Referee (Chris Jones):

This is an interesting manuscript that explores a potentially important feature of the land carbon cycle, namely an apparent shift in land carbon uptake in the late 1980s. This shift has been analysed before by some of the same authors and this paper seeks to bring new insights into the climate drivers of the shift. The results are that two distinct regions play the dominant role and for different process-reasons. The coincidence of these two regions (tropical northern Africa and Northern Eurasia) is enough to provide a global signal in NPP and by inference NEP.

In general the paper is carefully planned and the authors present a 3-fold attack based on data-driven model CASA, process based models from TRENDY, and a residual carbon budget analysis from GCP. For each they consider sources of uncertainty – for example my first reaction on reading the abstract was to question the precip dataset used, so I was pleased to see that multiple datasets were used to quantify the uncertainty from climate drivers.

The results and discussion are presented in a very convincing manner, although I was left wondering if the paper focused too much on the similarities in the different methods and neglected discussion of some of the differences. I've listed below a few areas I'd like reassurance that there isn't an underlying problem. For example, the change-point analysis on CASA output reveals quite clearly a shift in the 1980s, but to find the same shift in TRENDY runs, the long-term trend due to CO2 needs to be removed. This is very clear in table S2, but only gets a passing mention in the main text with no attempt to explain either WHY this is the case (maybe a slight hint that the authors don't believe the CO2 fertilisation response of the TRENDY models - but this is only speculation), or more importantly what this means for the analysis. What do we conclude about the 1980s shift if one approach finds it, but another approach needs more careful processing to reveal it?

Authors: We highly appreciate these thoughtful comments to increase the transparency in our analysis and also to discuss more about the differences in the modeling approaches. In our thoroughly revised version, we thus have extended the analyses to include all carbon component fluxes (NPP, Rh, and NEP) for both CASA and Trendy simulations. We added three additional figures (Fig. S2, S5-S6) and an additional table (Table S3) in the revised supplement that show corresponding results. In addition, we revised our estimate of a global late 1980s NEP shift using a combination of residual calculations (e.g. estimating Rh as residual of NPP from CASA and Trendy and NEP from GCP) and direct model estimates (see newly formed Section 3.4 and details are given in the Supplement as well).

Following this referee's suggestion, we also elaborated more on the differences between CASA and TRENDY (for example see added material in Section 3.3 about 'seasonal NPP carry-over effects'). In regards to CO2 fertilisation, however, our interpretation of the results are somewhat conservative. On one hand, the fact that the late 1980s shift is not seen in the TRENDY S2 experiments (neither in the component fluxes nor in the corresponding NEP data which are now included; see Table S1-S3 in the revised Supplement) may suggest the possibility that the sensitivity of process-based models in regards to CO2 may be too high. A brand new study that solely focused on a comparison of satellite-constrained NPP sensitivity to CO2 concentrations with those based on process-based models came to the

same conclusions (Smith et al. 2015; this reference is now also included in the revised ms).

On the other hand, the TRENDY ensembles also show a lower climate sensitivity when compared to CASA (e.g. temperature sensitivity over Northern Eurasia), which could be another explanation for the lack of a 1980s shift in the S2 experiments. We realize that our results and interpretations on this CO2 fertilisation issue do perhaps create more questions than they provide answers at this stage, but we also believe these are important points for the community to engage in.

Other more direct comparisons that the authors might show are:
- you compare and contrast 3 different solar and precip datasets for driving CASA. Can you also add to this comparison the radiation and precip datasets used in TRENDY. Is this the same data? and if not how does it differ? what happens if you drive CASA and TRENDY models with as close to possible all the same datasets (T, P, radiation etc)

Authors: In both the CASA and TRENDY simulations CRU monthly temperature and precipitation data play a primary role. For example, in TRENDY CRU climate data are exclusively used. Our analysis also shows that the differences between the 3 precip data sets that we investigated were actually rather small (see Fig. S11 in the Supplement). A substantial difference in the CASA and Trendy simulations, however, is that in CASA surface solar radiation is one of the driver variables, whereas in TRENDY it is not (see Sitch et al. 2015; ref. in ms). Yet, the two focus regions of our study (Northern Eurasia and Northern Africa) are both generally not considered radiation-limited, and this may suggest that the differences in model driver protocols may not play a substantial factor in explaining possible differences between simulated carbon fluxes from CASA and TRENDY.

- you mention that TRENDY models simulate FAPAR whereas CASA takes it as an (uncertain) input. How do these compare? the input FAPAR to CASA has a very marked jump visible in figure 2. What does FAPAR from the TRENDY models look like? Does it also follow this shift?

Authors: We also feel that it would be really informative to compare the satellite-based FAPAR (used in CASA) and the prognostic TRENDY-based FAPAR specifically in light of the marked late 1980s jump. However FAPAR outputs for the TRENDY models are as of yet not available and this precludes as unfortunately from looking into that in more detail. It should be noted, however, that comparisons between satellite-based and TRENDY ensemble leaf area indices (a variable closely related to FAPAR) generally show some broad consistency (see Fig. 7 in Sitch et al. 2015; ref. in ms). On the other hand, observation-model comparisons do indicate that the phenology responses in satellite-driven models are generally more robust than in process-based models (e.g. Raczka et al., 2013; this reference is now included in the revised ms in Section 3.3.).

