Three decades of simulated global terrestrial carbon fluxes from a data assimilation system confronted with different periods of observations.

Karel Castro-Morales[1*], Gregor Schürmann[1], Christoph Köstler[1], Christian Rödenbeck[1], Martin Heimann[1,3] and Sönke Zaehle[1,2]

[1] Max Planck Institute for Biogeochemistry, Jena, Germany
[2] Michael-Stifel-Center Jena for Data-Driven and Simulation Science, Jena, Germany
[3] Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, Helsinki, Finland
*Now at: Friedrich Schiller University, Institute of Biodiversity, Aquatic Geomicrobiology, Jena, Germany

Abstract

During the last decade, carbon cycle data assimilation systems (CCDAS) have focused on improving the simulation of seasonal and mean global carbon fluxes over a few years by simultaneous assimilation of multiple data streams. However, the ability of a CCDAS to predict longer-term trends and variability of the global carbon cycle, and the constraint provided by the observations, have not yet been assessed. Here, we evaluate two near-decade long assimilation experiments of the Max Planck Institute – Carbon Cycle Data Assimilation System (MPI-CCDAS v1) using spaceborne estimates of the fraction of absorbed photosynthetic active radiation (FAPAR) and atmospheric CO_2 concentrations from the global network of flasks measurements sites from either 1982-1990 or 1990-2000. We contrast these simulations with independent observations from the period 1982-2010, as well as a third MPI-CCDAS assimilation run using data from the full 1982-2010 period, and an atmospheric inversion covering the same data and time. With 30 years of data, MPI-CCDAS is capable of representing land uptake to a sufficient degree to make it compatible with the atmospheric CO_2 record. The long-term trend and seasonal amplitude of atmospheric CO_2 concentrations at station level over the period 1982 to 2010 is considerably improved after assimilating only the first decade (1982-1990) of observations. After 15-19 years of prognostic simulation, the simulated CO_2 mixing ratio in 2007-2010 diverges by only 2±1.3 ppm from the observations, the atmospheric inversion and the MPI-CCDAS assimilation run using observations from the full period. The long-term trend, phenological seasonality and interannual variability (IAV) of FAPAR in the Northern Hemisphere over the last one to two decades after the assimilation were also improved. Despite imperfections in the representation of the IAV in atmospheric CO_2, model-data fusion for a decade of data can already contribute to the prognostic capacity of land carbon cycle models at relevant time-scales.

Keywords: *Data assimilation, Global Carbon cycle, land biosphere modeling, atmospheric CO_2.*

1 Introduction

The observed contemporary increase in atmospheric CO_2 is driven by anthropogenic emissions from fossil fuels and land-use change (2007-2016 average: 11.1 ± 0.6 GtC yr^{-1}), and the concurrent net carbon uptake of the ocean and land from the atmosphere, which take up approximately 22.4 % and 28 % of the anthropogenic flux, respectively. Despite recent advances in atmospheric observations, ocean and land modeling, there is an imbalance of 5.6 % (0.6 GtC yr^{-1}) between the ocean and land sinks, carbon emissions and changes in the atmospheric CO_2 concentration (Le Quéré et al., 2018). Throughout past decades, notable progress has been made to improve the performance of terrestrial biosphere models, but the simulated global terrestrial carbon fluxes and the net land carbon balance still have the highest uncertainties from all of the components of the global carbon cycle (Friedlingstein et al., 2014; Le Quéré et al., 2018). Quantifying the magnitude and dynamics of the global terrestrial carbon cycle across different temporal scales, and their contribution to the global carbon cycle, is challenging because the substantial heterogeneity and complexity in land ecosystems, and challenges in the quantification of contemporary effects and response of these ecosystems to increasing post-industrial CO_2 concentrations (Lienert and Joos, 2018; Stocker et al., 2014; Wang et al., 2017).

One strategy to reduce the mismatch between carbon flux predictions from land surface models and measured atmospheric CO_2 concentrations is through data assimilation (DA) techniques, meaning to "train" the land models by confronting them systematically with observations of carbon-related variables (Raupach et al., 2005). During DA, process-parameters of land surface models are adjusted through numerical minimization techniques to reduce the misfit between model results and actual observations under consideration of the statistical properties of both data sets. While atmospheric transport inversions are a method used to infer the sinks and sources of CO_2 between the atmosphere and land, or ocean, from atmospheric CO_2 measurements (Newsam and Enting, 1988; Peylin et al., 2013; Rayner et al., 1999; Rödenbeck et al., 2003), the application of carbon cycle data assimilation systems (CCDAS) provides additional opportunities. In CCDAS, the process-based carbon cycle mechanisms in land surface models are informed with measurements to support a better estimate of the terrestrial carbon cycle, and improve the capacity to project its dynamics. With this purpose, several CCDAS have been developed in the past (e.g., Kaminski et al., 2012; Kaminski et al., 2013; Lienert and Joos, 2018; Peylin et al., 2016; Scholze et al., 2016). The difference among some of these models is the variational or sequential statistical

approach they follow during the data assimilation process (Montzka et al., 2012). A common characteristic in these models is their capacity for integrating long-term and time consistent global available observational records related to the carbon cycle such as: atmospheric CO_2 measurements from flask and in situ networks (Conway et al., 1994), as well as remote sensing products of canopy phenology properties such as MODIS-NDVI (Moderate Resolution Imaging Spectroradiometer - Normalized Difference Vegetation Index) (Rouse et al., 1974) and FAPAR (Disney et al., 2016; Pinty et al., 2011a).

Previous studies have analyzed the prognostic capability of the data assimilation systems (e.g., Rayner et al., 2011; Rayner et al., 2005; Scholze et al., 2007; Schürmann et al., 2016), but only for few years of prognosis after the assimilation. Scholze et al. 2007, concluded that the CCDAS built around BETHY (Biosphere Energy-Transfer Hydrology) is capable of providing a prognostic period of four years (2000-2003) of atmospheric CO_2 after data assimilation of 21 years (1979 to 1999) of CO_2 concentrations. Schürmann et al., (2016) discussed the prognosis capability of the Max Planck Institute - Carbon Cycle Data Assimilation System (MPI-CCDAS v1) for two years after a short assimilation period of five years. Rayner et al. (2011) showed that the uncertainty related to model parameters during the prediction of CO_2 fluxes with a CCDAS is considerably reduced when the model parameters are constrained with two decades of atmospheric measurements; however, these results were obtained with a model that ignores the interacting effects of water, energy, and phenology on the carbon cycle predictions.

The overarching aim of this work is to understand the ability of the MPI-CCDAS v1 to make decadal projections of the land C cycle when the assimilation is confronted to different temporal windows from two observational constraints: FAPAR from remote sensing data and atmospheric CO_2 concentrations from the global flask measurements network. For this, we present three decades of modeled land carbon fluxes with the MPI-CCDAS and investigate the effect of withholding information from recent decades in the projected carbon fluxes and the ability of the model to reproduce the observations during the period of data assimilation. We also analyze trends and seasonal variations in the simulated signals during the periods of the assimilation and compare to independent results to evaluate the model performance. With these results, we gain insights in the number of observations (in terms of decadal scale) necessary in data assimilation systems to improve the representation of the global terrestrial carbon cycle

109 components. These results open the possibility of including newly measured data in
110 CCDAS that are only available for periods of less than a decade.

111 **2 Methods**

112 **2.1 MPI-CCDAS**

113 The MPI-CCDAS was built around the Jena Scheme Biosphere-Atmosphere Coupling
114 in Hamburg (JSBACH) land-surface model (Dalmonech and Zaehle, 2013; Raddatz et
115 al., 2007; Reick et al., 2013) and follows a variational approach that simultaneously
116 reduces the model-data misfit for multiple independent carbon cycle data sets
117 (Kaminski et al., 2013). Since its first development based on the BETHY - CCDAS,
118 the MPI-CCDAS has undergone several code modifications and improvements, as well
119 as tests of the assimilation of new observational data sets (e.g. Kaminski et al., 2012;
120 Kaminski et al., 2013; Rayner et al., 2005; Scholze et al., 2016; Schürmann et al., 2016),
121 with the aim of further improving the representation of land carbon fluxes. The history
122 of the MPI-CCDAS and other current CCDAS is extensively discussed in Scholze et
123 al. (2017).

124 The code of the MPI-CCDAS version in this work is identical to the one used in
125 Schürmann et al. (2016). The model calculates the half-hourly storage and surface
126 fluxes of energy, water and carbon in terrestrial ecosystems at coarse spatial resolution
127 ($8° \times 10°$ grid) (Fig. 1). This horizontal resolution allows computational feasibility and
128 a realistic computational cost for the set of experiments presented in this work.
129 Furthermore, previous evidence has shown that a higher spatial resolution in global
130 vegetation models does not exert a considerable influence in the simulated carbon
131 fluxes at global or regional scales when compared to results obtained with a coarse grid
132 (Müller and Lucht, 2007). The lack of influence to improve the simulated global C
133 fluxes due to changes in the model spatial resolution might also apply to CCDAS
134 (Peylin et al., 2016).

135 The spatial distribution of the different plant-functional types (PFTs) in JSBACH is
136 shown in Fig. S1 (Supplement). The selected parameters for the assimilation procedure,
137 their prior and range of values were based on Schürmann et al. (2016), where an
138 extensive sensitivity study lead to retain those parameters with a substantial effect on
139 the simulated carbon and water fluxes, as well as in phenology. The majority of the
140 selected parameters for the optimization are linked to phenology, but also there are
141 parameters related to photosynthesis and global parameters that control the land carbon
142 turnover during the assimilation. The final list of parameters together with their initial

value obtained from an independent forward simulation of JSBACH 3.0 (see Sect. 2.3.1) is shown in Table 1.

The MPI-CCDAS starts with an initial guess for the model control vector (p_{pr}) of, e.g. carbon cycle properties, and model states, and their Gaussian uncertainty ("prior") with covariance C_{pr}. The model control vector p is iteratively updated to minimize a joint cost function J (Eq. 1) describing the misfit between observational data-streams (d; FAPAR and atmospheric CO_2, both with covariance C_d) and the corresponding simulated observation operators of the MPI-CCDAS $M(p)$, taking into account the uncertainties in the observational data assuming a Gaussian distribution and the information from the prior.

$$J(p) = \frac{1}{2}(M(p) - d)^T C_d^{-1}(M(p) - d) + (p - p_{pr})^T C_{pr}^{-1}(p - p_{pr}) \qquad (1)$$

During the optimization procedure, a new model trajectory is determined in each iteration (i.e. in every cycle when the model re-calculates the cost function for the difference between the model parameters and the observational constraint), such that energy and mass are conserved through the entire assimilation window (Kaminski and Mathieu, 2017). The gradient of the cost function with respect to the model control vector ($\frac{\partial J}{\partial p}$) is evaluated with a tangent-linear version of JSBACH 3.0, which was generated through automatic differentiation using a TAF (Transformation of Algorithms in Fortran) compiler tool (Giering and Kaminski, 1998). With this tangent-linear version of the model code, the derivatives for the parts of the model code where $J(p)$ is evaluated (i.e., code parts that depend on the control variables), are accurately calculated following the chain rule of calculus. Thus, the mathematical formulation of the code involved in the cost function must be differentiable. Since this was not the case for the phenological code of JSBACH 3.0, the phenology scheme was updated following Knorr et al. (2010) where the minimum and maximum calculations in the entire code were replaced by smoothing functions to avoid abrupt transitions (Schürmann et al., 2016).

2.2   Observational data sets

2.2.1   FAPAR

The fraction of the radiation that is absorbed by plants during photosynthesis (FAPAR) is a component of the land-surface radiation budget that dynamically indicates the status of the vegetation canopy over space and time (Gobron et al., 2006). In a previous study, MPI-CCDAS was constrained by MODIS-TIP (Two-stream Inversion Package) FAPAR (hereafter TIP-FAPAR) generated from the inversion of a 1-D radiation

177 transfer model (Pinty et al., 2006; Pinty et al., 2007) using the MODIS broadband
178 visible and near-infrared spectral white sky surface albedo as input (Clerici et al., 2010;
179 Pinty et al., 2011a; Pinty et al., 2011b). For this study, the TIP-FAPAR product was
180 available only from 2003 to 2011, making it unsuitable for the indented longer
181 assimilation period. While there are long-term remotely sensed proxies of FAPAR,
182 such as the NDVI (Rouse et al., 1974), it has been found previously that NDVI was less
183 reliable than TIP-FAPAR in terms of the seasonal cycle amplitude of vegetation
184 seasonality (Dalmonech and Zaehle, 2013; Dalmonech et al., 2015). Therefore, we used
185 as FAPAR proxy the Global Inventory Monitoring and Modeling System (GIMMS)
186 NDVI product for the period 1982 to 2006 (Tucker et al., 2005), and merged it with the
187 TIP-FAPAR product to provide a longer record of vegetation greenness. The maximum
188 and minimum NDVI values were rescaled at the pixel level to coincide with those from
189 the TIP-FAPAR for the overlapping periods (i.e., 2003 to 2006) following:

190 $$FAPAR_{\text{mod}} = \frac{NDVI - NDVI_{\text{min,x}}}{NDVI_{\text{max,x}} - NDVI_{\text{min,x}}} \times \left(TIP_{\text{max,x}} - TIP_{\text{min,x}}\right) + TIP_{\text{min,x}} \qquad (2)$$

191 Where x is the period 2003 to 2006 for each data set, NDVI is the full NDVI product
192 from 1982 to 2006, with minimum values given by $NDVI_{\text{min}}$ and maximum by
193 $NDVI_{\text{max}}$. TIP_{min} and TIP_{max} are the corresponding minimum and maximum values
194 from the TIP-FAPAR product. With this approach, the resulting merged product
195 maintains the maximum and minimum values from TIP-FAPAR while preserving the
196 temporal dynamics of NDVI. The median uncertainty of the available TIP-FAPAR data
197 was considered as the uncertainty for the entire time-series. Due to a technical failure
198 in the MPI-CCDAS, the final $FAPAR_{\text{mod}}$ product used in the assimilation procedure
199 only spans from 1982 to 2006 and the last four years from the TIP-FAPAR product
200 were not considered. For this study, this product was aggregated to match the model
201 grid horizontal resolution considering background snow-free and snow-covered
202 conditions separately (Schürmann et al., 2016).

203 To discard pixels in the global FAPAR data that might lead to bias during the
204 assimilation procedure, we applied a mask to the global FAPAR grid following three
205 criteria: 1) we masked out the grid cells with crop-dominating phenology of > 20 %
206 since no explicit crop phenology is described in JSBACH. This step has consequences
207 in areas where other relevant functional types are also present in the same grid cells,
208 such as deciduous broadleaves that are also abundant in the USA and Europe. As a
209 result, the parameters related to deciduous broadleaves are constrained from other
210 locations; 2) we further masked out pixels that hold a low correlation ($R^2 < 0.2$) when

211 compared the prior model result and the observations, as we had previously found that
212 the MPI-CCDAS is incapable of correcting such poor model behaviors (Schürmann et
213 al. 2016). Finally, 3) we masked out pixels located in areas where phenology abundance
214 is low, i.e. deserts, because they would influence the optimization causing significant
215 bias due to global compensating effects. The final FAPAR product used during the
216 assimilation contains 40 % of the original number of pixels after the applied mask,
217 resulting in more pixels distributed in the Northern Hemisphere compared to the
218 Southern areas. This observational data will be referred hereafter as $FAPAR_{obs}$ (see Fig.
219 1 for the global distribution of mean $FAPAR_{obs}$ from 1982 to 2006).

