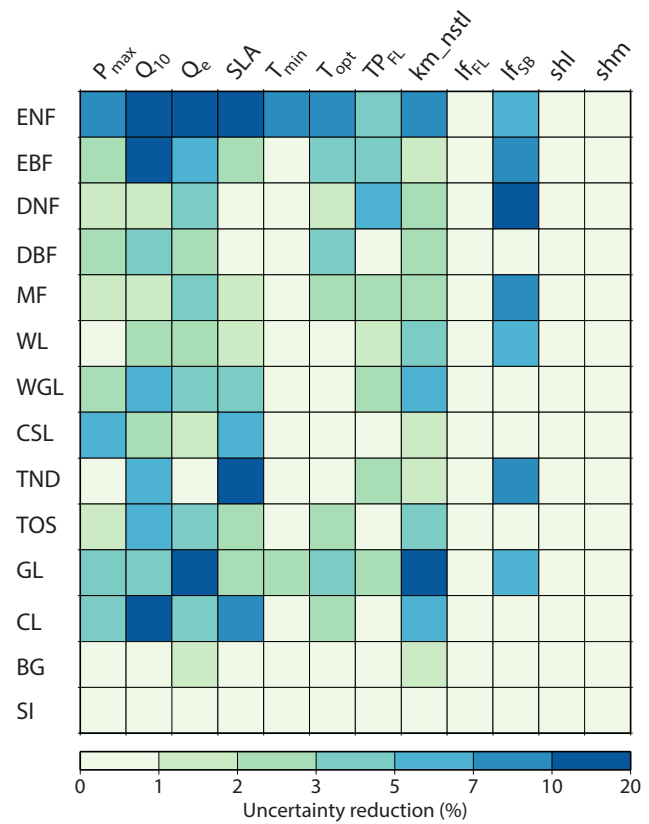
Figure 7. Differences between the posterior parameters (PRM_{pst}) and the prior parameters (PRM_{pri}), with uncertainties of prior parameters (UNC_{pri}) as ($PRM_{pst} - PRM_{pri}$) / UNC_{pri}.

Figure 8. As Fig. 7, but as the uncertainty reduction (%) from UNC_{pri} to uncertainties of posterior parameters (UNC_{pst}) for each parameter as ($UNC_{pri} - UNC_{pst}$) / $UNC_{pri} \times 100$.

model, SLA directly regulates the variability in GPP, along with changes in the LAI, which is similar to the function of P_{max}, which defines the potential maximum photosynthetic rate (Eq. 3). However, SLA is an ecological parameter that is closely related to canopy leaf nitrogen content, photosynthetic capacity, irradiance, and nutrient availability, and the relationship between P_{max} and SLA is non-linear (Pierce et al., 1994; Meziane and Shipley, 1999). Nevertheless, our optimization scheme does not account for the interactions among parameters or between parameters and environmental factors, and parameters that work similarly in the model are not explicitly excluded, resulting in similar changes in posterior parameters between P_{max} and SLA. Consequently, posterior SLA may be partly constrained by changes in P_{max}, leading to a misfit for ENF. This shortcoming in our system may extend to other parameters; therefore, further development of our system is required to include the interactions among parameters and to identify criteria for selecting optimized parameters.

The posterior parameters of temperature dependence of the respiration rate Q_{10}, light use efficiency Q_e, specific litter fall of stems and branches lf_{SB}, and the photosynthesis limitation factor in terms of soil moisture km_nstl show

large reductions in uncertainty relative to those of prior values (Fig. 8). This large reduction in uncertainty indicates that more information is integrated into the model parameters based on observational data. The causes of larger reductions in uncertainty in the four parameters are the relatively large uncertainties in the prior parameters (see Sect. 2.1) and the relatively high sensitivity of atmospheric CO_2, AGB, and NPP to changes in the parameters. However, very small reductions in uncertainty are observed in all parameters. This is attributed to indirect contributions of each parameter to observational variables and small uncertainties given in the prior parameters, resulting in a difficulty of integrating new information into the parameter (Rayner et al., 2005).