- you say there's no data-constrained RH product, so you use the output from CASA. But TRENDY will also produce RH of course, so what does this look like? Why neglect 9 models output to just show CASA?

Authors: In this revised version, we have included also an analysis on the Trendy Rh data and more (see also detailed response to the following comment as well as our first response above).

- in general, some direct comparison of the different methods would be nice - so a summary plot of all the different climate datasets in one place (those for driving CASA vs those for driving TRENDY), all the NPP outputs (CASA and TRENDY), all the RH outputs (CASA and TRENDY), all the NEP outputs (CASA, TRENDY and GCP). Figure 2 and 3 show there are similarities of course between CASA and TRENDY, but also some clear differences - how important are the differences? and how do would this comparison look at the seasonal scales (the TRENDY counterparts to fig 2 lower traces). I'm not suggesting there are any problems from the above comments nor that the authors have been selective in what they show. But some of the above comparisons seem obvious things to do, so it would be reassuring to know the authors have considered them and not found any issues. As long as we can be assured there are no hidden problems in any differences between the 3 datasets examined, then I recommend publication.

Authors: Again, we appreciate these comments to increase specifically the transparency of our analysis. In addition to our detailed response in regards to added elements in this revised version (see first response above), we also added the seasonal counterparts in the TRENDY NPP series for Northern Eurasia and Northern Africa as requested by this reviewers (Fig. 3 in revised ms). A discussion about the differences (in TRENDY and CASA) between seasonal responses is added in Section 3.3.

other minor points to address that may help improve the paper:
- I wondered about your pinup process for CASA. (a) why do you pinup to equilibrium for a period that almost certainly isn't (does this maybe explain your difference wrt TRENDY trends?). (b) why repeat a rather small period of forcing (1982-86) which itself is rather anomalous (containing both a big El Nino and Volcano - El Chichon). Have you tested the sensitivity of your results to the CASA initialisation?

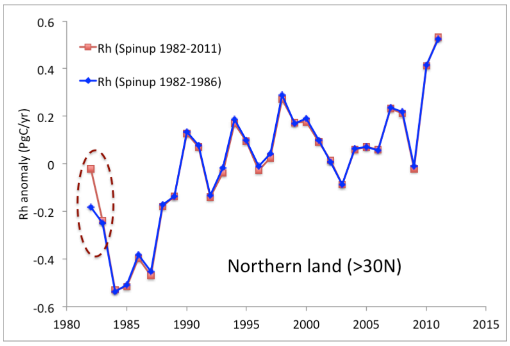


Fig. 1 Anomalies in CASA-based Rh flux for northern land and for two sets of model spinup initialisations. In one scenario, the model is spinup for 250 years using a driver climatology based on 1982-2011, whereas in the second scenario a climatology based on 1982-1986 is used.

Authors: This point is well taken, and when exploring this issue we discovered an error in our CASA spinup description. We actually did use a driver climatology based on the whole study period (1982-2011) and not as stated based on the shorter 1982-1986 period. We corrected this in our revised ms. In response to this reviewers' comment we did nevertheless test the robustness of our results in regards to spinup climatology. Results show that only in the case of Rh fluxes (and consequently NEP) and during the first year there is a some influence of spinup initializations (e.g. dashed circle in Fig. 1), suggesting that the CASA model is indeed close to equilibrium. Our relatively conservative approach in our change point

methodology (discounting the first and last 5 years in a series as possible change points) further ensures that spinup effects do not have any sizebale effects on our results.

- you mention land-use as another potential driver of a shift, but this is not included in the TRENDY S1 or S2 runs. Have you looked at the runs that DO include land-use? (I'm not up to date on whether these are published yet, but the last TRENDY protocol had land-use forcing)

Authors: Yes, the TRENDY S3 experiments DO include land use, but these simulations are unfortunately as of yet still under embargo from the TRENDY modeling team. But also recall that at least at global scale the land use is considered in calculations of the residual carbon sink from the GCB (e.g. see Table 1 in ms).

- you dismiss using the ISCCP solar radiation data as it is "biased high over the Amazon". But from figure S4a this doesn't jump out as being that different from the other datasets. And this looks like a region with very few obs. How subjective is this decision, and how important is a bias over the Amazon given your focus on other regions?

Authors: Yes, the comparison in Fig. S4 does not provide a clear picture, but a comparison of area-averaged radiation data representative of the Amazon basin shows a large positive 'bias' of roughly 15 W/m2 in the ISCCP data relative to the other 3 data sets that we evaluated (see Fig. 2). Including the ISCCP would increase the uncertainty in the CASA NPP simulations considerably, which would make identification of a potential shift even more challenging (but recall that we did not see evidence of robust prominent shifts in data-driven NPP over the Amazon basin).

