



1 Three decades of simulated global terrestrial carbon fluxes from a data 2 assimilation system confronted to different periods of observations.

3 Karel Castro-Morales^{1*}, Gregor Schürmann¹, Christoph Köstler¹, Christian
4 Rödenbeck¹, Martin Heimann^{1,3} and Sönke Zaehle^{1,2}

5
6 ¹ Max Planck Institute for Biogeochemistry, Jena, Germany

7 ² Michael-Stifel-Center Jena for Data-Driven and Simulation Science, Jena, Germany

8 ³ Institute for Atmospheric and Earth System Research, Faculty of Science, University
9 of Helsinki, Helsinki, Finland

10 *Now at: Friedrich Schiller University, Institute of Biodiversity, Aquatic
11 Geomicrobiology, Jena, Germany

12 13 Abstract

14 This paper presents global land carbon fluxes for the period 1982-2010 (gross primary
15 production, GPP, and net ecosystem exchange, NEE) estimated with the Max Planck
16 Institute – Carbon Cycle Data Assimilation System (MPI-CCDAS v1). The primary
17 aim of this work is to analyze the performance of the MPI-CCDAS when it is
18 confronted with three different time periods for data assimilation (DA), and thereby to
19 assess its prognostic capability. To this extend we assimilated nearly three decades
20 (1982-2010) of space borne measurements of the fraction of absorbed photosynthetic
21 active radiation (FAPAR) and atmospheric CO₂ concentrations from the global
22 network of flask and in situ measurements. Both data sets were incorporated with
23 different assimilation windows covering the periods 1982-1990, 1990-2000 and 1982-
24 2010. The assimilation results show a considerable improvement in the long-term
25 trend and seasonality of FAPAR in the Northern Hemisphere, as well as in the long-
26 term trend and seasonal amplitude of the atmospheric CO₂ concentrations when
27 compared to the observations in sites globally distributed. After the assimilation, the
28 global net land-atmosphere CO₂ exchange (NEE) was -1.2 PgC yr^{-1} , in agreement
29 with independent estimates, while gross primary production (GPP; 92.5 PgC yr^{-1}) was
30 somewhat below the magnitude of independent estimates. The NEE in boreal eastern
31 regions (Northeast Asia) increased on average by $-0.13 \text{ PgC yr}^{-1}$, which translated
32 into an intensification of the carbon uptake in those regions by nearly 30 % than the
33 contribution to the global annual average in the model before the assimilation.
34 Our results demonstrate that using information only over a decade already yielded a
35 large fraction of the overall model improvement, in particular for the simulation of
36 phenological seasonality, its interannual variability (IAV) and long-term trend.
37 Adding longer than decadal data did only lead to very moderate improvements in the
38 long-term trend of the FAPAR simulated by the model, which may be attributed to the



39 small model-data mismatch at the long timescales compared to the significantly larger
40 observational signal and model-data mismatch error at seasonal cycle time scale.
41 Decadal data also significantly improved the seasonality, IAV and long-term
42 simulated trend in atmospheric CO₂. Importantly, when running the MPI-CCDAS v1
43 with 30 years of data, the results remained in line with observations throughout this
44 period, suggesting that the model can represent land uptake to a sufficient degree to
45 make it compatible with the atmospheric CO₂ record. Using data from 1982 to 1990
46 in the assimilation yielded only a difference to the observations of 2±1.3 ppm for the
47 period 15 to 19 years after the end of the assimilation. This suggests that despite
48 imperfections in the representation of IAV, model-data fusion can increase the
49 prognostic capacity of land carbon cycle models at relevant time-scales.

50 Key words: *Data assimilation, Global Carbon cycle, modeling, atmospheric CO₂.*

51

52 1 Introduction

53 The observed contemporary in atmospheric CO₂ is driven by anthropogenic emissions
54 from fossil fuels and land-use change (2007-2016 average: 11.1±0.6 GtC yr⁻¹), and
55 the concurrent net carbon uptake of the ocean and land from the atmosphere, which
56 take up approximately 22.4 % and 28 % of the anthropogenic flux, respectively (Le
57 Quéré et al., 2018). Despite recent advances in atmospheric observations, ocean and
58 land modeling, there remains an imbalance between carbon emissions, ocean and land
59 sinks, and changes in the atmospheric CO₂ concentration of 5.6 % (0.6 GtC yr⁻¹).
60 Despite substantial progress in improving the performance of terrestrial biosphere
61 models over the past decades, the simulated global terrestrial carbon fluxes and the
62 net land carbon balance pose still the highest uncertainties from all of the components
63 of the global carbon cycle (Friedlingstein et al., 2014; Le Quéré et al., 2018).
64 Quantifying the magnitude and dynamics of the global terrestrial carbon cycle across
65 different temporal scales and their contribution to the global carbon cycle, is
66 challenging because the large heterogeneity and complexity of these ecosystems, in
67 addition to the quantification of contemporary effects and response of these
68 ecosystems to increasing post-industrial CO₂ concentrations (Lienert and Joos, 2018;
69 Stocker et al., 2014; Wang et al., 2017).

70 One strategy to reduce the mismatch between the carbon flux predictions of land
71 surface models and observed trends in atmospheric CO₂ concentrations is through
72 data assimilation (DA) techniques, meaning to “train” the land models by confronting



73 them systematically with observations of carbon-related variables (Raupach et al.,
74 2005). During DA, process parameters of land surface models are adjusted through
75 numerical minimization techniques to reduce the misfit between model results and
76 actual observations under consideration of the statistical properties of model and
77 observations. Contrary to the application of atmospheric transport inversion to infer
78 the sinks and sources of CO₂ between the atmosphere and land or ocean from
79 atmospheric CO₂ measurements (Newsam and Enting, 1988; Peylin et al., 2013;
80 Rayner et al., 1999; Rödenbeck et al., 2003), the application of these carbon cycle
81 data assimilation systems (CCDAS) provides the additional opportunity to inform the
82 process-based carbon cycle mechanisms in the land surface model to support a better
83 estimate and capacity to project dynamics of the terrestrial carbon cycle. Several
84 CCDAS have been developed for this purpose (e.g. Kaminski et al., 2012; Kaminski
85 et al., 2013; Lienert and Joos, 2018; Peylin et al., 2016; Scholze et al., 2016).
86 Although they rely in different statistical methods (i.e. variational or sequential data
87 assimilation) (Montzka et al., 2012), their common characteristic is integrating long-
88 term and time consistent global available observational records related to the carbon
89 cycle such as atmospheric CO₂ measurements from flask and in situ networks
90 (Conway et al., 1994), and remote sensing products of canopy phenology properties
91 such as MODIS-NDVI (Moderate Resolution Imaging Spectroradiometer -
92 Normalized Difference Vegetation Index) (Rouse et al., 1974) and FAPAR (Disney et
93 al., 2016; Pinty et al., 2011a).

94 In this work, we use the Max Planck Institute - Carbon Cycle Data Assimilation
95 System (MPI-CCDAS v1, Schürmann et al., 2016) that has been built around the Jena
96 Scheme Biosphere-Atmosphere Coupling in Hamburg (JSBACH) land-surface model
97 (Dalmonech and Zaehle, 2013; Raddatz et al., 2007; Reick et al., 2013). The MPI-
98 CCDAS follows a variational approach that iteratively reduces the model-data misfit
99 simultaneously for multiple observational and independent carbon cycle data sets
100 (Kaminski et al., 2013). Since its first development based on the BETHY (Biosphere
101 Energy-Transfer Hydrology) - CCDAS, the MPI-CCDAS has undergone several code
102 modifications and improvements, as well as tests of the assimilation of new
103 observational data sets (e.g. Kaminski et al., 2012; Kaminski et al., 2013; Rayner et
104 al., 2005; Scholze et al., 2016; Schürmann et al., 2016), with the aim of further



105 improving the representation of land carbon fluxes. The history of the MPI-CCDAS
106 and other current DA systems is extensively discussed in Scholze et al. (2017).

107 In this paper, we seek to analyze the extent to which the application of a CCDAS
108 leads to the improved representation of the contemporary land carbon cycle and its
109 prognostic capacity for subsequent years. To this extent, we analyze the estimated
110 major components of the terrestrial carbon cycle with the MPI-CCDAS in response to
111 the simultaneous assimilation of three decades of data from two observational
112 constraints: FAPAR from remote sensing data and atmospheric CO₂ concentrations
113 from the global flask measurements network. Our aim is to analyze the performance
114 of the MPI-CCDAS to: 1) how well the model is able to reproduce 30 years of
115 constraint data, 2) how sensitive is the assimilation success to the choice of different
116 temporal windows used for the assimilation (implying different amounts of
117 observational data during the assimilation), and 3) how good is the fit to the data in
118 the time period beyond the period constrained with observations.

119 2 Methods

120 2.1 MPI-CCDAS

121 The code of the MPI-CCDAS version in this work is identical to the one used in
122 Schürmann et al. (2016). The model calculates the half-hourly storage and surface
123 fluxes of energy, water and carbon in terrestrial ecosystems at a coarse spatial
124 resolution for computational feasibility (8° by 10° grid). The spatial distribution of the
125 different plant-functional types (PFTs, Table 1) in JSBACH is shown in Fig. S1
126 (Supplement). The selection of parameters for the assimilation procedure, their prior
127 values and range was based on Schürmann et al. (2016; Table 1). The initial state of
128 the parameters was obtained from an independent forward simulation of JSBACH 3.0
129 (see Sect. 2.3.1). As described in Schürmann et al. (2016), MPI-CCDAS starts with an
130 initial guess for the model control vector (\mathbf{p}_{pr}) of e.g. carbon cycle properties, and
131 model states, and their Gaussian uncertainty (“prior”) with covariance \mathbf{C}_{pr} . The model
132 control vector \mathbf{p} is iteratively updated to minimize a joint cost function J describing
133 the misfit between observational data-streams (\mathbf{d} ; FAPAR and atmospheric CO₂, both
134 with covariance \mathbf{C}_{d}) and the corresponding simulated observation operators of the
135 MPI-CCDAS $M(\mathbf{p})$, taking into account the uncertainties in the observational data
136 assuming a Gaussian distribution and the information from the prior.

$$137 \quad J(\mathbf{p}) = \frac{1}{2} (\mathbf{M}(\mathbf{p}) - \mathbf{d})^T \mathbf{C}_{\text{d}}^{-1} (\mathbf{M}(\mathbf{p}) - \mathbf{d}) + (\mathbf{p} - \mathbf{p}_{\text{pr}})^T \mathbf{C}_{\text{pr}}^{-1} (\mathbf{p} - \mathbf{p}_{\text{pr}}) \quad (1)$$



During the optimization procedure, a new model trajectory is determined in each iteration, 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, as described by Schürmann et al. (2016), was updated following Knorr et al. (2010) where the minimum and maximum calculations in the entire code were replaced by smoothing functions to avoid steep transitions.