2.2.2 Atmospheric CO_2 concentrations and observation operator

221 Measurements of atmospheric CO_2 mixing ratios were taken from the flask data
222 continuous record of 28 sites in the NOAA/CMDL station network (Conway et al.,
223 1994; Rödenbeck et al., 2003). The selection criteria included the length of the record
224 (on average 19 years) (Fig. A1) and focused on remote and ocean stations with low
225 impact of local carbon sources and sinks of carbon (Schürmann et al., 2016) (see the
226 location of CO_2 stations in Fig. 1). In the MPI-CCDAS, the atmospheric transport of
227 CO_2 is calculated by integrating the simulated half-hourly net CO_2 fluxes to monthly
228 values followed by the transport calculation with the Jacobian representation of the
229 atmospheric transport model TM3 that is driven with meteorology fields from NCEP
230 (National Centers for Environmental Prediction) reanalysis (Heimann and Körner,
231 2003; Rödenbeck et al., 2003). During the generation of the monthly transport matrices,
232 the precise sampling time of flask measurements as well as the 3-hourly atmospheric
233 transport was considered to minimize the representation error due to short-term
234 fluctuations in atmospheric transport and to minimize the impact of synoptic
235 atmospheric transport variability on the simulated seasonal and long-term dynamics of
236 atmospheric CO_2 at the monitoring stations. Through this approach, the non-linear
237 effect of weather anomalies on the surface fluxes were also taken into account. TM3
238 runs at horizontal "fine grid" (fg) resolution of $4° \times 5°$. Due to computational demands,
239 it is not possible at this stage to use the MPI-CCDAS at the same fine grid resolution
240 than in the TM3. The treatment of uncertainties is done in the same way as in the TM3
241 atmospheric inversion (Rödenbeck et al., 2003) but imposing a floor value of 1 ppm to
242 the uncertainties (Rayner et al., 2005) to allow a range for the comparison to the
243 observational operator.

244 We also compare the fluxes from the assimilation to fluxes obtained from an
245 atmospheric transport inversion (referred to as INV). Similar to the MPI-CCDAS, the

atmospheric transport inversion is constrained by atmospheric CO_2 data linked to surface fluxes through a tracer transport model, but the land surface CO_2 fluxes are adjusted directly rather than through changes in the parameters of a land-surface process model. The inversion set-up used in this study is similar to the Jena CarboScope v4.1 (Rödenbeck, 2005; Rödenbeck et al., 2003), involving the same TM3 model as in the MPI-CCDAS. To make the inversion results as comparable as possible to those from the MPI-CCDAS, we used in the inversion the same prior fluxes from fossil fuel emissions and ocean (Section 2.2.3), as well as the same CO_2 stations. This comparison also helps to gauge the impact of non-land surface fluxes on the ability to reproduce the observations.

2.2.3 Background carbon fluxes

To account for the total carbon balance during the comparison between the land fluxes from MPI-CCDAS and atmospheric concentrations, it is necessary to include background carbon fluxes (i.e., from fossil fuel emissions, use and change of land cover, and from the ocean).

Land-use and land-cover change: the LULCC fluxes were obtained from a transient simulation done with the JSBACH 3.0 forced with prescribed annual maps of modified cover fractions (Hurtt et al., 2006). These fluxes do not consider disturbances such as fluxes from fires.

Fossil fuel emissions: The FF emissions used for this work are the result of a merged product from various data sets to complete a long record of emissions, i.e., 1980 to 2012. This product was prepared for the GEOCARBON project (www.geocarbon.net) by P. Peylin after merging and harmonizing various data sets: 1) for the period 1980 to 1989, the CDIAC (Carbon Dioxide Information Analysis Center; http://cdiac.ess-dive.lbl.gov/) product prepared for the CMIP5 exercise (Andres et al., 2013; Andres et al., 2011; Andres et al., 1996); 2) for the period 1990 to 2009, the IER-EDGAR (Institute of Energy and Rational use of Energy, Stuttgart, Germany - Emission Database for Global Atmospheric Research; www.carbones.eu/wcmqs/project/) product where the FF emissions are constructed using the EDGAR v4.2 data set (http://edgar.jrc.ec.europa.eu/overview.php?v=42) and completed with profiles for different countries, emission sectors and time zones available for different temporal resolutions; and 3) for the period 2010 to 2012, the CarbonTracker product derived at NOAA-Climate Monitoring and Diagnostics Laboratory (CMDL; https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/).

Ocean fluxes: Two products were merged to account for the oceanic CO_2 fluxes: 1) results from the Jena CarboScope v3.4 for the period between 1990-2007 (Rödenbeck et al., 2013) (http://www.bgc-jena.mpg.de/CarboScope/?ID=s), and 2) annual ocean fluxes from the Global Carbon Budget 2014 (Le Quéré et al., 2015) (http://cdiac.ess-dive.lbl.gov/GCP/carbonbudget/2014/). The ocean fluxes for monthly resolution follow Takahashi et al. (2002), and the spatial distributions follow Mikaloff Fletcher et al. (2006).

2.3   Experimental setup

2.3.1   Spin up and preparation of initial files

The MPI-CCDAS was forced with meteorology from CRU-NCEP (the Climate Research Unit from the University of East Anglia, analysis of the NCEP reanalysis atmospheric forcing) version 6.1, available at daily resolution from 1901 to 2014 and a spatial resolution of 0.5° (Viovy and Ciais, 2015; last access July 2015). The atmospheric forcing fields (i.e., wind speed, air temperature, precipitation, downward short- and long-wave radiation and specific humidity) were remapped to the coarse (8° × 10°) model grid. Prescribed annual means (one annual global mean value) of atmospheric CO_2 were also included as part of the forcing fields for the model (https://www.esrl.noaa.gov/gmd/ccgg/trends/global.html, accessed July 2015).

Before the assimilation experiments, the JSBACH 3.0 model was spun up to equilibrium of the vegetation and soil carbon pools with 1901 atmospheric CO_2, land cover and 1901-1910 climate. The spin-up procedure was done for a model period of 1000 years with repeated cycles of atmospheric forcing data. After this period, a transient model simulation was also done with JSBACH 3.0 for the period 1901 to 2012. This transient simulation included a change in atmospheric CO_2, climate and land cover. The purpose of this simulation was: i) to obtain the initial conditions for the CCDAS experiments, and ii) to derive spatially resolved land-use emissions from a JSBACH 3.0 simulation as additional forcing (see section 2.2.2). Due to technical limitations, the cover fraction of each PFT is kept constant in MPI-CCDAS during data assimilation, and thus remained fixed through the simulation period to account for the imprint of the space-time dynamics of land-use change emissions on atmospheric CO_2 concentrations. After the spin-up procedure, an initial global scaling factor was set for the slowly varying carbon pool (f_{slow}, also selected as optimization parameter) to account for non-steady-state conditions at the beginning of the assimilation (Carvalhais et al., 2008; Schürmann et al., 2016).

314 **2.3.2 Assimilation experiments**

315 During the assimilation procedure, the model was forced with the same daily reanalysis

316 atmospheric data used during the model spin up. In this study we present the results of

317 three long-term experiments using the MPI-CCDAS, which differ in the timeframe of

318 the observational records used during the assimilation: 1) ALL, covers data in 1980-

319 2010 and includes the complete available timeframe of the two observational data sets,

320 i.e., for FAPAR is from 1982 to 2006 and for the atmospheric CO_2 concentrations from

321 1982 to 2010; 2) DEC1, covers observations from the two data sets available from 1982

322 to 1990; and 3) DEC2, covers measurements available from the two data sets from 1990

323 to 2000 (Fig. A2). Because of the different lengths of the CO_2 records for some stations,

324 this ultimately leads to a different number of observations used for each experiment

325 (Fig. A1).

326 The simulation period in the three assimilation experiments is from 1970 to 2010. The

327 first ten years (1970 to 1979) of the results are discarded because during this period the

328 phenology, vegetation productivity, and the fast land C pools adjust to the new model

329 control vector p. Through this adjustment any imprint of the initial conditions on the

330 calculation of the cost function is avoided. The soil C pool at the beginning of the

331 experiment was included in the model control vector. and only results from 1980 are

332 reported below. The results of the assimilation for the periods of time that fall within

333 the observational temporal window are considered for model diagnostic, whereas the

334 periods that fall outside the assimilation window on each experiment are periods of

335 model prognosis, i.e., the prognosis period in DEC1 is from 1991 to 2010, and in DEC2

336 for 2001 to 2010.

337 **3   Results**

338 We first evaluate the long-term trends, seasonal and spatial variability of the FAPAR

339 and carbon fluxes from the different assimilation experiments (Section 3.1 to 3.3), and

340 based on these analyze the prognostic ability of the MPI-CCDAS (Section 3.4). To

341 facilitate the analysis in some of our results, the global land is divided into eight regions:

342 Boreal West and East (BW and BE, for latitudes north of 60° N), temperate Northwest

343 and Northeast (TNW and TNE, between latitudes 20° N and 60° N); tropical West and

344 East (TW and TE, between latitudes 20° N and 20° S); temperate Southwest and

345 Southeast (TSW and TSE, for latitudes south of 20° S) (Fig. 1).

346 **3.1   Phenology**

347 In all assimilation experiments, the RMSE and the bias between the modeled and

348 observed FAPAR for 1982 to 2006 is reduced compared to the PRIOR (Table 2). One

349   important cause for this improvement is the change in the spatial distribution of the

350   yearly maximum leaf area index (LAI) due to the optimization of the PFT-specific

351   maximum LAI (Λ_{max}) parameter (Fig. S2) (see also section A1 and Fig. A3 in the

352   Appendix for more specific results of parameters changes due to the assimilation). The

353   improvement occurs in all regions (Fig. 2), despite notable differences between the

354   different assimilation experiments. In the decadal experiments DEC1 and DEC2, the

355   largest error reduction compared to the PRIOR is 19 % for boreal regions, while in the

356   temperate areas this reduction is about 16 %. In the ALL experiment, larger reductions

357   of 21 % on average are obtained in the tropical regions TE and TW.

358   One important factor in the error reduction is a substantial increase in the linear global

359   correlation (R^2) in FAPAR during spring and autumn in experiments DEC1 (0.42 and

360   0.48, respectively) and DEC2 (0.48 and 0.47, respectively) with respect to the PRIOR

361   (0.31 and 0.33, respectively), with changes mostly taking place in the Northern

362   Hemisphere (Fig. S3). An analysis for representative pixels (Fig. 1) shows that the

363   assimilation procedure results in a better representation of the timing and amplitude of

364   the mean seasonal cycle, particularly in the temperate and boreal zones of the Northern

365   Hemisphere (Fig. S4). As a result, the average global R^2 between modeled and observed

366   FAPAR increased with respect to the PRIOR experiment from 0.17 in the PRIOR to

367   0.20 for ALL and 0.34 for both DEC1 and DEC2 (Table 2, Fig. S3). Further details on

368   the pixel level analysis are presented in section A2 of the Appendix.

369   The observed FAPAR signal exhibits positive long-trends, indicating a greening trend

370   of vegetation for most of the regions, with the exception of the TSW region, where the

371   long-term trend indicates a decrease of FAPAR (i.e., browning). In most of the regions,

372   the assimilation the assimilation results agree on a positive long-term trend as in the

373   observations, the magnitude of this trend is in disagreement to the observations (Fig.

374   3). Particularly in the BE region, the PRIOR experiment overestimates the $FAPAR_{obs}$

375   trend by almost double. After the assimilation, the simulated FAPAR trend is reduced

376   leading instead to a slight underestimation of the growth rate in all of the posterior

377   experiments. In the TWS region, the assimilation improved the long-term trend from a

378   positive to a negative growth rate in the three posterior experiments. The most

379   substantial disagreement between $FAPAR_{obs}$ and $FAPAR_{mod}$ occurs in the TW region,

380   where the observations show a positive trend in FAPAR during the period of analysis,

381   whereas this is not captured in the PRIOR and all the posterior experiments. Despite

382   these trend adjustments, the model-data error (based on the four-years mean differences

to the observations at regional scale) remains more or less constant across the thirty-year period (Fig. 4).

The observed FAPAR signal also contains a small amount of interannual variability (Fig. S5). Compared to observations, the simulated IAV of FAPAR (obtained from the monthly signal for each experiment) is improved only in the predominantly temperature controlled boreal regions, whereas in temperate and tropical areas with a larger contribution of moisture-controlled phenology, the assimilation does not improve the variability (Fig. S5).

3.2 Atmospheric CO_2

To diagnose the performance of the MPI-CCDAS with respect to the atmospheric mole fractions of CO_2, we compare the observed and simulated values, in terms of the mean seasonal cycle, IAV and monthly growth rate, in three stations: 1) Alert (ALT) at the Northern Hemisphere, 2) Mauna Loa (MLO) at the Tropics, and 3) South Pole (SPO) at the Southern Hemisphere. Results of this comparison are shown in Fig. 5. For MLO and ALT, the timing of the seasonal cycle is already well reproduced in the PRIOR simulation, and the assimilation corrects errors in the amplitude of the seasonal cycle and the long-term trend. At SPO, there are also large relative differences between the model results and the observations, however, of a much smaller magnitude than for the two other stations. After the assimilation, the seasonal phase of CO_2 is shifted by approximately a month to better match the pattern in the measurements in the three experiments, and the amplitude of the seasonal cycle is in better agreement with the observations than compared to the PRIOR.

Figure 6 demonstrates that these examples are broadly representative of the global changes due to the assimilation. Fig. 6a shows a reduction in the CO_2 amplitude for stations of the Northern Hemisphere (> 40 °N) after the assimilation, which is in better agreement to the observations than the PRIOR simulation. The most substantial amplitude reduction occurs in the northernmost station (ALT), where the seasonal amplitude decreases from 23.5 ppm in the PRIOR experiment to 16.5 ppm in the ALL experiment, bringing it closer to the observed amplitude of 14.4 ppm. The latitudinal distribution of the linear correlation coefficient between the observed and simulated mean seasonal cycles is depicted in Fig. 6b, and demonstrates a very good agreement ($R^2 > 0.9$) in the Northern Hemisphere in all of the experiments (including the PRIOR simulation). In the tropics (specifically between 20 °N and 40 °N), the misfit of the phasing of the seasonal cycle is improved after the assimilation, as evidenced by an increased linear correlation. However, this is achieved at the expense of a considerable

reduction in the amplitude of the seasonal cycle compared to the observations. The results from the atmospheric inversions (INV) show a closer statistical agreement with the observations, as shown in Fig. 5 and Fig. 6.

