We appreciate the comments from two anonymous reviews. We answer below in blue text to each comment listed in black adding the explicit change made in the revised manuscript.

Reviewer #1

This paper develops an analysis of the global land C fluxes for three decades. The aim of the paper is to analyse the performance of a data assimilation system, over a historical period. The paper misses clear science questions that would broaden the interest of the study. I lack a clear understanding of the novelty of this work. The conclusions focus on the result that a single decade of data leads to similar results to 3 decades of assimilation, but I am not clear how robust this result is, and how significant it is. I suggest the authors focus more clearly on the novelty of their experimental results. It would be helpful to focus on science questions that are more broadly relevant to the C cycle science community.

We appreciate the reviewer's concerns and are grateful for the comments, which have helped us to formulate the manuscript more clearly. We present three decades of land C flux reanalysis with the aim to better understand the ability of MPI-CCDAS to make decadal projections of the land C cycle. Our approach was to assimilate different time-periods of data to understand the effect of data selection, but also the effect of withholding more recent information. In this way, we can assess the projection of carbon fluxes for the more recent period, and not only against the withheld observations, but also against a model run that has been constrained by this data, in order to disentangle model limitations (i.e. in reproducing an observation that the model was informed about through data assimilation) from prognostic uncertainty (i.e. in failing to reproduce an observation that it can reproduce when given this information). To our knowledge, this is a novel design and focus in data assimilation studies. Previous findings on the prognostic capacity of a CCDAS only looked at short timescales after the assimilation (Scholze et al. 2007; Schürmann et al. 2016), or in the case of Rayner et al. 2011, the authors used a much simpler model ignoring the interacting effects of water, energy and phenology on the carbon cycle predictions. In the revised manuscript, we are more explicit and clearer in the scientific contribution of our results that are relevant to the carbon cycle scientific community.

Our results demonstrate that it is not necessary to ingest more than one decade of observations to improve important features of the global carbon cycle over the following 1-2 decades. In particular, improvements were observed in the long-term trend and seasonal amplitude of atmospheric CO_2 concentrations at station level, as well as in the long-term trend, phenological seasonality and interannual variability of FAPAR. These results provide an insight into the amount of data that is necessary in data assimilation systems to improve the representation of the global carbon cycle components. This information might be decisive for opening the possibility of including newly measured data of other global indicators, such as SIF, that currently only exist for periods of less than a decade.

The new paragraph in the introduction of the revised ms where the aim of the ms is clarified reads:

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

It might be helpful to add experiments assimilating just FAPAR or just CO2, and test the outputs against the other (withheld) observational dataset. This experiment might indicate whether the model effectively couples canopy processes with atmospheric concentrations.
We appreciate the suggestion of the reviewer regarding the experiment of only assimilating one of the observational data sets at a time. However, we did not want to repeat this experiment because this has been previously done with the same system in Schürmann et al. 2016. In that work, 5 years of observational data (2005 – 2009) were assimilated in three independent experiments: only FAPAR, only atmospheric CO_2 concentrations and the two data sets simultaneously. The results showed that with this time period it was sufficient to demonstrate, in a nutshell, that when assimilating only FAPAR, the average growing season and vegetation seasonality (indicated by FAPAR) was considerably improved in the northern boreal areas. When atmospheric CO_2 concentrations were the only assimilated observations, the global gross and net carbon fluxes were overall improved, but the GPP in the tropics was significantly reduced when compared to the GPP after the FAPAR-only assimilation.
While we agree with the reviewer that in principle it would have been interesting to repeat these experiments to analyze their effect on long-term trends, this would require too much computational time to be included in this study and it was not the central focus of this work.

The abstract is too long, it needs to be reduced to 300 words and focused on the key outcome.
We significantly shortened the abstract and remain under 300 words, including only the key outcome regarding using different periods in the observational data during the assimilation.

Writing style is clunky, missing words – for instance the opening sentence is poorly comprehensible: "The observed contemporary in atmospheric CO2 is driven by anthropogenic emissions from fossil fuels and land-use change". I suggest the authors work together to improve the English, reduce sentence lengths, and shorten the more bloated paragraphs.
We realized that we missed the word "increase" in the opening sentence and appreciate the reviewer for pointing this out. The sentence should read: "The observed contemporary **increase** in atmospheric CO_2 …", we apologize for this. Besides this correction, the revised manuscript has undergone a thorough revision for the English language with the aim of shortening sentences and paragraphs.

Methods: The global grid is very coarse (8x10_) – what are the implications?
A previous work with the dynamic global vegetation model LPG-DGVM (Müller and Lucht, 2007), demonstrated that the effect of increasing the spatial resolution did not have a strong effect on the simulated regional and global carbon fluxes, whereas temporal dynamics are unaffected. To our knowledge, the effect on the simulated C fluxes after assimilation due to changes in the horizontal spatial resolution of the

model grid cells solely, has not yet been assessed with current data assimilation systems. One reason for this is that the computational cost of such approach is prohibitive. A study using ORCHIDAS (Peylin et al., 2016), suggested that the level of complexity of the ecosystem model in a data assimilation system, including its spatial resolution, does not guarantee improvements in the optimization of C fluxes. Such a conclusion is likely valid for the assimilation of atmospheric observations, but does decrease the ability of the CCDAS to adequately constrain phenology parameters given the mixed phenology in these large pixels. We would like to point out, however, that increasing the spatial resolution to a degree that would allow for a clean identification of PFT-wise parameters, exceeds the resolution of many state-of-the-art forward terrestrial biosphere models. The following paragraph in the revised ms is added in section 2.1 of Methods:
"This horizontal resolution allows computational feasibility and a realistic computational cost for the set of experiments presented in this work. Furthermore, previous evidence has shown that a higher spatial resolution in global vegetation models does not exert a considerable influence in the simulated carbon fluxes at global or regional scales when compared to results obtained with a coarse grid (Müller and Lucht, 2007). The lack of influence to improve the simulated global C fluxes due to changes in the model spatial resolution might also apply to CCDAS (Peylin et al., 2016)."

Why mix such a large spatial grid with such fine temporal scales? Is this valid or required? Why not have a grid that matches TM3?
We are unsure on the meaning of the question by the reviewer, because it combines both spatial and temporal scales. In the CCDAS, we integrate the simulated daily net CO_2 fluxes to monthly scale and transport them with the Jacobian representation of TM3. This approach allows us to account for the non-linear impact of weather anomalies on the surface fluxes, but removes the impact of synoptic atmospheric transport variability on the simulated seasonal and long-term dynamics of atmospheric CO_2 at the monitoring stations. We improved this explanation in the revised manuscript. Although it would be desirable to use MPI-CCDAS on the same grid as TM3, the land-surface model is currently numerically too expensive to allow these higher resolutions.

Explain what is meant by 'each iteration' (L. 139).
The term "each iteration" refers to every cycle when the model re-calculates the cost function for the difference between the model parameters and the observational constraint. In section 2.1 of methods we completed the following paragraph for clarification: "During the optimization procedure, a new model trajectory is determined in each iteration (i.e. in every cycle when the model re-calculates the cost function for the difference between the model parameters and the observational constraint), such that energy and mass are conserved through the entire assimilation window (Kaminski and Mathieu, 2017)."

It is unclear how FAPAR data are generated from equation 2, which seems to generate NDVI estimates. We require details on the FAPAR observation operator – I could not find any.
We apologize for the error in Equation 2. In section 2.2 of the revised ms, this is corrected and now it reads:
"Therefore, we used as FAPAR proxy the Global Inventory Monitoring and Modeling System (GIMMS) NDVI product for the period 1982 to 2006 (Tucker et al., 2005), and merged it with the TIP-FAPAR product to provide a longer record of vegetation

greenness. The maximum and minimum NDVI values were rescaled at the pixel level to coincide with those from the TIP-FAPAR for the overlapping periods (i.e., 2003 to 2006) following:

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

Where x is the period 2003 to 2006 for each data set, NDVI is the full NDVI product from 1982 to 2006, with minimum values given by $NDVI_{min}$ and maximum by $NDVI_{max}$. TIP_{min} and TIP_{max} are the corresponding minimum and maximum values from the TIP-FAPAR product. With this approach, the resulting merged product maintains the maximum and minimum values from TIP-FAPAR while preserving the temporal dynamics of NDVI. The median uncertainty of the available TIP-FAPAR data was considered as the uncertainty for the entire time-series."

Why are there no FAPAR data after 2006? (l. 278)

Unfortunately, this is an error during the assimilation procedure. Due to a technical fault, the CCDAS did not consider the remaining four years of data from the original TIP-FAPAR time-series as planned. We discovered this issue only during the post-processing phase, and we wrote it in the final sentence of the paragraph above. However, we believe that our results are still valid, because the main information gain of the CCDAS in terms of phenology stems from the seasonal cycle, with little effect on the overall trends between the three assimilation experiments with different time periods. This comment is added in section 2.1 of methods in the revised manuscript: "Due to a technical failure in the CCDAS, the final $FAPAR_{mod}$ product spans only from 1982 to 2006 and the last four years from the TIP-FAPAR product were not included."

And in the Discussion: "The technical error during the assimilation procedure to not include the period from 2007-2010 in the $FAPAR_{mod}$ product does not influence however the decadal results observed here, because the main information gain of the CCDAS in terms of phenology stems from the seasonal cycle, with little effect on the overall trends between the three assimilation experiments with different time periods."

What are the implications of not including fire emissions? Why were fire products such as GFED not used to provide this input?

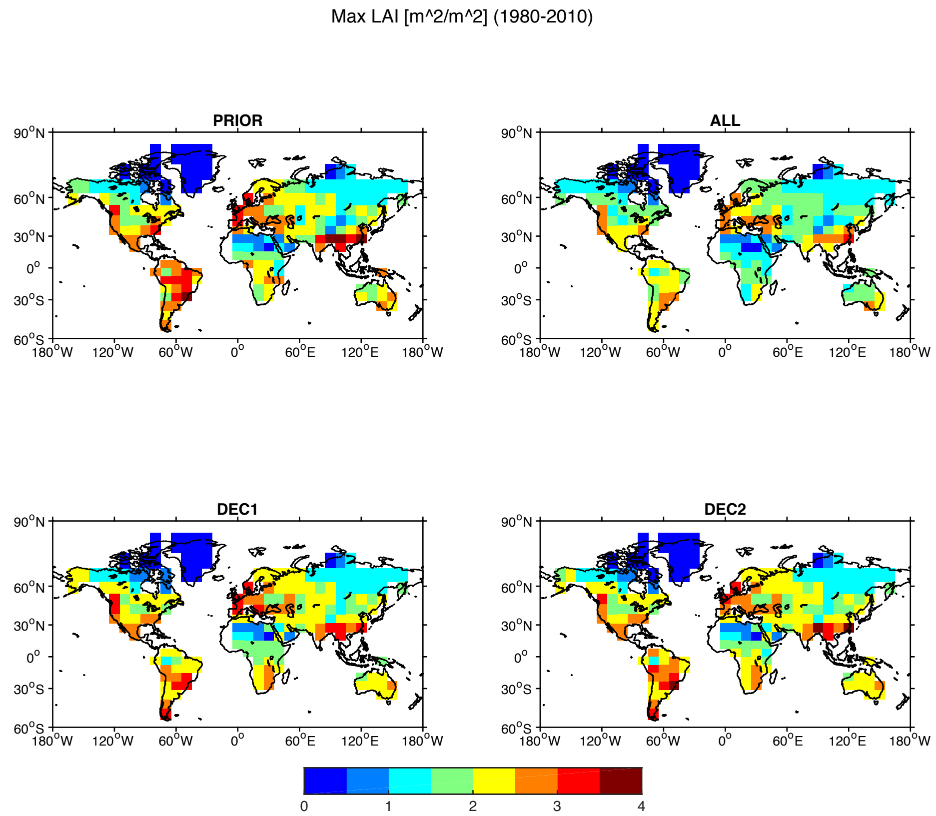
Omitting fire fluxes may impair the ability of the MPI-CCDAS to correctly infer the atmospheric growth rate of CO_2 in years with strong contribution of a fire flux, such as the 1998 El Niño. It was not possible to use GFED-like products in this particular assimilation experiment, because for example GFED4 data do not exist prior to 1997, whereas our inversion started in 1982. Adding them only starting in 1997 would have biased the assimilation procedure. In the discussion part of the revised manuscript, and based on what was already presented in the discussion ms, we completed this information to read: "Omitting fluxes in the current model configuration due to fire events may impair the ability of the model to infer the atmospheric growth rate of CO_2 associated with El Niño events (Frölicher et al., 2011; Frölicher et al., 2013). One way to overcome the IAV mismatch would be to include fire fluxes in the model by prescribing them from, e.g., the Global Fire Emissions Database (GFED, van der Werf et al., 2010), however the latest version of this data set (Version 4.0) is only available for years from 1997 which is a limiting factor for the timeframe of the simulations in this work. However, the contribution of these interannual variations to the overall CO_2 cost function is low in comparison to the signal contained in the seasonal cycle and deviations in the long-term trend, such that the MPI-CCDAS may simply not be sensitive enough to these aggregate system properties like the response of the tropical carbon cycle to El Niño events given the uncertainty in the atmospheric transport and the observational representation error."

I would like to know more about the process to determine which parameters were selected for optimisation? These seems to mostly phenological variables which will link to FAPAR. I would expect to see other parameters related to C turnover, e.g. mortality rates, decomposition rates.
The choices of parameters were done by an extensive parameter sensitivity study with a large set of MPI-CCDAS model parameters for a wide range of biomes. The retained parameters had a strong effect of the simulated carbon and water fluxes as well as in phenology. This selection process was extensively described in Schürmann et al. 2016. In Table 1 of the discussion and revised ms, we listed the parameters selected for the optimization. While the majority are indeed linked to phenology, we considered also parameters linked to photosynthesis and global parameters that control the land carbon turnover. Those are the last four parameters listed in Table 1: for the heterotrophic respiration the temperature sensitivity to respiration (Q_{10}) and a multiplier for initial slow pool (f_{slow}) to account for non-steady state conditions at the beginning of the assimilation; for the autotrophic respiration, the leaf fraction of maintenance respiration, and finally an initial atmospheric carbon concentration.