The biosphere model VISIT is designed to be used with biome-specific parameters that have constant values for each biome, many of which are independent of environmental influences. The parameters optimized in this study, such as the carbon translocation rate to foliage (TP_{FL}) and litter fall (lf_X) rate, have constant values for each biome, indicating that these parameters have been adjusted to represent the central parts of the distributions of the corresponding physiological processes on long-term and regional scales. However, for example, vegetation alters its biomass allocation patterns in

response to differences in climate to allow its survival under unfavorable conditions (Callaway et al., 1994; Maherali and DeLucia, 2001), and the amount of litter fall also varies considerably under different environmental conditions (Berg and Laskowski, 2006), but none of these factors are included in the model. Therefore, the model far from adequately represents the physiological processes under all conditions, which might result in large discrepancies in model results from the observed data at some points. If it is possible to describe the response of each parameter to the environmental conditions, a model that includes those variable parameters could improve the simulation of global carbon sequestration and its distribution, although adopting empirical rules also entails the possibility that the uncertainties in the modeling will increase (Saito et al., 2009).

In this study, we selected 12 key parameters in a prognostic biosphere model for model–data synthesis. Caution is required because it might be possible to obtain similar model results with different combinations of model parameters (Wang et al., 2009). Additional observations would be useful in optimizing many more model parameters and enhancing the performance of the model–data synthesis.

4.2   Model limitations

Our optimization scheme was tested by running it with only observational data on atmospheric CO_2 concentrations, in the same manner as performed using the three observational variables. The χ^2 value of 7.5 obtained using the prior values decreases to 4.6 after 10 iterations. Relative changes in prior and posterior parameters and the mean annual distributions of AGB and NPP for each biome show very similar patterns to those displayed in Figs. 5–7, despite the fact that any information on observational data for AGB and NPP is not integrated into the posterior parameters (not shown). The test results, and the discrepancies between observations and the results of posterior simulations for mean annual AGB and NPP (shown in Figs. 5 and 6), suggest that observational data of AGB and NPP might not contribute much to the minimization of the misfit between the observations and the model. A possible explanation for this observation is that the sizes of the observational uncertainties for AGB and NPP are too large to solve the minimization problem, resulting in an incomplete indication of variability. The large observational uncertainty reflects a lower degree of confidence in the accuracy of the AGB and NPP observations (Raupach et al., 2005), and improving the representation of the AGB and NPP distributions on large spatial scales would, thus, improve the results of the model simulations.

The model does not account for the effects of natural and anthropogenic disturbances, such as those due to fire, land use changes, and grazing and harvesting of forest products and croplands; thus, the mean global budget of NEP estimated from the posterior parameters is approximately $0\,\mathrm{Pg}\,\mathrm{C}\,\mathrm{yr}^{-1}$ on account of the equilibrium state of carbon pools. This result clearly underestimates the current carbon uptake rate of about $2.6\,\mathrm{Pg}\,\mathrm{C}\,\mathrm{yr}^{-1}$ (IPCC, 2007). Including the effects of disturbances and subsequent regrowth of vegetation would improve the discrepancies between observations and model simulations of the global carbon budgets.

5   Conclusions

This study has described a framework for optimizing a prognostic biosphere model and gives some typical results estimated from posterior parameters. We used a Bayesian inversion method to improve the model estimates against observed variabilities in atmospheric CO_2 concentrations, mean annual AGB, and NPP by adjusting the physiological model parameters. The simulation based on posterior parameters produces moderate seasonal variations and annual amplitudes in atmospheric CO_2; however, it is still not able to solve explicitly the distributions of the mean amounts of AGB and NPP. Nevertheless, by using the optimization scheme, it is possible to reproduce some features of the physiological parameters that characterize different biomes, suggesting that the optimization scheme could be useful in understanding the dynamics of the global carbon cycle. However, it should be noted that the following advances will be required for further improvement of carbon cycle models to be made: a greater number of observations and improvements in the model regarding physiological processes and global biome distributions; reductions in the uncertainties in the atmospheric transport model and the a priori information; and models with greater spatial and temporal resolution.