2.2 Observational data sets

2.2.1 FAPAR

FAPAR is the fraction of the radiation that is absorbed by plants during photosynthesis, thus 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 transfer model (Pinty et al., 2006; Pinty et al., 2007) using the MODIS broadband visible and near-infrared spectral white sky surface albedo as input (Clerici et al., 2010; Pinty et al., 2011a; Pinty et al., 2011b). For this study, the TIP-FAPAR product was available only from 2003 to 2011, making it unsuitable for the intended longer assimilation period. While there are long-term remotely sensed proxies of FAPAR, such as the NDVI (Rouse et al., 1974), it has been found previously that NDVI was less reliable than TIP-FAPAR in terms of the seasonal cycle amplitude of vegetation seasonality (Dalmonech and Zaehle, 2013; Dalmonech et al., 2015). We therefore merged the Global Inventory Monitoring and Modeling System (GIMMS) NDVI product, available from 1982 to 2006 (Tucker et al., 2005), with TIP-FAPAR to provide a longer record of vegetation greenness. The maximum and minimum NDVI values were rescaled at the pixel level to coincide



171 with those from the TIP-FAPAR for the overlapping time periods (2003 to 2006)
172 following:

$$173 \quad NDVI_{\text{mod}} = \frac{NDVI - NDVI_{\text{min}}}{NDVI_{\text{max}} - NDVI_{\text{min}}} \times (TIP_{\text{max}} - TIP_{\text{min}}) + TIP_{\text{min}} \quad (2)$$

174 where TIP is the TIP-FAPAR data. The median uncertainty of the available TIP-
175 FAPAR data was considered as the uncertainty for the entire time-series. For this
176 study, this product was aggregated to match the model grid horizontal resolution
177 considering separately background snow-free and snow-covered conditions
178 (Schürmann et al., 2016).

179 As in Schürmann et al. (2016), we applied a mask to the global FAPAR with the aim
180 of selecting useful pixels in the FAPAR global grid. The selection of pixels to be
181 removed from the global grid followed three criteria: 1) since no explicit crop
182 phenology is described in JSBACH, we masked out the grid cells with a crop-
183 dominated phenology of $> 20\%$. This step has consequences in areas where other
184 important functional types are also present in the same grid cells such as deciduous
185 broadleaves that are also abundant in the USA and Europe. As a result, the parameters
186 related to deciduous broadleaves are constrained from other locations; 2) we further
187 masked out pixels that hold a low correlation ($R^2 < 0.2$) when compared the prior
188 model result and the observations, as we had previously found that the MPI-CCDAS
189 is incapable of correcting such poor model behaviors (Schürmann et al. 2016); and 3)
190 finally, we masked out pixels located in areas where phenology abundance is low, i.e.
191 deserts, because they would influence the optimization causing a large bias due to
192 compensating effects. The final FAPAR product used during the assimilation contains
193 only 40 % of the initial number of pixels after the applied mask, resulting in more
194 pixels distributed in the Northern Hemisphere compared to the Southern areas. This
195 observational data will be referred hereafter as FAPAR_{obs} (see Fig. 1 for the global
196 distribution of mean FAPAR_{obs} from 1982 to 2006).

197 **2.2.2 Atmospheric CO₂ concentrations and observation operator**

198 Measurements of atmospheric CO₂ mixing ratios were taken from the flask data
199 continuous record of 28 sites in the NOAA/CMDL station network (Conway et al.,
200 1994; Rödenbeck et al., 2003). The selection criteria included length of the record (on
201 average 19 years) (Fig. A1) and focused on remote and ocean stations with low
202 impact of local carbon sources and sinks of carbon (Schürmann et al., 2016) (see
203 location of CO₂ stations in Fig. 1). The atmospheric transport of CO₂ is calculated in



204 MPI-CCDAS through the Jacobian representation of the TM3 atmospheric transport
205 model driven by meteorology fields from NCEP (National Centers for Environmental
206 Prediction) reanalysis. TM3 is run at horizontal “fine grid” (fg) resolution of $4^{\circ} \times 5^{\circ}$
207 (Heimann and Körner, 2003; Rödenbeck et al., 2003). During the generation of this
208 matrix representation, the precise sampling time of flask measurements was
209 considered to minimize the representation error due to short-term fluctuations in
210 atmospheric transport. The treatment of uncertainties is done in the same way as in
211 the TM3 atmospheric inversion (Rödenbeck et al., 2003) but, as in Schürmann et al.
212 (2016), imposing a floor value of 1 ppm to the uncertainties (Rayner et al., 2005) to
213 allow a range for the comparison to the observational operator.

214 In order to compare the land fluxes from MPI-CCDAS to atmospheric concentrations,
215 background carbon fluxes (from fossil fuel emissions, use and change of land cover,
216 and from the ocean) are necessary to account for the total carbon balance.

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

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

236 *Ocean fluxes:* Two products were merged to account for the oceanic CO₂ fluxes: 1)
237 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) ocean annual 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

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. The CRU-NCEP v6.1 reanalysis data is available at daily resolution from 1901 to 2014 with a spatial resolution of 0.5° (Viovy and Ciais, 2015; last access July 2015). These atmospheric forcing fields (i.e. wind speed, air temperature, precipitation, downward short- and long-wave radiation and specific humidity) were remapped to the coarse ($8^\circ \times 10^\circ$) model grid. In addition, prescribed annual means (one global mean annual 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).

Prior to the assimilation experiments, the JSBACH 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 done also with the forward JSBACH model for the period 1901 to 2012. This transient simulation included 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 JSBACH 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 in order 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).



271 2.3.2 Assimilation experiments

272 During the assimilation procedure, the model was forced with the same daily
273 reanalysis atmospheric data used during the model spin up. The simulation period is
274 from 1970 to 2010 for all the experiments. In this study we present the results of three
275 long-term experiments using the MPI-CCDAS, which differ only in the timeframe of
276 the observational records used during the assimilation: 1) experiment ALL, covers
277 data in 1980-2010 and includes the complete available timeframe of the two
278 observational data sets, i.e. for FAPAR is from 1982 to 2006 and for the atmospheric
279 CO₂ concentrations is from 1982 to 2010; 2) DEC1, covers observations available
280 from 1982 to 1990; and 3) DEC2, covers observations available from 1990 to 2000.
281 Because of the different lengths of the CO₂ records for some stations, this ultimately
282 leads to different number of observational data used for each experiment (Fig. 2).
283 In all of the experiments the first ten years of the simulation (1970 to 1979) are to
284 allow phenology, vegetation productivity and the fast land C pools to adjust to the
285 new model control vector \mathbf{p} and avoid any imprint of the initial conditions on the
286 calculation of the cost function. The soil C pool at the beginning of the experiment
287 was included in the model control vector. Thus, the initial period is discarded and
288 only results from 1980 are reported. The periods of time that fall outside the
289 assimilation window of the observational constraints on each experiment are thus
290 periods of model prognosis, i.e. the prognosis period in DEC1 is from 1991 to 2010,
291 and in DEC2 for 2001 to 2010.

292 3 Results

293 3.1 Mean seasonal phenology

294 We analyzed the global distribution of FAPAR before and after the assimilation
295 against the observations. To facilitate the analysis, we also divided the global land
296 into eight regions: Boreal West and East (BW and BE, for latitudes north of 60 °N),
297 subtropical Northwest and Northeast (STNW and STNE, between latitudes 20 °N and
298 60 °N); tropical West and East (TW and TE, between latitudes 20 °N and 20 °S);
299 subtropical Southwest and Southeast (STSW and STSE, for latitudes south of 20 °S)
300 (Fig. 1). The normalized RMSE (NRMSE = RMSE / mean(FAPAR_{obs})) and bias
301 between the modeled and observed FAPAR for 1982 to 2006 are somewhat reduced
302 by all assimilation experiments compared to the PRIOR (Table 2). One cause for this
303 decreased model-data misfit is the change in the spatial distribution of LAI, primarily
304 caused by the optimization of the PFT-specific maximum LAI (Λ_{\max}) parameter (see



305 Fig. A2 in the appendix for results of parameters changes after the assimilation).
306 Compared to the PRIOR experiment, the assimilation leads to substantial changes in
307 the LAI of the tropical forest area, with general reductions of LAI in all three
308 assimilation experiments. There is less agreement for the extra-tropical areas, with the
309 ALL experiment suggesting small reductions in LAI, whereas the experiments DEC1
310 and DEC2 see slight increases (Fig. 3, left panels) relative to the PRIOR experiment.
311 The second reason for the reduced misfit is an improved representation of the FAPAR
312 interannual variability (IAV) at a regional scale (Fig. 4) and seasonal cycle at the
313 pixel level (Fig. S2), particularly in the temperate and boreal zones of the Northern
314 Hemisphere. This is also evidenced by the increase in linear correlation coefficient R^2
315 between modeled and observed FAPAR with respect to the PRIOR experiment (Fig.
316 3, right panels). However, it is important to note that the fit remains far from perfect,
317 likely owing to model structural errors in the way that the meteorological triggers of
318 phenological events adjust to local climatic conditions. The average global correlation
319 between model and observations increased moderately in all the assimilation
320 experiments compared to the PRIOR experiment (Table 2 and Fig. 3). This is
321 particularly true for the DEC1 and DEC2 experiments ($R^2=0.34$ in both experiments)
322 than in the experiment that includes all the window of assimilation, i.e. ALL
323 ($R^2=0.20$). The improved correlation is primarily the consequence of an increased
324 ability of the model to simulate the timing of green-up and brown-down, and its IAV
325 at regional scale (Fig. 4). Interestingly, the model fit is better if the model is only
326 subjected to 10 years of data, instead of exposing it to the entire time series.
327 To further analyze the effect of the assimilation procedure on the simulated
328 seasonality and monthly growth rate of the FAPAR, we also selected six pixels that
329 are distributed in locations characterized by a dominant PFT (see Fig. 1 for the
330 geographic location of the pixels). A clear improvement after the assimilation is in
331 pixels P1, P2, and P6, where the magnitude of the mean seasonal cycle is better
332 represented when compared to the observations (Fig. S2). Also, the timing of the
333 mean seasonal cycle is corrected e.g. in pixels with large seasonal amplitude such as
334 in P1 (located in Eastern Siberia) and in P6 (located in Canada). While in the PRIOR
335 experiment (and ALL experiment) the onset and peak of the growing season in P1 and
336 P6 are delayed by up to two months, in the results from experiments DEC1 and DEC2
337 this delay is reduced to only one month. This correction might be partially due to
338 changes in some optimized parameters: increase in the day length at leaf shedding (t_c)



339 and reduction in the temperature at leaf onset T_{ϕ} , detected for both the CD and CE
340 phenotypes (as well as for ETD and TeCr) (Fig. A2); this is because these parameters
341 control the onset and end of the vegetation activity. This temporal shift however, is
342 less evident in other pixels such as in P2, despite changes in T_{ϕ} and t_c after the
343 assimilation in TrH, and this is because the amplitude of the seasonal cycle is small
344 and only changes in the magnitude of the amplitude are evident (Fig. S2).

345 In the results of DEC1 and DEC2 for pixel P3 (dominated by TeCr), the water stress
346 tolerance time (τ_w) and T_{ϕ} were largely reduced, whereas the leaf shedding timescale
347 ($1/\tau_l$; earlier shedding) increased. These changes allowed a considerable improvement
348 in the timing and duration of the FAPAR in the growing season. The seasonal spatial
349 distribution of the correlation coefficient R^2 for the period 1982-2006 (considering
350 only the period of available FAPAR_{obs}), obtained after the linear correlation between
351 the FAPAR_{obs} and the model output, is shown in Fig. S3, and the mean global values
352 are also listed in Table 2. The R^2 increased mostly in the Northern Hemisphere where
353 is evident the spatial extent of the improvement in FAPAR after the assimilation
354 during spring and autumn in the experiments with a shorter window of assimilation
355 (DEC1 and DEC2).