During the nearly thirty years of atmospheric CO_2 data available, the time series of the CO_2 mole fractions in the PRIOR model results, strongly underestimate the long-term trend, and start to deviate in the first five years of the time series. In all the assimilation experiments, the long-term atmospheric CO_2 trend is in much better agreement to the thirty-years trend of the measurements in the entire period of the assimilation (leftmost panels of Fig. 5 and Fig. 6c). The mean growth rate calculated from the results of the ALL experiment is in good agreement with the results in the observations (0.15 ppm month^{-1} in both cases) compared to the PRIOR model (0.087 ppm month^{-1}). Despite the moderate improvement, the MPI-CCDAS is incapable of improving the IAV of the atmospheric CO_2 concentration substantially; with the most notable deviations from the observed signals remaining unchanged after the assimilation procedure (Fig. 5).

3.3 Global and regional carbon pools and fluxes

We next compare the simulated land carbon cycle in the PRIOR and posterior experiments to independent data. In the posterior experiments, the vegetation C pool decreased between 14 and 20 % of the value in the PRIOR but remaining within the range of the literature estimate (442±146 PgC). The global soil C stock showed significant changes after the assimilation. In all the posterior experiments, the soil C pool decreased by 45, 43 and 53 % with respect to the value in the PRIOR. Still, the total C in the soil (1362 PgC) in the ALL experiment after the assimilation is in closer agreement to the estimate from the Harmonized World Soil Database (http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML; last access January 2015) of 1343 PgC (Table 3). As for the total global vegetation C stock, the PRIOR and assimilation are in closer agreement to the lower end of the estimate by Carvalhais et al., 2014 (296 PgC).

The simulated latitudinal GPP values agree well with the data-driven Model Tree Ensemble (MTE) estimate from Jung et al. (2011) for the period 1982 to 2010 north of 30 °N. However, the assimilation results are low biased in the tropics, which propagated into lower estimates of global GPP in all the posterior results (Fig. 7d and Table 3). After the assimilation, the global GPP and NPP are reduced in the three posterior experiments compared to the PRIOR (118.8 PgC yr^{-1} and 54.5 PgC yr^{-1}, respectively). In contrast to the posterior global mean of GPP, the value in the PRIOR simulation compares favorably well to the global mean from the MTE product (118.9 PgC yr^{-1})

453 for the same period of analysis. The global mean GPP is reduced by up to 26 PgC yr^{-1}
454 on average in the three posterior experiments compared to the PRIOR experiment.
455 Simulation DEC1 experienced the largest reduction in the global photosynthetic C
456 uptake (83.1 PgC yr^{-1}) relative to the PRIOR value (Table 3 and spatial distribution of
457 the GPP difference to the PRIOR in Fig. S6d, f, and h).
458 At large-scale, the variation of the NBE (net biome exchange of CO_2 with the
459 atmosphere, calculated as the Net Ecosystem Exchange (NEE) minus the flux related
460 to land use change) from all of the simulations through the time series is similar to that
461 from the Global Carbon Project 2017 (GCP17; Le Quéré et al., 2018) and INV, with
462 the significant anomalies collocated in time (Fig. 7a, Fig. A4). Contrary to the PRIOR
463 simulation, the total annual NBE from the three posterior experiments falls within the
464 uncertainty (shadow green area in Fig. A4d calculated as ±1 standard deviation) of the
465 mean NBE from the terrestrial ecosystem models in the GCP17. However, the 1982-
466 2010 mean net biome exchange in all of the assimilation experiments through the time
467 series is on average 1.4 PgC yr^{-1} lower than the flux in the PRIOR simulation (−2.06
468 PgC yr^{-1}) and 0.8 PgC yr^{-1} less than the GCP17 value (−1.27±0.97 PgC yr^{-1}) (Table 3,
469 Fig. A4d and Fig. S7 for summary of C balance).
470 In all MPI-CCDAS simulations, the NEE is reduced relative to the PRIOR in most of
471 the Southern Hemisphere, while it is increased in the Northern Hemisphere (Fig. S6c,
472 e, and g). Temperate northern areas contribute the most to the global net CO_2 uptake.
473 In the boreal east and west regions (BE and BW), the net land C emissions increased in
474 all of the posterior experiments compared to the PRIOR (Fig. S6c, e and g) with the
475 most significant increase in BE for DEC2 (−0.29 PgC yr^{-1}) relative to the corresponding
476 value in the PRIOR (−0.09 PgC yr^{-1}). The decrease in GPP in the tropics is depicted in
477 the latitudinal gradient of NBE shown in Fig. 7c and in the spatial distribution of the
478 NEE difference between the PRIOR and the posterior experiments (Fig. S6c, e, and g).
479 As in the tropics, the NEE in the southern temperate region is consistently reduced after
480 the assimilation experiments, also switching the NEE of the TSE region from a C sink
481 of −0.18 PgC yr^{-1} in the PRIOR to a mean C source to the atmosphere of 0.016 PgC
482 yr^{-1} in the DEC2 experiment.
483 The magnitude of the global NBE and GPP is smaller in the posterior experiments than
484 in the PRIOR. However, the trend in the anomaly of these fluxes (calculated relative to
485 the temporal mean of each time series) is comparable in all the experiments (Fig. 7a
486 and b), suggesting that the response to the environmental conditions is similar through
487 the simulation period also after the assimilation. This robust response shows, e.g., in

488   GPP a similar and gradual increasing C uptake (positive trend) during the period of

489   analysis, only with a slightly reduced slope in the PRIOR experiment (Fig. 7b).

490   **3.4   Prognostic capability of MPI-CCDAS**

491   Finally, we evaluate the goodness of the model-data fit of the decadal assimilation runs

492   with respect to their long-term carbon cycle simulation relative to: i) that of the prior

493   and ii) that of the assimilation run using data from the 30 years-experiment as a

494   reference case for "best possible" model-data match given the structural limitations of

495   the MPI-CCDAS to match the observations (as evaluated in Sections 3.1 and 3.2). We

496   calculate the four-years mean differences between the atmospheric CO_2 mole fraction

497   measurements and the CO_2 model results and also the INV results, for all of the stations

498   (Fig. 8). In the ALL assimilation experiment, the atmospheric CO_2 concentration

499   consistently matches the observations across the entire assimilation period (that also

500   corresponds to the window of assimilation) with a -0.03 ± 1 ppm average bias to the

501   observations (Fig. 8). This is comparable to the trend (Fig. 6c), and four-years mean

502   differences inferred by the inversions, where the four-years mean results in the ALL

503   fall within the standard deviation of the four-years mean of the INV (Fig. 8). This is in

504   striking contrast to the PRIOR experiment, where the four-years mean of the CO_2 mole

505   fraction at the end of this simulation is 18.8 ppm lower than observed. For the DEC1

506   experiment, the four-year mean difference among the measurements and the model

507   results is between -0.3 and 0.3 ppm in the 1980s. This level of model-data agreement

508   remains for the 1990s, where the experiment did not see any observations. After the

509   year 2000, the fit increasingly degrades, with an underestimation of the CO_2 mole

510   fraction by 1.6 ppm for the last four-years average. However, this is still a 90 %

511   reduction in misfit compared to the PRIOR experiment.

512   We next quantify the RMSE between the CO_2 measurements and model results for each

513   station for four different periods: 1982-1990, 1990-2000, 2000-2010 and 1982-2010

514   (Fig. 9 and Fig. A2). The large bias of the PRIOR is reflected in the RMSE where the

515   results of this experiment have the most substantial error in all of the stations and

516   periods (between 2.8 and 18.7 ppm) (Fig. 9). For the posterior experiments with a

517   decadal window of assimilation (DEC1 and DEC2), the performance of the assimilation

518   of CO_2 mole fraction also improves substantially across all time periods. Within the

519   period of the assimilation, the difference to the measurements and RMSE is most

520   strongly reduced, and the error increases somewhat outside of the window of

521   assimilation. The model results show that when only the first decade of data is

522   assimilated (DEC1), a more significant deviation to the long-term trend of atmospheric

523   CO_2 occurs between 2000 and 2010 compared to DEC2 and ALL (Fig. 9c). Similarly,

524   a larger bias is also observed in the results from DEC2 where the lowest four-years

525   mean difference between the observations and the assimilation results takes place in the

526   period of the window of assimilation for this experiment (1990-2000) (Fig. 8 and Fig.

527   9b for RMSE). During this period, the model overestimates the CO_2 atmospheric

528   concentration only by 0.15 ppm on average whereas, for the periods outside the window

529   of assimilation, the CO_2 concentration is underestimated by 0.64 ppm in the period

530   1982-1990, and by 1.04 ppm in the period 2000-2010. Thus, also in experiment DEC2

531   the prognostic skill of CCDAS is reduced outside the window of assimilation, and the

532   long-term trend is less well reproduced than in the ALL experiment.

533   The analysis of the four-year mean differences for the period 1982-2006 between

534    $FAPAR_{obs}$ and the results of the PRIOR and assimilation experiments at the regional

535   scale (areas in Fig. 1) reveals, contrary to the CO_2 observations, a near constant four-

536   years mean FAPAR difference within the time series and each of the experiments (Fig.

537   4). In general terms, the decadal experiments are better able to reproduce the mean

538   FAPAR across all regions. The largest difference between posterior results to the

539   observations is in the tropical regions, where the FAPAR four-years mean difference

540   showed that the observations remained consistently larger than the ALL results by on

541   average 0.042 in TE and 0.095 in TW (Fig. 4). Importantly, however, the trend

542   correction for the boreal and temperate areas (Fig. 3) are similar across the different

543   assimilation experiments, suggesting that important biases of the JSBACH 3.0 model,

544   including the tendency to simulate too strong boreal greening, can be readily reduced

545   with only 10 years of data, as the further improvement with the 30 years record is small.

546   **4   Discussion**

547   The parameter optimization with a simultaneous assimilation of long-term spaceborne

548   FAPAR and atmospheric CO_2 measurements in the MPI-CCDAS, resulted in a

549   considerable reduction in the cost function and norm of the gradient, which can be seen

550   as an overall improvement in the modeled global carbon fluxes with a decrease in the

551   root mean squared error of the MPI-CCDAS compared to the CO_2 and FAPAR and

552   observations (Fig. 9 and Fig. 2). The trajectory of model parameters involved in the

553   optimization differed for each experiment and each phenotype. While some parameters

554   were consistently retrieved after the assimilation, such as the maximum leaf area of

555   grasses and shrubs and the correction parameter for the initial soil pool size, some final

556   parameter estimates varied considerably between the three experiments, e.g., the

557   tropical maximum leaf area index and some of the parameters controlling the

558   seasonality of the phenology (Fig. A3). These variations lead to regional differences in

559   the simulated compartment fluxes GPP and ecosystem respiration, which are not well

560   constrained from the observations. Interestingly, these differences result in very similar

561   absolute values in global carbon fluxes and their trends. This demonstrates a certain

562   degree of equifinality in the results and cautions a too stringent interpretation of the

563   MPI-CCDAS outcome in terms of improving understanding about biosphere processes

564   and their long-term trends.

565   **4.1   FAPAR**

566   MPI-CCDAS is capable of extracting information about the seasonal cycle and the

567   long-term trends from the FAPAR observations. Using decade-long FAPAR data

568   during the assimilation (DEC1 and DEC2), already leads to notable improvement of

569   the simulated seasonal phenology of the land surface within and outside the window of

570   assimilation, i.e., maintaining these changes during the prognosis periods. This

571   improvement is predominantly the result of the ability in the model to simulate the

572   timing of green-up and brown-down in spring and summer through the optimization of

573   parameters that regulate the onset and end of the growing season (i.e., parameters for

574   temperature and day-length thresholds). The greening effect is especially taking place

575   in the Northern Hemisphere, dominated by the phenotypes deciduous and evergreen

576   needle leaf and extra-tropical deciduous trees.

577   The long-term greening trend in the vegetation of boreal regions previously observed

578   in spaceborne data (Forkel et al., 2016; Lucht et al., 2002), was captured in the results

579   of MPI-CCDAS before the assimilation, but it was mostly overestimated in northern

580   regions and underestimated in the Southern Hemisphere. After the assimilation

581   experiments, the greening trend was improved primarily in the boreal areas and is in

582   closer agreement to the reported satellite FAPAR data. The modest improvements

583   achieved in the simulated greening trend of temperate areas in the western hemisphere

584   are associated with a decreased performance in the eastern hemisphere, indicating that

585   the model structure of MPI-CCDAS is incapable of reconciling regional differences.

586   This difference could be an indicator of the need to parameterize both hemispheres

587   differently in terms of their phenological response to the underlying driving factors

588   (such as temperature, moisture availability and day-length); also, this could be due to

589   the lack of process to account for the land-use or vegetation dynamics in the MPI-

590   CCDAS.

591   Despite these broad-scale improvements, the MPI-CCDAS does not reproduce the

592   magnitude of the greening trend and its interannual variability in all the posterior

593   experiments at pixel and regional levels. This is likely a result of the MPI-CCDAS
594   structure, which relies on few globally relevant PFT-level parameters. Although some
595   of the phenological parameters in CCDAS adapt to local mean growing season
596   temperature, other thresholds are only globally applicable, causing a trend to
597   temperature grasslands that cover a wide climatological range. For example, some of
598   the global parameters such as f_{aut_leaf} and f_{slow}, imply that improvements of modeled
599   fluxes in the boreal regions directly affect fluxes in the tropics, inducing parameter
600   changes to compensate for the altered C fluxes. Defining instead such global parameters
601   per PFT would alleviate this issue but will compromise the computational cost and
602   might not necessarily reduce the overall uncertainty. Another technical challenge is the
603   use of a spatially mixed signal at the coarse spatial model resolution (due to the high
604   computational requirements necessary to increase model resolution) to infer PFT-
605   specific parameters. A likely better strategy for constraining PFT-specific parameters
606   would be to resample the highly resolved satellite product to PFT-specific FAPAR
607   classes per pixel before the aggregation into a global grid. This change would allow
608   finding more spatially refined classes and provide PFT-specific FAPAR maps to the
609   CCDAS to reduce issues in the identification of phenological parameters for different
610   climatic regions.

611   Except for the tropical latitudes, the difference between the regional IAV of the
612   observations and model output is small compared to seasonal variations. The modeled
613   signal remains within a range of 0.05 (dimensionless) $FAPAR_{obs}$. The signal and the
614   model-data difference is also smaller than the global mean retrieval error of the FAPAR
615   product, which is ±0.2088 (Schürmann et al., 2016). This error was used to quantify the
616   observational FAPAR uncertainty in the assimilation, thereby reducing the ability of
617   the MPI-CCDAS to detect and correct such smaller variation. Overall, the lacking
618   match of the IAV may therefore be of little overall concern. Nevertheless, the lower
619   than observed IAV in the tropical bands may be indicative of too weak drought response
620   in the maximum leaf area index of the model. Although the assimilation procedure
621   allows changes in the phenology response to water stress (given by parameter τ_w), the
622   assimilation procedure decreased the drought sensitivity of tropical phenology given
623   the entire spatially explicit FAPAR time series, and therefore did not allow capturing
624   the regional drought events that could be in principle linked to changes in LAI.