References

Barrett, D. J.: Steady state turnover time of carbon in the Australian terrestrial biosphere, Global Biogeochem. Cy., 16, 1108, doi:10.1029/2002GB001860, 2002.

Berg, B. and Laskowski, R.: Litter decomposition: a guide to carbon and nutrient turnover, Adv. Ecol. Res., 38, 1–423, 2006.

Berry, J. and Björkman, O.: Photosynthetic response and adaptation to temperature in higher plants, Annu. Rev. Plant Phys., 31, 491–543, 1980.

Bousquet, P., Peylin, P., Ciais, P., Le Quéré, C., Friedlingstein, P., and Tans, P. P.: Regional Changes in Carbon Dioxide Fluxes of Land and Oceans Since 1980, Science, 290, 1342–1346, 2000.

Braswell, B. H., Sacks, W. J., Linder, E., and Schimel, D. S.: Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, Glob. Change Biol., 11, 335–355, 2005.

Brooks, A. and Farquhar, G. D.: Effect of temperature on the CO_2/O_2 specificity of ribulose-1,5-bisphosphate carboxylase/oxygenase and the rate of respiration in the light, Planta, 165, 397–406, 1985.

Callaway, R. M., DeLucia, E. H., and Schlesinger, W. H.: Biomass allocation of montane and desert ponderosa pine: an analog for response to climate change, Ecology, 75, 1474–1481, 1994.

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G. J., Mahecha, M. D., Montagnani, L., Papale, D., Rambal, S., and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, Global Change Biol., 16, 2813–2829, 2010.

Ciais, P., Tans, P. P., Trolier, M., White, J. W. C., and Francey, R. J.: A Large Northern Hemisphere Terrestrial CO_2 Sink Indicated by the $^{13}C/^{12}C$ Ratio of Atmospheric CO_2, Science, 269, 1098–1102, 1995.

Cramer, W., Kicklighter, D. W., Bondeau, A., Moor III, B., Churkina, G., Nemry, B., Ruimy, A., Schloss, A. L., and the participants of the Potsdam NPP Model intercomparison: Comparing global models of terrestrial net primary productivity (NPP): overview and key results, Global Change Biol., 5, 1–15, 1999.

Enting, I. G., Trudinger, C. M., and Francey, R. J.: A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO_2, Tellus, 47, 35–52, 1995.

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P.: Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components, Science, 281, 237–240, 1998.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K.-G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison, J. Climate, 19, 3337–3353, 2006.

GLOBALVIEW-CO_2: Cooperative Atmospheric Data Integration Project–Carbon Dioxide, NOAA ESRL, 2010.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y. H., Ciaisand, 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.

Howell, J. F. and Mahrt, L.: An adaptive decomposition: Application to turbulence, in: Wavelets in Geophysics, edited by: Foufoula-Georgiou, E. and Kumar, P., Vol. 1, 107–128, Academic Press, 1994.

Ichii, K., Suzuki, T., Kato, T., Ito, A., Hajima, T., Ueyama, M., Sasai, T., Hirata, R., Saigusa, N., Ohtani, Y., and Takagi, K.: Multi-model analysis of terrestrial carbon cycles in Japan: limitations and implications of model calibration using eddy flux observations, Biogeosciences, 7, 2061–2080, doi:10.5194/bg-7-2061-2010, 2010.

IPCC: Climate Change 2007: The Physical Science Basis, Cambridge University Press, 2007.

Ito, A.: Changing ecophysiological processes and carbon budget in East Asian ecosystems under near-future changes in climate: Implications for long-term monitoring from a process-based model, J. Plant Res., 123, 577–588, 2010.

Ito, A. and Oikawa, T.: A simulation model of the carbon cycle in land ecosystems (Sim-CYCLE): a description based on dry-matter production theory and plot-scale validation, Ecol. Model., 151, 143–176, 2002.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sens. Environ., 101, 534–553, 2006.