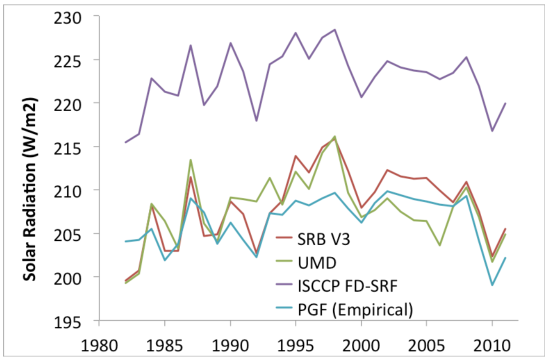


Fig. 2 Time series of various total solar radiation data products representative of the Amazon basin. For single product descriptions see Table S4 in the Supplement of the ms.

- if the UMD data is based on ISSCP, how different is it and why is it acceptable?

Authors: The UMD data use the same ISCCP-DX cloud, water vapor and ozone data, but the radiative transfer model and supporting data (e.g. aerosols, land cover) differ. Comparisons with surface observations indicate that UMD has indeed lower biases than the other products (Ma and Pinker, 2012; ref. in the Supplement) and does not suffer from discontinuities in the record (Vinokullo et al., 2011).

- you mention that some radiation datasets include the effects of aerosols - do you mean on the total radiation? or also on the diffuse fraction? if only on the total then aerosols lead to reduced light and productivity, if you include the diffuse effect productivity can be enhanced - so can you specify which is included?

Authors: Yes, some radiation data sets include the effects of aerosols on the total radiation. All solar radiation data sets included in our analysis do not separate diffuse and direct components. The CASA model has no explicit diffuse light response parameterizations, which may be considered a limitation. On the other hand, some diffuse effects on productivity may be captured through the satellite-based fAPAR.

In general the inclusion of aerosol effects in available solar radiation data is somewhat limited. For example, the SRB (V3) radiation product uses an aerosol climatology that does not include the impact of biomass burning which is clearly important in the Amazon and other tropical regions (Stackhouse et al. 2011; ref. in the Supplement). The UMD also uses a aerosol climatology, but has a more detailed treatment of aerosol properties that includes variations in single scattering albedo and asymmetry factor (Ma and Pinker, 2012; ref. in supplement).

Anonymous Referee #1:

Buermann et al. assess abrupt increases in NPP, their relations with shifts in global NBP, and the drivers of the shifts in NPP. The manuscript is very interesting and well written. Assessing potential abrupt changes in NPP is novel. While reading this manuscript, I was wondering if the reported shifts in NPP (especially in boreal Eurasia) might relate to the observed increase in the seasonal amplitude of atmospheric CO2 (Graven et al., 2013). This increase in CO2 amplitude originates from northern ecosystems and had the largest increase within the last years (Graven et al., 2013). I also was wondering how the results relate to previous findings of the same author (Buermann et al., 2014) which emphasize the role of drought on decreasing NDVI in boreal regions.

Authors: We also highly appreciate the thoughtful comments of this referee. The atmospheric CO2 seasonal amplitude is an interesting and potentially useful metric to provide further independent evidence for the unraveled northern land attribution of the late 1980s shift in plant carbon uptake. Identifying a corresponding signal in the CO2 amplitude is however not straightforward due to the presence of a strong increasing trend (Graven et al. 2013) with an apparent substantial contribution from agricultural expansion and intensification (Zeng et al. 2014). Given these complications we decided not to include it in this analysis. But recall that one additional atmospheric CO2 metric is used; changes in the slope of the North-South atmospheric CO2 gradient which are consistent with a sizeable northern land contribution of the late 1980s shift in plant carbon uptake (see first paragraph in discussion).

In Buermann et al. (2014), we find a change in the interannual summer NDVI response to temperature in a region (the Urals) of the Northen Eurasian boreal forests in the mid 1990s. While this change in the temperature-NDVI link may indicate emerging responses of the boreal systems (e.g. increased summer drought sensitivity) the more regional and seasonally localized character of this phenomena appears to be not strong enough to influence continental to global scale pattern that we unraveled in the present study.

1.1 Uncertainty from FAPAR datasets
A thoughtful assessment of uncertainties from different temperature, precipitation and radiation datasets is done in this study. However, the major contribution to the temporal dynamic of CASA-modelled NPP originates from the FAPAR dataset. Several studies have shown large difference between different FAPAR (or NDVI) datasets (Fensholt and Proud, 2012; McCallum et al., 2010; Scheftic et al., 2014; Tian et al., 2015). Although the used GIMMS3g FAPAR dataset is the most reliable long-term dataset, the difference in more recent periods in comparison with datasets from modern sensors, highlights the need to account for FAPAR-related uncertainties. In order to convince the reader about the reliability of the reported NPP changes, it is necessary to #1 evaluated the reported changes to sensor changes in the underlying GIMMS NDVI3g record (similar as in (Tian et al., 2015)), and #2 to assess the uncertainty in NPP estimates also based on alternative FAPAR datasets at least for the overlapping period with newer sensors (e.g. MODIS).