356 3.2 Mean characteristics of atmospheric CO₂

357 We next analyze the performance of MPI-CCDAS with respect to the atmospheric
358 mole fractions of CO₂. As example, we compare observed and simulated CO₂ mole
359 fractions at three stations: 1) at the Northern Hemisphere (Alert, ALT), 2) at the
360 Tropics (Mauna Loa, MLO), and 3) at the Southern Hemisphere (South Pole, SPO) in
361 terms of the mean seasonal cycle, IAV and monthly growth rate. We also compare the
362 fluxes from the assimilation to fluxes obtained from an atmospheric transport
363 inversion (referred to as INV). Very similarly to the MPI-CCDAS, the atmospheric
364 transport inversion is constrained by atmospheric CO₂ data linked surface fluxes
365 through a tracer transport model, but it adjusts the land surface CO₂ fluxes directly
366 rather than through adjustments to the parameters of a land-surface process model.
367 The inversion set-up used here is similar to the Jena CarboScope v4.1 (Rödenbeck,
368 2005; Rödenbeck et al., 2003), involving the same transport model (TM3) as in the
369 MPI-CCDAS. To make the inversion results as comparable as possible to the results
370 of this study, the same prior fluxes from fossil fuel emissions and ocean as in MPI-
371 CCDAS were used, as well as the same CO₂ stations. This comparison also helps to



gauge the impact of non-land surface fluxes on the ability to reproduce the observations. Results on this comparison are shown in Fig. 5. For MLO and ALT the timing of the seasonal cycle was already well reproduced in the PRIOR simulation, whereas the assimilation mainly corrects for the amplitude of the seasonal cycle and the long-term trend. In the SPO there are larger relative differences between the model results and the observations, however of a much smaller magnitude than for the two other stations. After the assimilation in the three experiments, the phase in the seasonal CO₂ is shifted by approximately a month to better match the pattern in the measurements, and the amplitude of the seasonal cycle after the assimilation 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 that the amplitude for stations located in the Northern Hemisphere (> 40 °N) is reduced after the assimilation, and in closer agreement to the observations than in the PRIOR simulation. The largest reduction took place in the Northernmost Station (ALT) where the seasonal amplitude decreased from 23.5 ppm in the PRIOR experiment to 16.5 ppm in the ALL experiment after the assimilation, 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, i.e. values of $R^2 > 0.9$ in the Northern Hemisphere in all of the experiments (including the PRIOR simulation). In the tropics (specifically between 40 °N and 20 °N) the MPI-CCDAS achieves an improvement of the model performance by reducing the misfit of the phasing of the seasonal cycle, as evidenced by an increased linear correlation (Fig. 6b), however at the expense of reducing the seasonal cycle amplitude stronger than the observed one. The INV results show a closer agreement to the observations in the statistical analysis shown in Fig. 5 and 6.

3.3 Global and regional carbon fluxes

We next analyzed the spatiotemporal changes of the simulated land surface gross and net carbon fluxes in the posterior experiments relative to the PRIOR compared to independent data. At large-scale, the variation of the NBE (net biome exchange of CO₂ with the atmosphere, referred to as the Net Ecosystem Exchange, NEE plus the land use change related flux) from all of the simulations through the time series is similar to that from the Global Carbon Project 2017 (GCP17; Le Quéré et al., 2018) and INV, with the major anomalies collocated in time (Fig. 7a). A comparison of the



406 fluxes from the ocean and fossil fuels from this data set to the corresponding fluxes
407 that are prescribed in CCDAS is shown in Fig. 8. The total annual NBE from the three
408 posterior experiments falls within the spread (shadow green area in Fig. 8d calculated
409 as ± 1 standard deviation) of the NBE mean of the terrestrial ecosystem models in the
410 GCP17, contrary to the PRIOR simulation. However, the 1982-2010 mean net biome
411 exchange in all of the assimilation experiments through the time series is on average
412 1.6 PgC yr^{-1} lower than the flux in the PRIOR simulation ($-2.07 \text{ PgC yr}^{-1}$) and 0.6
413 PgC yr^{-1} lower than the GCP17 value ($-1.23 \pm 0.98 \text{ PgC yr}^{-1}$) (Table 3, Fig. 8d and
414 Fig. S4 for summary of C balance).

415 In all MPI-CCDAS simulations, the net land-atmosphere C exchange is reduced
416 relative to the PRIOR result in most of the Southern Hemisphere, while NEE
417 increased in the Northern Hemisphere (Fig. S5c, e and g). The analysis per regions
418 illustrates that the extra-tropical northern areas contribute the most to the global net
419 CO_2 flux. The increase in respiration (more CO_2 emissions to the atmosphere) in the
420 tropics is clearly depicted in the latitudinal gradient of NBE shown in Fig. 7c and in
421 the spatial distribution of the NEE difference between the PRIOR and the posterior
422 experiments (Fig. S5c, e and g). As in the tropics, the NEE in the southern subtropical
423 regions was consistently reduced after the assimilation experiments, also switching
424 the NEE of the STSE region from a C sink of $-0.18 \text{ PgC yr}^{-1}$ in the PRIOR to a mean
425 C source to the atmosphere of $0.016 \text{ PgC yr}^{-1}$ in the DEC2 posterior experiment.

426 In the boreal east and west regions (BE and BW), the net land C emissions increased
427 in all of the posterior experiments relative to the PRIOR (Fig. S5c, e and g) with the
428 largest increase in BE for DEC2 ($-0.29 \text{ PgC yr}^{-1}$) relative to the corresponding value
429 in the PRIOR ($-0.09 \text{ PgC yr}^{-1}$).

430 The simulated latitudinal GPP values agree well with the data-driven Model Tree
431 Ensemble (MTE) estimate from Jung et al. (2011) for the period 1982 to 2010 north
432 of 30°N . However, the assimilation results are low biased in the tropics, which
433 propagated into low estimates of global GPP in all the posterior results (Fig. 7d and
434 Table 3). After the assimilation, the global GPP and NPP are reduced in the three
435 posterior experiments compared to the PRIOR ($118.5 \text{ PgC yr}^{-1}$ and 54.5 PgC yr^{-1} ,
436 respectively). In contrast to the posterior global mean of GPP, the value in the PRIOR
437 simulation compares favorably well to the global mean from the MTE product (118.9
438 PgC yr^{-1}) for the same period of analysis. The global mean GPP is reduced by up to
439 26 PgC yr^{-1} on average in the three posterior experiments compared to the PRIOR



experiment, but simulation DEC1 experienced the largest reduction in the global photosynthetic C uptake (mean global GPP of 82.9 PgC yr^{-1}) relative to the PRIOR value (Table 3 and spatial distribution in Fig. S5d, f, and h).

Although the magnitude of the global NBE and GPP is smaller in the posterior experiments than in the prior, the similar slope detected between the prior and posterior experiments in the anomaly of these fluxes (calculated relative to the temporal mean of each time series) (Fig. 7a and b), suggests that the response to the environmental conditions remains the same through the simulation period even after the assimilation. This robust response shows e.g. in GPP a similar and gradual increasing C uptake (positive trend) during the period of analysis, only with a slightly reduced slope in the PRIOR experiment (Fig. 7b). These trends of the GPP anomaly differ from the one in the MTE GPP, which is only driven by trends in the remote sensing FAPAR and climate parameters and it does not consider increases in photosynthetic light-use efficiency due to CO_2 fertilisation.

3.4 Interannual variability and long-term phenology trend

We analyzed the long-term trend in the FAPAR signal at regional scale with the purpose of evaluating if the model is capable of reproducing the observed long-term greening or browning (i.e. trend to increase or decrease of FAPAR throughout time, respectively) over large regions (Fig. 1). Compared to the $\text{FAPAR}_{\text{obs}}$, the IAV of FAPAR (obtained from the monthly signal for each experiment) is improved only in the Boreal regions after the assimilation, whereas in the tropical and subtropical regions, the assimilation does not improve the variability (Fig. 4). This is also identified in the monthly growth rate of FAPAR, which is a representation of the long-term trend for each region after the simulation experiments (Fig. 9). A positive monthly growth rate indicates a trend for the vegetation to greening, and this is occurring in all of the regions according to the FAPAR satellite observations, except in STSW where the long-term trend indicates a decrease of FAPAR (i.e. browning). In this region, the assimilation improved the long-term trend from a positive to a negative growth rate in the three posterior experiments (Fig. 9). Despite in most of the regions the assimilation results agree on a positive long-term trend as in the observations, the magnitude of this trend is in large disagreement to the observations. Particularly in the Boreal East region, the PRIOR experiment overestimates the FAPAR trend by almost double when compared to the observations, and after the assimilation the trend is reduced leading instead to a slight underestimation of the



474 growth rate in all of the experiments. The largest disagreement between $FAPAR_{obs}$
475 and $FAPAR_{mod}$ after the assimilation is in the TW region, where the observations
476 show a larger and positive trend in FAPAR during the period of analysis, whereas this
477 trend is not captured in the PRIOR and in all the posterior experiments (Fig. 9).
478 Despite the lack of correction of variability and long-term trend in most of the
479 regions, the RMSE between the observations and the regional model results (for
480 1982-2006) is reduced in all of the areas after the assimilation (Fig. A3). The error
481 reduction is evident in experiments DEC1 and DEC2 (the average error reduction
482 after assimilation, in comparison to the error in the PRIOR, for Boreal regions is 19
483 %, and for subtropical northern and southern regions is about 16 %), except for the
484 tropical regions TE and TW where the largest error reduction took place after the
485 ALL experiment (21 % on average) in comparison to the PRIOR experiment.

486 3.5 Interannual variability and long-term trends in atmospheric CO₂

487 During the nearly thirty years of atmospheric CO₂ data available, the time series of
488 the CO₂ mole fractions for the PRIOR model underestimate the long-term trend, and
489 start to deviate in the first five years of the time series; whereas for all the assimilation
490 experiments, the trend is in closer agreement to the long-term trend of the
491 measurements during the entire period of the assimilation (leftmost panels of Fig. 5).
492 This correction to the long-term trend is depicted also in the rest of the stations,
493 expressed as the linear monthly trend in simulated or observed CO₂ concentrations
494 (rightmost panels of Fig. 5 for ALT, MLO and SPO, and Fig. 6c for summary of all
495 28 stations). The mean of the growth rate calculated from the results of the ALL
496 experiment matches very well with the result of the same calculation in the
497 observations (0.15 ppm month⁻¹ in both cases) compared to the PRIOR model (0.087
498 ppm month⁻¹). The results of the experiments using only 10 years of CO₂ data show
499 marked improvements compared to the PRIOR, but tend to underestimate the
500 atmospheric CO₂ trend (thus overestimate the terrestrial land uptake) with DEC1
501 (0.14 ppm month⁻¹) and DEC2 (0.145 ppm month⁻¹). Despite moderate improvement,
502 the MPI-CCDAS is incapable of improving the IAV of the atmospheric CO₂
503 concentration substantially; remaining unchanged the most notable deviations from
504 the observed signals after the assimilation procedure (Fig. 5).