625   The technical error during the assimilation procedure to not include the period from
626   2007-2010 in the $FAPAR_{mod}$ product does not influence however the decadal results
627   observed here, because the main information gain of the CCDAS in terms of phenology

628 stems from the seasonal cycle, with little effect on the overall trends between the three
629 assimilation experiments with different time periods.
630 Bearing in mind the different spatial resolution of methods (i.e., model grids and remote
631 sensing pixels), a robust comparison between the mean and maximum LAI values
632 before and after the assimilation per region are presented in Table A1 of the Appendix.
633 The results fall within LAI values from MODIS and LiDAR reported in the literature.
634 Ground-based observations in the tropical Amazon-Savanna transition forest between
635 2005 and 2008 show an annual mean LAI value for the total canopy of 7.4 ± 0.6 m^2 m^{-2},
636 and for the seasonally flooded forest the value of 3.4 ± 0.1 m^2 m^{-2}. For the remote
637 sensing data from MODIS, the reported values are 6.2 ± 0.2 m^2 m^{-2} and 5.8 ± 0.3 m^2 m^{-2},
638 respectively (Biudes et al., 2014). In the eastern Amazon forest, the remote sensing-
639 based LAI has been reported as 6.2 m^2 m^{-2} from LiDAR, and 4.8 m^2 m^{-2} with a low
640 end of 2.0 m^2 m^{-2} from MODIS (Qu et al., 2011).

641 **4.2   Atmospheric CO$_2$**

642 The considerable improvement of the seasonal amplitude and the long-term trend of
643 atmospheric CO$_2$ at Northern Hemisphere stations is independent of the different
644 periods of data used for the assimilation. However, the MPI-CCDAS consistently fails
645 to resolve some of the features of the year-to-year variability detected in the measured
646 atmospheric CO$_2$ stations, which translates into an acceptable, but far from perfect fit
647 to the inferred annual carbon budget of the GCP17 (Le Quéré et al., 2018). We
648 compared the performance to the results from an atmospheric CO$_2$ inversion (INV) with
649 the same input fields and atmospheric transport model than MPI-CCDAS, to illustrate
650 that these deviations do not reflect uncertainties in the representation of the atmospheric
651 transport. It needs to be mentioned that both the choice of the atmospheric transport
652 model (and associated imperfections at resolving the vertical and lateral atmospheric
653 transport of CO$_2$), and the method to aggregate atmospheric observations to obtain an
654 estimate of the annual growth rate in the global carbon budget introduce some error in
655 any forecast of the interannual variability. As a consequence, only the occurrence of
656 more significant model-data mismatches can be interpreted as an actual result of the
657 MPI-CCDAS' inability to correctly resolve the carbon flux variation.

658 Notably, the model lacks the representation of some key processes that contribute to
659 climate induced interannual variability of the carbon cycle, such as the possibility to
660 dynamically account for fire disturbance (Lasslop et al., 2014), ENSO related tropical
661 peat-land fires (van der Werf et al., 2008), or the increase of terrestrial carbon uptake
662 after large-scale volcanic eruptions such as for Mt. Pinatubo in 1991 (Lucht et al., 2002;

663   Mercado et al., 2009). Omitting fluxes in the current model configuration due to fire

664   events may impair the ability of the model to infer the atmospheric growth rate of CO_2

665   associated with El Niño events (Frölicher et al., 2011; Frölicher et al., 2013). One way

666   to overcome the IAV mismatch would be to include fire fluxes in the model by

667   prescribing them from, e.g., the Global Fire Emissions Database (GFED, van der Werf

668   et al., 2010), however the latest version of this data set (Version 4.0) is only available

669   for years from 1997 which is a limiting factor for the timeframe of the simulations in

670   this work. However, the contribution of these interannual variations to the overall CO_2

671   cost function is low in comparison to the signal contained in the seasonal cycle and

672   deviations in the long-term trend, such that the MPI-CCDAS may simply not be

673   sensitive enough to these aggregate system properties like the response of the tropical

674   carbon cycle to El Niño events given the uncertainty in the atmospheric transport and

675   the observational representation error.

4.3   Carbon-cycle simulation

677   Independent of the amount of data used in the assimilation window, our results show

678   that the GPP and NEE were consistently reduced globally compared to the PRIOR run,

679   i.e., less carbon uptake by plants leading to the model results to be in closer agreement

680   to other independent estimates such as the GCP17. The MPI-CCDAS suggests a

681   somewhat lower average annual atmospheric CO_2 growth rate (calculated by the sum

682   of the net C fluxes from the ocean, land and fossil fuel emissions) than the one estimated

683   in the GCP17 (Le Quéré et al., 2018), even if the MPI-CCDAS estimate falls within the

684   uncertainty of the GCP17 (Fig. 7 and S7). Most of the difference stems from small

685   differences in the assumed fossil and ocean carbon fluxes. In the case of the carbon

686   fluxes from fossil fuels, the data prescribed in MPI-CCDAS does not contain fluxes

687   due to, e.g., cement and flaring, thus the magnitude of the annual carbon sources

688   through the time series is consistently lower but still within the ±5 % uncertainty of the

689   GCP17 data (Le Quéré et al., 2018) (Fig. A4). As for the ocean carbon sink, the annual

690   mean values prescribed in MPI-CCDAS are also of lower magnitude than the mean

691   value in the GCP17 but falling in the lower limit of the uncertainty value (Fig. A4c and

692   S7). The flux due to LULCC prescribed in MPI-CCDAS is also of smaller magnitude

693   than that one from the GCP17 because the simulation made by JSBACH 3.0 does not

694   consider disturbances like fires and gross transitions, which might have also contributed

695   to the lower land C sink obtained in the assimilation experiments compared to the total

696   land C sink in GCP17 (Fig. A4d).

697   The MPI-CCDAS GPP matches well the observation-based product MTE-GPP (Jung

698   et al., 2007) in regions with a distinct, light- and temperature-driven seasonal cycle (i.e.,

699   north of approx. 30 °N), translating to a reduction in modeled GPP by 0.7 PgC yr^{-1} in

700   boreal regions. However, similar to the results in Schürmann et al. (2016) with only

701   five years of assimilation, the tropical productivity is strongly reduced by the

702   assimilation to estimates that are substantially lower than independent estimates such

703   as MTE. This feature is likely the result of a global compensating effect to

704   heterotrophic respiration, and this effect is observed in the drop of the photosynthetic

705   capacity (f_{photos}) in the tropical evergreen and deciduous PFTs, as well as in the

706   reduction of the maximum tropical LAI in the three assimilation experiments compared

707   to the PRIOR. In addition, another critical factor influencing the global reduction of

708   GPP and the tropical uptake of C appears to be related to the difference in data

709   availability of CO_2 stations between the defined assimilation windows. Specifically,

710   this is evident in the results of the data-poorer experiment DEC1, where the topical GPP

711   is substantially lower than in the independent estimates and in the assimilation

712   experiments that use more stations (DEC2 and ALL). As a result, the mean tropical

713   land C source to the atmosphere in the prior experiment (mean NBE value of 0.12 PgC

714   yr^{-1}, and minimum value of −0.07 PgC yr^{-1}, reflecting C uptake in the 4 °S latitudinal

715   band) was increased to 0.37±0.17 PgC yr^{-1} on average for all the posterior results.

716   The NPP:GPP ratio in ALL and DEC2 decreased to 0.35 and 0.31, respectively, when

717   compared to the PRIOR value (0.45). This reduction might be mainly because the NPP

718   is not well constrained from the atmospheric record. In JSBACH 3.0, autotrophic

719   respiration (Ra) is directly coupled to GPP, hence the fraction of GPP partitioned to Ra

720   leads to an increase in the seasonal cycle of the ecosystem respiration. An increase in

721   Ra with respect to the PRIOR (which is only visible in the global average value in

722   DEC2; Table 3), results in a reduced net land carbon uptake, masking the smaller

723   changes in the vegetation turnover.

724   The reduction in the soil C pool after the assimilation can be explained due to an

725   unavoidable effect in the model. The MPI-CCDAS was initially spun-up until the soil

726   C pools reached equilibrium considering pre-industrial forcing; however, this new

727   initial state does not consider climate variability. To compensate for this and to reduce

728   the respiration when the MPI-CCDAS is confronted with contemporary changes in the

729   climate, the model creates an artificial C sink that leads to a reduction in the soil C

730   stocks. It is important noting that the JSBACH 3.0 version used in this MPI-CCDAS

731 does not include permafrost processes; therefore, the global soil C stock might still be
732 underestimated.

4.4 Value of long-term data sets to constrain CCDAS

734 Notwithstanding the MPI-CCDAS conceptual issues, the set-up of this study enables to
735 test by how much the quality of the data-model agreement is reduced after exposing the
736 MPI-CCDAS to shorter observational time-series. In terms of FAPAR, there is no
737 apparent degradation of fit with time, despite that in general terms, the trend in the data
738 is best matched with the ALL experiment. This is foremost a consequence of
739 comparatively small trends in the observed FAPAR, implying that extracting the mean
740 seasonal patterns and amplitude for few years, is essential for simulating current and
741 near-term FAPAR. Issues with model structure and with the assimilation set-up prevent
742 a better model-data fit irrespective of the length of the record. This would suggest that
743 a focus of assimilation on high-quality and highly spatially resolved FAPAR should be
744 a priority over the use of long-term data sets. The results are different for the case of
745 projecting atmospheric CO_2, where a long record of atmospheric CO_2 measurements
746 favorably contributes to a better representation of the long-term values after the
747 assimilation, whereas a shorter window leads to deviations to the observations in the
748 periods outside the assimilation years. The model-data agreement is of approximately
749 ±0.5 ppm during the assimilation period and starts to deviate for the DEC1 experiment
750 later than 10 years after the end of the assimilation window, whereas in the DEC2
751 experiment, the degradation of the model-data match already starts after approximately
752 5 years. Still, the average deviation to the observations by using shorter assimilation
753 periods do not deviate far from the upper limit of the uncertainty when using the longest
754 record. Nonetheless, with the caveat that MPI-CCDAS does not fully explain the
755 interannual variability of the land net carbon flux, this suggests a reasonable short-term
756 (for a small number of years) forecasting skill of atmospheric CO_2.

5 Conclusion

758 The MPI-CCDAS is capable of simultaneously integrating two independent
759 observational data sets over three consecutive decades at the global scale to estimate
760 global carbon fluxes. The results demonstrate that assimilating only one decade of
761 observations, for two observational data (FAPAR and atmospheric CO_2
762 concentrations), leads to broadly comparable results and trends in the global carbon
763 cycle components than using the entire time series of available observations (thirty
764 years). Currently, the system can confidently predict the carbon fluxes in short time
765 scales (up to 5 years after the end of the window of assimilation), e.g., for atmospheric

CO$_2$ concentrations at the site level, and the mean prediction remains within the uncertainty of the observations. However, long-term forecasts with CCDAS are less certain, as the observational record does not sufficiently constrain the interannual variability of the simulated land carbon fluxes, and longer-term changes in the decadal net carbon uptake. Nevertheless, the comparatively small error of 2±1.3 ppm after 15-19 years of prognostic simulation shows the potential for mid-term carbon cycle predictions constrained using the CCDAS approach.

The MPI-CCDAS is a computationally expensive system, and the demonstration that large-scale carbon fluxes can be improved by only using a limited period of observations increases the feasibility of using data assimilation systems to constrain the land carbon budget in land surface models. However, we also show that there are considerable variations in the estimated parameters and the regional distribution of the land C uptake suggesting that further improvements in the land-surface model, especially in the current structure and design, must be first solved to improve the model and computational efficiencies of the system. This is recommended to be done before an attempt to include another observational stream or other modifications aiming to test an enhancement on the prognostic skill in the full MPI-CCDAS.

6 Code availability

The code of the JSBACH model is available upon request to S. Zaehle (szaehle@bgc-jena.mpg.de). The TM3 model code is available upon request to C. Rödenbeck (christian.roedenbeck@bgc-jena.mpg.de). The TAF-generated derivative code is not available and it is subject to license restrictions.

Acknowledgements

The European Space Agency supported this research through the STSE Carbonflux (contract no. 4000107086/12/NL/Fv0), the 7th Framework program of the European Commission (grant no. GEOCARBON FP7-283080), as well as the Max Planck Society for the Advancement of Science, e.V., through the ENIGMA project. The authors thank P. Peylin for providing the fossil fuel emission data and T. Thum for the constructive comments during the preparation of the manuscript.

Table 1 – Model parameters selected for the optimization: No. 1 to 6: related to phenology; No. 7 to photosynthesis and No. 8 to 11 to land-carbon turnover. The values in the table for each PFT (where applies only) are for the prior conditions: $p_{pr} \pm C_{pr}$. [&]Values in f_{photos} are the photosynthetic parameters Vc_{max} / J_{max} (µmol CO_2 m^{-2} s^{-1} / µmol m^{-2} s^{-1}). In Λ_{max} the values marked with * are multiplied in the model by a factor of 1±0.2 and those with ^ (in Λ_{max} and in f_{photos}) by a factor of 1±0.1; in f_{photos} values with [a] are multiplied by 1±0.02, [b] by 1±0.03 and [c] by 1±0.06; these operations allowed a change in the standard values in the model. Letters in parenthesis below each PFT name are the predominant environmental controls that influence each group: T, temperature; D, daylight; W, water. No. 6 and 8 to 11 are global parameters and apply to all PFTs.

#	Parameter	Description	TrBe (W)	TrBD (W)	ETD (T,D)	CE (T,D)	CD (T,D)	RS (W)	TeH (T,W)	TeCr (T,W)	TrH (T,W)	TrCr (T,W)
1	Λ_{max}	Maximum LAI (m^2 m^{-2})	7.0*	7.0*	5.0*	1.7*	5.0*	2.0*	3.0^	4.0^	3.0^	4.0^
2	$1/\tau_l$	Leaf shedding timescale (d^{-1})	-	-	0.07±0.01	5e-4±1e-4	0.07±0.01	0.07±0.01	0.07±0.01	0.07±0.01	0.07±0.01	0.07±0.01
3	τ_w	Water stress tolerance time (d)	300±30	114±10	-	-	-	50±5	250±25	250±25	250±25	250±25
4	T_ϕ	Temperature at leaf onset (°C)	-	-	9.21±1	9.21±1	9.21±1	-	1.92±0.5	1.92±0.5	1.92±0.5	1.92±0.5
5	t_c	Day length at leaf shedding (h)	-	-	13.37±1	13.37±1	13.37±1	-	-	-	-	-
6	ξ	Initial leaf growth state (d^{-1})					0.37±0.03					
7	f_{photos}[&]	Photosynthesis rate modifier	39.0/ 74.1^	31.0/ 58.9^	66.0/ 125.4[a]	62.5/ 118.8[b]	39.1/ 74.3[c]	61.7/ 117.2^	78.2/ 148.6^	100.7/ 191.3^	8.0/ 140.0^	39.0/ 700.0^
8	Q_{10}	Temperature sensitivity to resp.					1.8±0.15					
9	f_{slow}	Multiplier for initial slow pool					1.0±0.1					
10	f_{aut_leaf}	Leaf fraction of maintenance resp.					0.4±0.1					
11	CO_2^{offset}	Initial atmospheric carbon (ppm)					0±3					

TrBE, Tropical evergreen trees; TrBD, Tropical deciduous trees; RS, Rain-green shrubs;
CE, Coniferous evergreen trees; ETD, Extra-tropical deciduous trees; CD, Coniferous deciduous trees; TeH, C3 grasses; TeCr, C3 crops; TrH, C4 grasses; TrCr, C4 crops.