Results:
What are the substantial changes in tropical LAI (l. 306) – how do these match in situ measurements?
As shown in Table 1 of the revised ms, the maximum LAI value is one of the optimization parameters and it was prescribed for each PFT. In the case of the tropical evergreen and deciduous trees, this equals to 7 $m^2\ m^{-2}$. To provide a numerical context to the LAI changes, we summarize in the table below (added in the revised ms as Table A1 in appendix) the mean and maximum LAI values per regions of Fig. 1 for each experiment.

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

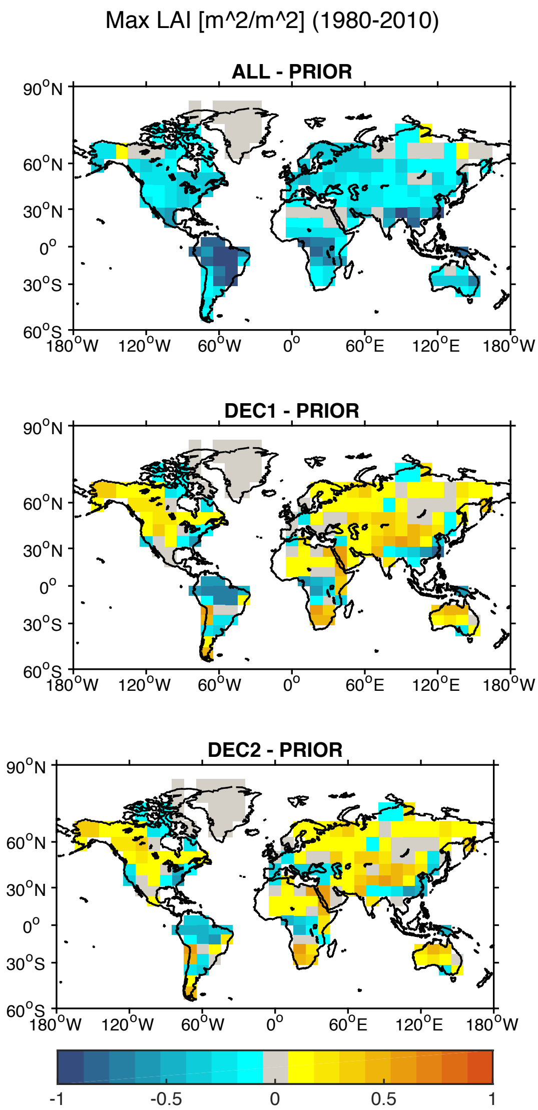
We also show in the figure below, the average maximum LAI for each experiment for the period 1980-2010.



When we compare the LAI mean values between the experiments and the PRIOR results (Fig. 3 in the discussion manuscript), we observe that the largest change in LAI values was in the tropical west area (TW) comprising Brazil, with a decrease in LAI values of up to 24 % in the ALL experiment with respect to the PRIOR, as a response of the maximum LAI decay in the tropical evergreen PFT (visible in the figure above). We argue that this decrease is a response of a global compensating effect to heterotrophic respiration, leading to the lower GPP tropical value.
Ground based observations in the tropical Amazon-Savanna transition forest have been reported with an annual mean LAI value for the total canopy between 2005 and 2008 of 7.4±0.6 m^2 m^{-2}, and for the seasonal flooded forest a value of 3.4±0.1 m^2 m^{-2}. For the remote sensing data from MODIS, the reported values are 6.2±0.2 m^2 m^{-2} and 5.8±0.3 m^2 m^{-2}, respectively (Biudes et al., 2014).
In the eastern Amazon forest, the remote sensing-based LAI has been reported as 6.2 m^2 m^{-2} from LiDAR, and 4.8 m^2 m^{-2} with a low end of 2.0 m^2 m^{-2} from MODIS (Qu et al., 2011). The maximum LAI values from our model results before and after the assimilation (see table above) fall within the values from MODIS and LiDAR. However, this comparison is robust because of the spatial resolution of the different methods: a coarse model grid cell resolution vs. the resolution of ground-based measurements and the resolution of the remote sensing pixels (50x50 m for ground-based and LiDAR data, and 463 x 463 m for MODIS). This discussion is added
The following paragraph is now added with this information in the discussion of the revised ms:
"Bearing in mind the different spatial resolution of methods (i.e., model grids and remote sensing pixels), a robust comparison between the mean and maximum LAI values before and after the assimilation per region are presented in Table A1 of the Appendix. The results fall within LAI values from MODIS and LiDAR reported in

the literature. Ground-based observations in the tropical Amazon-Savanna transition forest between 2005 and 2008 show an annual mean LAI value for the total canopy of 7.4±0.6 m^2 m^{-2} and for the seasonally flooded forest the value of 3.4±0.1 m^2 m^{-2}. For the remote sensing data from MODIS, the reported values are 6.2±0.2 m^2 m^{-2} and 5.8±0.3 m^2 m^{-2}, respectively (Biudes et al., 2014). In the eastern Amazon forest, the remote sensing-based LAI has been reported as 6.2 m^2 m^{-2} from LiDAR, and 4.8 m^2 m^{-2} with a low end of 2.0 m^2 m^{-2} from MODIS (Qu et al., 2011)."

In addition, former Figure 3 was removed and replaced into the supplement by a figure (now Fig. S2 presented below) showing the differences between experiments for the average maximum LAI. This is for the sake of keeping the revised manuscript more focused on the main aim, and because the reference to the LAI results were done only sporadically through the manuscript.



What is the increase in R2 (l, 314) – please report in the text.
The R^2 values between the FAPAR observations and model results are: 0.1638 for PRIOR, 0.1984 for ALL, 0.3412 for DEC1 and 0.3402 for DEC2. The values are given in Lines 318-323 of the discussion ms. In the revised ms, we improved these paragraphs to make clearer that the values are given in those lines by moving the values to the lines above.

The key result seems to be shifts in the timing of CO2 exchanges – it would be interesting to focus more on the shifts in model process representation (parameters) required to allow these changes. I do not find Figure A2 very helpful in this regard.
We appreciate the comment from the reviewer, however we believe that an interpretation of the results at the level of changes in parameters only adds an incomplete picture to the analysis and it is difficult to conclude the effect of the changes. For this reason, we only focus on the relative changes summarized in Fig. A2 (now A3). In the revised manuscript we add a section in the Appendix regarding the assimilation performance where we discuss in more detail the response of each optimization parameter.

The large drop in GPP in the posteriors is significant – what process is this traced to in the parameter adjustment?
This results primarily from a reduction in tropical leaf area index. This is observed in the drop in the photosynthetic capacity (see larger change in now Fig. A3 and formerly A2, for parameter f_{photos} in the tropical evergreen and deciduous PFTs) after the assimilation. Also, this is observed in the new Fig. S2 (as replacement of former Figure 3 in the discussion manuscript) in the revised manuscript and presented above, where are shown the differences between experiments for the maximum LAI. In a new section in the Appendix, we explicitly add these observations as part of the assimilation performance analysis.

I did not find fig 4 helpful in regard to identifying improvements in IAV modelling – it would help to have some statistics to support the statements here (l. 461).
In the revised manuscript, we provide summary statistics of FAPAR which are also presented in Fig. 3: "In the decadal experiments DEC1 and DEC2, the largest error reduction compared to the PRIOR is 19 % for boreal regions, while in the temperate areas this reduction is about 16 %. In the ALL experiment, larger reductions of 21 % on average are obtained in the tropical regions TE and TW". We also moved former Fig. 4 to the supplementary material (now Fig. S4) since it is only referred briefly in the main ms.

Discussion
"the mismatch between observations and model output is small, and thus of little concern". This statement needs to be more rigorous – how is 'small' determined? What is the threshold for concern? The lack of tropical IAV suggests too weak an ENSO response – were any relevant parameters included in the CCDAS that would have allowed identification of drought response? Some of the discussion seems circular – that using FAPAR data in the assimilation improves modelling of FAPAR seasonality.
In the revised ms, we clarify this statement a) to note that this corresponds to the IAV of FAPAR, b) that the observed signal is small compared to seasonal variations, and c) the retrieval error or the FAPAR product, which as a global average corresponds to ±0.2088 (relative units, Schürmann et al. 2016). The assimilation procedure allows changes in the phenology response to water stress (τ_{W}). However, the assimilation

tended to decrease rather than increase the drought sensitivity of tropical phenology given the entire spatially explicit FAPAR time series, and therefore did not allow to capture these excursions (assuming that they are actually driven by drought related changes in LAI). The modified paragraphs in the discussion are:

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

It would be good to clarify the text to the explicit calibration and validation results to strengthen this section. It seems this result has already been identified in an earlier CCDAS publication, so it is not clear what is novel here.
In the revised manuscript, we clarify the text and reduce redundancy to the earlier publication. However, since we are using new data and run a different set-up, we believe that it is important to establish the baseline performance of the CCDAS before looking into the novel results of long-term trends. We condensed this section to the absolute necessary to give more space to the novel results (see also our response to the selection of Figures below).

The authors identify problems "results from the structural dependence of the MPICCDAS on few, globally applicable PFT-level parameters, and challenges in using the spatial mixed signal at the model resolution to infer PFT-specific parameters." It would help to develop these ideas some more – do we expect these issues to specifically affect the current analysis and in what ways?
In the revised manuscript, we expand this discussion. Firstly, this relates to the problem mentioned by the reviewer above with respect to the impact of coarse spatial resolution. We believe that aggregating the remote sensing data into PFT-specific classes per pixel from a high-resolution grid would allow reducing the problems in the identification of phenological parameters. Secondly, although some of the phenological parameters adapt to mean growing season temperature, some of the thresholds are globally applicable, which causes mainly a problem for temperate grasslands, which cover a wide climatological range. Finding appropriate means to cluster grasslands into more spatially refined classes would further reduce the errors of the MPI-CCDAS to simulating boreal, temperate and tropical phenology. Finally, some of the global parameters (such as f_{aut_leaf} and f_{slow}) imply that improvements of modeled fluxes in the boreal regions directly affect fluxes in the tropics, inducing parameter changes to compensate for the altered C fluxes. Such dependency is typical in biosphere models, but may not be ecologically and eco-physiologically correct. Defining these parameters per PFT would reduce such a problem. However, any of these changes would inflate the inverse problem to be solved, therefore increasing

computational costs and would not necessarily reduce overall uncertainty (equifinality).
The new paragraphs in the discussion section reads:
"Although some of the phenological parameters in CCDAS adapt to mean growing season temperature, other thresholds are only globally applicable, causing a trend to temperature grasslands that cover a wide climatological range. For example, some of the global parameters such as faut_leaf and fslow, imply that improvements of modeled fluxes in the boreal regions directly affect fluxes in the tropics, inducing parameter changes to compensate for the altered C fluxes. Defining instead such global parameters per PFT would alleviate this issue but will compromise the computational cost and might not necessarily reduce the overall uncertainty."
And:
"A likely better strategy for constraining PFT-specific parameters would be to resample the highly resolved satellite product to PFT-specific FAPAR classes per pixel before the aggregation into a global grid. This change would allow finding more spatially refined classes and provide PFT-specific FAPAR maps to the CCDAS to reduce issues in the identification of phenological parameters for different climatic regions."

Table 3. The posteriors suggest a Ra:GPP ratio of _65% - it would be useful to discuss this value which seems high for a global estimate.
We believe that this feature is a result of the fact that net primary production itself is not well constrained from the atmospheric record. We suspect that two factors contribute to the low NPP:GPP ratio: i) the observed fast coupling between GPP and both autotrophic and heterotrophic respiration, which cannot be reproduced by a state-of-the-art first-order-decay soil carbon turnover model. Since autotrophic respiration in MPI-CCDAS is directly coupled to GPP, increasing the fraction of GPP partitioned to it increases the seasonal cycle of ecosystem respiration; ii) Increasing Ra reduces the net land carbon uptake, and may mask changes in vegetation carbon turnover, which were excluded from the analysis, because their effect on carbon storage was much lower than that of changing $f_{\mathrm{aut\_leaf}}$. We note that accounting for the vegetation carbon turnover parameters without any further constraint on NPP would likely not have increased the confidence in the CCDAS outcome because of equifinality.

The added paragraph in the discussion of the revised ms reads:
"The NPP:GPP ratio in ALL and DEC2 decreased to 0.35 and 0.31, respectively, when compared to the PRIOR value (0.45). This reduction might be mainly because the NPP is not well constrained from the atmospheric record. Also, the instantaneous coupling between GPP and both autotrophic and heterotrophic respiration (Ra and Rh), cannot yet be reproduced by a state-of-the-art first-order-decay soil carbon turnover model. Because Ra is directly coupled to GPP in MPI-CCDAS, increasing the fraction of GPP partitioned to Ra leads to an increase in the seasonal cycle of the ecosystem respiration. An increase in Ra with respect to the PRIOR (which is only visible in the global average value in DEC2; Table 3), leads to a reduction in the net land carbon uptake, masking the smaller changes in the vegetation turnover."

The very large reduction in soil C stocks from the prior needs further discussion – JSBACH was spun up to steady state for the prior, so I am not clear how the experiments generated 50% drops in this value. How far is the model from steady state with such a reduction in soil C?
The model was spun-up initially until the soil carbon pools reached equilibrium considering pre-industrial forcing. However, this new "initial state" for the model is

not on steady state when considering climate variability, hence to compensate this, the CCDAS creates an artificial sink of C, leading to a reduction in the soil C stocks, in order to reduce the respiration. Unfortunately, this is unavoidable and is rather a model effect to compensate by contemporary climate changes.