505 3.6 Prognostic capability of MPI-CCDAS

506 Finally, we analyze the prognostic capability of CCDAS by comparing the model-
507 data fit of the decadal assimilation runs to that of the assimilation runs using all data



508 as a reference case for “best possible” model-data match given the structural
509 limitations of the MPI-CCDAS to match the observations. To achieve this, we
510 calculated four-year mean differences between the atmospheric CO₂ mole fraction
511 measurements and the CO₂ model results for all of the stations (for the period 1982-
512 2010) (Fig. 10), and also between the FAPAR satellite data and the monthly FAPAR
513 model results (for the period 1982-2006) (Fig. 11). We also calculated the RMSE
514 between the CO₂ measurements and model results for each station for four different
515 periods: 1982-1990, 1990-2000, 2000-2010 and 1982-2010 (Fig. A4). The choice of a
516 four-year window was made because with Fig. 5 it was established that the capacity
517 of the MPI-CCDAS to improve the representation of observed interannual variability
518 was very moderate.

519 In the ALL assimilation experiment, the atmospheric CO₂ concentration is
520 consistently matched across the entire assimilation period with a -0.03 ± 1 ppm
521 average bias to the observations (Fig. 10). This is in striking contrast to the PRIOR
522 experiment, which fails in all of the stations to reproduce the long-term trend (as
523 discussed earlier). The four-year mean CO₂ mole fraction at the end of this simulation
524 is 18.8 ppm lower than in the observations. This is also recognized in the RMSE
525 results where the PRIOR results have the largest error in all of the stations and periods
526 (between 2.8 and 18.7 ppm) (Fig. A4).

527 As for the posterior experiments, the performance of the assimilation of CO₂ mole
528 fraction improves, and mostly during the period of the window of assimilation.
529 During those periods, the difference to the measurements and RMSE is reduced,
530 whereas the error increases during the periods of time outside of the window of
531 assimilation. For the DEC1 experiment, the four-year mean difference among the
532 measurements and the model results is between -0.3 and 0.3 ppm in the 80's, a level
533 at which it remains for the 1990s, where the experiment did not see any observations,
534 but the fit increasingly degrades after year 2000, with an underestimate of the CO₂
535 mole fraction by 1.6 ppm on a four-year average (still a 90 % reduction in misfit), and
536 the RMSE is also higher than in DEC2 and ALL for the period 2000-2010 (Fig. A4c).
537 The model results show that when only the first decade of data is assimilated (i.e. in
538 DEC1), a larger deviation to the long-term trend of atmospheric CO₂ is identified
539 after 2000. This is also identified in the results from DEC2 where the lowest four-year
540 mean difference between the observations and the assimilation results takes place in
541 the period of the window of assimilation for this experiment (1990-2000) (Fig. 10 and



Fig. A4b for RMSE). During this period, the model overestimates the CO₂ atmospheric concentration only by 0.15 ppm on average whereas for the periods of time outside the window of assimilation, the CO₂ concentration is underestimated by 0.64 ppm (in the period 1982-1990) and by 1.04 ppm (in the period 2000-2010). Thus, also in experiment DEC2 the prognostic skill of CCDAS is reduced outside the window of assimilation, and the long-term trend is less well reproduced than in the ALL experiment.

In the ALL assimilation experiment, the atmospheric CO₂ concentration is 0±1 ppm lower than the average value in the observations for the entire simulation period (that corresponds also to the window of assimilation). This suggests that a longer record of atmospheric CO₂ measurements favorably contributes to a better representation of the long-term values after the assimilation, but the average deviation to the observations by using shorter assimilation periods do not deviate far from the upper limit of the uncertainty when using the longest record.

We also calculated the four-year mean differences between the satellite FAPAR and the results of the PRIOR and assimilation experiments at regional scale (Fig. 11). In this case, the prognostic skill of CCDAS for the periods outside the windows of assimilation is less evident, with a consistent four-year mean difference within the time series and between experiments.

4 Discussion

The simultaneous assimilation of long-term space borne FAPAR and atmospheric CO₂ measurements in the MPI-CCDAS leads to an overall improvement in the modeled global carbon fluxes (as summarized in Fig. A3 and A4). The MPI-CCDAS is capable of extracting information about the seasonal cycle and the long-term trend from the FAPAR observations. However, the imprint of the interannual variability (IAV) on the cost function of the MPI-CCDAS is comparatively low. Therefore, the IAV remains largely unchanged in the posterior. With the exception of the tropical latitudes, the mismatch between observations and model output is small, and thus of little concern. The lacking ability of the MPI-CCDAS to reproduce the higher IAV in the tropical bands, may be indicative of a too weak drought response in the maximum leaf area index of the model. However, the modeled signal remains within 0.05 FAPAR (dimensionless) of the observations, and the importance of this mismatch should thus not be too interpreted too strongly.



575 The use of decade-long FAPAR data (DEC1 and DEC2) already leads to notable
576 improvement of the simulated seasonal phenology of the land surface. This
577 improvement is predominantly the result of adjustments in the Northern Hemisphere
578 dominated by phenotypes controlled by parameters for temperature and day-length
579 thresholds such as deciduous and evergreen needle leaf and extra-tropical deciduous
580 trees. Thus the optimization of parameters that regulate the onset and end of the
581 growing season improved the timing of the Northern Hemisphere FAPAR during
582 spring and autumn. This finding is generally consistent with the previous application
583 of the MPI-CCDAS for only 5 years (Schürmann et al., 2016).

584 A long-term greening trend in vegetation, especially in boreal regions has been
585 previously observed in analysis of space borne data (Forkel et al., 2016; Lucht et al.,
586 2002). While this enhanced vegetation greening was captured in the model already
587 before the assimilation, it was mostly overestimated in northern regions and
588 underestimated in the Southern Hemisphere. At regional scale, the assimilation in all
589 of the posterior experiments improved the growth rate of FAPAR, reflecting a
590 greening trend, and is in closer agreement to the satellite FAPAR data. This was
591 mostly achieved in boreal regions. However, the moderate improvements in the
592 simulated trend in temperate regions of the western hemisphere are associated with a
593 decreased performance in the eastern hemisphere, indicating that the model structure
594 of MPI-CCDAS is incapable of reconciling regional differences. It is unclear whether
595 this is an indication of the need to parameterize these hemispheres differently in terms
596 of their phenological response to the underlying driving factors (such as temperature,
597 moisture availability and day-length), or whether land-use or vegetation dynamics
598 processes not considered by MPI-CCDAS are the reason for this mismatch. Despite
599 these broad-scale improvements, at pixel level the MPI-CCDAS does not necessarily
600 reproduce the magnitude of the greening trend and its interannual variability in all the
601 posterior experiments, which results from the structural dependence of the MPI-
602 CCDAS on few, globally applicable PFT-level parameters, and challenges in using
603 the spatial mixed signal at the model resolution to infer PFT-specific parameters. A
604 likely better strategy for constraining these PFT-specific parameters would be to
605 resample the highly resolved satellite product to PFT-specific FAPAR maps prior to
606 aggregation, and provide PFT-specific FAPAR maps to the CCDAS.

607 Our results also demonstrate that the long-term trend of atmospheric CO₂ and of its
608 seasonal amplitude in the Northern Hemisphere and at station level is considerably



improved. This is independent of the different periods of data used for the assimilation. However, the MPI-CCDAS consistently fails to resolve some of the features of the year-to-year variability of the measured atmospheric CO₂ stations, which translates into an acceptable, but far from perfect fit to the inferred annual carbon budget of the global carbon project (Le Quéré et al., 2018). We compared the performance at this time-scale to the results from an atmospheric CO₂ inversion (INV) with the same input fields and atmospheric transport model than MPI-CCDAS, to illustrate that these deviations do not reflect uncertainties in the representation of the atmospheric transport. It needs to be borne in mind that both the choice of the atmospheric transport model (and associated imperfections at resolving the vertical and lateral atmospheric transport of CO₂) and the method to aggregate atmospheric observations to obtain an estimate of the annual growth rate in the global carbon budget introduce some error in any estimate of the interannual variability. As a consequence, only the occurrence of larger model-data mismatches is of concern and can be interpreted as a genuine result of the MPI-CCDAS' inability to correctly resolve the carbon flux variation.

Particularly, the model is not able to capture large-scale relevant climatic disturbances that influence the interannual variability of the carbon cycle like fires, or the decrease in atmospheric CO₂ growth after explosive volcanic eruptions such as for Mt. Pinatubo in 1991, or increase in atmospheric CO₂ concentration due to fire occurrence associated with El Niño events (Frölicher et al., 2011; Frölicher et al., 2013). MPI-CCDAS lacks a representation of the effect of diffusive radiation on photosynthesis that likely contributed to the post-Pinatubo increase in terrestrial carbon uptake (Mercado et al., 2009). Other important limitations in the current MPI-CCDAS structure that influence the results are: the possibility of only prescribe annual non-dynamic LUCC fields, limiting the performance of the model for long-term dynamic changes in vegetation (Reick et al., 2013) and the possibility to dynamically account for fire disturbance (Lasslop et al., 2014) and peatland fires.

Independently of the amount of data used in the assimilation window, our results show that the GPP and NEE were consistently reduced globally compared to the prior run, i.e. less carbon uptake by plants leading to the model results to be in closer agreement to other independent estimates such as the GCP17. The MPI-CCDAS suggests a somewhat lower average annual atmospheric CO₂ growth rate (calculated by the sum of the net C fluxes from the ocean, land and fossil fuel emissions) than the



one estimated in the GCP17 (Le Quéré et al., 2018), even if the MPI-CCDAS estimate falls within the uncertainty of the GCP17 (Fig. 8 and S4). Most of the difference stems from small differences in the assumed fossil and ocean carbon fluxes. In case of the carbon fluxes from fossil fuels, the data prescribed in MPI-CCDAS does not contain fluxes due to e.g. cement and flaring, thus the magnitude of the annual carbon sources through the time series is consistently somewhat lower but still within the ± 5 % uncertainty of the GCP17 data (Le Quéré et al., 2018). As for the ocean carbon sink, the annual mean values prescribed in MPI-CCDAS are also of lower magnitude than the mean value in the GCP17 but falling in the lower limit of the uncertainty value (Fig. 8c and S4). The flux due to LULCC prescribed in MPI-CCDAS is also of lower magnitude than that one from the GCP17 because the simulation made by JSBACH 3.0 does not consider disturbances like fires and gross transitions, which might have also contributed to the lower land C sink obtained in the assimilation experiments compared to the total land C sink in GCP17.

Compared to independent estimates of GPP (Jung et al., 2007), the MPI-CCDAS GPP matches well in regions with a distinct, light and temperature driven seasonal cycle (i.e. north of approx. 30 °N), translating to a reduction in modeled GPP by 0.7 PgC yr⁻¹ in boreal regions. However, as in Schürmann et al. (2016), the tropical productivity is strongly reduced by the assimilation to estimates that are substantially lower than independent estimates (Jung et al., 2007). An important factor influencing the global reduction of GPP and the tropical uptake of C appears to be related to the difference in data availability of CO₂ stations between the assimilation windows, specifically the fact that in the data-poor period DEC1, tropical GPP is substantially lower than estimated independently and compared to the assimilation runs with more stations in DEC2 and ALL. As a result, the mean tropical land C source to the atmosphere in the prior experiment (mean NBE value of 0.12 PgC yr⁻¹, and minimum value of -0.07 PgC yr⁻¹, reflecting C uptake in the 4 °S latitudinal band) was increased to 0.37±0.17 PgC yr⁻¹ on average for all the posterior results.