Jung, M., Vetter, M., Herold, M., Churkina, G., Reichstein, M., Zaehle, S., Ciais, P., Viovy, N., Bondeau, A., Chen, Y., Trusilova, K., Feser, F., and Heimann, M.: Uncertainties of modeling gross primary productivity over Europe: A systematic study on the effects of using different drivers and terrestrial biosphere models, Global Biogeochem. Cy., 21, GB4021, doi:10.1029/2006GB002915, 2007.

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, 1066, doi:10.1029/2001GB001463, 2002.

Kindermann, G. E., McCallum, I., Fritz, S., and Obersteiner, M.: A global forest growing stock, biomass and carbon map based on FAO statistics, Silva Fennica, 42, 387–396, 2008.

Kuroiwa, S.: Dry matter production by plants, Modern Biology, 9, 71–100, 1966.

Law, R. M., Rayner, P. J., Denning, A. S., Erickson, D., Fung, I. Y., Heimann, M., Piper, S. C., Ramonet, M., Taguchi, S., Taylor, J. A., Trudinger, C. M., and Watterson, I. G.: Variations in modeled atmospheric transport of carbon dioxide and the consequences for CO_2 inversions, Global Biogeochem. Cy., 10, 783–796, 1996.

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., and Schimel, D. S.: Ecological forecasting and data assimilation in a data-rich era, Ecol. Appl., 21, 1429–1442, 2011.

Maherali, H. and DeLucia, E. H.: Influence of climate-driven shifts in biomass allocation on water transport and storage in ponderosa pine, Oecologia, 129, 481–491, 2001.

Maksyutov, S., Patra, P. K., Onishi, R., Saeki, T., and Nakazawa, T.: NIES/FRCGC Global Atmospheric Tracer Tnansport Model: description, validation, and surface sources and sinks inversion, J. Earth Sim., 9, 3–18, 2008.

Meziane, D. and Shipley, B.: Interacting determinants of specific leaf area in 22 herbaceous species: effects of irradiance and nutrient availability, Plant Cell Environ., 22, 447–459, 1999.

Monsi, M. and Saeki, T.: Über den Lichtfaktor in den Pflanzengesellschaften und seine Bedeutung für die Stoffproduktion, Jpn. J. Bot., 14, 22–52, 1953.

Monson, R. K., Stidham, M. A., Williams, G. J., Edwards, G. E., and Uribe, E. G.: Temperature dependence of photosynthesis in *Agropyron smithii* Rydb. I. Factors affecting net CO_2 uptake in intact leaves and contribution from ribulose-1, 5-bisphosphate carboxylase measured *in vivo and in vitro*, Plant Physiol., 69, 921–928, 1982.

Oda, T. and Maksyutov, S.: A very high-resolution (1 km × 1 km) global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543–556, doi:10.5194/acp-11-543-2011, 2011.

Olson, R. J., Johnson, K. R., Zheng, D. L., and Scurlock, J. M. O.: Global and Regional Ecosystem Modeling: Databases of Model Drivers and Validation Measurements, Ork Ridge National Laboratory, 2001.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 Reanalysis, J. Meteorol. Soc. Jpn., 85, 369–432, 2007.

Peylin, P., Bousquet, P., Le Quéré, C., Sitch, S., Friedlingstein, P., McKinley, G., Gruber, N., Rayner, P., and Ciais, P.: Multiple constraints on regional CO_2 flux variations over land and oceans, Global Biogeochem. Cy., 19, GB1011, doi:10.1029/2003GB002214, 2005.

Pierce, L. L., Running, S. W., and Walker, J.: Regional-scale relationships of leaf area index to specific leaf area and leaf nitrogen content, Ecol. Appl., 4, 313–321, 1994.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Moonet, H. A., and Klooster, S. A.: Terrestrial ecosystem production: a process model based on global satellite and surface data, Global Biogeochem. Cy., 7, 811–841, 1993.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S., and Schmullius, C. C.: Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, Global Change Biol., 11, 378–397, 2005.