Authors: These are valid points. We actually put considerable thoughts into that but feel that any accounting of uncertainties in FAPAR would lead to additional biases. In more detail, we also thought about comparing GIMMS FAPAR and MODIS FAPAR as suggested by this referee, however, MODIS data

are already utilized in the generation of the GIMMS FAPAR3g that is used in this study (for details see Zhu et al. (2013; ref. in ms)) and hence MODIS data would not provide an independent metric. A further complication is that, based on our own recently published work in Guay et al. (2014), NDVI trends from more modern sensors (e.g SPOT, MODIS, SeaWiFS) are also not necessarily consistent with one another. While this is clearly not a desirable situation it is the status quo, and it is also clear that much more work (and resources) should be allocated to improve/validate long-term satellite vegetation time series. In fact this status quo was one of the motivations for us to also exploit process-based models that do not rely on prescribed satellite vegetation inputs. This follows the overall logic that each independent approach in carbon cycle science has their own (considerable) set of uncertainties, but together may provide a consistent picture in regards to our key findings.

1.2 Model evaluation

The results are praised by saying in the abstract "using (...) models constrained by observations". However the only constrain is the use of GIMMS3g FAPAR within CASA. No further constraints are used for modelled NPP. Model results are not at all evaluated against independent data. In order to be more convincing, it is necessary to evaluate model results against independent data, e.g. NPP databases (Luyssaert et al., 2007), GPP site-level time series (FLUXNET), upscaled fields (Jung et al., 2011), C stock maps (Carvalhais et al., 2014), or long-term changes in the seasonality of atmospheric CO2 that might be indicative of changes in northern terrestrial productivity (Graven et al., 2013).

Authors: Here, we do not fully agree with these comments, and our choice of for example the data-driven CASA model was exactly motivated by the fact that this model has been extensively tested by the community since its birth in the early 1990s (e.g. van der Werf et al. 2006; ref. in ms). In this regard, we also have included a key reference in the revised manuscript (Raczka et al. 2013; see Section 3.3) where multiple biosphere models including CASA where tested against EC flux time series (GPP, Respiration and NEP) at seasonal and interannual time scales. The results of Raczka et al. (2013) show that data-driven models like CASA tend to simulate carbon fluxes more robustly most likely because of a more realistic phenology response. In regards to TRENDY, all of the 9 participating models can be considered state-of-the-art models and have been tested/validated for example in Piao et al. (2013; ref. in ms) against a number of metrics including upscaled flux measurements, FACE CO2 and global carbon budget data.

Taken together, while models do still have severe deficiencies (e.g. carbon turnover times as pointed out in Carvalhais et al. (2014; ref. in ms and also discussed in the ms) we do feel that the models have been sufficiently validated to justify their use in our study (see also our first response in regards to seasonal CO2 amplitude).

1.3 Change point detection algorithm

The statistical analysis is very valuable. Especially, I very much appreciate that the authors evaluate several alternative statistical models by means of SIC and the uncertainty analysis for the change points is also a necessary step given the low robustness of such change detection methods (Forkel et al., 2013). However, many studies report changes in trends on NDVI datasets, such as greening to browning (de Jong et al., 2011, 2013; Piao et al., 2011). A good overview of potential changes *in given in de Jong et al. (2013). Trend changes as further option in change detection was not assessed in this study. Trend changes might be here therefore either represented as changes in mean or as continuous long-term trend. I'm wondering if ignoring the trend change-option results in an overestimation of abrupt changes and thus affects the main conclusions of the study. In my opinion it is necessary to additionally account for the trend change option in the statistical analysis (Verbesselt et al., 2010a). The author's fear of overfitting time series with additional parameters as in trend change models (p.13774, l. 1-3) can be easily handled by again using SIC on the trend change option. Further it was not clear to me how the seasonality of NPP time series was treated in the change point algorithm*.

Authors: We appreciate these comments, but feel that these are really two separate objectives: It is quite different to find the statistical model that best fits an observed time series at a given locality (as hinted at by this referee) or to design a statistical framework that seeks consistency between potential shifts in modeled carbon fluxes and previously observed shifts in carbon fluxes from global carbon budgets. In order to unravel corresponding consistent large-scale pattern, we thus feel that our relatively simple three statistical model approach is justified (see also Section 2.2: first paragraph). All NPP, Rh and NEP time series that were exposed to our change point algorithm represent annual means.