The total global vegetation C stock in all of the experiments, including the PRIOR, is in closer agreement to the lower end of the estimate by Carvalhais et al., (2014) (296 PgC). In the posterior experiments, the vegetation C pool decreased between 14 and 20 % of the value in the PRIOR but still remaining within the range of the literature estimate (442±146 PgC). The global soil C stock showed a more drastic change after the assimilation. In all the posterior experiments, the soil C pool decreased by 45, 43



677 and 53 % with respect to the value in the PRIOR. Particularly after the assimilation,
678 the total C in the soil (1362 PgC) in the ALL experiment is in closer agreement to the
679 estimate from the Harmonized World Soil Database
680 (<http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML>;
681 last access January 2015) of 1343 PgC (Table 3). It is important noting that the
682 JSBACH 3.0 version used in this MPI-CCDAS does not include permafrost
683 processes; therefore the global soil C stock might be underestimated.

684 The parameter optimisation resulted in considerable reduction in the cost function and
685 norm of gradient, which can be clearly seen as a reduction in the root mean squared
686 error of the MPI-CCDAS compared to the FAPAR and CO₂ observations (Fig. A3
687 and A4). The trajectory of model parameters involved in the optimization differed for
688 each experiment and each phenotype. While some parameters such as the maximum
689 leaf area of grasses and shrubs and the correction parameter for the initial soil pool
690 size were consistently retrieved, some final parameter estimates varied considerably
691 between the three experiments, e.g., the tropical maximum leaf area index and some
692 of the parameters controlling the seasonality of the phenology (Figure A2). The
693 consequence of these variations are regional differences in the simulated compartment
694 fluxes GPP and ecosystem respiration, which apparently are not well constrained
695 from the observations. Interestingly, these differences lead to very similar absolute
696 values in global carbon fluxes and their trends. This clearly demonstrates a certain
697 degree of equifinality in the results, and cautions a too stringent interpretation of the
698 outcome of the MPI-CCDAS in terms of improving understanding about biosphere
699 processes and their long-term trends.

700 Notwithstanding these conceptual issues, the set-up of this study enables to test by
701 how much the quality of the data-model agreement is reduced by exposing the MPI-
702 CCDAS to shorter observational time-series. This can be done by comparing the
703 results of the ALL experiment to the years of 1990-2010 for the DEC1 experiment,
704 and for 2000-2010 for the DEC2 experiment. In terms of FAPAR, there is no clear
705 degradation of fit with time even though in general terms the trend in the data are best
706 matched with the ALL experiment. This is foremost a consequence of comparatively
707 small trends in observed FAPAR, implying that extracting mean seasonal patterns and
708 amplitude for a few years is most essential for simulating current and near-term
709 FAPAR. This would suggest that a focus of assimilation on high-quality and highly
710 spatially resolved FAPAR should be a priority over the use of long-term data sets.



711 The results are different for the case of projecting atmospheric CO₂, where the model-
712 data agreement of approximately ± 0.5 ppm during the assimilation period starts to
713 deviate for the DEC1 experiment later than 10 years after the end of the assimilation
714 window, whereas in the DEC2 experiment, the degradation of the model-data match
715 already starts after approximately 5 years. Nonetheless, with the caveat that MPI-
716 CCDAS does not fully explain the interannual variability of the land net carbon flux,
717 this suggests a reasonable short-term forecasting (for a small number of years) skill of
718 atmospheric CO₂.

719 5 Conclusion

720 The MPI-CCDAS is capable of simultaneously integrating two independent
721 observational data sets over three consecutive decades at the global scale to estimate
722 global carbon fluxes. The results demonstrate that assimilating only one decade of
723 observations, for two observational data (FAPAR and atmospheric CO₂
724 concentrations), leads to broadly comparable results and trends in the global carbon
725 cycle components than using the full time series of available observations (thirty
726 years). Currently the system is able to confidently predict the carbon fluxes in short
727 time scales (up to 5 years after the end of the window of assimilation) e.g. for
728 atmospheric CO₂ concentrations at site level, and the mean prediction remains within
729 the uncertainty of the observations. However, long-term predictions with CCDAS are
730 more uncertain, as the observational record does not fully constrain the long-term land
731 net C uptake in the current phase of rising atmospheric CO₂ and gradually changing
732 climate. The MPI-CCDAS is a computational expensive system, and the
733 demonstration that large-scale carbon fluxes can be improved by only using a limited
734 period of observations increases the feasibility of using DA to constrain the land
735 carbon budget in land surface models. However, we also show that there is
736 considerable variations in the estimated parameter space and regional distribution of
737 the land C uptake suggesting that further improvements in the land-surface model,
738 especially in the current structure and design, must be first solved to improve the
739 model and computational efficiencies of the system before an attempt to include
740 another observational stream can be made to potentially improve its prognostic skill.

741

742

743

744



745 **6 Code availability**

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

750

751 **Acknowledgements**

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



Table 1 – Model parameters selected for the optimization: rows 1 to 6: related to phenology, row 7 to photosynthesis and rows 8 to 11 to land-carbon turnover. The values in the table for each PFT (where applies only) are for the prior conditions: $p_{\text{pr}} \pm C_{\text{pr}}$. Values in f_{photos} are the photosynthetic parameters $I_{C_{\text{max}}} / J_{\text{max}}$ ($\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1} / \mu\text{mol m}^{-2} \text{ s}^{-1}$). 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.

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)
Λ_{max}	Maximum LAI ($\text{m}^2 \text{ m}^{-2}$)	7.0*	7.0*	5.0*	1.7*	5.0*	2.0*	3.0^	4.0^	3.0^	4.0^
$1/\eta$	Leaf shedding timescale (d^{-1})			0.07 ± 0.01	$5e-4 \pm 1e-4$	0.07 ± 0.01	0.07 ± 0.01	0.07 ± 0.01		0.07 ± 0.01	
τ_w	Water stress tolerance time (d)	300 ± 30	114 ± 10	-	-	-	50 ± 5	250 ± 25		250 ± 25	
T_ϕ	Temperature at leaf onset ($^{\circ}\text{C}$)	-	-	9.21 ± 1	9.21 ± 1	9.21 ± 1	-	1.92 ± 0.5		1.92 ± 0.5	
t_c	Day length at leaf shedding (h)	-	-	13.37 ± 1	13.37 ± 1	13.37 ± 1	-	-		-	-
ξ	Initial leaf growth state (d^{-1})					0.37 ± 0.03					
$f_{\text{photos}}^{\text{a}}$	Photosynthesis rate modifier	$39.0/74.1^{\wedge}$	$31.0/58.9^{\wedge}$	$66.0/125.4^{\text{a}}$	$62.5/118.8^{\text{b}}$	$39.1/74.3^{\text{c}}$	$61.7/117.2^{\wedge}$	$78.2/148.6^{\wedge}$	$100.7/191.3^{\wedge}$	$8.0/140.0^{\wedge}$	$39.0/700.0^{\wedge}$
Q_{10}	Temperature sensitivity to resp.					1.8 ± 0.15					
f_{slow}	Multiplier for initial slow pool					1 ± 0.1					
$f_{\text{aut_leaf}}$	Leaf fraction of maintenance resp.					0.4 ± 0.1					
$\text{CO}_2^{\text{offset}}$	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.

	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 (1980-2010), and other independent estimates (see table foot for description).

	PRIOR	ALL	DEC1	DEC2	INV	Literature
GPP (PgC yr^{-1})	118.5	96.8	82.9	97.0	-	118.9 ^a
NPP (PgC yr^{-1})	54.5	34.2	37.2	30.2	-	-
NEE (PgC yr^{-1})	-2.65	-1.14	-1.32	-1.17	-1.17 ^c	-2.25±1.17 ^b
NBE (NEE + LUCC) (PgC yr^{-1})	-2.07	-0.56	-0.74	-0.59	-	-1.23±0.98 ^b
ER (PgC yr^{-1})	115.5	95.0	80.9	95.1	-	-
Ra (PgC yr^{-1})	64.1	62.6	45.7	66.8	-	-
Rh (PgC yr^{-1})	51.4	32.4	65.2	28.3	-	-
Soil C (PgC)	2480	1362	1422	1165	-	1343 ^d
Vegetation C (PgC)	392	311	335	312	-	442±146 ^c
Litter C (PgC)	228	167	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 1980-2010. The NBE values include the LULCC reported for each individual model.

^c Inversion result is the average for 1980-2009

^d <http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML>

^e Carvalhais et al. (2014).



References

- Andres, R. J., Boden, T. A., and Marland, G.: Monthly fossil-fuel CO₂ emissions: mass of emissions gridded by one degree latitude by one degree longitude, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, 2013, Oak Ridge, Tenn., U.S.A., 10.3334/CDIAC/ffe.MonthlyMass.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: Chemical and physical meteorology*, 63, 309-327, 2011, 10.1111/j.1600-0889.2011.00530.x.
- Andres, R. J., Marland, G., Fung, I., and Matthews, E.: A 1° x 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950-1990, *Global Biogeochemical Cycles*, 10, 419-429, 1996, 10.1029/96GB01523.
- 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, 2014, 10.1038/nature13731.
- 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 Biogeochemical Cycles*, 22, GB2007, 2008, 10.1029/2007GB003033.
- 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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 3, 286-295, 2010, 10.1109/JSTARS.2010.2046626.
- 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, *Journal of Geophysical Research*, 99, 22831-22855, 1994, 10.1029/94JD01951.
- Dalmonech, D. and Zaehle, S.: Towards a more objective evaluation of modelled land-carbon trends using atmospheric CO₂ and satellite-based vegetation activity observations, *Biogeosciences*, 10, 4189-4210, 2013, 10.5194/bg-10-4189-2013.
- 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, *Journal of Climate*, 28, 2015, 10.1175/JCLI-D-13-00593.1.
- 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 Sensing*, 8, 1-29, 2016, 10.3390/rs8040275.

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., Zaehle, S., and Reichstein, M.: Enhanced seasonal CO₂ exchange caused by amplified plant productivity in northern ecosystems, *Science*, 351, 696-699, 2016, 10.1126/science.aac4971.

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, *Journal of Climate*, 27, 511-526, 2014, 10.1175/JCLI-D-12-00579.1.

Frölicher, T. L., Joos, F., and Raible, C. C.: Sensitivity of atmospheric CO₂ and climate to explosive volcanic eruptions, *Biogeosciences*, 8, 2317-2339, 2011, 10.5194/bg-8-2317-2011.

Frölicher, T. L., Joos, F., Raible, C. C., and Sarmiento, J. L.: Atmospheric CO₂ response to volcanic eruptions: the role of ENSO, season, and variability, *Global Biogeochemical Cycles*, 27, 239-251, 2013, 10.1002/gbc.20028.