Table 2 – Statistical analysis of FAPAR for 1982 – 2006 in all of the experiments, and also for the periods of the window of assimilation only for DEC1 and DEC2. R^2 is obtained from the linear correlation between $FAPAR_{obs}$ and $FAPAR_{mod}$ calculated for the entire period and by seasons. NRMSE is the normalized root mean squared error, defined as RMSE / mean ($FAPAR_{obs}$).

	Bias	NRMSE	R^2				
			All year	DJF	MAM	JJA	SON
PRIOR	0.37	0.95	0.16	0.14	0.31	0.21	0.33
ALL	0.10	0.76	0.20	0.14	0.34	0.20	0.37
DEC1	0.08	0.64	0.34	0.15	0.39	0.18	0.41
DEC2	0.09	0.65	0.34	0.14	0.39	0.18	0.41
Only for the period of the assimilation window							
DEC1 (1980-1990)	0.09	0.66	0.34	0.18	0.42	0.21	0.48
DEC2 (1990-2000)	0.05	0.48	0.34	0.18	0.41	0.21	0.47

Table 3 – Global average of the terrestrial carbon cycle components and carbon stocks in results from the assimilation experiments and prior (1982-2010), and other independent estimates (see table foot for description).

	PRIOR	ALL	DEC1	DEC2	INV	Literature
GPP (PgC yr^{-1})	118.8	96.9	83.1	97.2	-	118.9[a]
NPP (PgC yr^{-1})	54.5	34.2	37.3	30.3	-	-
NEE (PgC yr^{-1})	−2.64	−1.13	−1.32	−1.18	−1.20[c]	−2.52±0.98[b]
NBE (NEE + LUCC) (PgC yr^{-1})	−2.06	−0.54	−0.74	−0.60	-	−1.27±0.97[b]
ER (PgC yr^{-1})	115.7	95.2	81.0	95.3	-	-
Ra (PgC yr^{-1})	64.2	62.7	45.8	66.9	-	-
Rh (PgC yr^{-1})	51.5	32.4	35.2	28.4	-	-
Root Exudates (PgC yr^{-1})	3.3	2.0	2.2	1.7	-	-
Soil C (PgC)	2481	1364	1423	1167	-	1343[d]
Vegetation C (PgC)	394	310	335	311	-	442±146[e]
Litter C (PgC)	228	166	171	158	-	-

[a] Model Tree Ensemble data-driven product; Jung et al., 2011; average for 1982-2010,
[b] Global Carbon Project 2017; Le Quéré et al., 2018; average for 1982-2010. The NBE values include the LULCC reported for each individual model.
[c] Inversion result is the average for 1982-2009
[d] http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML
[e] Carvalhais et al. (2014).

Appendix

A1. Assimilation performance

Fig. A3 depicts the final posterior value (X_f) for each optimization parameter I and in each assimilation experiment. The last parameter value is normalized to its corresponding prior value (X_p, shown in Table 1), i.e. (X_{fi}/X_p)-1; this is done to make a comparison between parameters on their response to the assimilation because each parameter holds a different range of values. The normalized result is also shown for each phenotype for the phenology and photosynthesis-related parameters, and also for the initial leaf growth rate (ξ), CO_2 initial offset and land carbon turnover parameters that are applied globally.

More significant changes in some phenology parameter values are observed, e.g. the maximum LAI (Λ_{max}) decreased in almost all PFT's and in all experiments, except for the phenotypes CE (coniferous evergreen) in the ALL experiment, ETD (temperate broadleaf evergreen and deciduous; mostly dominating in Europe and eastern USA and Asia). In CD (coniferous-deciduous trees; located in Northeast Asia, specifically in the east Siberian Taiga) the Λ_{max} value increased notably in the DEC1 and DEC2 experiments (Fig. A3e).

In the tropical forest areas, the reduction of the Λ_{max} was from 3.17 in the PRIOR experiment to 2.27 (33 %) in ALL for the TW area, and from 3.27 in the PRIOR to 2.43 (26 %) in ALL for the TE area respectively. For the other assimilation experiments the average maximum LAI moderately decreased in TW from 3.17 in the PRIOR to 2.89 (8.8 %) in DEC1 and from 3.17 in the PRIOR to 3.00 (5.3 %) in DEC2.

In other extra-tropical areas results from experiments DEC1 and DEC2 experienced an average increase in Λ_{max} by 5.6 % in BE (from 2.29 in the PRIOR to 2.42), 24 % in BW (from 1.62 in the PRIOR to 2.01), and 3.8 % in TNW (from 3.11 in the PRIOR to 3.23). As a result, the temperature and daylight-related parameters were modulated such that the largest decrease with respect to the prior value in the temperature at leaf onset (T_ϕ) was also observed for these two PFT's, especially for CD in the DEC1 and DEC2 experiments. Also, the day length at leaf shedding (t_c) and the timescale of leaf senescence (leaf shedding timescale, $1/\tau_l$) primarily increased for CD. As for the PFT's influenced by temperature and water (TeH, TeCr, TrH and TrCr), the most significant change with respect to the prior value took place in the posterior value for the C3 crops (TeCr; distributed in Europe, USA and East Asia) whose value decreased considerably for the water stress tolerance (τ_w) in experiments DEC1 and DEC2, whereas the value

of the timescale of leaf senescence (leaf shedding timescale, $1/\tau_1$) also increased considerably for the same experiments; these changes seemed to be a response of the large decrease in the foliar area Λ_{max} for this PFT which took place in all three experiments. The value of the photosynthesis rate modifier (f_{photos}) influences the productivity at leaf-level. Thus, a lower value of f_{photos} will lead to lower GPP (less carbon uptake and a potential increase in NEE). Our results show that after the assimilation experiments the value of f_{photos} decreased with respect to the PRIOR experiment, mainly for the C3 grasses and pasture (TeH; distributed mostly in the Northern Hemisphere) as well as for the tropical evergreen and deciduous trees (TrBE and TrBD), and this is more noticeable in the DEC1 experiment.

As for the global parameters, significant deviations from the prior value are observed in the parameter that controls the initial size of the slow soil C pool (f_{slow}) and also in the parameter that defines the initial atmospheric CO_2 mole fraction ($CO2_{offset}$) which is globally set to be constant. The posterior value of both of these parameters decreased in the three posterior experiments. Variations in f_{slow} induce changes in the global heterotrophic respiration, controlling in this way the disequilibrium between GPP and the ecosystem respiration. Because JSBACH tends to overestimate the soil C pool, optimizing f_{slow} is a mean to improve this estimation; however, the spatial distribution of the carbon pools remains unchanged, and the prior value controls the prior value, meaning that the GPP and ER relation remains similar in the posterior experiments to that in the PRIOR experiment. Since the magnitude of the initial slow carbon pool was set, this might influence the other modeled carbon pools to the soil carbon pool, leading to biased soil and vegetation carbon stocks; therefore, the assessment on the predicted pools should be done with care. We compare the resulting global total soil and vegetation carbon pools robustly to independent estimates from the literature or other products, and results are shown in the main text of the Discussion section.

Table A1 – Regional mean and maximum Leaf Area Index in prior and posterior experiments.

Region	PRIOR (LAI mean ; max) (m^2 m^{-2})	ALL (LAI mean ; max) (m^2 m^{-2})	DEC1 (LAI mean ; max) (m^2 m^{-2})	DEC2 (LAI mean ; max) (m^2 m^{-2})
BE	0.61 ; 2.29	0.60 ; 1.94	0.70 ; 2.42	0.69 ; 2.42
BW	0.31 ; 1.62	0.30 ; 1.44	0.35 ; 2.01	0.35 ; 2.02
TNE	1.28 ; 4.28	1.17 ; 3.33	1.31 ; 3.49	1.32 ; 3.79
TNW	1.26 ; 3.11	1.15 ; 2.84	1.30 ; 3.23	1.30 ; 3.21
TE	1.62 ; 3.27	1.30 ; 2.43	1.63 ; 3.20	1.67 ; 3.33
TW	2.21 ; 3.17	1.68 ; 2.27	2.00 ; 2.89	2.08 ; 3.00
TSE	1.54 ; 2.72	1.43 ; 2.51	1.86 ; 2.77	1.83 ; 2.68
TSW	2.42 ; 3.69	2.04 ; 2.71	2.38 ; 3.47	2.43 ; 3.66

A2. Pixel level phenology analysis

The FAPAR analysis at the pixel level, shows that in pixels P1 (located in Eastern Siberia), P2 (located in eastern Brazil), and P6 (located in Canada), the magnitude of the mean seasonal cycle is better represented when compared to the observations (Fig. S4). Also, the timing of the mean seasonal cycle is corrected, e.g., in pixels with large seasonal amplitude such as in P1 and in P6. While in the PRIOR experiment (and ALL experiment) the onset and peak of the growing season in P1 and P6 are delayed by up to two months, in the results from experiments DEC1 and DEC2 this delay is reduced to only one month. This correction might be partially due to changes in some optimized parameters: increase in the day length at leaf shedding (t_c) and reduction in the temperature at leaf onset T_ϕ detected for both the CE and CD, as well as for ETD and TeCr/TeH phenotypes (Fig. A3 panels c, d, e and g); this is because these parameters control the onset and end of the vegetation activity. Despite changes in T_ϕ and t_c after the assimilation in TrH, this temporal shift is less evident in P2. In this pixel, the amplitude of the seasonal cycle is small, and only changes in the magnitude of the amplitude are visible after the assimilation (Fig. S4). In the results of DEC1 and DEC2 for pixel P3 (located in USA and dominated by TeCr), the water stress tolerance time (τ_w) and T_ϕ were primarily reduced, whereas the leaf shedding timescale ($1/\tau_l$; earlier shedding) increased.

References

Andres, R. J., Boden, T. A., and Marland, G.,Monthly fossil-fuel CO_2 emissions: mass of emissions gridded by one degree latitude by one degree lingitude, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, https://doi.org/10.3334/CDIAC/ffe.MonthlyMass.2013, 2013.

Andres, R. J., Gregg, J. S., Losey, L., Marland, G., and Boden, T. A.: Monthly, global emissions of carbon dioxide from fossil fuel consumption, Tellus B, 63, 309-327, https://doi.org/10.1111/j.1600-0889.2011.00530.x, 2011.

Andres, R. J., Marland, G., Fung, I., and Matthews, E.: A 1º x 1º distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950-1990, Global Biogeochem. Cycles, 10, 419-429, https://doi.org/10.1029/96GB01523, 1996.

Biudes, M. S., Machado, N. G., de Morais Danelichen, V. H., Caldas Souza, M., Vourlitis, G., and Nogeuira, J. d. S.: Ground and remote sensing-based mesurements of leaf area index in a transitional forest and seasonal flooded forest in Brazil, Int. J. Biometeorol., 58, 1181-1193, https://doi.org/10.007/s00484-013-0713-4, 2014.

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., Mu, M., Saatchi, S., Santoro, M., Thurner, M., Weber, U., Ahrens, B., Beer, C., Cescatti, A., Randerson, J. T., and Reichstein, M.: Global covariation of carbon turnover times with climate in terrestrial ecosystems, Nature, 514, 213-217, https://doi.org/10.1038/nature13731, 2014.

Carvalhais, N., Reichstein, M., Seixas, J., Collatz, G. J., Pereira, J. S., Berbigier, P., Carrara, A., Granier, A., Montagnani, L., Papale, D., Rambal, S., Sanz, M. J., and Valentini, R.: Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval, Global Biogeochem. Cycles, 22, GB2007, https://doi.org/10.1029/2007GB003033, 2008.

Clerici, M., Voßbeck, M., Pinty, B., Kaminski, T., Taberner, M., Lavergne, T., and Andreadakis, I.: Consolidating the Two-stream Inversion Package (JRC-TIP) to retrieve land surface parameters from albedo products, IEEE J. Sel. Top. Appl., 3, 286-295, https://doi.org/10.1109/JSTARS.2010.2046626, 2010.

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

Dalmonech, D. and Zaehle, S.: Towards a more objective evaluation of modelled land-carbon trends using atmospheric CO_2 and satellite-based vegetation activity observations, Biogeosciences, 10, 4189-4210, https://doi.org/10.5194/bg-10-4189-2013, 2013.

Dalmonech, D., Zaehle, S., Schürmann, G. J., Brovkin, V., Reick, C., and Schnur, R.: Separation of the effects of land and climate model errors on simulated contemporary land carbon cycle trends in the MPI Earth System Model version 1, J. Clim., 28, https://doi.org/10.1175/JCLI-D-13-00593.1, 2015.

Disney, M., Muller, J.-P., Kharbouche, S., Kaminski, T., Voßbeck, M., Lewis, P., and Pinty, B.: A new global fAPAR and LAI data set derived from optimal albedo estimates: comparison with MODIS products, Remote Sens.-Basel, 8, 1-29, https://doi.org/10.3390/rs8040275, 2016.

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., and Reichstein, M.: Enhanced seasonal CO_2 exchange caused by amplified plant productivity in northern ecosystems, Science, 351, 696-699, https://doi.org/10.1126/science.aac4971, 2016.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks, J. Clim., 27, 511-526, https://doi.org/10.1175/JCLI-D-12-00579.1, 2014.

Frölicher, T. L., Joos, F., and Raible, C. C.: Sensitivity of atmospheric CO_2 and climate to explosive volcanic eruptions, Biogeosciences, 8, 2317-2339, https://doi.org/10.5194/bg-8-2317-2011, 2011.

Frölicher, T. L., Joos, F., Raible, C. C., and Sarmiento, J. L.: Atmospheric CO_2 response to volcanic eruptions: the role of ENSO, season, and variability, Global Biogeochem. Cycles, 27, 239-251, https://doi.org/10.1002/gbc.20028, 2013.

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, ACM TOMS, 24, 437-474, https://doi.org/10.1145/293686.293695, 1998.

Gobron, N., Pinty, B., Taberner, M., Mélin, F., Verstraete, M. M., and Widlowski, J.-L.: Monitoring the photosynthetic activity of vegetation from remote sensing data, Adv. Space Res., 38, 2196-2202, https://doi.org/10.106/j.asr.2003.07.079, 2006.

Heimann, M. and Körner, S., The global atmospheric tracer model TM3: model description and user's manual release 3.8a, Max-Planck-Institut für Biogeochemie, Jena, Germany, Technical Report 5, 2003, 10.4126/98-004424387.