Another reason for why we chose not to fit more complex models (e.g. combining a trend and a change-point as suggested by this referee) is due to the limited record length of the data analyzed. Even if we use the SIC to penalize for these additional parameters, we are concerned that we would detect spurious shifts as shorter time series have increased false detection rates (e.g. Beaulieu et al., 2012b; ref. provided in ms). As for the reviewer's concern that we are overestimating abrupt changes by avoiding the more complex models: we rather anticipate an overestimation of the number of abrupt changes if we add more models – fitting 2 additional models with shifts should lead to the detection of additional shifts (some real and some spurious). To illustrate our point, we ran the analysis also including two more statistical models: a model with a trend and a shift as well as a model with trend change (Fig. 3). This shows that our original results are robust – we still find a predominant shift in the late 1980s in Northern Africa and Northern Eurasia. It also demonstrates our point that adding models tend to overestimate shifts, as more regions exhibit shifts now. Following the principle of parsimony, we prefer to stick with our 3 models approach since the signal we discuss in this manuscript is robust to the number of models used.

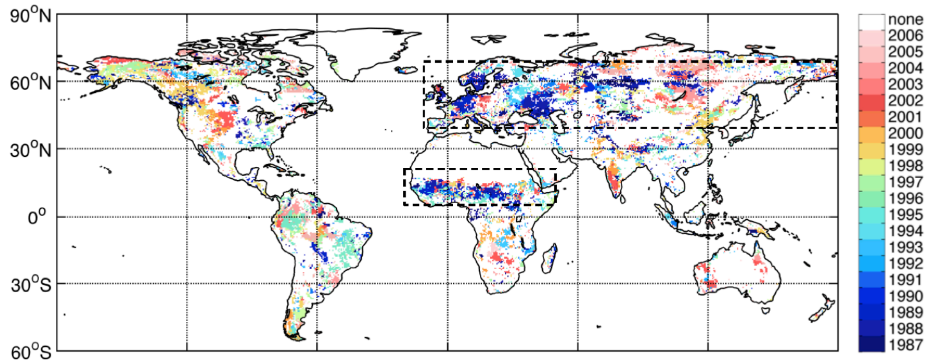


Fig. 3 **Spatial pattern of abrupt shifts in data-driven NPP**. Map shows timing of abrupt shifts in data-driven (CASA) annual NPP for the satellite period 1982-2011. Here results are shown using a 5-model approach (see disussion). All shifts shown are statistically significant ($P < 0.05$) based on Monte Carlo simulations that take into account explicit uncertainties.

2. Specific comments
These are comments to specific parts of the manuscript. However some of these comments will be resolved by addressing the major comments.

p. 13770, l. 6: I don't understand why forest regrowth and fire suppression where used as examples for land-use patterns. Forest regrowth is a dynamic in land cover, the corresponding change in land use could be rather named reforestation or afforestation as the term land use usually implies human management.

Authors: Yes, we were a bit sloppy here and exchanged the term 'land-use patterns' with 'land-use and land-cover change' in the revised ms.

p. 13771, l. 16-18: But this study applies only to arid grasslands. Are you aware of any references that try to quantify the CO2 fertilization effect in FAPAR data for forest ecosystems?

Authors: This point is well taken and to our knowledge the only studies that touch on this issue and also include forests is Los (2013) and Forkel et al. (2015). Corresponding results diverge to some extent with one study suggesting a more minor effect of increasing CO2 on FAPAR (Forkel et al. 2015) and the other study a more sizeable contribution (Los, 2013). We included both references in this context in the revised ms (see Section 2.1 and along these lines see also our reply to the following comment). Also recall that our original discussion in this context in how far CASA may capture a CO2 fertilisation effect was already quite conservative (see Section 2.1).

p. 13771: How is the CO2 fertilization effect on photosynthesis considered in CASA? Only through the FAPAR forcing dataset or is there an additional module that accounts for CO2 fertilization? FAPAR might be not sensitive enough to the CO2 fertilization effect especially in forest ecosystems that have upper FAPAR values. Based on modeling experiments it has been shown that the CO2 fertilization effect contributes only minor to changes in FAPAR (Forkel et al., 2015). Therefore it might be possible to underestimate the CO2 fertilization effect if the NPP model relies just on FAPAR.

Authors: In CASA, the CO2 fertilization effect is only included through the FAPAR forcing data (see equation 1 in van der Werf et al. 2006; ref in ms). In how far FAPAR captures a CO2 fertilization effect is, however, an open question (see reply in previous comment).

The change point results from the S2 TRENDY experiments show that in the presence of a strong trend in the carbon flux time series (resulting from CO2 fertilization), a climate shift signal may be masked (for both component fluxes NPP and Rh as well as for NEP). We can only speculate at this point, but for CASA the late 1980s climate NPP appears to be stronger (when compared with TRENDY) and may be still detectable even in the presence of an additional trend component from CO2 fertilization.

p. 13771, l. 20: Is this really land surface (i.e. skin) temperature? I thought CRU provides air temperature at 2 m?

Author: Yes, it is air temperature at 2m, and we removed the 'surface' portion.

p. 13771, l. 25: I agree but at least it would be possible to assess FAPAR-dataset uncertainty for the overlapping period with MODIS or you could based on the FAPARNDVI relation you could try to use other long-term NDVI datasets (Marshall et al., 2015). An assessment of the findings in relation to potential uncertainty sources from different FAPAR datasets seems necessary given the striking differences in these datasets regarding trends and inter-annual variability (Fensholt and Proud, 2012; Tian et al., 2015).