Giering, R. and Kaminski, T.: Recipes for adjoint code construction, *ACM Transactions on Mathematical Software (TOMS)*, 24, 437-474, 1998, 10.1145/293686.293695.

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, *Advances in Space Research*, 38, 2196-2202, 2006, 10.106/j.asr.2003.07.079.

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.

Hurt, 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 Biology*, 12, 1208-1229, 2006, 10.1111/j.1365-2486.2006.01150.x.

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, *Journal of Geophysical Research*, 116, G00J07, 2011, 10.1029/2010JG001566.

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 Biogeochemical Cycles*, 21, GB4021, 2007, 10.1029/2006GB002915.

Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., and Mathieu, P.-P.: Consistent assimilation of MERIS FAPAR and atmospheric CO₂ into a terrestrial vegetation model and interactive mission benefit analysis, *Biogeosciences*, 9, 3173-3184, 2012, 10.5194/bg-9-3173-2012.

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., Sebal, 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, 2013, 10.1002/jgrg.20118.

Kaminski, T. and Mathieu, P.-P.: Reviews and syntheses: 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, 2017, 10.5194/bg-14-2343-2017.

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, *Journal of Geophysical Research*, 115, G04017, 2010, 10.1029/2009JG001119.

Lasslop, G., Thonicke, K., and Kloster, S.: SPITFIRE within the MPI Earth system model: model development and evaluation, *Journal of Advances in Modeling Earth Systems*, 6, 740-755, 2014, 10.1002/2013MS000284.

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., Lombardozi, 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, 2018, 10.5194/essd-10-405-2018.

Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneeth, 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., Segsneider, 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, 2015, 10.5194/essd-7-47-2015.

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

Lucht, W., Prentice, 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, 2002, 10.1126/science.1071828.

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, 2009, 10.1038/nature07949.

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 CO₂ uptake, transport, and storage by the ocean, *Global Biogeochemical Cycles*, 20, GB2002, 2006, 10.1029/2005GB002530.

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, 2012, 10.3390/s121216291.

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, 1988, 10.1088/0266-5611/4/4/008.

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, *Geoscientific Model Development*, 9, 3321-3346, 2016, 10.5194/gmd-9-3321-2016.

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 CO₂ inversions, *Biogeosciences*, 10, 6699-6720, 2013, 10.5194/bg-10-6699-2013.

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, *Journal of Geophysical Research*, 116, D09105, 2011a, 10.1029/2010JD015372.



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, *Journal of Geophysical Research*, 116, D09106, 2011b, 10.1029/2010JD015373.

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, *Journal of Geophysical Research*, 111, D02116, 2006, 10.1029/2005JD005952.

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, *Journal of Geophysical Research*, 112, D10116, 2007, 10.1029/2006JD008105.

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?, *Climate Dynamics*, 29, 565-574, 2007, 10.1007/s00382-007-0247-8.

Raupach, M. R., Rayner, P. J., Barrett, D. J., DeFries, R. S., Heimann, M., Ojima, D. S., Quegan, S., and Schimmlus, C. C.: Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications, *Global Change Biology*, 11, 378-397, 2005, 10.1111/j.1365-2486.2005.00917.x.

Rayner, P. J., Enting, I. G., Francey, R. J., and Langenfelds, R.: Reconstructing the recent carbon cycle from atmospheric CO₂, δ¹³C and O₂/N₂ observations, *Tellus B: Chemical and Physical Meteorology*, 51, 213-232, 1999, 10.3402/tellusb.v51i2.16273.

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 Biogeochemical Cycles*, 19, GB2026, 2005, 10.1029/2004GB002254.

Reick, C., Raddatz, T., Brovkin, V., and Gayler, V.: Representation of natural and anthropogenic land cover change in MPI-ESM, *Journal of Advances in Modeling Earth Systems*, 5, 459-482, 2013, 10.1002/jame.20022.

Rödenbeck, C.: Estimating CO₂ 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 Heimann, M.: CO₂ flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmospheric Chemistry and Physics*, 3, 1919-1964, 2003, 10.5194/acp-3-1919-2003.



Rödenbeck, C., Keeling, R. F., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean $p\text{CO}_2$ and sea-air CO_2 flux variability from an observation-driven ocean mixed-layer scheme, *Ocean Science*, 9, 193-216, 2013, 10.5194/os-9-193-2013.

Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W.: Monitoring vegetation system in the great plains with ERTS, Greenbelt, USA, 1974, 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, 2017, 10.5194/bg-14-3401-2017.

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 Sensing of Environment*, 180, 334-345, 2016, 10.1016/j.rse.2016.02.058.

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, *Geoscientific Model Development*, 9, 2999-3026, 2016, 10.5194/gmd-9-2999-2016.

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, 2014, 10.3402/tellusb.v66.23188.

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 $p\text{CO}_2$, and seasonal biological and temperature effects, *Deep-Sea Research 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, *International Journal of Remote Sensing*, 26, 4485-4498, 2005, 10.1080/01431160500168686.

Viovy, N. and Ciais, P.: CRUNCEP data set for 1901-2014, Version 6.1, https://esgf.extra.cea.fr/thredds/catalog/store/p529viovy/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, *Nature Climate Change*, 7, 345-350, 2017, 10.1038/nclimate3263.



Figures.

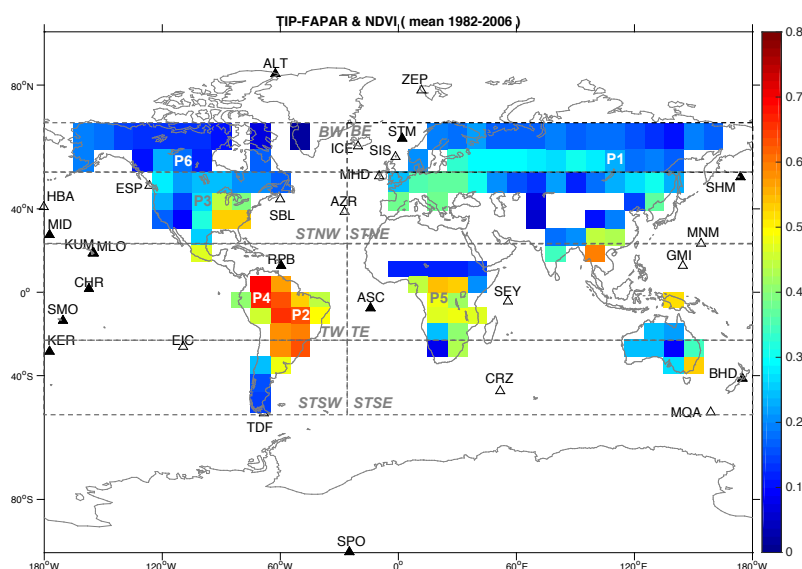


Figure 1 – Global distribution of the temporal mean (1982-2006) of the merged satellite FAPAR product used in the assimilation procedure. It shows also the spatial coverage of eight regions globally distributed: Boreal West and East (BW and BE, for latitudes north of 60 °N), subtropical Northwest and Northeast (STNW and STNE, between latitudes 20 °N and 60 °N); tropical West and East (TW and TE, between latitudes 20 °N and 20 °S); subtropical Southwest and Southeast (STSW and STSE, 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 CO₂ network measurements (filled triangles, stations only included in DEC1; empty triangles, stations included also in ALL and DEC2) for analysis of the assimilation results.

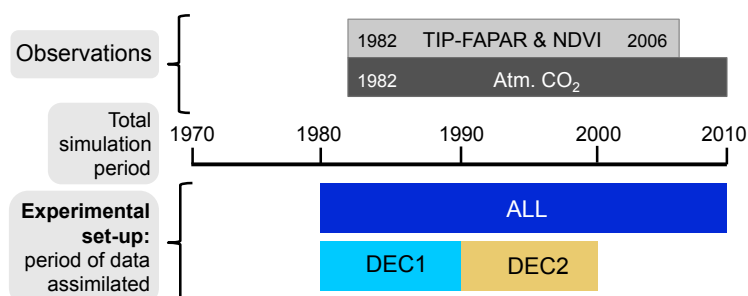


Figure 2 – Experimental set up for posterior experiments ALL, DEC1 and DEC2 that use different temporal windows for the assimilation of observations of FAPAR and molar fractions of atmospheric CO₂.

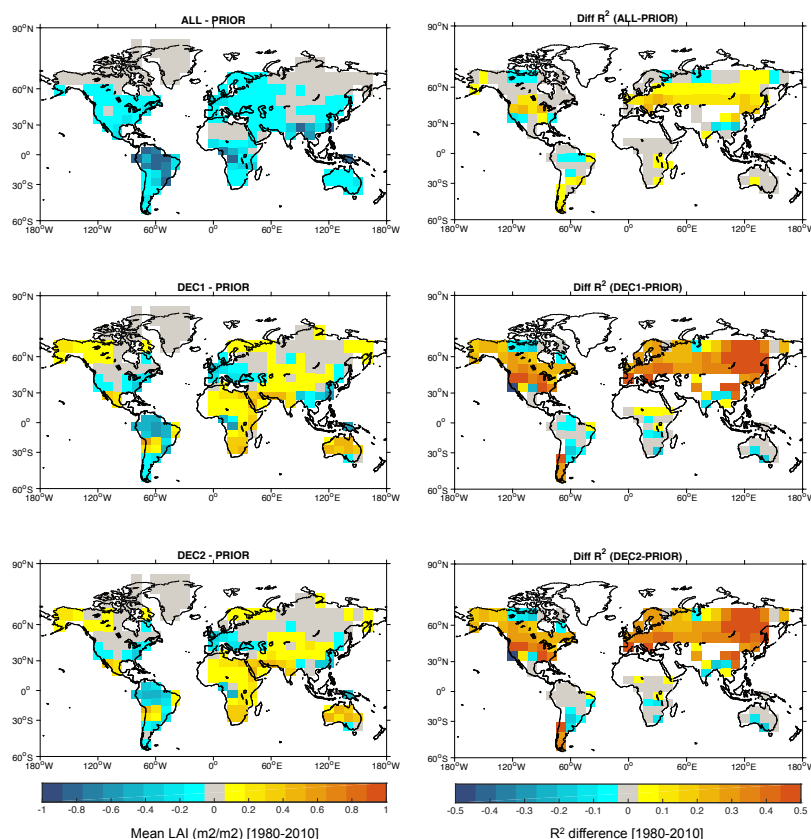


Figure 3 – Spatial difference between the results from the posterior and the PRIOR experiments for the total period of the simulation (1980-2010) of the mean Leaf Area Index (LAI) (left panels) and the correlation coefficient (R^2) of FAPAR between the model and the observations (right panels).