We added this discussion and the new paragraph reads:
"The reduction in the soil C pool after the assimilation can be explained due to an unavoidable effect in the model. The MPI-CCDAS was initially spun-up until the soil C pools reached equilibrium considering pre-industrial forcing; however, this new initial state does not consider climate variability. To compensate for this and to reduce the respiration when the MPI-CCDAS is confronted with contemporary changes in the climate, the model creates an artificial C sink that leads to a reduction in the soil C stocks. It is important noting that the JSBACH 3.0 version used in this MPI-CCDAS does not include permafrost processes; therefore, the global soil C stock might still be underestimated."

We hear again that a result here repeats an earlier CCDAs result (l. 660), which reduces the novelty of the analysis.
To investigate the mechanisms that influence the patterns observed in the simulated global GPP or NEP after the assimilation is out of the scope of the presented manuscript. The reference to the work of Schürmann et al. 2016 in this line, as well as in previous others throughout the ms, serves as a point of comparison to previous results with the same model but obtained under a different experimental design (5 years only of assimilation), which also contributes to set the preceding performance of the current set up (see comment above).

Other comments:
Figures – there are too many figures, some of which are of low value, and this distracts from the key message of the paper. Some of the figures in the appendix are referred to several times, so why are they not in the main text (replacing those referenced less).
With the aim of shortening the manuscript and avoid distraction to the main findings, we removed some figures in the revised version. Specifically, we removed those figures that are less referenced in the main manuscript such as: Fig. 2, on the experimental design and mentioned only once in the main text, it is now Fig. A2 of the Appendix. Figures 3 and 4 were moved to the supplement: Fig. 3, showing the spatial distribution of mean LAI before and after the assimilation, but the manuscript does not focus on the specific changes on LAI after the assimilation, instead R^2 values are given in Table 2 and a replacement figure (Fig. S2) shows the differences of the average maximum LAI between experiments. Former Fig. 4 shows the interannual variability of FAPAR for the different sub-regions, and since it is only mentioned briefly in the results it is now moved to the supplement as Fig. S4.
As from figures from the Appendix, former Figures A3 and A4 were mentioned more frequently in the text, hence they are now Fig. 2 and 9 in the main revised text. Former Fig. 8 (showing the comparison of global C fluxes to GCP17 models results) is now in Appendix as Fig. A4. In new Fig. 8 (on the four-years mean atm. CO_2 difference to the observations) and Fig. 9 (RMSE between atm. CO_2 observations and model results), we added the inversion results.
In total, 9 figures are shown in the main text, 4 in the appendix and 7 in the supplement.

L. 59. Citation needed for this 5.6% value
The reference for this value is LeQueré et al., 2018 which is cited in the lines below.

L. 297. The zone between 20-60_ is not well described as "sub-tropical"
We refer in the revised ms to this range of latitudes as north and south temperate zones (TNW and TNE for west and east northern hemisphere, and TSW and STE for west and east southern hemisphere).

Table 1: Add row numbers
Ok, in a first column of Table 1 we added row numbers.

Reviewer #2

This is an interesting and useful paper, albeit of more technical than scientific interest. There are a number of factors that reduce the scientific impact of the paper, while focusing more on the interaction of observations with a model of this type when used in assimilation mode.
The reanalysis is limited by a number of factors, the very low spatial resolution dictated (I suppose) by the resolution of the atmospheric inverse model, the limited data fields assimilated (just carbon fluxes and FAPAR) and the lack of potentially important processes, such as fire.
Reviewer #2 rightly points out the relevant insights and limitations of our study, which some of those were also mentioned by anonymous Reviewer #1. In the revised ms, we improve considerably the focus of our main aim at the same time of discussing the limitations in more detail. Our aim centers in the use and effect of long-term data sets for assimilation and the question on how long the improved model/data agreement can last. We hope that this study will inspire the future use of CCDAS systems to integrate further data streams (such as SIF or VOD), for which a CCDAS is uniquely suited given its ability to use data at different resolutions and for different time-periods. We discuss this also further in the revised manuscript.

The new paragraph in the revised ms with the overarching aim reads: "The overarching aim of this work is to understand the ability of the MPI-CCDAS v1 to make decadal projections of the land C cycle when the assimilation is confronted to different temporal windows from two observational constraints: FAPAR from remote sensing data and atmospheric CO2 concentrations from the global flask measurements network. For this, we present three decades of modeled land carbon fluxes with the MPI-CCDAS and investigate the effect of withholding information from recent decades in the projected carbon fluxes and the ability of the model to reproduce the observations during the period of data assimilation. We also analyze trends and seasonal variations in the simulated signals during the periods of the assimilation and compare to independent results to evaluate the model performance. With these results, we gain insights in the number of observations (in terms of decadal scale) necessary in data assimilation systems to improve the representation of the global terrestrial carbon cycle components. These results open the possibility of including newly measured data in DAS that are only available for periods of less than a decade."

As in our response to reviewer #1, the spatial resolution is indeed dictated by the computational setup. Increasing resolution would of course allow for a better integration of remote sensing data as well as the current sub-grid scale variability in

climate. However, previous studies (Müller and Lucht, 2007; Peylin et al., 2016), have suggested that increased resolution would not necessarily have a strong effect on the overall performance of the model against global carbon cycle observations. In this regard, we added the following paragraph in the 2.1 section of methods of the revised ms: "This horizontal resolution allows computational feasibility and a realistic computational cost for the set of experiments presented in this work. Furthermore, previous evidence has shown that a higher spatial resolution in global vegetation models does not exert a considerable influence in the simulated carbon fluxes at global or regional scales when compared to results obtained with a coarse grid (Müller and Lucht, 2007). The lack of influence to improve the simulated global C fluxes due to changes in the model spatial resolution might also apply to DAS models (Peylin et al., 2016)."
Further, MPI-CCDAS does not include all processes as in any model study. As noted in our response to reviewer #1, using data sets to account for processes such as fluxes due to fires is not possible given the lack of data before 1997. While there is a fire module in MPI-CCDAS (Lasslop et al., 2014), a number of issues have been identified with that module that would need to be addressed before attempting its use into DAS. Also, the effect of these issues on the spatio-temporal dynamics of the land carbon balance would need to be clarified before it is possible to include it into these long-term and computationally expensive MPI-CCDAS simulations. We agree that the addition of disturbance processes due to fires is an interesting aspect for future MPI-CCDAS developments and may contribute to an improved representation of the interannual variability. However, we note that some of the major fluxes (deforestation, peatland fires) are not considered by this, and many other, fire models. In the discussion of the revised manuscript, we included the following paragraph containing the potential implications to our results of not having explicitly included fire emissions in our experiments:
"Notably, the model lacks the representation of some key processes that contribute to climate induced interannual variability of the carbon cycle, such as the possibility to dynamically account for fire disturbance (Lasslop et al., 2014), ENSO related tropical peat-land fires (van der Werf et al., 2008), or the increase of terrestrial carbon uptake after large-scale volcanic eruptions such as for Mt. Pinatubo in 1991 (Lucht et al., 2002; Mercado et al., 2009). Omitting fluxes in the current model configuration due to fire events may impair the ability of the model to infer the atmospheric growth rate of CO_2 associated with El Niño events (Frölicher et al., 2011; Frölicher et al., 2013). One way to overcome the IAV mismatch would be to include fire fluxes in the model by prescribing them from, e.g., the Global Fire Emissions Database (GFED, van der Werf et al., 2010), however the latest version of this data set (Version 4.0) is only available for years from 1997 which is a limiting factor for the timeframe of the simulations in this work. However, the contribution of these interannual variations to the overall CO_2 cost function is low in comparison to the signal contained in the seasonal cycle and deviations in the long-term trend, such that the MPI-CCDAS may simply not be sensitive enough to these aggregate system properties like the response of the tropical carbon cycle to El Niño events given the uncertainty in the atmospheric transport and the observational representation error."

The author's assessment of model skill is ambivalent, they point to low errors in some places, while noting that the El Nino cycle is not well-captured, a time scale that others have argued provides a critical clue to climate sensitivity (eg Cox et al).
In the revised manuscript we target this point in the paragraph above. The core issue of many model-data inter-comparison studies relies in the absolute misfit calculation through cost functions, which limits the representation of these individual climatic

events because they are not specifically weighted in the long-term scales. By including other fluxes such as fires, either prescribed or by adding specific modules into the model, that are indicators of such climatic events, may support its representation.

By contrast, the advanced methods used and the useful assessment of the impact of the duration of the assimilation experiment, as well as other technical innovations provides a useful update to their prior paper, as the scientific conclusions are overlapping. As assimilation becomes more prevalent, and as data records lengthen (for this study, of a 30-year time scale these really are the most relevant global fields) with SIF, radar-constrained biomass, and water variables such as vegetation optical depth becoming available for > 10 years, this paper provides encouraging news about the utility and impact of records of decadal length.
We appreciate the comments and support from the reviewer to this manuscript and for valuing the scientific contribution of our work.
I'd suggest rewriting the paper modestly to emphasize the lessons learned about the impact of assimilation, and the time horizons, and placing less emphasis on the carbon cycle results, especially as the authors note (and correctly) the conclusions broadly overlap their earlier paper. I note that papets of the paper are awkwardly written and could use a careful edit, and there are a lot of figures I found them helpful in reviewing the paper but several of the figures could clearly be moved to supplemental material.
We believe that it is important to demonstrate that the carbon cycle results of a 30 years and a 5 years experiment (as in Schürmann et al. 2016) are broadly comparable to set the stage for the impact of the different time horizons. However, we agree with both reviewers that in the previous version this obscured the key innovation of the study and we therefore revised the manuscript to make this clearer. More concretely, we focus our results and discussions of the time horizons and the evaluation material was shortened with some associated text moved to the appendix (e.g. pixel level FAPAR analysis) or supplementary material.
Specifically, for the figure's changes: we removed those figures that are less referenced in the main manuscript such as: Fig. 2, on the experimental design and mentioned only once in the main text, it is now Fig. A2 of the Appendix. Figures 3 and 4 were moved to the supplement: Fig. 3, showing the spatial distribution of mean LAI before and after the assimilation, but the manuscript does not focus on the specific changes on LAI after the assimilation, instead R^2 values are given in Table 2 and a replacement figure (Fig. S2) shows the differences of the average maximum LAI between experiments. Former Fig. 4 shows the interannual variability of FAPAR for the different sub-regions, and since it is only mentioned briefly in the results it is now moved to the supplement as Fig. S4.
As from figures from the Appendix, former Figures A3 and A4 were mentioned more frequently in the text, hence they are now Fig. 2 and 9 in the main revised text. Former Fig. 8 (showing the comparison of global C fluxes to GCP17 models results) is now in Appendix as Fig. A4. In new Fig. 8 (on the four-years mean atm. CO_2 difference to the observations) and Fig. 9 (RMSE between atm. CO_2 observations and model results), we added the inversion results. In total, 9 figures are shown in the main text, 4 in the appendix and 7 in the supplement.
With the aim of delivering a clearer message in our manuscript, the revised version underwent a thorough English revision before re-submission.

References listed in these responses.

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

Frölicher, T. L., Joos, F., and Raible, C. C.: Sensitivity of atmospheric CO_2 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 CO_2 response to volcanic eruptions: the role of ENSO, season, and variability, Global Biogeochemical Cycles, 27, 239-251, 2013, 10.1002/gbc.20028.

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.

Müller, C. and Lucht, W.: Robustness of terrestrial carbon and water cycle simulations against variations in spatial resolution, Journal of Geophysical Research, 112, D06105, 2007, 10.1029/2006JD007875.

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.

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

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.

van der Werf, G. R., Dempewolf, J., Trigg, S. N., Randerson, J. T., Kasibhatla, P. S., Giglio, L., Murdiyarso, D., Peters, W., Morton, D. C., Collatz, G. J., Dolman, A. J., and DeFries, R. S.: Climate regulation of fire emissions and deforestation in equatorial Asia, Proceedings of the National Academy of Sciences of the United States of America, 105, 20350-20355, 2008, 10.1073/pnas.0803375105.