Authors: Please see our corresponding response to main comment 1.1

p. 13773, l. 10-12: Is there a reason why the option "change in trend" (i.e. stable to positive, positive to stable, positive to negative etc.) was not considered? Several studies have shown that such trend changes exist in satellite-derived NDVI data (de Jong et al., 2011, 2013; Verbesselt et al., 2010a). Such changes were also detected in the GIMMS3g NDVI and thus are likely also present in the GIMMS3g FAPAR data. I assume by ignoring the "trend change" option, there is the risk of over-selecting option 2 (change in mean) as the preferred statistical model.

Authors: Please see response to main comment 1.3

p. 13773, l. 26-28: : : : and this might result in an over-estimation of abrupt shifts. I think it could be worth-while to check alternative change detection algorithms that also account for smooth changes by considering trends (e. g. (Verbesselt et al., 2010a, 2010b)). I think the risk of overfitting is low by adding two more slope parameters to the statistical model as you could use the SIC as well. Given the large use of trend change detection methods on NDVI time series it seems not plausible to my why this should not be done for NPP data. Furthermore, based on the large uncertainty of trend change detection methods (Forkel et al., 2013) it is necessary to consider several methods.

Authors: This point is well taken, but our approach is more conservative in regards to risks of overfitting and detection of spurious shifts (see response to main comment 1.3)

p. 13774, l. 6-26: The approach of the uncertainty assessment is very valuable. However in order to fully understand it but not to overload the average reader, I would suggest to extent the description of this approach (maybe incl. some illustrative figures or equations) and move it to the supplement.

Authors: This point is well taken, and there is always a fine line in keeping descriptions brief but providing enough detail that an independent researcher can reproduce the analysis. While somewhat brief, we do feel we have provided enough information in Section 2.2 in regards to uncertainty assessment and would rather refrain from adding another chapter in the already large Supplement.

p. 13775, l. 14-15: Do the numbers represent the magnitudes of the shifts? Please clarify.

Authors: Yes, this is correct and we indicated that in the revised ms.

Results section: I suggest to have some sub-chapters (e.g. 1. NPP shifts, 2. drivers)

Authors: We agree, and formed sub-chapters in the Results section to improve the organization in the revised ms.

p. 13776, l. 8-24: I think you cannot separate the driving factors on NPP with CASA. As you are admitting in the caption of Fig. 2a, the FAPAR dataset already integrates changes in temperature and precipitation and other drivers. Thus, the FAPAR dataset explains most of the dynamic in NPP. Even if you try to separate these factors, we still don't know about the temperature or precipitation effects. In my opinion, this separation of drivers cannot be insightful done with CASA but only with the TRENDY results.

Authors: We do not fully agree with this interpretation. If changes in FAPAR and climate coincide in space and time (as it is the case for our 2 target regions Northern Eurasia and Northern Africa) one can make a compelling argument that climate is the primary driver and that vegetation cover (inferred from FAPAR) is responding to that and that these two drivers together influence NPP in the CASA model.

p. 13776, l. 10-15: Recent studies suggest that changes in spring FAPAR and the begin of the growing season in boreal ecosystems are related to changes in water availability from changing snow cover (Barichivich et al., 2014) and to water supply from changes in permafrost dynamics (Forkel et al., 2015). I'm wondering if and how these processes are represented in CASA and if you are seeing similar relations on spring NPP.

Authors: Admittedly, we have not explored this in great detail in the context of the present study, but from our own previous research on boreal systems (Buermann et al. 2013), we found that at very large spatial scales spring temperatures do seem to be the most important driver for the onset of growing season consistent with numerous other studies. These findings, together with results from the factorial analysis in the present study (see Section 3.2 in revised ms) was the reason why there was a focus on temperature. At more regional scales, we also found evidence in our previous study for the importance of water availability in determining start of season (e.g., Fig S5 in Buermann et al. 2013), which we believe is consistent with the mentioned findings from Barichivich et al. (2014) and Forkel et al. (2015).

In regards to CASA soil processes, this model has a very simple 'one-layer bucket' scheme for soil moisture. Soil moisture storage typically depends on precipitation (both rainfall and snow melt), evaporation, snow storage and melting, storage of water in the soil and runoff (for details see Potter et al. 1993). For example, if monthly average air temperature is above freezing, it initiates spring snow-melt processes. Processes unique to permafrost dynamics such as freeze–thaw within the soil layer are not represented in the CASA model and this is a limitation. However, permafrost related phenomena are also more regional-scale phenomena (whereas the focus in our study is at very large spatial scales) and it is also not established how well such complex below-ground processes are represented in models that do consider permafrost.