Hurtt, G. C., Frolking, S., Fearon, M. G., Moore, B., Shevliakova, E., Malyshev, S., Pacala, S. W., and Houghton, R. A.: The underpinnings of land-use history: three centuries of global gridded land-use transitions, wood-harvest activity, and resulting secondary lands, Global Change Biol., 12, 1208-1229, https://doi.org/10.1111/j.1365-2486.2006.01150.x, 2006.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Alaf Arain, M., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res., 116, G00J07, https://doi.org/10.1029/2010JG001566, 2011.

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. Cycles, 21, GB4021, https://doi.org/10.1029/2006GB002915, 2007.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO_2 into a terrestrial vegetation model and interactive mission benefit analysis, Biogeosciences, 9, 3173-3184, https://doi.org/10.5194/bg-9-3173-2012, 2012.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Howeling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P.-P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, S., Scipal, K., Sebald, C., Stacke, T., Terwisscha van Scheltinga, A., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: experiences and challenges, Biogeosciences, 118, 1414-1426, https://doi.org/10.1002/jgrg.20118, 2013.

Kaminski, T. and Mathieu, P.-P.: Reviews and sysntheses: flying the satellite into your model: on the role of observation operators in constraining models of the Earth system and the carbon cycle, Biogeosciences, 14, 2343-2357, https://doi.org/10.5194/bg-14-2343-2017, 2017.

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

Lasslop, G., Thonicke, K., and Kloster, S.: SPITFIRE within the MPI Earth system model: model development and evaluation, J. Adv. Model Earth Sy., 6, 740-755, https://doi.org/10.1002/2013MS000284, 2014.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S. I., Nojiri, Y., Padin, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D.: Global Carbon Budget 2017, Earth Syst. Sci. Data, 10, 405-448, https://doi.org/10.5194/essd-10-405-2018, 2018.

Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneth, A., Boden, T. A., Bopp, L., Bozec, Y., Canadell, J. G., Chini, L. P., Chevallier, F., Cosca, C. E., Harris, I., Hoppema, M., Houghton, R. A., House, J. I., Jain, A. K., Johannessen, T., Kato, E., Keeling, R. F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C. S., Landschützer, P., Lenton, A., Lima, I. D., Marland, G., Mathis, J. T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Salisbury, J. E., Schuster, U., Schwinger, J., Séférian, R., Segschneider, J., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Werf, G. R., Viovy, N., Wang, Y. P., Wanninkhof, R., Wiltshire, A., and Zeng, N.:

Global carbon budget 2014, Earth Syst. Sci. Data, 7, 47-85, https://doi.org/10.5194/essd-7-47-2015, 2015.

Lienert, S. and Joos, F.: A Bayesian ensemble data assimilation to constrain model parameters and land-use carbon emissions, Biogeosciences, 15, 2909-2930, https://doi.org/10.5194/bg-15-2909-2018, 2018.

Lucht, W., Prentince, I. C., Myneni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W., and Smith, B.: Climatic control of the high-latitude vegetation greening trend and Pinatubo effect, Science, 296, 1687-1689, https://doi.org/10.1126/science.1071828, 2002.

Mercado, L. M., Bellouin, N., Stich, S., Boucher, O., Huntingford, C., Wild, M., and Cox, P. M.: Impact of changes in diffuse radiation on the global land carbon sink, Nature, 458, 1014-1018, https://doi.org/10.1038/nature07949, 2009.

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J. L.: Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean, Global Biogeochem. Cycles, 20, GB2002, https://doi.org/10.1029/2005GB002530, 2006.

Montzka, C., Pauwels, V. R. N., Franssen, H.-J. H., Han, X., and Vereecken, H.: Multivariate and multiscale data assimilation in terrestrial systems: a review, Sensors, 12, 16291-16333, https://doi.org/10.3390/s121216291, 2012.

Müller, C. and Lucht, W.: Robustness of terrestrial carbon and water cycle simulations against variations in spatial resolution, J. Geophys. Res., 112, D06105, https://doi.org/10.1029/2006JD007875, 2007.

Newsam, G. N. and Enting, I. G.: Inverse problems in atmospheric constituent studies. I. Determination of surface sources under a diffusive transport approximation, Inverse Problems, 4, 1037-1054, https://doi.org/10.1088/0266-5611/4/4/008, 1988.

Peylin, P., Bacour, C., MacBean, N., Leonard, S., Rayner, P., Kuppel, S., Koffi, E., Kane, A., Maignan, F., Chevallier, F., Ciais, P., and Prunet, P.: A new stepwise carbon cycle data assimilation system using multiple data streams to constrain the simulated land surface carbon cycle, Geosci. Model Dev., 9, 3321-3346, https://doi.org/10.5194/gmd-9-3321-2016, 2016.

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

Pinty, B., Andredakis, I., Clerici, M., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 1. Effective leaf area index, vegetation, and soil properties, J. Geophys. Res., 116, D09105, https://doi.org/10.1029/2010JD015372, 2011a.

Pinty, B., Clerici, M., Andredakis, I., Kaminski, T., Taberner, M., Verstraete, M. M., Gobron, N., Plummer, S., and Widlowski, J.-L.: Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 2. Fractions of transmitted and absorbed fluxes in the vegetation and soil layers, J. Geophys. Res., 116, D09106, https://doi.org/10.1029/2010JD015373, 2011b.

Pinty, B., Lavergne, T., Dickinson, R. E., Widlowski, J.-L., Gobron, N., and Verstraete, M. M.: Simplifying the interaction of land surfaces with radiation for relating remote sensing products to climate models, J. Geophys. Res., 111, D02116, https://doi.org/10.1029/2005JD005952, 2006.

Pinty, B., Lavergne, T., Voßbeck, M., Kaminski, T., Aussedat, O., Giering, R., Gobron, N., Taberner, M., Verstraete, M. M., and Widlowski, J.-L.: Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS)-Multiangle Imaging Spectroradiometer (MISR) albedo products, J. Geophys. Res., 112, D10116, https://doi.org/10.1029/2006JD008105, 2007.

Qu, Y., Shaker, A., Silva, C. A., Klauberg, C., and Rangel Pinagé, E.: Remote sensing of Leaf Area Index from LiDAR height percentile metrics and comparison with MODIS product in a selectivley logged tropical forest area in Eastern Amazonia, Remote Sens.-Basel, 10, 1-23, https://doi.org/10.3390/rs10060970, 2011.

Raddatz, T., Reick, C., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P., and Jungclaus, J.: Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, Clim. Dynam., 29, 565-574, https://doi.org/10.1007/s00382-007-0247-8, 2007.

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, https://doi.org/10.1111/j.1365-2486.2005.00917.x, 2005.

Rayner, P., Koffi, E., Schilze, M., Kaminski, T., and Dufresne, J.: Constraining predictions of the carbon cycle using data, Philos. Trans. R. Soc. London, Ser. A, 369, 1955-1966, https://doi.org/10.1098/rsta.2010.0378, 2011.

Rayner, P. J., Enting, I. G., Francey, R. J., and Langenfelds, R.: Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, Tellus B, 51, 213-232, https://doi.org/10.3402/tellusb.v51i2.16273, 1999.

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. Cycles, 19, GB2026, https://doi.org/10.1029/2004GB002254, 2005.

Reick, C., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, J. Adv. Model Earth Sy., 5, 459-482, https://doi.org/10.1002/jame.20022, 2013.

Rödenbeck, C., Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Technical Report, Max Planck Institute for Biogeochemistry, Jena, Germany, 2005.

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

Rödenbeck, C., Keeling, R. F., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean pCO_2 and sea-air CO_2 flux variability from an observation-driven ocean mixed-layer scheme, Ocean Sci., 9, 193-216, https://doi.org/10.5194/os-9-193-2013, 2013.

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W.: Monitoring vegetation system in the great plains with ERTS, Greenbelt, USA1974, 3010-3017.

Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L., and Quegan, S.: Reviews and syntheses: systematic earth observations for use in terrestrial carbon cycle data assimilation systems, Biogeosciences, 14, 3401-3429, https://doi.org/10.5194/bg-14-3401-2017, 2017.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J. P., and Scipal, K.: Simultaneous assimilation of SMOS soil moisture and atmospheric CO_2 in-situ observations to constrain the global terrestrial carbon cycle, Remote Sens. Environ., 180, 334-345, https://doi.org/10.1016/j.rse.2016.02.058, 2016.

Scholze, M., Kaminski, T., Rayner, P., Knorr, W., and Giering, R.: Propagating uncertainty through prognostic CCDAS simulations, J. Geophys. Res., 112, D17305, https://doi.org/10.1029/2007JD008642, 2007.

Schürmann, G. J., Kaminski, T., Köstler, C., Carvalhais, N., Voßbeck, M., Kattge, J., Giering, R., Rödenbeck, C., Heimann, M., and Zaehle, S.: Constraining a land-surface model with multiple observations by application of the MPI-Carbon Cycle Data Assimilation System V1.0, Geosci. Model Dev., 9, 2999-3026, https://doi.org/10.5194/gmd-9-2999-2016, 2016.

Stocker, B. D., Feissli, F., Strassmann, K., Spahni, R., and Joos, F.: Past and future carbon fluxes from land use change, shifting cultivation and wood harvest, Tellus B, 1, 1-15, https://doi.org/10.3402/tellusb.v66.23188, 2014.

Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R. A., Sabine, C., Olafsson, J., and Nojiri, Y.: Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, Deep Sea Res. II, 49, 1601-1622, 2002.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F., and El Saleous, N.: An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetations NDVI data, Int. J. Remote Sens., 26, 4485-4498, https://doi.org/10.1080/01431160500168686, 2005.

van der Werf, G. R., Dempewolf, J., Trigg, S. N., Randerson, J. T., Kasibhatla, P. S., Giglio, L., Murdiyarso, D., Peters, W., Morton, D. C., Collatz, G. J., Dolman, A. J., and DeFries, R. S.: Climate regulation of fire emissions and deforestation in equatorial Asia, Proc. Natl. Acad. Sci. USA, 105, 20350-20355, https://doi.org/10.1073/pnas.0803375105, 2008.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997-2009), Atmos. Chem. Phys., 10, 11707-11735, https://doi.org/10.5194/acp-10-11707-2010, 2010.

Viovy, N. and Ciais, P.: CRUNCEP data set for 1901-2014, Version 6.1, https://esgf.extra.cea.fr/thredds/catalog/store/p529viov/cruncep/V7_1901_2015/catalog.html, 2015.

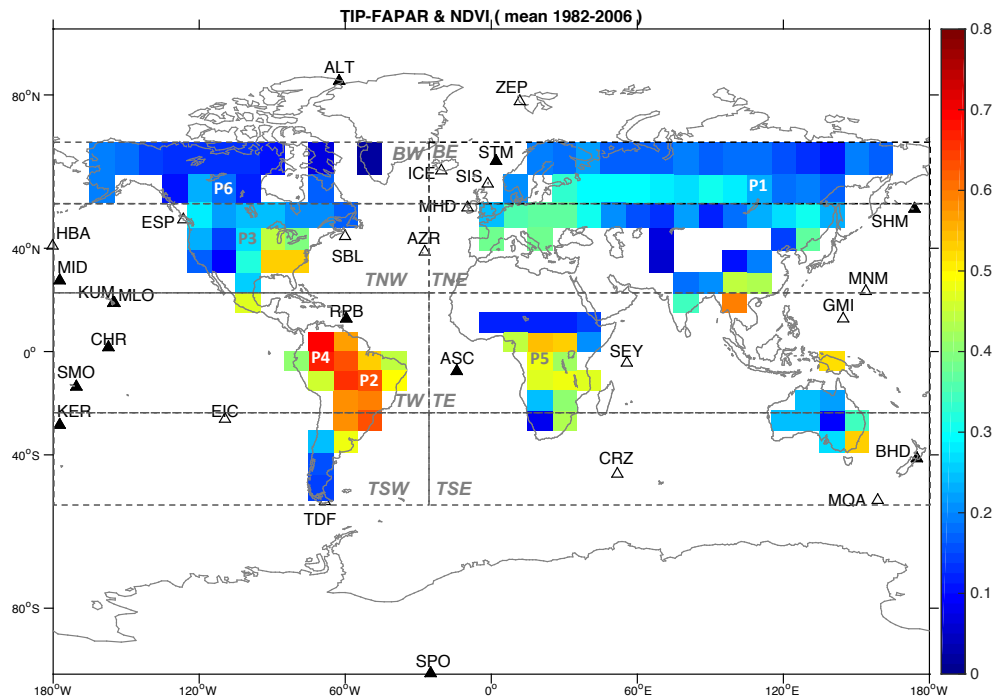
Wang, Z., Hoffmann, T., Six, J., Kaplan, J. O., Govers, G., Doetterl, S., and Van Oost, K.: Human-induced erosion has offset one-third of carbon emissions from land cover change, Nat. Clim. Change, 7, 345-350, https://doi.org/10.1038/nclimate3263, 2017.

Figures.

Figure 1 – Global distribution of the temporal mean (1982-2006) of the merged satellite FAPAR product used in the assimilation procedure. It shows also the spatial coverage of eight regions globally distributed: Boreal West and East (BW and BE, for latitudes north of 60 °N), Temperate Northwest and Northeast (TNW and TNE, between latitudes 20 °N and 60 °N); tropical West and East (TW and TE, between latitudes 20 °N and 20 °S); Temperate Southwest and Southeast (TSW and TSE, for latitudes south of 20 S). Also shown six selected pixels: P1, for the coniferous deciduous (CD) phenotype in the East Siberian Taiga; P2, for the C4 pastures and grasses (TrH) of central Brazil; P3, for the C3 and C4 crops, pastures and grasses (TeCr and TeH) of Northern USA; P4 and P5, for tropical evergreen trees (TrBe) situated in Northwestern Brazil and central Africa; and P6, for coniferous evergreen (CE) located in Canada; and the location of 28 stations of the CO_2 network measurements (filled triangles, stations only included in DEC1; empty triangles, stations included also in ALL and DEC2) for analysis of the assimilation results.

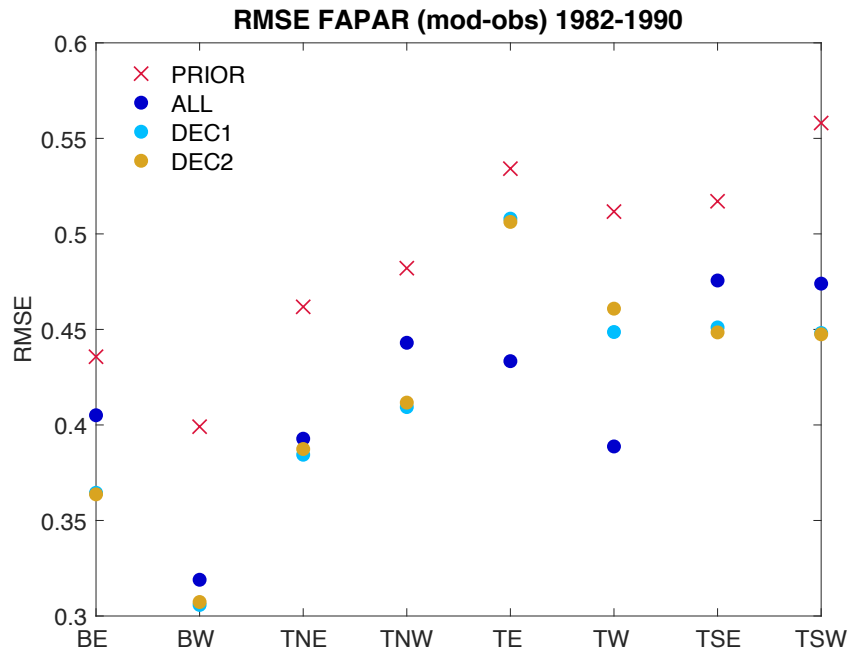



Figure 2 – RMSE for FAPAR from the model results and observations for the period 1982-2006 and for different regions.