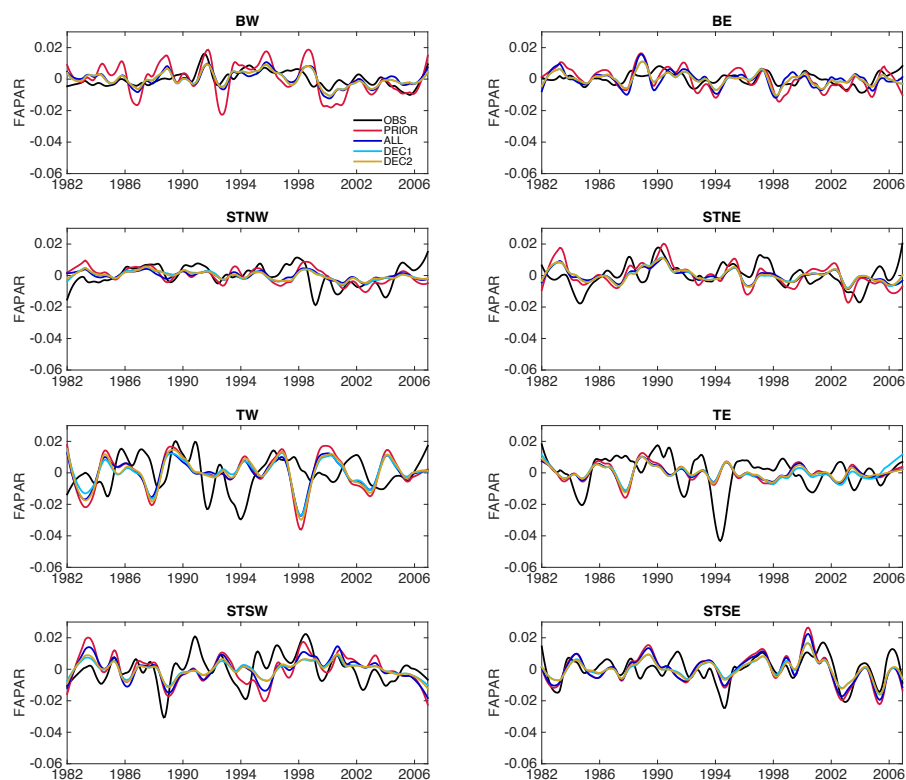


Figure 4 – Interannual variability of FAPAR in the satellite observations and model experiments for the six selected regions.

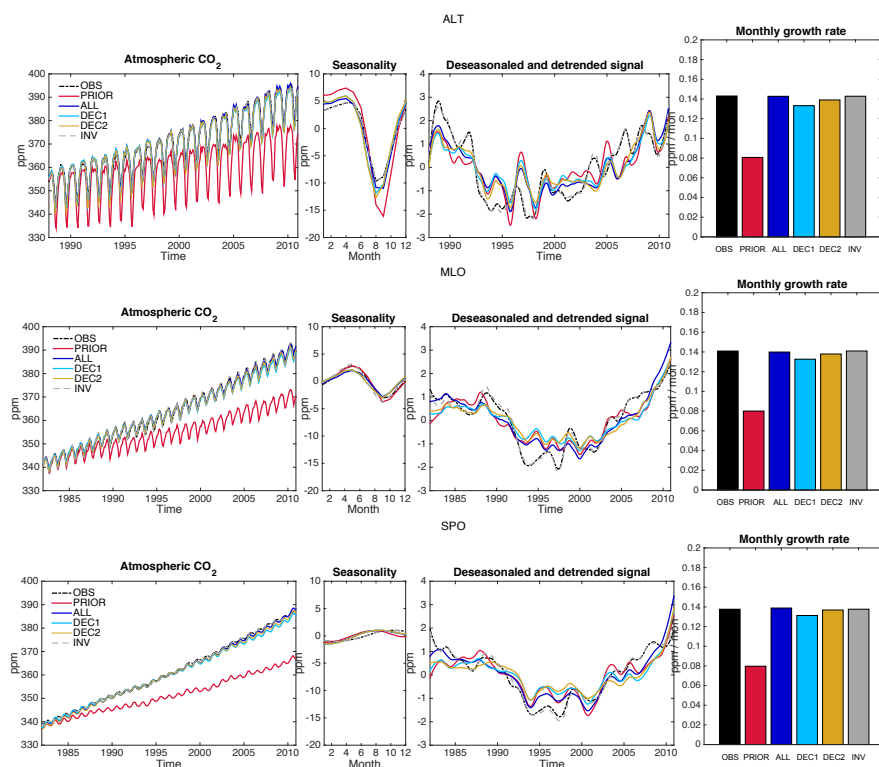


Figure 5 – Statistical analysis of atmospheric CO₂ in three flask measurement sites: Alert (ALT; top panels), Mauna Loa (MLO, center panels) and South Pole (SPO, bottom panels), from the measurements, PRIOR, posterior experiments (ALL, DEC1 and DEC2) and inversion (INV1). For each station the panels 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).

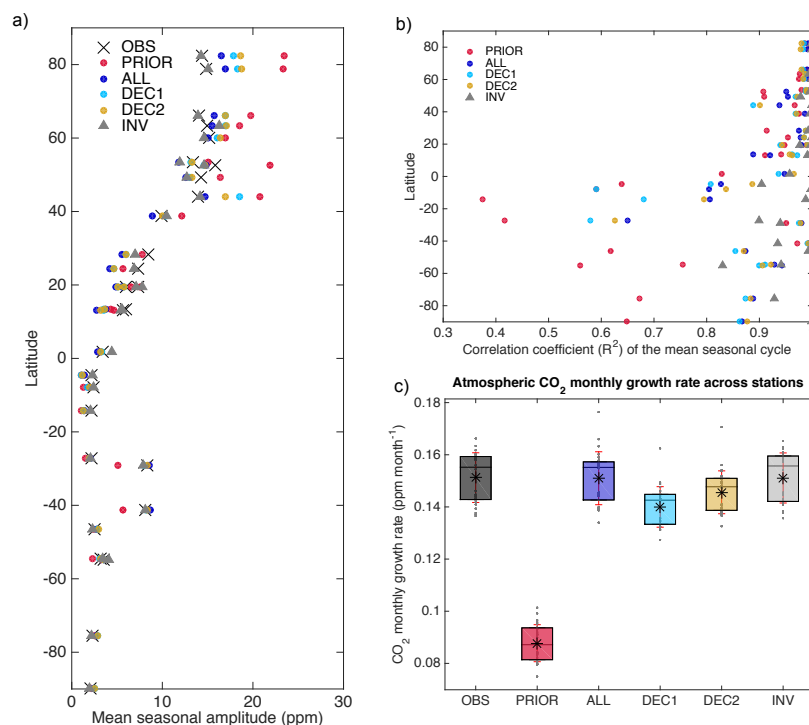


Figure 6 – a) Latitudinal distribution of the mean CO₂ seasonal amplitude for the 28 flask-measurement stations from the observations, PRIOR and posterior experiments; b) Latitudinal distribution of R^2 obtained from the correlation between the observations and each simulation results of the mean atm. CO₂ seasonal cycle and c) average atmospheric CO₂ monthly growth rate across stations for the observations and model results. The star on each bar is the mean of the atm. CO₂ monthly growth rate, the horizontal middle black line on each box is the median, the red whiskers depict the error as $\pm 1\sigma$, and the grey dots on each box are the actual monthly growth rate values for all the stations in each data set.

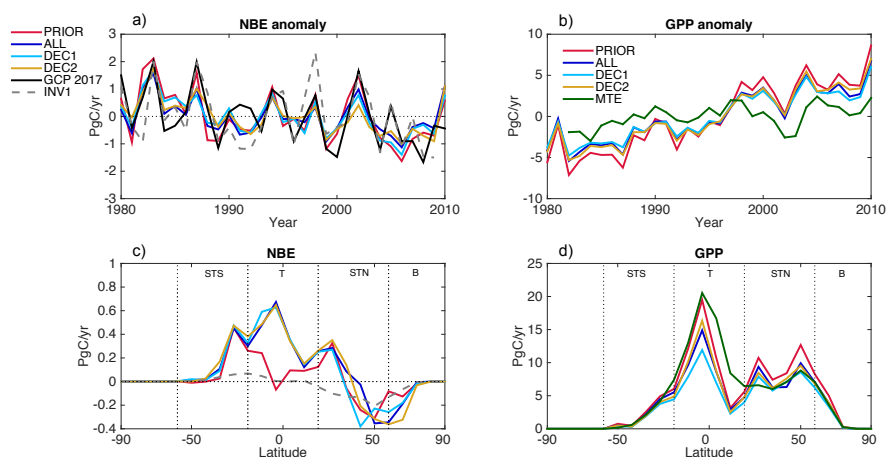


Figure 7 – Time series of the anomaly to the temporal mean of the time series (a and b), and latitudinal gradient (c and d) of the total Net Ecosystem Exchange (NEE including the influence of LULCC) (left) and Gross Primary Production (right) for the results of each model simulation. NEE from the model 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).

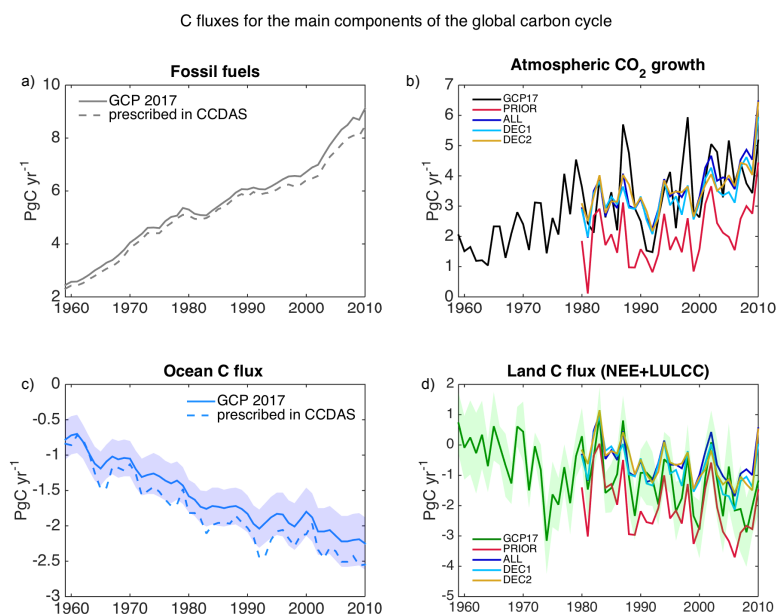


Figure 8 – Time series of the annual mean of the major components of the C cycle used as background fluxes in CCDAS compared to those from the GCP 2017. The atm. CO₂ growth from the model output is the result of the sum of fossil fuel, ocean, and land C fluxes. The blue shadow in the ocean C sink of 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.

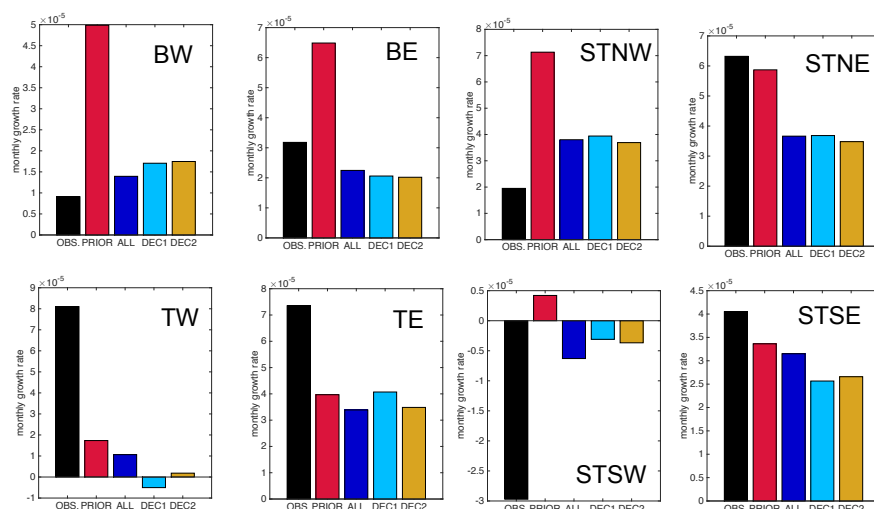


Figure 9 – Mean monthly growth rate of FAPAR for 1982-2006 on each analyzed geographical region for the satellite observations and results of PRIOR and the posterior experiments.