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

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

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

[1] Max Planck Institute for Biogeochemistry, Jena, Germany

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

[3] Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, Helsinki, Finland

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

Abstract

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

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

Deleted: to

Deleted: ¶

Deleted: This paper presents global land carbon fluxes for the period 1982-2010 (gross primary production, GPP, and net ecosystem exchange, NEE

Formatted: Subscript

Deleted:) estimated with the Max Planck Institute – Carbon Cycle Data Assimilation System (MPI-CCDAS v1). The primary aim of this work is to analyze the performance of the MPI-CCDAS when it is confronted with three different time periods for data assimilation (DA), and thereby to assess its prognostic capability. To this extend we assimilated nearly three decades (1982-2010) of space borne measurements of the fraction of absorbed photosynthetic active radiation (FAPAR) and atmospheric CO_2 concentrations from the global network of flask and in situ measurements. Both data sets were incorporated with different assimilation windows covering the periods 1982-1990, 1990-2000 and 1982-2010. The assimilation results show a considerable improvement in

Deleted: the long-term trend and seasonality of FAPAR

Deleted: , as well as in the long-term trend and seasonal amplitude of the atmospheric CO_2 concentrations when compared to the observations in sites globally distributed

Deleted: Using data from 1982 to 1990 in the assimilation yielded only a difference to the observations of 2±1.3 ppm for the period 15 to 19 years after the end of the assimilation. This suggests that

Deleted: d

Formatted: Subscript

Deleted: increase

Deleted:

Deleted:

Deleted: ¶

Formatted: Wrap Around

1 Introduction

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

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

Deleted: (Le Quéré et al., 2018)
Formatted: English (US)
Deleted: remains
Deleted: ,
Deleted: ocean and land sinks,
Deleted: of 5.6 % (0.6 GtC yr^{-1})
Deleted: Despite substantial
Deleted: in
Deleted: ing
Deleted: over the past decades
Deleted: pose
Deleted: large
Deleted: of
Deleted: these
Deleted: in addition
Deleted: to
Deleted: the
Deleted: of
Deleted: observed
Deleted: trends in
Deleted:
Deleted: model and observations
Deleted: Contrary to the application of a
Deleted: ,
Deleted: these
Deleted: the
Deleted: y
Deleted: to inform
Deleted: the
Deleted: of the terrestrial carbon cycle
Deleted: S
Deleted: for this purpose
Deleted: . Although they rely in different
Deleted: statistical methods (i.e.
Formatted: Wrap Around

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

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

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

Deleted:)

Deleted: , their common characteristic is

Deleted: and

Formatted: Subscript

Formatted: Subscript

Formatted: Wrap Around

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

2 Methods

2.1 MPI-CCDAS

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

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

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

Deleted: In this work, we use the Max Planck Institute - Carbon Cycle Data Assimilation System (MPI-CCDAS v1, Schürmann et al., 2016)

Moved down [3]: The MPI-CCDAS follows a variational approach that iteratively reduces the model-data misfit simultaneously for multiple observational and independent carbon cycle data sets (Kaminski et al., 2013). Since its first development based on the BETHY (Biosphere Energy-Transfer Hydrology) - CCDAS, the MPI-CCDAS has undergone several code modifications and improvements, as well as tests of the assimilation of new observational data sets (e.g. Kaminski et al., 2012; Kaminski et al., 2013; Rayner et al., 2005; Scholze et al., 2016; Schürmann et al., 2016), with the aim of further improving the representation of land carbon fluxes. The history of the MPI-CCDAS and other current DA systems is extensively discussed in Scholze et al. (2017).

Deleted: that has been built around the Jena Scheme Biosphere-Atmosphere Coupling in Hamburg (JSBACH) land-surface model (Dalmonech and Zaehle, 2013; Raddatz et al., 2007; Reick et al., 2013). The MPI-CCDAS follows a variational approach that iteratively reduces the model-data misfit simultaneously for multiple observational and independent carbon cycle data sets (Kaminski et al., 2013). Since its first development based on the BETHY (Biosphere Energy-Transfer Hydrology) - CCDAS, the MPI-CCDAS has undergone several code modifications and improvements, as well as tests of the assimilation of new observational data sets (e.g. Kaminski et al., 2012; Kaminski et al., 2013; Rayner et al., 2005; Scholze et al., 2016; Schürmann et al., 2016), with the aim of further improving the representation of land carbon fluxes. The history of the MPI-CCDAS and other current DA systems is extensively discussed in Scholze et al. (2017).¶
In this paper, we seek to analyze the extent to which the application of a CCDAS leads to the improved representation of the contemporary land carbon cycle and its prognostic capacity for subsequent years. To this extent, we analyze the estimated major components of the terrestrial carbon cycle with the MPI-CCDAS in response to the simultaneous assimilation of three decades of data from two observational constraints: FAPAR from remote sensing data and atmospheric CO_2 concentrations from the global flask ... [1]

Moved (insertion) [3]

Deleted: The MPI-CCDAS

Deleted: iteratively

Deleted: simultaneously

Deleted: observational and

Deleted: (Biosphere Energy-Transfer Hydrology)

Field Code Changed

Deleted: systems

Deleted: a

Deleted: for computational feasibility

Deleted: by

Deleted: , Table 1

Deleted: selection of

Deleted: and range

Deleted: as

Deleted: ; Table 1)

Formatted: Wrap Around

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

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

$$J(\boldsymbol{p}) = \frac{1}{2}(M(\boldsymbol{p}) - \boldsymbol{d})^{T}\boldsymbol{C}_{\mathrm{d}}^{-1}(M(\boldsymbol{p}) - \boldsymbol{d}) + (\boldsymbol{p} - \boldsymbol{p}_{\mathrm{pr}})^{T}\boldsymbol{C}_{\mathrm{pr}}^{-1}(\boldsymbol{p} - \boldsymbol{p}_{\mathrm{pr}}) \qquad (1)$$

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

2.2 Observational data sets

2.2.1 FAPAR

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

Deleted: . The initial state of the parameters was obtained from an independent forward simulation of JSBACH 3.0 (see Sect. 2.3.1).

Deleted: As described in Schürmann et al. (2016),

Deleted:

Deleted: d

Deleted: ,

Deleted: as described by Schürmann et al. (2016),

Deleted: steep

Deleted: FAPAR is

Deleted: the

Deleted: ,

Deleted: thus

Formatted: Wrap Around

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 indented 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). Therefore, we used as FAPAR proxy the Global Inventory Monitoring and Modeling System (GIMMS) NDVI product for the period 1982 to 2006 (Tucker et al., 2005), and merged it with the TIP-FAPAR product to provide a longer record of vegetation greenness. The maximum and minimum NDVI values were rescaled at the pixel level to coincide with those from the TIP-FAPAR for the overlapping periods (i.e., 2003 to 2006) following:

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

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

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

- **Deleted:** t
- **Deleted:** We t
- **Deleted:** merged
- **Deleted:** , available from
- **Deleted:** time
- **Deleted:** →
- **Deleted:** *NDVI*
- **Deleted:** *NDVI*
- **Deleted:** *NDVI*□
- **Deleted:** *NDVI*□
- **Deleted:** *NDVI*□
- **Formatted:** Subscript
- **Formatted:** Subscript
- **Formatted:** Subscript
- **Formatted:** Subscript
- **Deleted:** is the TIP-FAPAR data
- **Deleted:**
- **Formatted:** Subscript
- **Deleted:** separately
- **Deleted:** .
- **Deleted:** As in Schürmann et al. (2016), w
- **Deleted:** with the aim of
- **Deleted:** selecting
- **Deleted:** useful pixels in the FAPAR global grid. The selection of pixels to be removed from the global grid followed …
- **Deleted:** , we masked out the grid cells with a crop-dominated phenology of > 20 %
- **Deleted:** important
- **Deleted:** ;
- **Formatted:** Wrap Around

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

2.2.2 Atmospheric CO_2 concentrations and observation operator

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

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

Deleted: ; and

Deleted: finally,

Deleted: a large

Deleted: only

Deleted: initial

Formatted: Subscript

Formatted: Subscript

Deleted: The atmospheric transport of CO_2 is calculated in MPI-CCDAS through the Jacobian representation

Deleted: atmospheric transport model

Deleted: by

Deleted: is run

Deleted: (Heimann and Körner, 2003; Rödenbeck et al., 2003)…

Deleted: During the generation of this matrix representation, the precise sampling time of flask measurements was considered to minimize the representation error due to short-term fluctuations in atmospheric transport.

Deleted: ,

Deleted: as in Schürmann et al. (2016),

Formatted: Wrap Around

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

2.2.3 Background carbon fluxes

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

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

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

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

Deleted:

Deleted: (Hurtt et al., 2006)

Deleted: with the aim

Deleted:

Deleted: (Andres et al., 2013; Andres et al., 2011; Andres et al., 1996)

Formatted: Wrap Around

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

2.3 Experimental setup

2.3.1 Spin up and preparation of initial files

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

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

Deleted: annual

Deleted: . The CRU-NCEP v6.1 reanalysis data is

Deleted: with

Deleted: se

Deleted: In addition, p

Deleted: annual

Deleted: Prior to

Deleted:

Deleted: also

Deleted: the forward

Deleted: model

Deleted: JSBACH

Deleted: in order

Formatted: Wrap Around

2.3.2 Assimilation experiments

During the assimilation procedure, the model was forced with the same daily reanalysis atmospheric data used during the model spin up. In this study we present the results of three long-term experiments using the MPI-CCDAS, which differ in the timeframe of the observational records used during the assimilation: 1) ALL, covers data in 1980-2010 and includes the complete available timeframe of the two observational data sets, i.e., for FAPAR is from 1982 to 2006 and for the atmospheric CO_2 concentrations from 1982 to 2010; 2) DEC1, covers observations from the two data sets available from 1982 to 1990; and 3) DEC2, covers measurements available from the two data sets from 1990 to 2000 (Fig. A2). Because of the different lengths of the CO_2 records for some stations, this ultimately leads to a different number of observations used for each experiment (Fig. A1).

The simulation period in the three assimilation experiments is from 1970 to 2010. The first ten years (1970 to 1979) of the results are discarded because during this period the phenology, vegetation productivity, and the fast land C pools adjust to the new model control vector ***p***. Through this adjustment any imprint of the initial conditions on the calculation of the cost function is avoided. The soil C pool at the beginning of the experiment was included in the model control vector. and only results from 1980 are reported below. The results of the assimilation for the periods of time that fall within the observational temporal window are considered for model diagnostic, whereas the periods that fall outside the assimilation window on each experiment are periods of model prognosis, i.e., the prognosis period in DEC1 is from 1991 to 2010, and in DEC2 for 2001 to 2010.

3 Results

We first evaluate the long-term trends, seasonal and spatial variability of the FAPAR and carbon fluxes from the different assimilation experiments (Section 3.1 to 3.3), and based on these analyze the prognostic ability of the MPI-CCDAS (Section 3.4). To facilitate the analysis in some of our results, the global land is divided into eight regions: Boreal West and East (BW and BE, for latitudes north of 60° N), temperate Northwest and Northeast (TNW and TNE, between latitudes 20° N and 60° N); tropical West and East (TW and TE, between latitudes 20° N and 20° S); temperate Southwest and Southeast (TSW and TSE, for latitudes south of 20° S) (Fig. 1).

3.1 Phenology

In all assimilation experiments, the RMSE and the bias between the modeled and observed FAPAR for 1982 to 2006 is reduced compared to the PRIOR (Table 2). One

Moved down [1]: The simulation period is from 1970 to 2010 for all the experiments.
Deleted: only
Deleted: experiment
Deleted: is
Deleted: observations
Deleted: al data
Deleted: 2
Moved (insertion) [1]
Deleted: for all the experiments
Deleted: In all of the experiments the first ten years of the simulation (1970 to 1979)
Deleted: are to allow
Deleted: to
Deleted: and avoid
Deleted: Thus, the initial period is discarded
Deleted: s
Deleted: of the observational constraints
Deleted: thus
Deleted: ¶
Formatted: Normal, Justified
Deleted: 2
Deleted: Mean seasonal p
Deleted: We analyzed the global distribution of FAPAR before and after the assimilation against the observations. To facilitate the analysis, we also divided the global land into eight regions: 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) (Fig. 1). T
Deleted: normalized
Deleted: (NRMSE = RMSE / mean($FAPAR_{obs}$))
Deleted: are somewhat reduced by
Deleted: all assimilation experiments compared to the PRIOR …
Formatted: Wrap Around

important cause for this improvement is the change in the spatial distribution of the yearly maximum leaf area index (LAI) due to the optimization of the PFT-specific maximum LAI (Λ_{max}) parameter (Fig. S2) (see also section A1 and Fig. A3 in the Appendix for more specific results of parameters changes due to the assimilation). The improvement occurs in all regions (Fig. 2), despite notable differences between the different assimilation experiments. In the decadal experiments DEC1 and DEC2, the largest error reduction compared to the PRIOR is 19 % for boreal regions, while in the temperate areas this reduction is about 16 %. In the ALL experiment, larger reductions of 21 % on average are obtained in the tropical regions TE and TW.

One important factor in the error reduction is a substantial increase in the linear global correlation (R^2) in FAPAR during spring and autumn in experiments DEC1 (0.42 and 0.48, respectively) and DEC2 (0.48 and 0.47, respectively) with respect to the PRIOR (0.31 and 0.33, respectively), with changes mostly taking place in the Northern Hemisphere (Fig. S3). An analysis for representative pixels (Fig. 1) shows that the assimilation procedure results in a better representation of the timing and amplitude of the mean seasonal cycle, particularly in the temperate and boreal zones of the Northern Hemisphere (Fig. S4). As a result, the average global R^2 between modeled and observed FAPAR increased with respect to the PRIOR experiment from 0.17 in the PRIOR to 0.20 for ALL and 0.34 for both DEC1 and DEC2 (Table 2, Fig. S3). Further details on the pixel level analysis are presented in section A2 of the Appendix.

The observed FAPAR signal exhibits positive long-trends, indicating a greening trend of vegetation for most of the regions, with the exception of the TSW region, where the long-term trend indicates a decrease of FAPAR (i.e., browning). In most of the regions, the assimilation the assimilation results agree on a positive long-term trend as in the observations, the magnitude of this trend is in disagreement to the observations (Fig. 3). Particularly in the BE region, the PRIOR experiment overestimates the $FAPAR_{obs}$ trend by almost double. After the assimilation, the simulated FAPAR trend is reduced leading instead to a slight underestimation of the growth rate in all of the posterior experiments. In the TWS region, the assimilation improved the long-term trend from a positive to a negative growth rate in the three posterior experiments. The most substantial disagreement between $FAPAR_{obs}$ and $FAPAR_{mod}$ occurs in the TW region, where the observations show a positive trend in FAPAR during the period of analysis, whereas this is not captured in the PRIOR and all the posterior experiments. Despite these trend adjustments, the model-data error (based on the four-years mean differences

Deleted: decreased model-data misfit

Deleted: , primarily caused by

Deleted: Fig. A2 in the appendix

Deleted: after

Formatted: Font: (Default) Times New Roman

Formatted: Subscript

Formatted: Wrap Around

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

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

3.2 Atmospheric CO_2

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

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

Moved down [2]: However, it is important to note that the fit remains far from perfect, likely owing to model structural errors in the way that the meteorological triggers of phenological events adjust to local climatic conditions.