p. 13777, l. 3: Trends in LAI from TRENDY models have been also evaluated against GIMM3g LAI showing diverging regional patterns of greening and browning trends, *especially also in boreal Eurasia (Murray-Tortarolo et al., 2013). Therefore, I would expect similar diverging results for NPP changes from these models. I think it's worthwhile to provide results for individual TRENDY models, and assess their outputs against the observed FAPAR and your NPP estimate in order to draw conclusions only from those models with realistic changes.*

Authors: On the other hand, trend comparisons of LAI from satellite (LAI3g) and the TRENDY ensemble show a broadly consistent picture (see Fig. 7 in Sitch et al. 2015; ref. in ms). In addition, results from Schwalm et al. (2015) show that multi model ensembles weighted by model skill (as suggested by this referee) or 'naïve' ensembles (ignoring model skill) do not show stark contrasting pattern in simulations of the global carbon cycle. This may be explained by the fact that each model has its strengths and weaknesses, and in fact the latter point would make it quite difficult to select a subset of models, as suggested by this referee.

p. 13777, l. 13-16: How does this sentence relate to recent findings about the importance of semi-arid ecosystems for the interannual variability of the net land uptake (Ahlström et al., 2015)?

Authors: This is not straightforward to compare since interannual variability and identification of major shifts in time are not exactly the same. However, that one largely semi-arid region (Northern Africa) provides a significant contribution to a global NPP shift is not inconsistent with the findings in Ahlström et al. (2015).

p. 13778, l. 23-29: I was already wondering before how fire was treated in your models. Did the CASA version use the "GFED mode" to simulate fire dynamics? How is postfire succession modelled? Given the large importance of fire activity on ecosystem dynamics in the two focus regions, it should be worth to assess the potential role of fire on NPP changes. The discussed relation between spring warming/greening and summer fire emissions targets in my opinion to the wrong effect. Although the fire season peaks usually in summer in boreal regions, the years with large fires often show different seasonalities. Moreover, it seems important that the increasing fire activity in boreal regions (Kasischke and Turetsky, 2006) resulted in a larger growth of deciduous trees (Beck et al., 2011) which might result in increasing NPP.

Authors: Clearly fire do play an important role in the carbon balance of ecosystems though typically not through changes in NPP but as carbon losses from standing biomass. Yet, for Northern African woody savannas/savannas the fraction of burned area is on the order of 30% (whereas for most other ecosystems it is in the lower single digits; see van der Werf et al. 2010) and one may thus expect a sizeable impact also on NPP. In CASA such effects of fire on NPP are included through the satellite-based FAPAR.

In regard to carbon losses from fires, we did not use the GFED capacity in CASA since this would require satellite burned area as model driver, which is not available prior 1997. In our original submission, we did mention this important caveat (see Discussion).

p. 13780, l. 1-6: Are these changes in NAO and AO detected and are also significant based on the change detection method?

Authors: We did apply our change point framework also on the AO and NAO indices as well as on climate and FAPAR driver data for the satellite period 1982-2011. We found significant shifts in winter (JFM) AO (1989, p=0.07), Northern Eurasian spring temperature (1989, p<0.001) as well as Northern Eurasian spring FAPAR (1990, p<0.001) further providing evidence for one of our key finding that a substantial shift in the AO played a major role in the large-scale shift in plant carbon uptake across northern Eurasia. We included these results in in the Discussion of the revised ms.

Fig. 2 and Fig. 3: I suggest combining the CASA and TRENDY results in one figure for a better comparability of results.

Authors: Good suggestion, but we do feel that it would display too much information in one Figure, especially in light of added information (seasonal time series) in the revised Figure 3 (see above).

Fig. S1: I'm wondering what is causing the abrupt decreases in NPP over moist tropical Africa and SE-Asia. Can you provide any explanation?

Authors: We do not have an explanation for these more local to regional shifts, and this would require further exploration also in regards to assessing their level of robustness in a first step. Note that there are multiple shifts at more local levels (e.g. across tropical South America), but only a limited number of them are also statistically significant (e.g. compare Fig. 1 in ms and Fig. S1).

Literature:

Buermann,W., Bikash, P. R., Jung, M., Burn, D. H. & Reichstein (2013), M. Earlier springs decrease peak summer productivity in North American boreal forests. *Environ. Res. Lett*. 8, 024027.

Guay KC, Beck PSA, Berner LT, Goetz SJ, Baccini A, Buermann W (2014), Vegetation productivity patterns at high northern latitudes: a multi-sensor satellite data assessment. *Global Change Biolog*y, 20, 3147–3158.

Los, S. (2013), Analysis of trends in fused AVHRR and MODIS NDVI data for 1982-2006: Indication for a CO2fertilization effect in global vegetation. *Global Biogeochemical Cycle*s 27(2), 318-330.

Potter, C. S. et al. (1993), Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, 7, 811-841.

Schwalm, C. R., et al. (2015), Toward "optimal" integration of terrestrial biosphere models. *Geophys. Res. Lett*., 42, 4418–4428. doi: 10.1002/2015GL064002.

Van der Werf, G. R. et al. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997_2009). Atmos. Chem. Phys. 10, 11707_11735, 2010.