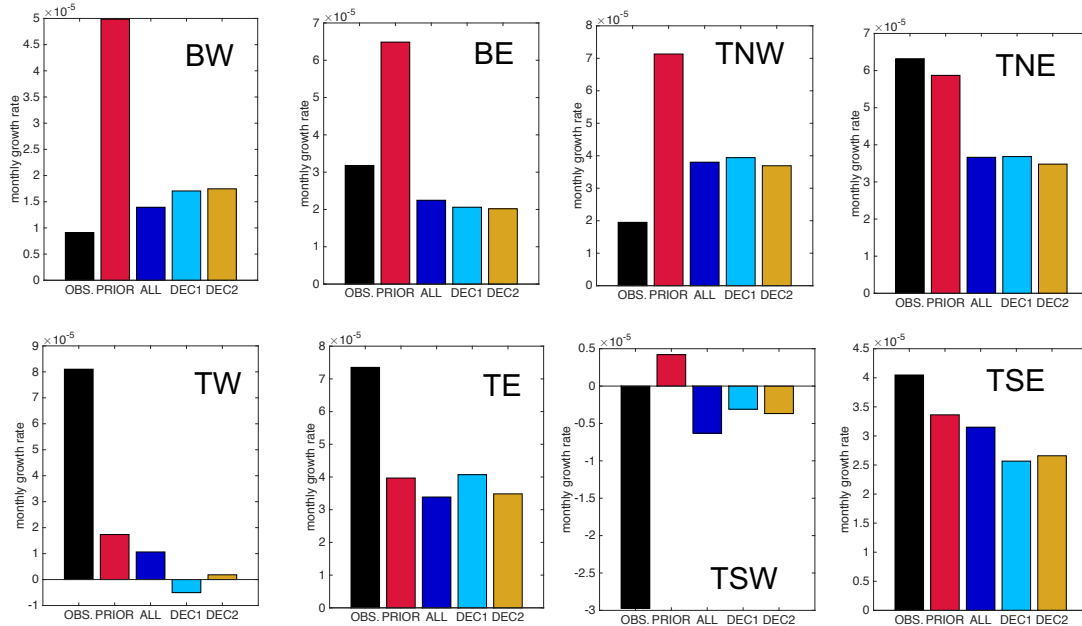



Figure 3 – Mean monthly growth rate of FAPAR for 1982-2006 on each analyzed geographical region for the satellite observations and results of PRIOR and the posterior experiments.

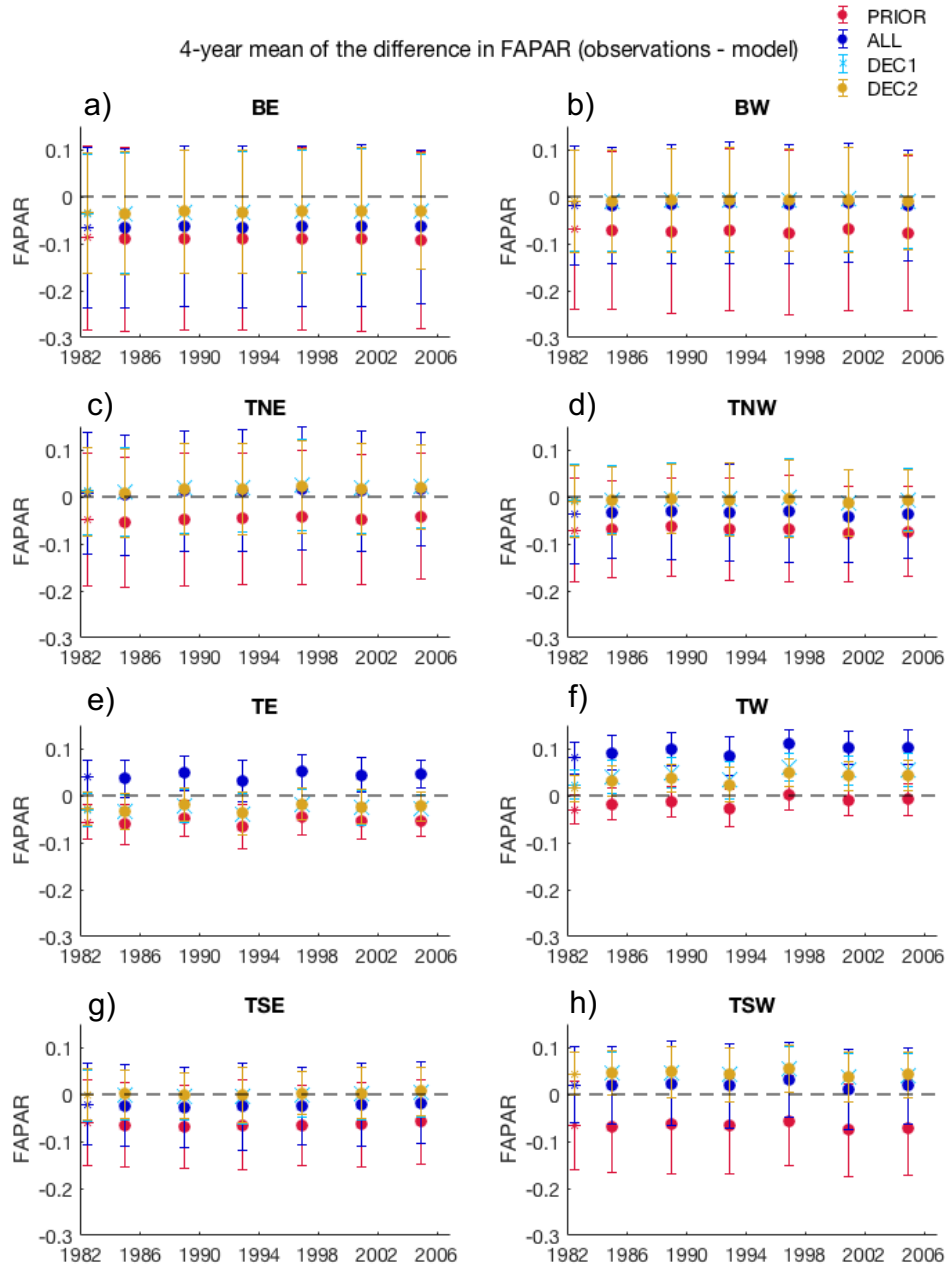



Figure 4 – Time series of the four-years mean of the FAPAR anomaly to the satellite data for each model experiment in six selected model pixels. The error bar indicates the +/− 1 standard deviation of the four-years differences. The first marker (as asterisk) in the time series is the single value for 1982.

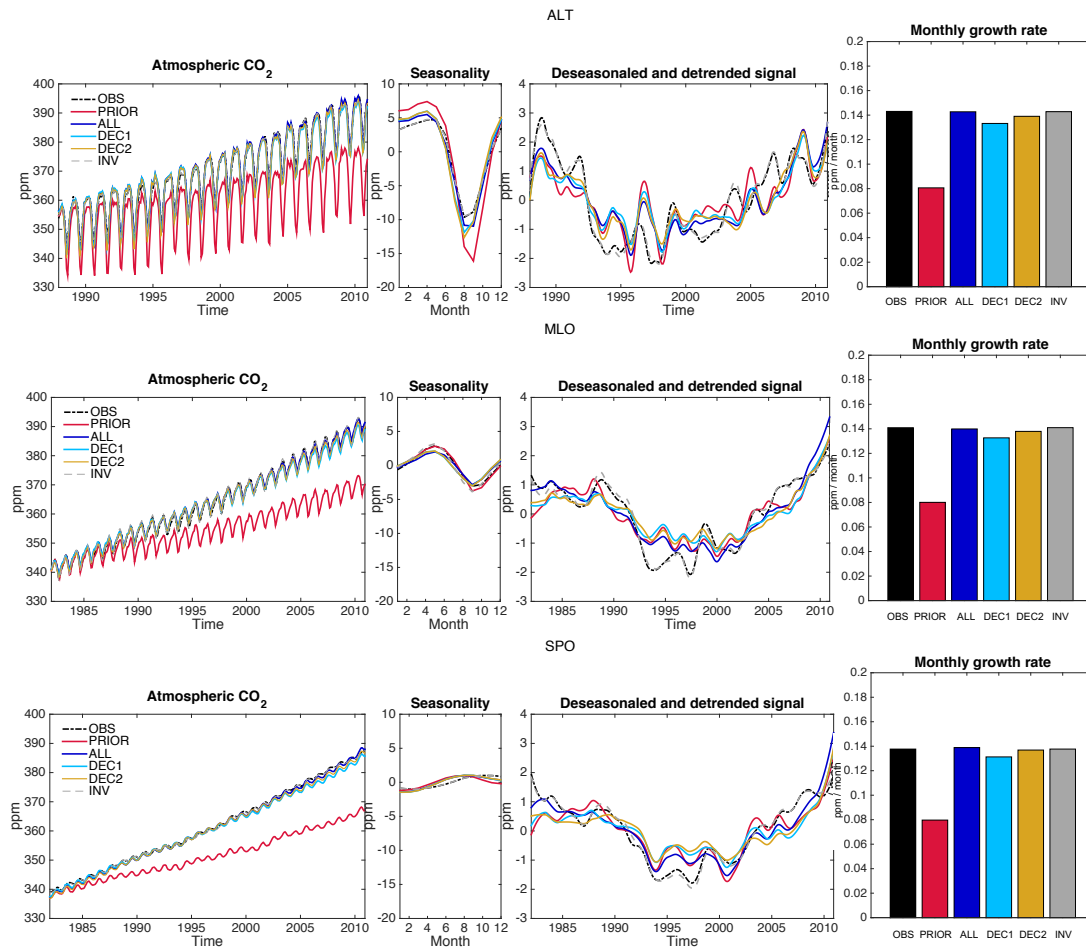



Figure 5 – Statistical analysis of atmospheric CO_2 in three flask measurement sites: Alert (ALT; top panels), Mauna Loa (MLO, center panels) and South Pole (SPO, bottom panels), from the measurements, PRIOR, posterior experiments (ALL, DEC1 and DEC2) and inversion (INV1). For each station the panels show the time series of the mean monthly values, the mean seasonal cycle, the interannual variability and the monthly growth rate for the entire period of the simulation (1980-2010).

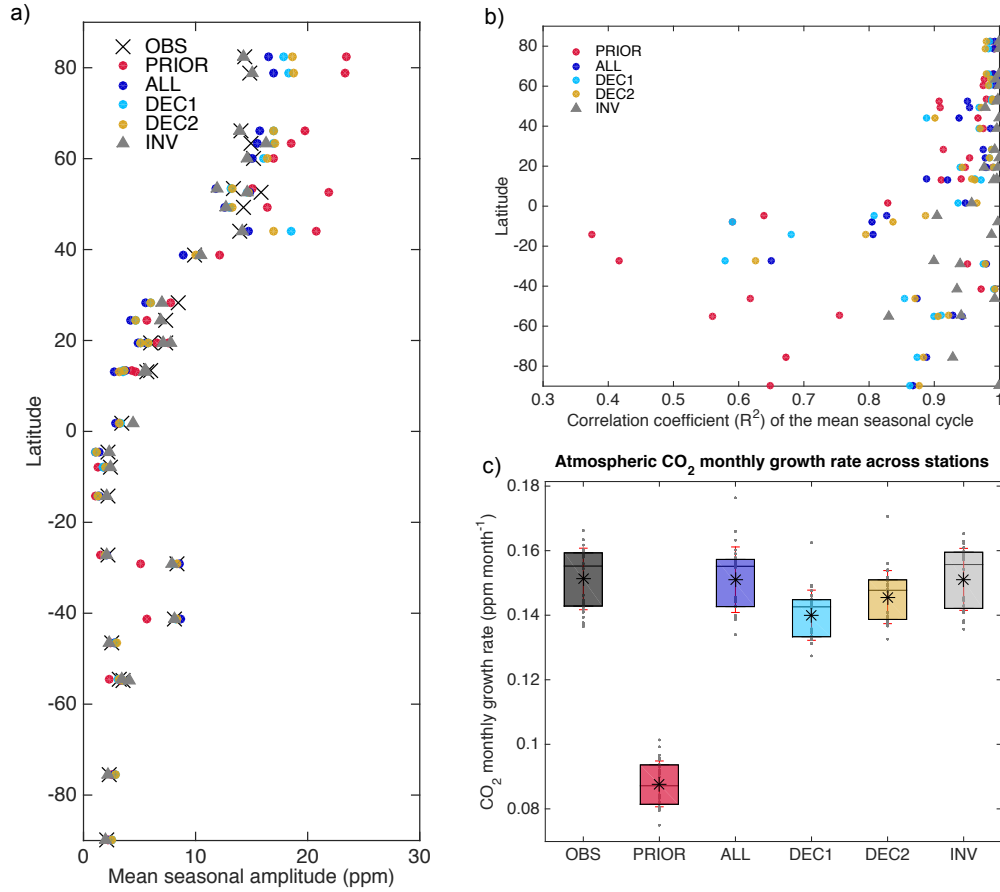



Figure 6 – a) Latitudinal distribution of the mean CO_2 seasonal amplitude for the 28 flask-measurement stations from the observations, PRIOR and posterior experiments; b) Latitudinal distribution of R^2 obtained from the correlation between the observations and each simulation results of the mean atm. CO_2 seasonal cycle and c) average atmospheric CO_2 monthly growth rate across stations for the observations and model results. The star on each bar is the mean of the atm. CO_2 monthly growth rate, the horizontal middle black line on each box is the median, the red whiskers depict the error as one standard deviation, and the grey dots on each box are the actual monthly growth rate values for all the stations in each data set.