Deleted: Compared to the PRIOR experiment, the assimilation leads to substantial changes in the LAI of the tropical forest area, with general reductions of LAI in all three assimilation experiments. There is less agreement for the extra-tropical areas, with the ALL experiment suggesting small reductions in LAI, whereas the experiments DEC1 and DEC2 see slight increases (Fig. 3, left panels) relative to the PRIOR experiment. ¶
The second reason for the reduced misfit is an improved representation of the FAPAR interannual variability (IAV) at a regional scale (Fig. 4) and seasonal cycle at the pixel level (Fig. S2), particularly in the temperate and boreal zones of the Northern Hemisphere. This is also evidenced by the increase in linear correlation coefficient R^2 between modeled and observed FAPAR with respect to the PRIOR experiment (Fig. 3, right panels). However, it is important to note that the fit remains far from perfect, likely owing to model structural errors in the way that the meteorological triggers of phenological events adjust to local climatic conditions. The average global correlation between model and observations increased moderately in all the assimilation experiments compared to the PRIOR experiment (Table 2 and Fig. 3). This is particularly true for the DEC1 and DEC2 experiments (R^2=0.34 in both experiments) than in the experiment that includes all the window of assimilation, i.e. ALL (R^2=0.20). The improved correlation is primarily the consequence of an increased ability of the model to simulate the timing of green-up and brown-down, and its IAV at regional scale (Fig. 4). Interestingly, the model fit is better if the model is only subjected to 10 years of data, instead of exposing it to the entire time series. However, it is important to note that the fit remains far from perfect, likely owing to model structural errors in the way that the meteorological triggers of phenological events adjust to local climatic conditions.¶
To further analyze the effect of the assimilation procedure on the simulated seasonality and monthly growth rate of the FAPAR, we also selected six pixels that are distributed in locations characterized by a dominant PFT (see Fig. 1 for the geographic location of the pixels). A clear improvement after the assimilation is in ¶ ... [2]

Formatted: Highlight

Moved (insertion) [2]

Deleted: We next…o analyze …iagnose the performance of the MPI-CCDAS with respect to the atmospheric mole fractions of CO_2,…As example, …e compare the observed and simulated CO_2 mole fractions…alues, in terms of the mean seasonal cycle, IAV and monthly growth rate, inat…three stations: 1) Alert (ALT) at the Northern Hemisphere (Alert, ALT)… 2) Mauna Loa (MLO) at the ... [3]

Deleted: 6 …demonstrates that these examples are broadly representative of the global changes due to the assimilation. Fig.…66… shows a reduction that ... [4]

Formatted: Subscript

Deleted: located …fin…the Northern Hemisphere (> 40 °N) after the assimilation, whichis reduced after the assimilation,…and …s in closer …etter agreement to the observations than in …he PRIOR simulation. The most substantial largest…amplitude reduction occurs took place the Northernmost …orthernmost Station …tation (ALT), where the seasonal amplitude decreasesd…from 23.5 ppm in ... [5]

Formatted: Wrap Around

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

3.3 Global and regional carbon pools and fluxes

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

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

Deleted: reducing
Deleted: amplitude stronger
Deleted: than
Deleted: ed one
Deleted: INV
Deleted: to
Deleted: in the statistical analysis
Deleted: 5
Deleted: 6
Formatted: Font: (Default) Times New Roman
Deleted: W
Deleted: next analyzed the spatiotemporal changes of the
Deleted: surface gross and net carbon fluxes in the posterior experiments relative to
Deleted: compared
Field Code Changed
Deleted: 7
Deleted: 5
Formatted: Wrap Around

for the same period of analysis. The global mean GPP is reduced by up to 26 PgC yr^{-1} on average in the three posterior experiments compared to the PRIOR experiment. Simulation DEC1 experienced the largest reduction in the global photosynthetic C uptake (83.1 PgC yr^{-1}) relative to the PRIOR value (Table 3 and spatial distribution of the GPP difference to the PRIOR in Fig. S6d, f, and h).

At large-scale, the variation of the NBE (net biome exchange of CO_2 with the atmosphere, calculated as the Net Ecosystem Exchange (NEE) minus the flux related to land use change) 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 significant anomalies collocated in time (Fig. 7a, Fig. A4). Contrary to the PRIOR simulation, the total annual NBE from the three posterior experiments falls within the uncertainty (shadow green area in Fig. A4d calculated as ±1 standard deviation) of the mean NBE from the terrestrial ecosystem models in the GCP17. However, the 1982-2010 mean net biome exchange in all of the assimilation experiments through the time series is on average 1.4 PgC yr^{-1} lower than the flux in the PRIOR simulation (−2.06 PgC yr^{-1}) and 0.8 PgC yr^{-1} less than the GCP17 value (−1.27±0.97 PgC yr^{-1}) (Table 3, Fig. A4d and Fig. S7 for summary of C balance).

In all MPI-CCDAS simulations, the NEE is reduced relative to the PRIOR in most of the Southern Hemisphere, while it is increased in the Northern Hemisphere (Fig. S6c, e, and g). Temperate northern areas contribute the most to the global net CO_2 uptake. In the boreal east and west regions (BE and BW), the net land C emissions increased in all of the posterior experiments compared to the PRIOR (Fig. S6c, e and g) with the most significant increase in BE for DEC2 (−0.29 PgC yr^{-1}) relative to the corresponding value in the PRIOR (−0.09 PgC yr^{-1}). The decrease in GPP in the tropics is depicted in the latitudinal gradient of NBE shown in Fig. 7c and in the spatial distribution of the NEE difference between the PRIOR and the posterior experiments (Fig. S6c, e, and g). As in the tropics, the NEE in the southern temperate region is consistently reduced after the assimilation experiments, also switching the NEE of the TSE region from a C sink of −0.18 PgC yr^{-1} in the PRIOR to a mean C source to the atmosphere of 0.016 PgC yr^{-1} in the DEC2 experiment.

The magnitude of the global NBE and GPP is smaller in the posterior experiments than in the PRIOR. However, the trend in the anomaly of these fluxes (calculated relative to the temporal mean of each time series) is comparable in all the experiments (Fig. 7a and b), suggesting that the response to the environmental conditions is similar through the simulation period also after the assimilation. This robust response shows, e.g., in

Deleted: , but s…mulation DEC1 experienced the largest reduction in the global photosynthetic C uptake (mean global GPP of …3.12.9…PgC yr^{-1}) relative to the PRIOR value (Table 3 and spatial distribution of the GPP difference to the PRIOR in Fig. S65 ... [6]

Deleted: referred to …s the Net Ecosystem Exchange (, NEE) minusplus…the flux related to 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 …ignificant anomalies collocated in time (Fig. 77…, Fig. A4). A comparison of the fluxes from the ocean and fossil fuels from this data set to the corresponding fluxes that are prescribed in CCDAS is shown in Fig. 8. …ontrary to the PRIOR simulation, T…he total annual NBE from the three posterior experiments falls within the spread …ncertainty (shadow green area in Fig. A48… calculated as ±1 standard deviation) of the mean NBE mean of…rom the terrestrial ecosystem models in the GCP17, contrary to the PRIOR simulation… However, the 1982-2010 mean net biome exchange in all of the assimilation experiments through the time series is on average 1.6 … PgC yr^{-1} lower than the flux in the PRIOR simulation (−2.067…PgC yr^{-1}) and 0.86…PgC yr^{-1} lessower…than the GCP17 value (−1.23…7±0.98 …7 PgC yr^{-1}) (Table 3, Fig. A48… and Fig. S74 ... [7]

Formatted: Font color: Auto

Deleted: net land-atmosphere C exchange …EE is reduced relative to the PRIOR result …n most of the Southern Hemisphere, while NEE …t is increased in the Northern Hemisphere (Fig. S65…, e, and g). The analysis per regions illustrates that the extra-tropical…mperate northern areas contribute the most to the global net CO_2 uptakeflux… The increase in respiration (more CO_2 emissions to the atmosphere) in the tropics is clearly depicted in the latitudinal gradient of NBE shown in Fig. 7c and in the spatial distribution of the NEE difference between the PRIOR and the posterior experiments (Fig. S5c, e and g). As in the tropics, the NEE in the southern subtropical regions was consistently reduced after the assimilation experiments, also switching the NEE of the STSE region from a C sink of −0.18 PgC yr^{-1} in the PRIOR to a mean C source to the atmosphere of 0.016 PgC yr^{-1} in the DEC2 posterior experiment.¶
In the boreal east and west regions (BE and BW), the net land C emissions increased in all of the posterior experiments relative to…ompared to the PRIOR (Fig. S65…, e and g) with the largest increase ... [8]

Deleted: Although …t…e magnitude of the global NBE and GPP is smaller in the posterior experiments than in the prior…RIOR. However, the similar slope…rend in the anomaly of these fluxes (calculated relative to the temporal mean of each time series) is comparable in all the detected between the prior and posterior …xperiments in the anomaly of these fluxes (calculated relative to the temporal mean of each time series) …Fig. 77… and b), suggestings…that the response to the environmental conditions remains the same…s similar through the simulation period also even... [9]

Formatted: Wrap Around

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).

3.4 Prognostic capability of MPI-CCDAS

Finally, we evaluate the goodness of the model-data fit of the decadal assimilation runs with respect to their long-term carbon cycle simulation relative to: i) that of the prior and ii) that of the assimilation run, using data from the 30 years-experiment as a reference case for “best possible” model-data match given the structural limitations of the MPI-CCDAS to match the observations (as evaluated in Sections 3.1 and 3.2). We calculate the four-years mean differences between the atmospheric CO_2 mole fraction measurements and the CO_2 model results and also the INV results, for all of the stations (Fig. 8). In the ALL assimilation experiment, the atmospheric CO_2 concentration consistently matches the observations across the entire assimilation period (that also corresponds to the window of assimilation) with a −0.03±1 ppm average bias to the observations (Fig. 8). This is comparable to the trend (Fig. 6c), and four-years mean differences inferred by the inversions, where the four-years mean results in the ALL fall within the standard deviation of the four-years mean of the INV (Fig. 8). This is in striking contrast to the PRIOR experiment, where the four-years mean of the CO_2 mole fraction at the end of this simulation is 18.8 ppm lower than observed. For the DEC1 experiment, the four-year mean difference among the measurements and the model results is between −0.3 and 0.3 ppm in the 1980s. This level of model-data agreement remains for the 1990s, where the experiment did not see any observations. After the year 2000, the fit increasingly degrades, with an underestimation of the CO_2 mole fraction by 1.6 ppm for the last four-years average. However, this is still a 90 % reduction in misfit compared to the PRIOR experiment.

We next quantify the RMSE between the CO_2 measurements and model results for each station for four different periods: 1982-1990, 1990-2000, 2000-2010 and 1982-2010 (Fig. 9 and Fig. A2). The large bias of the PRIOR is reflected in the RMSE where the results of this experiment have the most substantial error in all of the stations and periods (between 2.8 and 18.7 ppm) (Fig. 9). For the posterior experiments with a decadal window of assimilation (DEC1 and DEC2), the performance of the assimilation of CO_2 mole fraction also improves substantially across all time periods. Within the period of the assimilation, the difference to the measurements and RMSE is most strongly reduced, and the error increases somewhat outside of the window of assimilation. The model results show that when only the first decade of data is assimilated (DEC1), a more significant deviation to the long-term trend of atmospheric

Deleted: 7

Deleted: 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.

Deleted: Finally, …e analyze the prognostic capability of CCDAS by comparing …valuate the goodness of the model-data fit of the decadal assimilation runs with respect to their long-term carbon cycle simulation relative to: i) that of the prior and ii) that of the assimilation runs…using all …ata from the 30 years-experiment as a reference case for “best possible” model-data match given the structural limitations of the MPI-CCDAS to match the observations (as evaluated in Sections 3.1 and 3.2). To achieve this,… w…e calculated…the four-years mean differences between the atmospheric CO_2 mole fraction measurements and the CO_2 model results and also the INV results, for all of the stations (for the period 1982-2010)…(Fig. 810….,…and also between the FAPAR satellite data and the monthly FAPAR model results (for the period 1982-2006) (Fig. 11). We also calculated the RMSE between the CO_2 measurements and model results for each station for four different periods: 1982-1990, 1990-2000, 2000-2010 and 1982-2010 (Fig. A4). The choice of a four-year window was made because with Fig. 5 it was established that the capacity of the MPI-CCDAS to improve the representation of observed interannual variability was very moderate.¶
In the ALL assimilation experiment, the atmospheric CO_2 concentration is …onsistently matchesd…the observations across the entire assimilation period (that also corresponds to the window of assimilation) with a −0.03±1 ppm average bias to the observations (Fig. 810…. This is comparable to the trend (Fig. 6c), and four-years mean differences inferred by the inversions, where the four-years mean results in the ALL fall within the standard deviation of the four-years mean of the INV (Fig. 8). This is in striking contrast to the PRIOR experiment, which fails in all of the stations to reproduce the long-term trend (as discussed earlier). …here tT…e four-years mean of the CO_2 mole fraction at the end of this simulation is 18.8 ppm lower than observed. For the DEC1 experiment, the four-year mean difference among the measurements and the model results is between −0.3 and 0.3 ppm in the 1980’…. This level of model-data agreement , a level at which it …emains for the 1990s, where the did not see any observations. A, but the fit increasingly degrades a…ter the year year …000, the fit increasingly degrades, with an underestimatione…of the CO_2 mole fraction by 1.6 ppm for the last on a …our-years average. However, this is (…till a 90 % reduction in misfit compared to the PRIOR experiment) ... [10]

Deleted: 3… and , Appendix …ig. A2). The large bias of the PRIOR is also recognize…eflected in the RMSE results where the PRIOR …esults of this experiment have the most substantiallargest…error in all of the stations and periods (between 2.8 and 18.7 ppm) (Fig. A…4….¶
As f…or the posterior experiments with a decadal window of assimilation (DEC1 and DEC2), the performance of the assimilation of CO_2 mole fraction also improves substantially across all time periods. W, and mostly during …thin the period of the window of …ssimilation, . During those periods, …he difference to the measurements and RMSE is reduced…ost strongly reduced, whereas …nd the error increases during the periods of time outside…omewhat outside of the window of assimilation. , and the RMSE is also higher than in DEC2 and ALL for the period 2000-2010 (Fig. ... [11]

Formatted: Wrap Around

CO_2 occurs between 2000 and 2010 compared to DEC2 and ALL (Fig. 9c). Similarly, a larger bias is also observed in the results from DEC2 where the lowest four-years mean difference between the observations and the assimilation results takes place in the period of the window of assimilation for this experiment (1990-2000) (Fig. 8 and Fig. 9b for RMSE). During this period, the model overestimates the CO_2 atmospheric concentration only by 0.15 ppm on average whereas, for the periods outside the window of assimilation, the CO_2 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.