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, GB2026, doi:10.1029/2004GB002254, 2005.

Reich, P. B., Walters, M. B., and Ellsworth, D. S.: From tropics to tundra: global convergence in plant functioning, P. Natl. Acad. Sci. USA, 94, 13730–13734, 1997.

Reich, P. B., Walters, M. B., Ellsworth, D. S., Vose, J. M., Volin, J. C., Gresham, C., and Bowman, W. D.: Relationships of leaf dark respiration to leaf nitrogen, specific leaf area and leaf lifespan: a test across biomes and functional groups, Oecologia, 114, 471–482, 1998.

Reichstein, M., Tenhunen, J., Roupsard, O., Ourcival, J. M., Rambal, S., Miglietta, F., Peressotti, A., Pecchiari, M., Tirone, G., and Valentini, R.: Inverse modeling of seasonal drought effects on canopy CO_2/H_2O exchange in three Mediterranean ecosystems, J. Geophys. Res., 108, 4726, doi:10.1029/2003JD003430, 2003.

Running, S. W. and Hunt, E. R.: Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global scale models, in: Scaling Physiological Processes: Leat to Globe, edited by: Ehleringer, J. R. and Field, C. B., 141–158, Academic Press, 1993.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., and Morel, A.: Benchmark map of forest carbon stocks in tropical regions across three continents, P. Natl. Acad. Sci. USA, 108, 9899–9904, 2011.

Sacks, W. J., Schimel, D. S., Monson, R. K., and Braswell, B. H.: Model-data synthesis of diurnal and seasonal CO_2 fluxes at Niwot Ridge, Global Change Biol., 12, 240–259, 2006.

Saito, M., Maksyutov, S., Hirata, R., and Richardson, A. D.: An empirical model simulating diurnal and seasonal CO_2 flux for diverse vegetation types and climate conditions, Biogeosciences, 6, 585–599, doi:10.5194/bg-6-585-2009, 2009.

Saito, M., Ito, A., and Maksyutov, S.: Evaluation of biases in JRA-25/JCDAS precipitation and their impact on the global terrestrial carbon balance, J. Climate, 24, 4109–4125, 2011.

Schimel, D. S.: Terrestrial ecosystems and the carbon cycle, Glob. Change Biol., 1, 77–91, 1995.

Schubert, S. D., Rood, R. B., and Pfaendtner, J.: An Assimilated Dataset for Earth Science Applications, B. Am. Meteorol. Soc., 74, 2331–2342, 1993.

Scurlock, J. M. O., Cramer, W., Olson, R. J., Parton, W. J., and Prince, S. D.: Terrestrial NPP: toward a consistent data set for global model evaluation, Ecol. Appl., 9, 913–919, 1999.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate–carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), Global Change Biol., 14, 2015–2039, 2008.

Tans, P. P., Fung, I. Y., and Takahashi, T.: Observational Contrains on the Global Atmospheric CO_2 Budget, Science, 247, 1431–1438, 1990.

Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Indstrial and Applied Mathematics, 2005.

Valsala, V. and Maksyutov, S.: Simulation and assimilation of global ocean pCO_2 and air-sea CO_2 fluxes using ship observations of surface ocean pCO_2 in a simplified biogeochemical offline model, Tellus B, 62, 821–840, 2010.

van der Werf, G. R., Morton, D. C., DeFries, R. S., Olivier, J. G. J., Kasibhatla, P. S., Jackson, R. B., Collatz, G. J., and Randerson, J. T.: CO_2 emissions from forest loss, Nat. Geosci., 2, 737–738, 2009.

Wang, Y. P., Trudinger, C. M., and Enting, I. G.: A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales, Agr. Forest Meteorol., 149, 1829–1842, 2009.

Zhou, T. and Luo, Y.: Spatial patterns of ecosystem carbon residence time and NPP-driven carbon uptake in the conterminous United States, Global Biogeochem. Cy., 22, GB3032, doi:10.1029/2007GB002939, 2008.