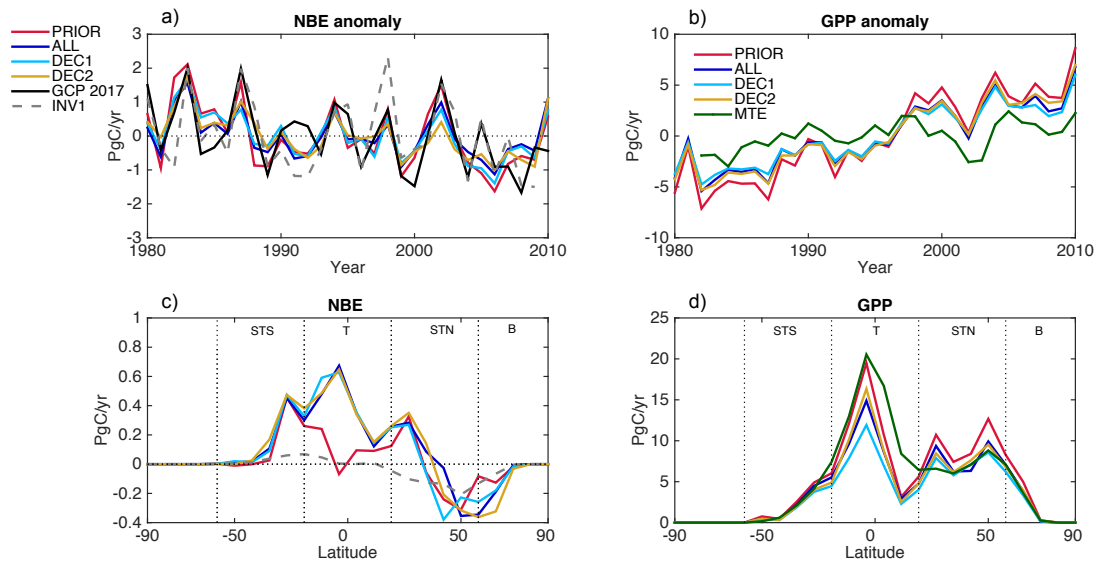



Figure 7 – Time series of the anomaly to the temporal mean of the time series (a and b), and latitudinal gradient (c and d) of the total Net Ecosystem Exchange (NEE including the influence of LULCC) (left) and Gross Primary Production (right) for the results of each model simulation. NEE from the model is compared to the GCP 2017 and INV data set (a and c). GPP is compared to the MTE data-data driven estimate of Jung et al., (2011) (b and d).

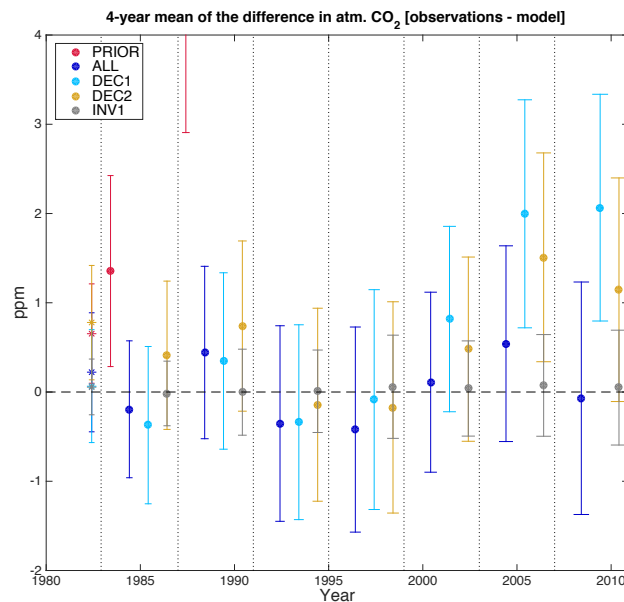



Figure 8 – Time series of the four-years mean of the atm. CO_2 anomaly to the observations for each model experiment and inversion results, for all the stations. The y-axis is limited to the results in the posterior experiments. The error bar is one standard deviation to the four-years mean of the differences to the observations. The first marker to the left in the time series (as asterisk) is the single value for 1982 not included in the subsequent four-years means.

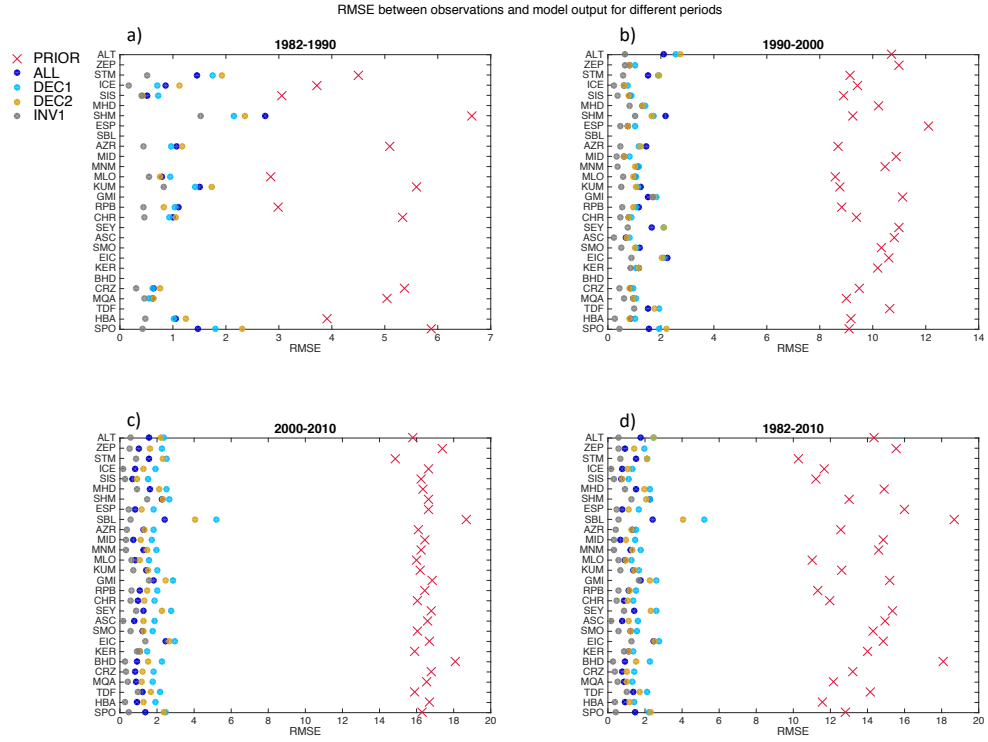



Figure 9 – RMSE for different periods between CO_2 atm. concentrations from measurements and model results for the different assimilation experiments and inversion results for each of the flask measurement stations.

Appendix

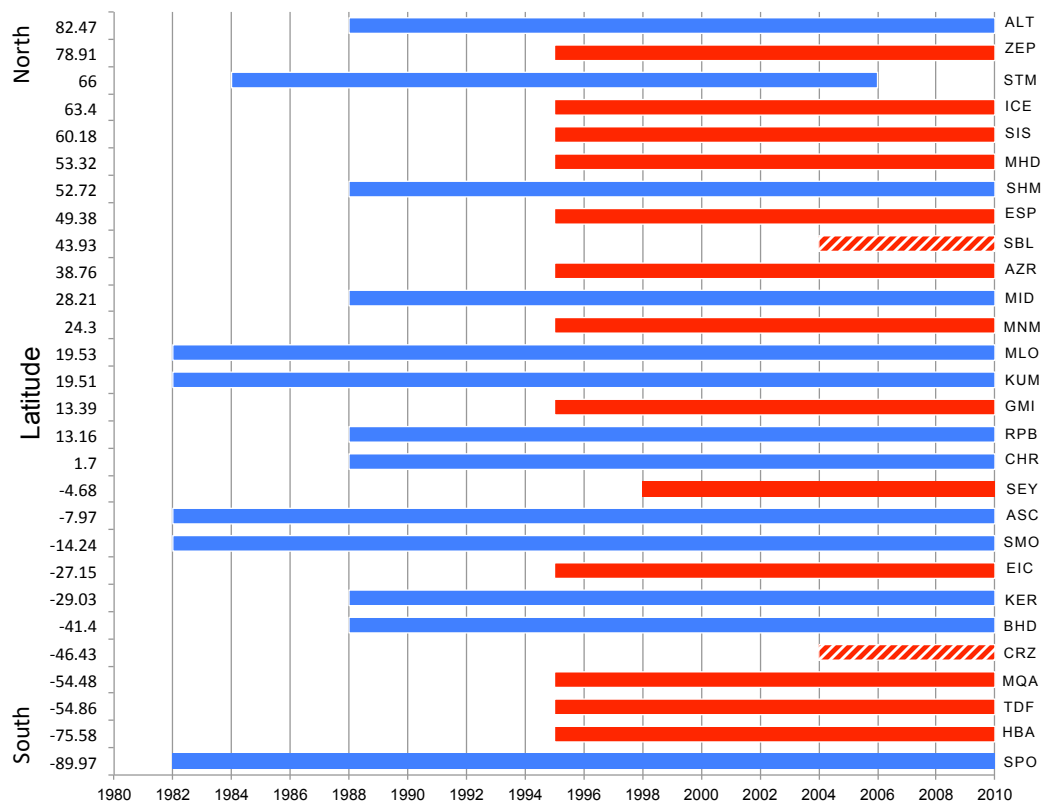



Figure A1 – Data availability and latitudinal location of the 28 stations where the long-term flask measurements of atmospheric CO_2 mole fractions were taken for assimilation in CCDAS. ALL experiment used all the stations of the time series (blue and red bars) (1980-2010); DEC1 used data only from stations with blue bars (1980-1990), and DEC2 used also the data in the stations with red bars (1990-2000) (except stations SBL and CRZ marked with patterned bar).

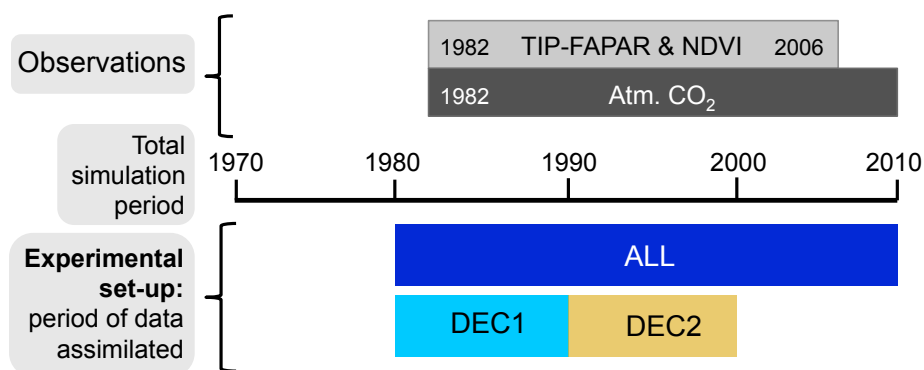



Figure A2 – Experimental set up for posterior experiments ALL, DEC1 and DEC2 with different temporal windows for the assimilation of FAPAR and molar fractions of atmospheric CO_2.

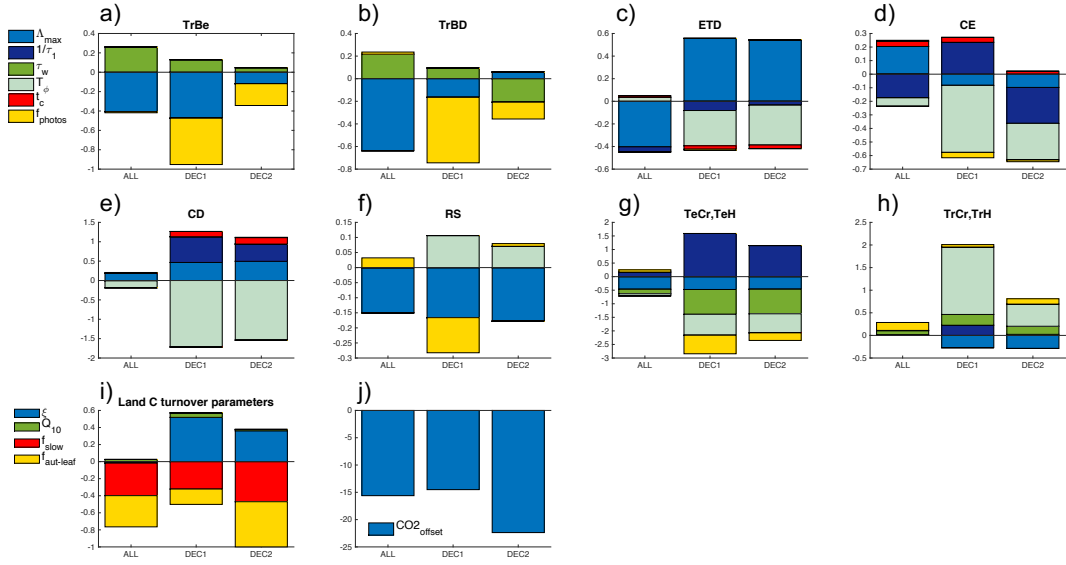



Figure A3 – Final value for each parameter p at the end of the assimilation experiments, normalized to the prior value (p_{pr}), i.e. (p/p_{pr})-1. This is shown for each model plant functional type (a to h) and globally for the land C turnover parameters (i and j).

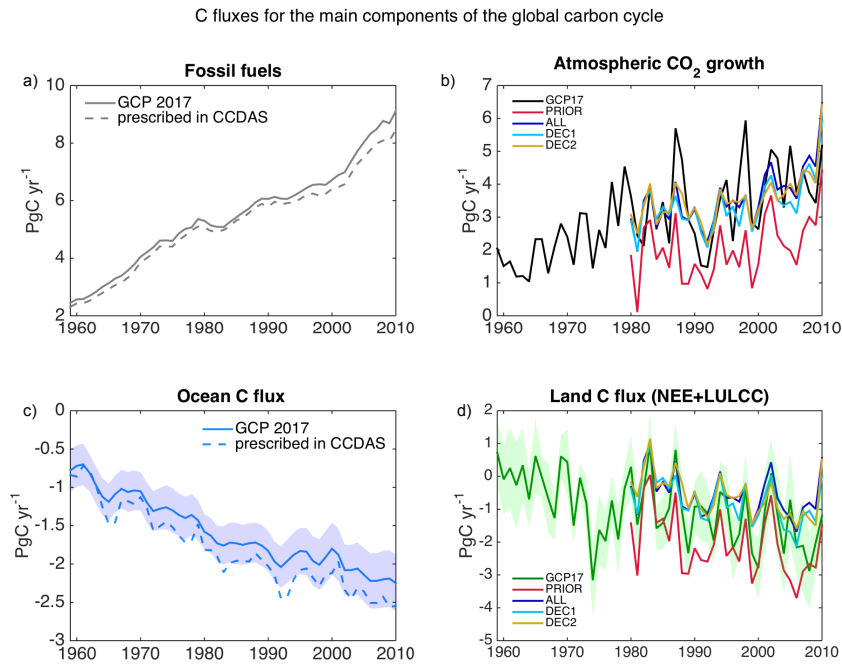



Figure A4 – Time series of the annual mean of the major components of the C cycle used as background fluxes in CCDAS compared to those from the GCP 2017. The atm. CO_2 growth from the model output is the result of the sum of fossil fuel, ocean, and land C fluxes. The blue shadow in the ocean C sink of the GCP 2017 data is the standard deviation of the mean sink from the models that contributed to the GCP. The land C flux is the total NEE with contribution of the flux due to LULCC. The green shadow area is the standard deviation of the mean land C flux from the terrestrial models that contributed to the GCP.