The analysis of the four-year mean differences for the period 1982-2006 between $FAPAR_{obs}$ and the results of the PRIOR and assimilation experiments at the regional scale (areas in Fig. 1) reveals, contrary to the CO_2 observations, a near constant four-years mean FAPAR difference within the time series and each of the experiments (Fig. 4). In general terms, the decadal experiments are better able to reproduce the mean FAPAR across all regions. The largest difference between posterior results to the observations is in the tropical regions, where the FAPAR four-years mean difference showed that the observations remained consistently larger than the ALL results by on average 0.042 in TE and 0.095 in TW (Fig. 4). Importantly, however, the trend correction for the boreal and temperate areas (Fig. 3) are similar across the different assimilation experiments, suggesting that important biases of the JSBACH 3.0 model, including the tendency to simulate too strong boreal greening, can be readily reduced with only 10 years of data, as the further improvement with the 30 years record is small.

4 Discussion

The parameter optimization with a simultaneous assimilation of long-term spaceborne FAPAR and atmospheric CO_2 measurements in the MPI-CCDAS, resulted in a considerable reduction in the cost function and norm of the gradient, which can be seen as an overall improvement in the modeled global carbon fluxes with a decrease in the root mean squared error of the MPI-CCDAS compared to the CO_2 and FAPAR and observations (Fig. 9 and Fig. 2). The trajectory of model parameters involved in the optimization differed for each experiment and each phenotype. While some parameters were consistently retrieved after the assimilation, such as the maximum leaf area of grasses and shrubs and the correction parameter for the initial soil pool size, some final parameter estimates varied considerably between the three experiments, e.g., the tropical maximum leaf area index and some of the parameters controlling the

Deleted: is identified after

Deleted: This

Deleted: identified

Deleted: 10

Deleted: A

Deleted: 4

Deleted: of time

Deleted: (

Deleted:)

Deleted: (

Deleted:)

Deleted: In the ALL assimilation experiment, the atmospheric CO_2 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 CO_2 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 t

Deleted: the satellite

Formatted: Subscript

Deleted: (Fig. 11).

Formatted: Subscript

Deleted: In this case, the prognostic skill of CCDAS for the periods outside the windows of assimilation is less evident, with a consistent

Deleted: between

Deleted: ¶

Deleted: pixels P1, P2, and P6, where the magnitude of the mean seasonal cycle is better represented when compared to the observations (Fig. S2). Also, the timing of the mean seasonal cycle is corrected e.g. in pixels with large seasonal amplitude such as in P1 (located in Eastern Siberia) and in P6 (located in Canada). While in the PRIOR experiment (and ALL experiment) the onset and peak of the growing season in P1 and P6 are delayed by up to two months, in the results from experiments DEC1 and DEC2 this delay is reduced to only one month. This correction might be partially due to changes in some optimized parameters: increase in the day length at leaf shedding (t_c) and reduction in the temperature at leaf onset $T\phi$ detected for both the CD and CE phenotypes (as well as for ETD and TeCr) (Fig. A2); this is because these parameters control the onset and end of the vegetation activity. This temporal shift however, is less evident in other pixels such as in P2, despite changes in $T\phi$ and t_c after the assimilation in TrH, and this is because the amplitude of the seasonal cycle is small and only changes in the magnitude of the amplitude are evident (Fig. S2).¶ ... [12]

Formatted: Heading 1, Line spacing: 1.5 lines

Formatted: Subscript

Deleted: ¶

Formatted: Wrap Around

seasonality of the phenology (Fig. A3). These variations lead to regional differences in the simulated compartment fluxes GPP and ecosystem respiration, which are not well constrained from the observations. Interestingly, these differences result in very similar absolute values in global carbon fluxes and their trends. This demonstrates a certain degree of equifinality in the results and cautions a too stringent interpretation of the MPI-CCDAS outcome in terms of improving understanding about biosphere processes and their long-term trends.

4.1 FAPAR

MPI-CCDAS is capable of extracting information about the seasonal cycle and the long-term trends from the FAPAR observations. Using decade-long FAPAR data during the assimilation (DEC1 and DEC2), already leads to notable improvement of the simulated seasonal phenology of the land surface within and outside the window of assimilation, i.e., maintaining these changes during the prognosis periods. This improvement is predominantly the result of the ability in the model to simulate the timing of green-up and brown-down in spring and summer through the optimization of parameters that regulate the onset and end of the growing season (i.e., parameters for temperature and day-length thresholds). The greening effect is especially taking place in the Northern Hemisphere, dominated by the phenotypes deciduous and evergreen needle leaf and extra-tropical deciduous trees.

The long-term greening trend in the vegetation of boreal regions previously observed in spaceborne data (Forkel et al., 2016; Lucht et al., 2002), was captured in the results of MPI-CCDAS before the assimilation, but it was mostly overestimated in northern regions and underestimated in the Southern Hemisphere. After the assimilation experiments, the greening trend was improved primarily in the boreal areas and is in closer agreement to the reported satellite FAPAR data. The modest improvements achieved in the simulated greening trend of temperate areas in the western hemisphere are associated with a decreased performance in the eastern hemisphere, indicating that the model structure of MPI-CCDAS is incapable of reconciling regional differences. This difference could be an indicator of the need to parameterize both hemispheres differently in terms of their phenological response to the underlying driving factors (such as temperature, moisture availability and day-length); also, this could be due to the lack of process to account for the land-use or vegetation dynamics in the MPI-CCDAS.

Despite these broad-scale improvements, the MPI-CCDAS does not reproduce the magnitude of the greening trend and its interannual variability in all the posterior

Deleted: The simultaneous assimilation of long-term space borne FAPAR and atmospheric CO_2 measurements in the MPI-CCDAS leads to an overall improvement in the modeled global carbon fluxes (as summarized in Fig. A3 and A4). The

Deleted: 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. 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.
The use of decade-long FAPAR data (DEC1 and DEC2) already leads to notable improvement of the simulated ... [13]

Moved up [4]: The lacking ability of the MPI-CCDAS to

Moved (insertion) [4]

Deleted: ,

Deleted: especially in

Deleted: has been

Deleted: in analysis of

Deleted:

Deleted: . While this enhanced vegetation greening

Deleted: the

Deleted: model

Deleted: already

Deleted: At regional scale, the assimilation in all of the ... [14]

Deleted: This was mostly achieved in boreal regions.

Deleted: However,

Deleted: t

Deleted: rate

Deleted: in

Deleted: regions

Deleted: of

Deleted: It is unclear whether this is an indication

Deleted: these

Deleted: ,

Deleted: or

Deleted: whether

Deleted: processes

Deleted: not considered by

Deleted: are the reason for this mismatch

Formatted: Wrap Around

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

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

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

Deleted: the MPI-CCDAS does not necessarily reproduce the magnitude of the greening trend and its interannual variability in all the posterior experiments

Deleted: , which results from the structural dependence

Deleted: ,

Deleted: applicable

Deleted: , and challenges in using the spatial mixed signal at the model resolution to infer PFT-specific parameters. A likely better strategy for constraining these PFT-specific parameters would be to resample the highly resolved satellite product to PFT-specific FAPAR maps prior to aggregation, and provide PFT-specific FAPAR maps to the CCDAS.

Formatted: Font: Italic

Formatted: Subscript

Formatted: Font: Italic

Formatted: Subscript

Formatted: Subscript

Formatted: Wrap Around

stems from the seasonal cycle, with little effect on the overall trends between the three assimilation experiments with different time periods.

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

4.2 Atmospheric CO_2

The considerable improvement of the seasonal amplitude and the long-term trend of atmospheric CO_2 at Northern Hemisphere stations 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 detected in the measured atmospheric CO_2 stations, which translates into an acceptable, but far from perfect fit to the inferred annual carbon budget of the GCP17 (Le Quéré et al., 2018). We compared the performance to the results from an atmospheric CO_2 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 mentioned that both the choice of the atmospheric transport model (and associated imperfections at resolving the vertical and lateral atmospheric transport of CO_2), 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 forecast of the interannual variability. As a consequence, only the occurrence of more significant model-data mismatches can be interpreted as an actual result of the MPI-CCDAS' inability to correctly resolve the carbon flux variation.

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

Formatted: Superscript
Formatted: Superscript
Deleted: ¶
Formatted: Font: Bold
Formatted: Font: Bold
Formatted: Font: Bold, Subscript
Formatted: Font: Bold
Deleted: Our results also demonstrate that
Deleted: and of its seasonal amplitude in the
Deleted: and at
Deleted: level
Deleted: considerably improved. This is
Deleted: of
Deleted: global carbon projec
Deleted: t
Deleted: at this time-scale
Deleted: borne in mind
Deleted: estimate
Deleted: larger
Deleted: is of concern and a
Deleted: genuine
Deleted: Particularly
Deleted: is not able to capture large-scale relevant climatic disturbances that influence the
Deleted: like fires
Deleted: decrease
Deleted: in atmospheric CO_2 growth after explosive
Formatted: Wrap Around

Mercado et al., 2009). Omitting fluxes in the current model configuration due to fire events may impair the ability of the model to infer the atmospheric growth rate of CO_2 associated with El Niño events (Frölicher et al., 2011; Frölicher et al., 2013). One way to overcome the IAV mismatch would be to include fire fluxes in the model by prescribing them from, e.g., the Global Fire Emissions Database (GFED, van der Werf et al., 2010), however the latest version of this data set (Version 4.0) is only available for years from 1997 which is a limiting factor for the timeframe of the simulations in this work. However, the contribution of these interannual variations to the overall CO_2 cost function is low in comparison to the signal contained in the seasonal cycle and deviations in the long-term trend, such that the MPI-CCDAS may simply not be sensitive enough to these aggregate system properties like the response of the tropical carbon cycle to El Niño events given the uncertainty in the atmospheric transport and the observational representation error.

4.3 Carbon-cycle simulation

Independent 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 CO_2 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. 7 and S7). Most of the difference stems from small differences in the assumed fossil and ocean carbon fluxes. In the 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 lower but still within the ±5 % uncertainty of the GCP17 data (Le Quéré et al., 2018) (Fig. A4). 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. A4c and S7). The flux due to LULCC prescribed in MPI-CCDAS is also of smaller 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 (Fig. A4d).

Formatted: Subscript

Deleted: , or increase in atmospheric CO_2 concentration due to fire occurrence associated with El Niño events

Formatted: English (US)

Deleted: such as

Deleted: ly

Deleted: prior

Deleted: 8

Deleted: 4

Deleted: somewhat

Deleted: .

Deleted: 8

Deleted: 4

Deleted: lower

Formatted: Wrap Around

The MPI-CCDAS GPP matches well the observation-based product MTE-GPP (Jung et al., 2007) 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, similar to the results in Schürmann et al. (2016) with only five years of assimilation, the tropical productivity is strongly reduced by the assimilation to estimates that are substantially lower than independent estimates such as MTE. This feature is likely the result of a global compensating effect to heterotrophic respiration, and this effect is observed in the drop of the photosynthetic capacity (f_{photos}) in the tropical evergreen and deciduous PFTs, as well as in the reduction of the maximum tropical LAI in the three assimilation experiments compared to the PRIOR. In addition, another critical factor influencing the global reduction of GPP and the tropical uptake of C appears to be related to the difference in data availability of CO_2 stations between the defined assimilation windows. Specifically, this is evident in the results of the data-poorer experiment DEC1, where the topical GPP is substantially lower than in the independent estimates and in the assimilation experiments that use more stations (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 NPP:GPP ratio in ALL and DEC2 decreased to 0.35 and 0.31, respectively, when compared to the PRIOR value (0.45). This reduction might be mainly because the NPP is not well constrained from the atmospheric record. In JSBACH 3.0, autotrophic respiration (Ra) is directly coupled to GPP, hence the fraction of GPP partitioned to Ra leads to an increase in the seasonal cycle of the ecosystem respiration. An increase in Ra with respect to the PRIOR (which is only visible in the global average value in DEC2; Table 3), results in a reduced net land carbon uptake, masking the smaller changes in the vegetation turnover.

The reduction in the soil C pool after the assimilation can be explained due to an unavoidable effect in the model. The MPI-CCDAS was initially spun-up until the soil C pools reached equilibrium considering pre-industrial forcing; however, this new initial state does not consider climate variability. To compensate for this and to reduce the respiration when the MPI-CCDAS is confronted with contemporary changes in the climate, the model creates an artificial C sink that leads to a reduction in the soil C stocks. It is important noting that the JSBACH 3.0 version used in this MPI-CCDAS

Deleted: Compared to

Deleted: independent estimates of

Deleted:

Deleted: ,

Deleted: the MPI-CCDAS GPP matches well

Deleted:

Deleted: as in Schürmann et al. (2016)

Deleted: (Jung et al., 2007).