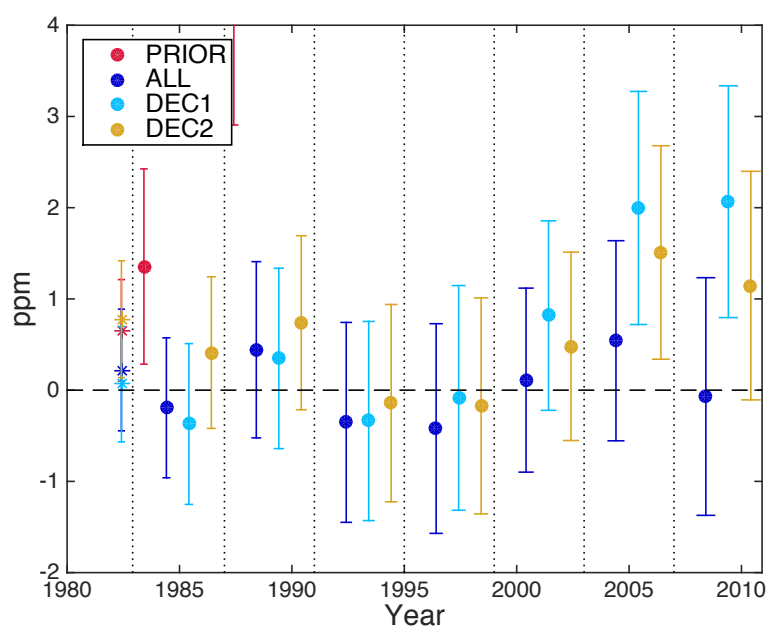


Figure 10 – Time series of the four-year mean of the atm. CO₂ anomaly to the observations for each model experiment and for all the stations. The y-axis is limited to the results in the posterior experiments. The error bar indicates ± 1 standard deviation of the four-year mean of the differences to the observations. The first marker in the time series (in asterisk) is the single value for 1982.

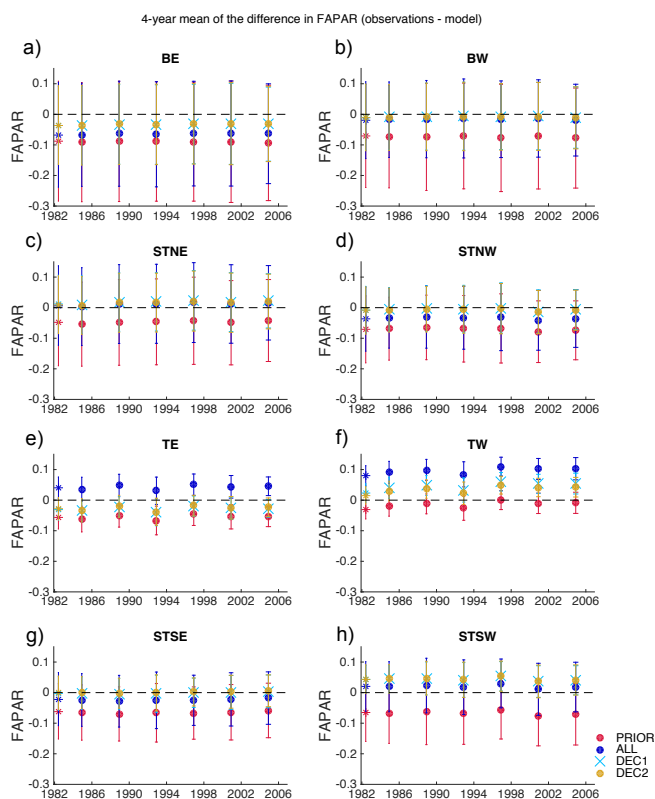


Figure 11 – Time series of the four-year mean of the FAPAR anomaly to the satellite data for each model experiment in six selected model pixels. The error bar indicates the ± 1 standard deviation of the four-year differences. The first marker (in asterisk) in the time series is the single value for 1982.



Appendix

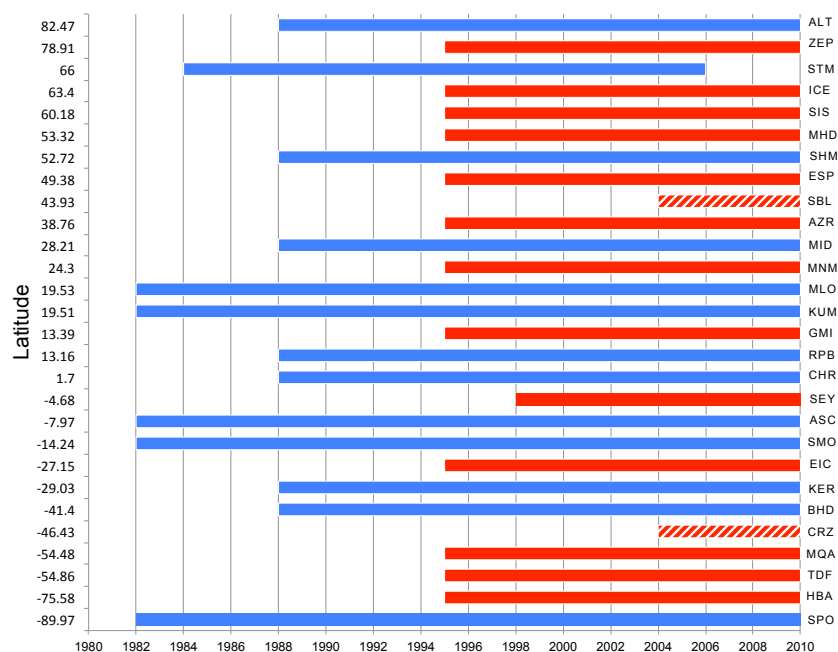


Figure A1 – Data availability and latitudinal location of the 28 stations where the long-term flask measurements of atmospheric CO₂ mole fractions were taken for assimilation in CCDAS. ALL experiment used all the stations of the time series (blue and red bars) (1980-2010); DEC1 used data only from stations with blue bars (1980-1990), and DEC2 used also the data in the stations with red bars (1990-2010) (except stations SBL and CRZ marked with patterned bar).

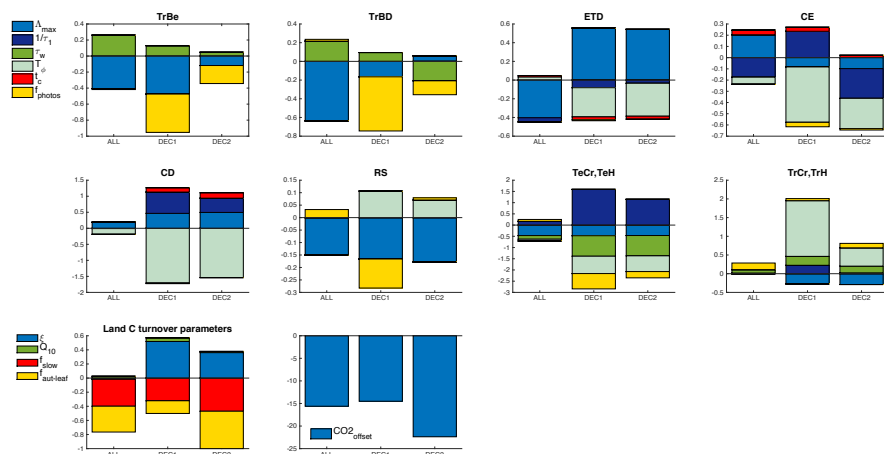


Figure A2 – Final value for each parameter p at the end of the assimilation experiments, normalized to the prior value (p_{pr}), i.e. $(p/p_{pr})-1$. This is shown for each model plant functional type and globally for the land C turnover parameters.

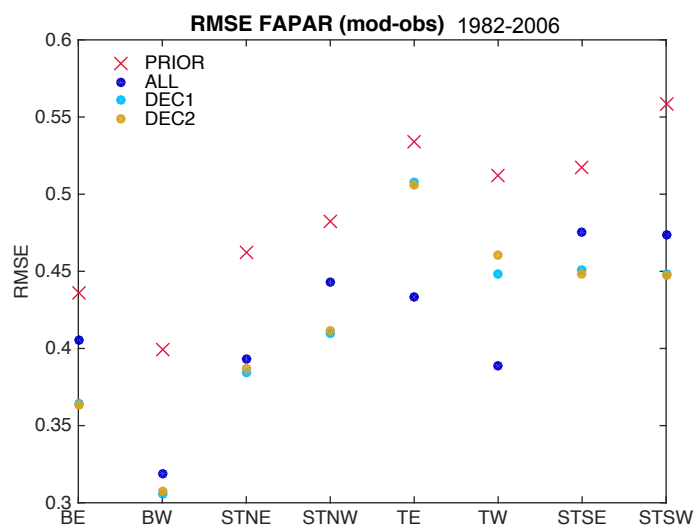


Figure A3 – RMSE for FAPAR from the model results and observations for the period 1982-2006 and for different regions.

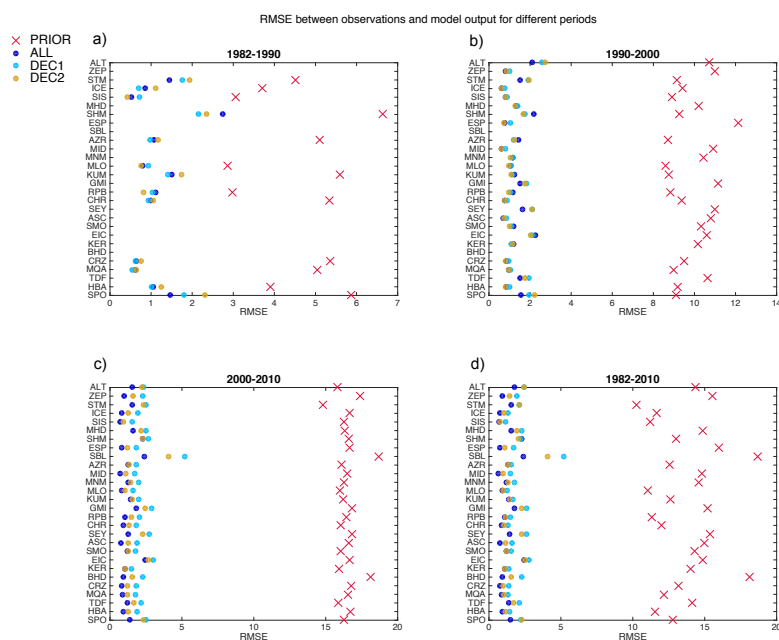


Figure A4 – RMSE for different periods between CO₂ atm. concentrations from measurements and model results for the different assimilation experiments for each of the station.