Formatted: Font: Italic

Formatted: Subscript

Deleted: An important

Deleted:

Deleted: , s

Deleted: the fact that in the

Deleted: period

Deleted: estimated

Deleted: ly

Deleted: compared to the

Deleted: runs

Deleted: with

Deleted: in

Deleted:

Deleted: 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 and 53 % with respect to the value in the PRIOR.

Deleted: Particularly after the assimilation, the total C in the soil (1362 PgC) in the ALL experiment is in closer agreement to the estimate from the Harmonized World Soil Database (http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML; last access January 2015) of 1343 PgC (Table 3). ...

Formatted: English (US)

Formatted: English (US)

Formatted: Wrap Around

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

4.4 Value of long-term data sets to constrain CCDAS

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

5 Conclusion

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

Deleted: therefore

Formatted: Font: Bold

Deleted: The parameter optimisation resulted in considerable reduction in the cost function and norm of gradient, which can be clearly seen as a reduction in the root mean squared error of the MPI-CCDAS compared to the FAPAR and CO_2 observations (Fig. A3 and A4). The trajectory of model parameters involved in the optimization differed for each experiment and each phenotype. While some parameters such as the maximum leaf area of grasses and shrubs and the correction parameter for the initial soil pool size were consistently retrieved, some final parameter estimates varied considerably between the three experiments, e.g., the tropical maximum leaf area index and some of the parameters controlling the seasonality of the phenology (Figure A2). The consequence of these variations are regional differences in the simulated compartment fluxes GPP and ecosystem respiration, which apparently are not well constrained from the observations. Interestingly, these differences lead to very similar absolute values in global carbon fluxes and their trends. This clearly demonstrates a certain degree of equifinality in the results, and cautions a too stringent interpretation of the outcome of the MPI-CCDAS in terms of improving understanding about biosphere processes and their long-term trends.¶

Deleted: se

Deleted: by

Deleted: This can be done by comparing the results of the ALL experiment to the years of 1990-2010 for the DEC1 experiment, and for 2000-2010 for the DEC2 experiment.

Deleted: clear

Deleted: even though

Deleted: are

Deleted: a

Deleted: most

Deleted: .

Deleted: t

Deleted: (for a small number of years)

Formatted: Font:

Deleted: full

Deleted:

Deleted: is able

Deleted: to

Formatted: Wrap Around

CO_2 concentrations at the site level, and the mean prediction remains within the uncertainty of the observations. However, long-term forecasts with CCDAS are less certain, as the observational record does not sufficiently constrain the interannual variability of the simulated land carbon fluxes, and longer-term changes in the decadal net carbon uptake. Nevertheless, the comparatively small error of 2±1.3 ppm after 15-19 years of prognostic simulation shows the potential for mid-term carbon cycle predictions constrained using the CCDAS approach.

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

6 Code availability

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

Acknowledgements

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

Deleted: predictions

Deleted: more un

Deleted: fully

Deleted: long-term land net C uptake in the current phase of rising atmospheric CO_2 and gradually changing climate.

Deleted: DA

Deleted: is

Deleted: space

Deleted: can be made

Deleted: to potentially improve

Deleted: its

Deleted: ¶ ¶ ¶ ¶

Deleted: ¶

Deleted: T

Deleted: was supported by the European Space Agency

Deleted: ¶

Formatted: Wrap Around

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

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

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

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

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

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

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

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

Deleted: 0
Deleted: 5
Deleted: 8
Deleted: 2.9
Deleted: 0
Deleted: 2
Deleted: 2
Deleted: 5
Deleted: 4
Deleted: 7
Deleted: 17
Deleted: 2
Deleted: 1.17
Deleted: 7
Deleted: 6
Deleted: 3
Deleted: 8
Formatted: Font: (Default) Times New Roman
Formatted: Font: (Default) Times New Roman
Deleted: .59
Deleted: 5
Deleted: 0
Deleted: 0
Deleted: 9
Deleted: 1
Deleted:
Formatted Table
Deleted: 1
Deleted: 6
Deleted: 7
Deleted: 8
Deleted:
Deleted: 4
Deleted: 6
Deleted: 3
Deleted: 0
Deleted: 2
Deleted: 2
Deleted: 5
Deleted: 2
Deleted: 1
Deleted: 2
Deleted: 7
Deleted: 0
Deleted: 0

Appendix

A1. Assimilation performance

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

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

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

In other extra-tropical areas results from experiments DEC1 and DEC2 experienced an average increase in Λ_{max} by 5.6 % in BE (from 2.29 in the PRIOR to 2.42), 24 % in BW (from 1.62 in the PRIOR to 2.01), and 3.8 % in TNW (from 3.11 in the PRIOR to 3.23). As a result, the temperature and daylight-related parameters were modulated such that the largest decrease with respect to the prior value in the temperature at leaf onset (T_ϕ) was also observed for these two PFT's, especially for CD in the DEC1 and DEC2 experiments. Also, the day length at leaf shedding (t_c) and the timescale of leaf senescence (leaf shedding timescale, $1/\tau_1$) primarily increased for CD. As for the PFT's influenced by temperature and water (TeH, TeCr, TrH and TrCr), the most significant change with respect to the prior value took place in the posterior value for the C3 crops (TeCr; distributed in Europe, USA and East Asia) whose value decreased considerably for the water stress tolerance (τ_w) in experiments DEC1 and DEC2, whereas the value

Formatted: Font: Bold

Formatted: Not Highlight

Formatted: Not Highlight

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

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

Deleted: ¶

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

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

Formatted Table

A2. Pixel level phenology analysis

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

References

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

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

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

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

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

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

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

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

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

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

Deleted: ¶
¶
¶
¶
Page Break
¶

Deleted: :

Deleted: 2013, 10.3334/CDIAC/ffe.MonthlyMass.2013.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R., and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Global Biogeochem. Cycles, 19, GB2026, https://doi.org/10.1029/2004GB002254, 2005.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Figures.

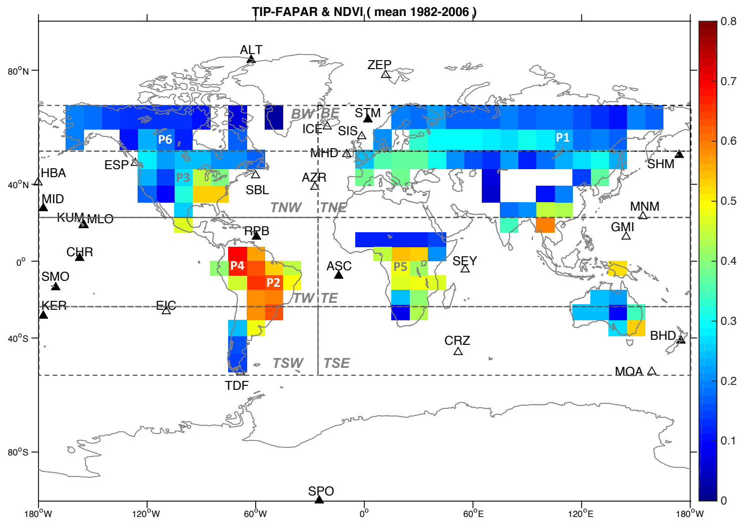


Figure 1 – Global distribution of the temporal mean (1982-2006) of the merged satellite FAPAR product used in the assimilation procedure. It shows also the spatial coverage of eight regions globally distributed: Boreal West and East (BW and BE, for latitudes north of 60 °N), Temperate Northwest and Northeast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

Deleted: subtropical

Deleted: S

Deleted: S

Deleted: subtropical

Deleted: S

Deleted: S

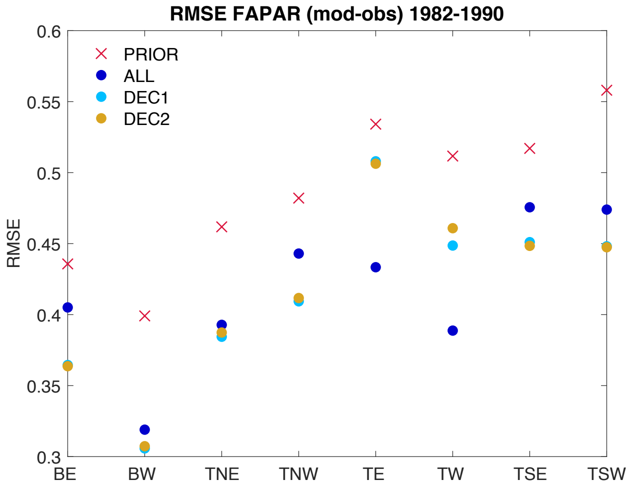


Figure 2 – RMSE for FAPAR from the model results and observations for the period 1982-2006 and for different regions.

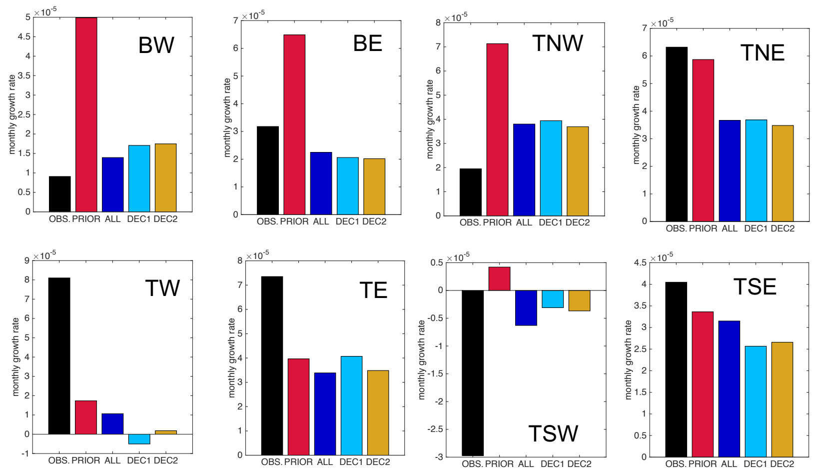


Figure 3 – Mean monthly growth rate of FAPAR for 1982-2006 on each analyzed geographical region for the satellite observations and results of PRIOR and the posterior experiments.

Deleted: ¶
¶

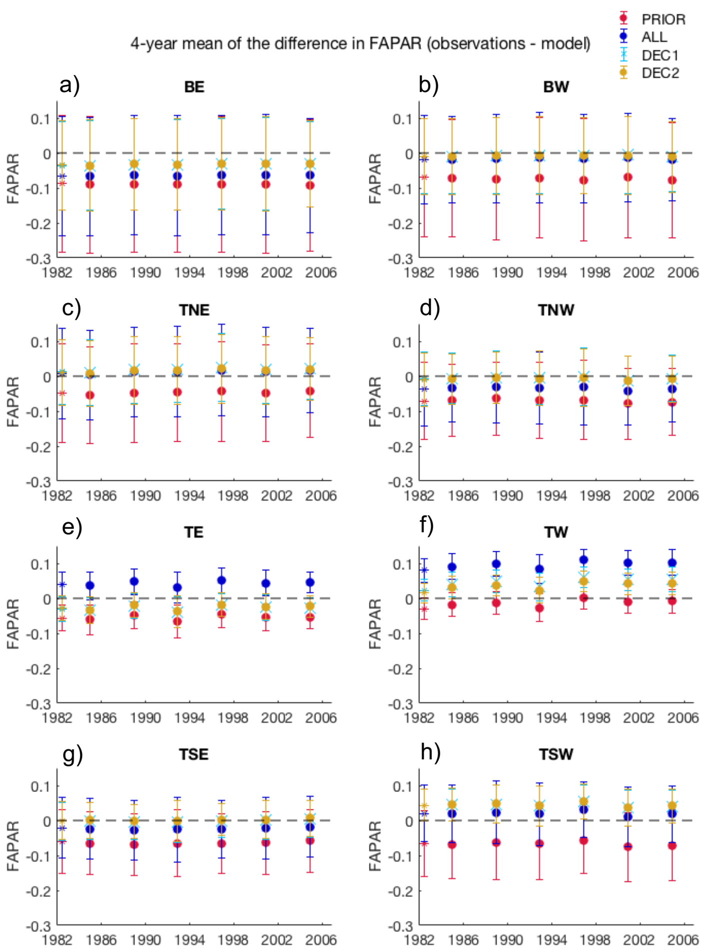


Figure 4 – Time series of the four-years mean of the FAPAR anomaly to the satellite data for each model experiment in six selected model pixels. The error bar indicates one standard deviation of the four-years differences. The first marker to the left (as asterisk) in the time series is the single value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

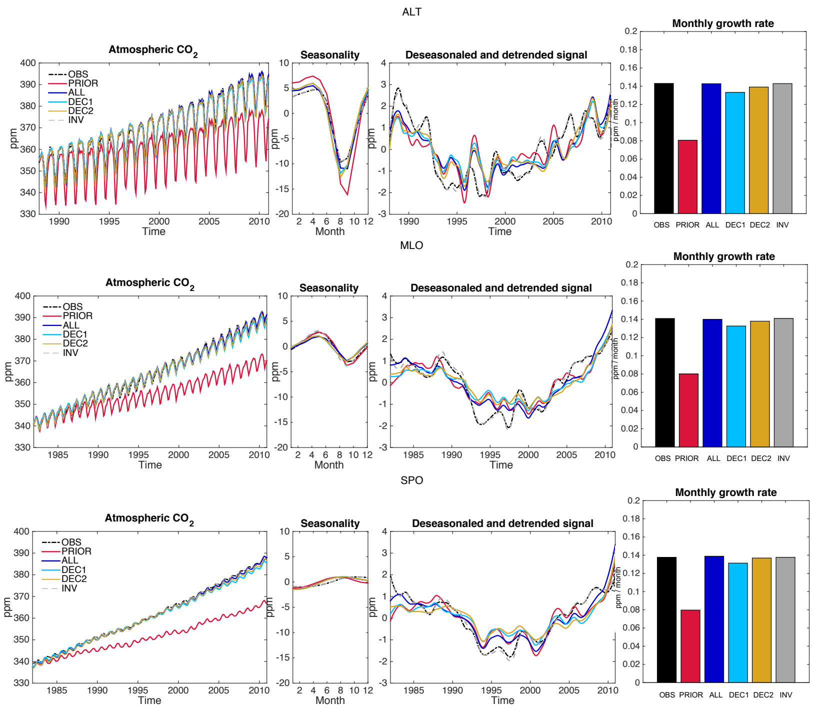


Figure 5 – Statistical analysis of atmospheric CO_2 in three flask measurement sites: Alert (ALT; top panels), Mauna Loa (MLO, center panels) and South Pole (SPO, bottom panels), from the measurements, PRIOR, posterior experiments (ALL, DEC1 and DEC2) and inversion (INV1). For each station the panels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

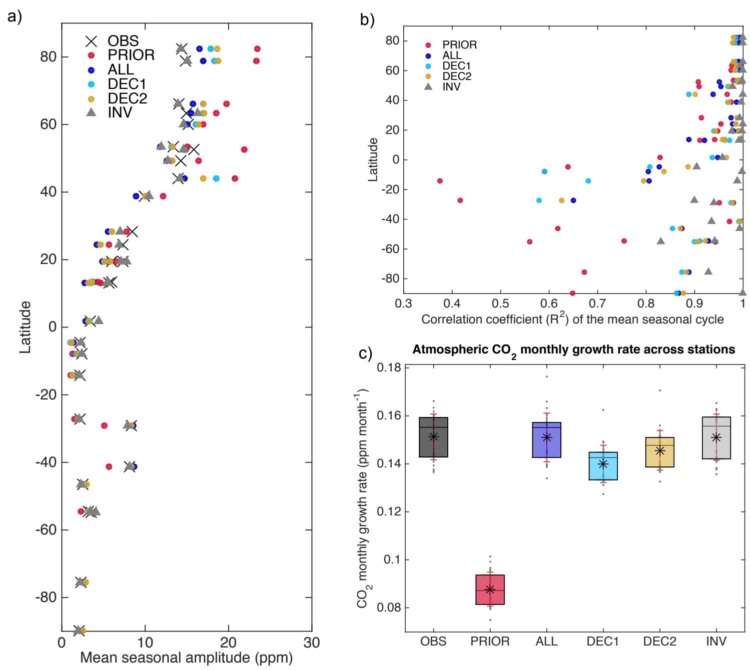


Figure 6 – a) Latitudinal distribution of the mean CO_2 seasonal amplitude for the 28 flask-measurement stations from the observations, PRIOR and posterior experiments; b) Latitudinal distribution of R^2 obtained from the correlation between the observations and each simulation results of the mean atm. CO_2 seasonal cycle and c) average atmospheric CO_2 monthly growth rate across stations for the observations and model results. The star on each bar is the mean of the atm. CO_2 monthly growth rate, the horizontal middle black line on each box is the median, the red whiskers depict the error as $+/- 1\sigma$, and the grey dots on each box are the actual monthly growth rate values for all the stations in each data set.

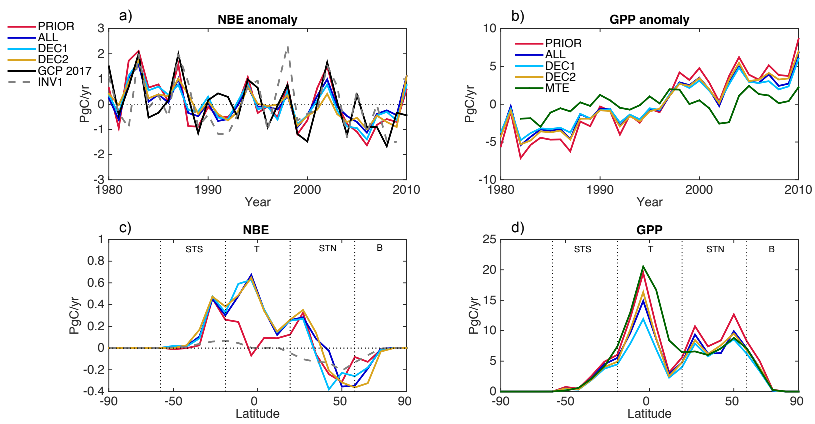


Figure 7 – Time series of the anomaly to the temporal mean of the time series (a and b), and latitudinal gradient (c and d) of the total Net Ecosystem Exchange (NEE including the influence of LULCC) (left) and Gross Primary Production (right) for the results of each model simulation. NEE from the model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

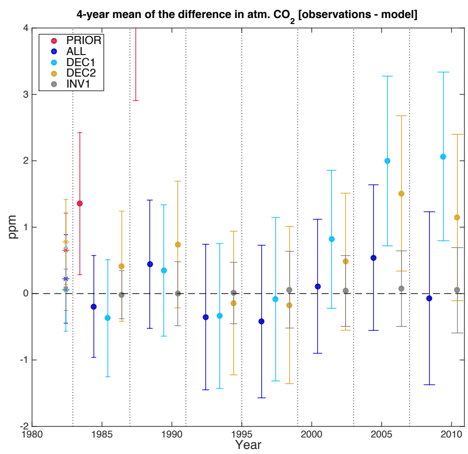


Figure 8 – Time series of the four-years mean of the atm. CO_2 anomaly to the observations for each model experiment and inversion results, for all the stations. The y-axis is limited to the results in the posterior experiments. The error bar is one standard deviation to the four-years mean of the differences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

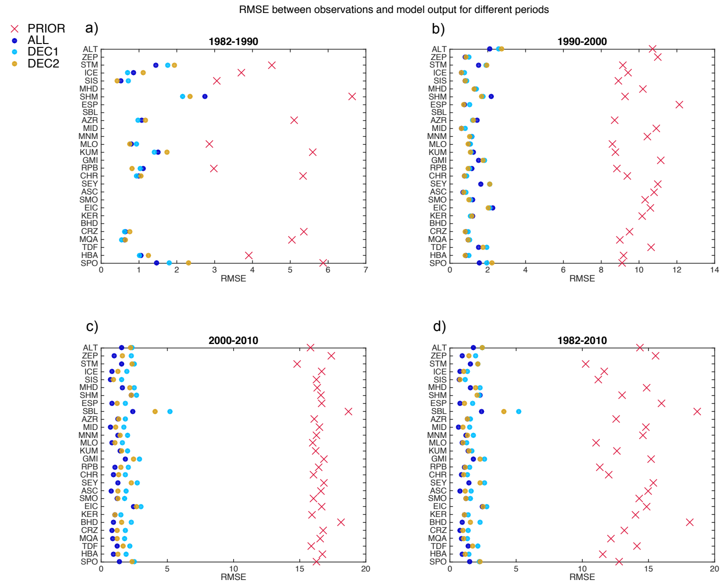


Figure 9 – RMSE for different periods between CO_2 atm. concentrations from measurements and model results for the different assimilation experiments for each of the flask measurement stations.

Appendix

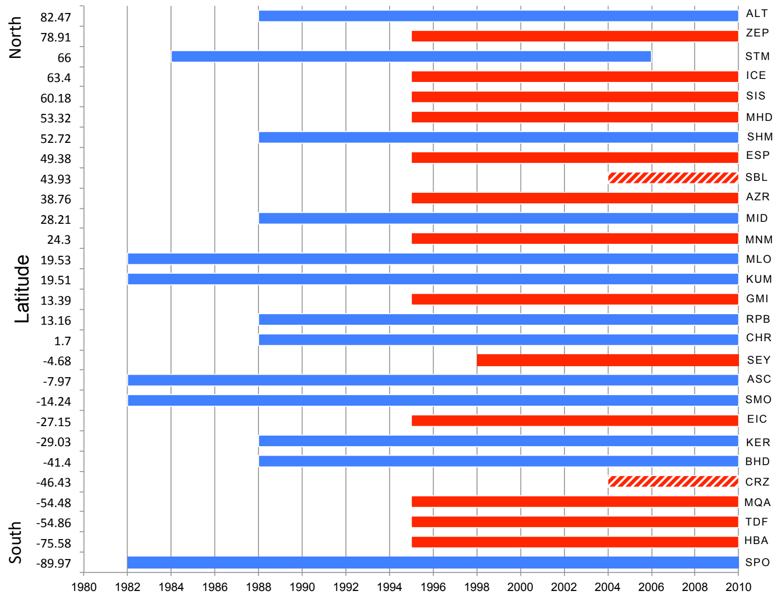


Figure A1 – Data availability and latitudinal location of the 28 stations where the long-term flask measurements of atmospheric CO_2 mole fractions were taken for assimilation in CCDAS. ALL experiment used all the stations of the time series (blue and red bars) (1980-2010); DEC1 used data only from stations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

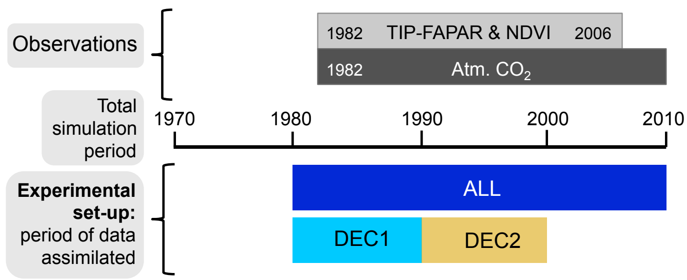


Figure A2 – Experimental set up for posterior experiments ALL, DEC1 and DEC2 with different temporal windows for the assimilation of FAPAR and molar fractions of atmospheric CO_2.

Deleted: ¶

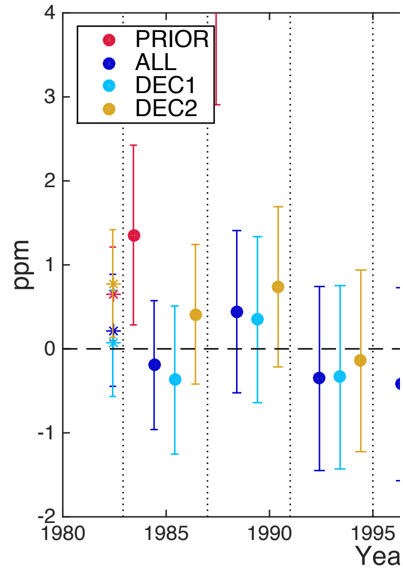


Figure 10 – Time series of the four-year mean of the atm. CO_2 anomaly to the observations for each model experiment and for all the stations. The y-axis is limited to the results in the posterior experiments. The error bar indicates +/− 1 standard deviation of the four-year mean of the differences to the observations. The first marker in the time series (as asterisk) is the single value for 1982.

<object>

... [1]

Deleted: that use

Deleted: of observations of

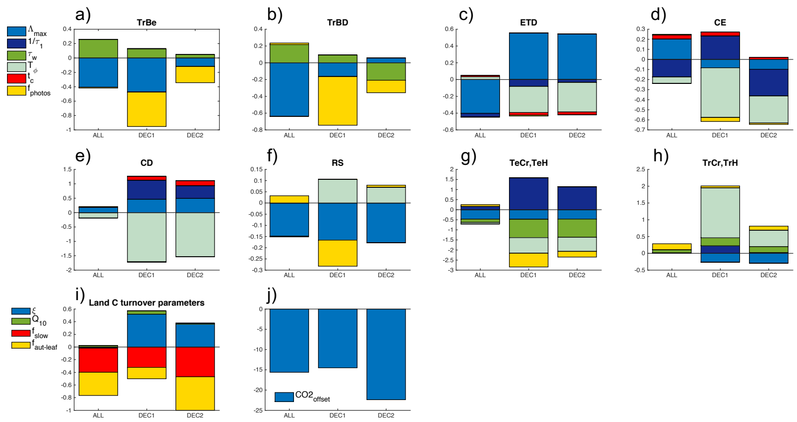


Figure A3 – Final value for each parameter p at the end of the assimilation experiments, normalized to the prior value (p_{pr}), i.e. (p/p_{pr})-1. This is shown for each model plant functional type (a to h) and globally for the land C turnover parameters (i and j).

Deleted: 2

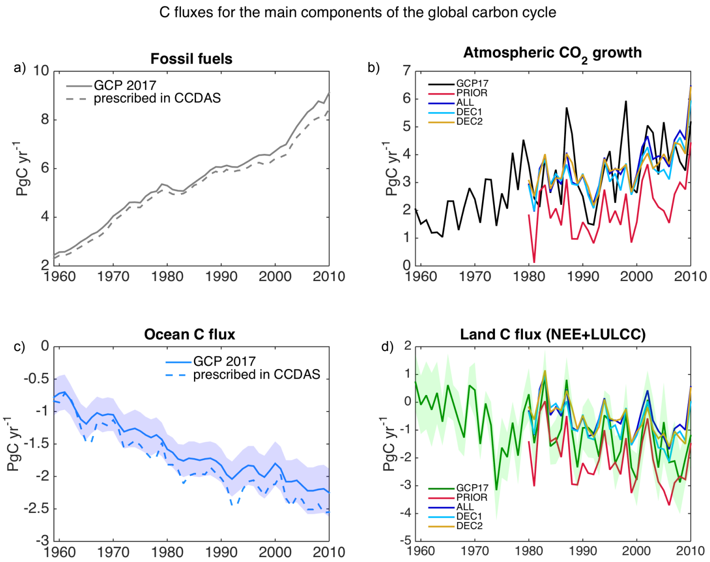


Figure A4 – Time series of the annual mean of the major components of the C cycle used as background fluxes in CCDAS compared to those from the GCP 2017. The atm. CO_2 growth from the model output is the result of the sum of fossil fuel, ocean, and land C fluxes. The blue shadow in the ocean C sink of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