Vinukollu, R. K., R. Meynadier, J. Sheffield, and E. F. Wood (2011), Multi-model, multi-sensor estimates of global evapotranspiration: Climatology, uncertainties and trends. *Hydrol. Processes*, 25, 3993–4010, doi:10.1002/hyp.8393.

Zeng, N., F. Zhao, G. J. Collatz, E. Kalnay, R. J. Salawitch, T. O. West, and L. Guanter (2014), Agricultural Green Revolution as a driver of increasing atmospheric CO2 seasonal amplitude, *Nature*, 515, 394–397.

List of all relevant changes

1) Extension of the analyses to include all carbon component fluxes (NPP, Rh, and NEP) for both CASA and Trendy simulations. We added three additional figures (Fig. S2, S5-S6) and an additional table (Table S3) in the revised supplement that show corresponding results.

2) Revision of our estimate of a global late 1980s NEP shift using a combination of residual calculations (e.g. estimating Rh as residual of NPP from CASA and Trendy and NEP from GCP) and direct model estimates (see newly formed Section 3.4 and details are given in the revised Supplement as well).

3) Revised Fig. 3: Added seasonal time series as requested by Ref. Chris Jones

Climate-Driven Shifts in Continental Net Primary Production Implicated as a Driver of a Recent Abrupt Increase in the Land Carbon Sink

W. Buermann[1], C. Beaulieu[2], B. Parida[3] ,D. Medvigy[4], G.J. Collatz[5], J. Sheffield[6], and J.L. Sarmiento[4]

[1]Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK.

[2]Ocean and Earth Science, National Oceanography Centre Southampton University of Southampton, Southampton SO14 3ZH, UK

[3]Civil Engineering Department, Shiv Nadar University, Dadri, 203207 UP, India

[4]Atmospheric and Oceanic Sciences Program, Princeton University, 08540 Princeton, New Jersey, USA.

[5]National Aeronautics and Space Administration, Goddard Space Flight Center, 20771 Greenbelt, Maryland, USA.

[6]Department of Civil and Environmental Engineering, Princeton University, 08540 Princeton, New Jersey, USA.

Correspondence to W. Buermann (w.buermann@leeds.ac.uk)

Abstract

The World's ocean and land ecosystems act as sinks for anthropogenic CO_2, and over the last half century their combined sink strength grew steadily with increasing CO_2 emissions. Recent analyses of the global carbon budget, however, uncovered an abrupt, substantial (~1 PgC yr^{-1}) and sustained increase in the land sink in the late 1980s whose origin remains unclear. In the absence of this prominent shift in the land sink, increases in atmospheric CO_2 concentrations since the late 1980s would have been ~30% larger than observed (or ~12 ppm above current levels). Global data analyses are limited in regards to attributing causes to changes in the land sink because different regions are likely responding to different drivers. Here, we address this challenge by using terrestrial biosphere models constrained by observations to determine if there is independent evidence for the abrupt strengthening of the land sink. We find that net primary production has significantly increased in the late 1980s (more so than heterotrophic respiration) consistent with the inferred increase in the global land sink, and that large-scale climate anomalies are responsible for this shift. We identify two key regions in which climatic constraints on plant growth have eased: northern Eurasia experienced warming, and northern Africa received increased precipitation. Whether these changes in continental climates are connected is uncertain, but North Atlantic climate variability is important. Our findings suggest that improved understanding of climate variability in the North Atlantic may be essential for more credible projections of the land sink under climate change.

1 Introduction

The world's land ecosystems act as a major sink in the contemporary global carbon cycle and, hence, alleviates the rise of atmospheric CO_2 concentrations from global CO_2 emissions and as a consequence climate change (IPCC, 2013). Yet, while critical for society our present understanding of the evolution of the land carbon sink under global change is still severely limited (Le Quéré et al., 2009). This is in part because multiple complex factors can influence the carbon balance of terrestrial ecosystems, including climate change, land-use and land-cover change (forest regrowth, fire suppression etc.), nitrogen deposition, and CO_2 fertilization (Ciais et al., 2013). In this regard, it has been well documented that the land carbon sink (typically inferred as the 'residual' in the global carbon mass balance of fossil fuel and net land use change (LUC) emissions, the atmospheric CO_2 growth rate and oceanic uptake) is quite variable at decadal time scales (Denman et al., 2007). But, in recent global carbon budget (GCB) studies with longer (~ last 5 decades) and annually resolved records a rather abrupt, substantial (~1 PgC yr^{-1}) and sustained strengthening of the 'residual' land carbon sink in the late 1980s has been identified (Sarmiento et al., 2010; Beaulieu et al., 2012). Our overall confidence in this prominent shift, however, is somewhat limited since the 'residual' land sink is the most uncertain term in the GCB. This is because uncertainties embedded in the individual budget terms (e.g. LUC emissions and oceanic uptake) propagate into estimates of the 'residual' land carbon sink (Le Quéré et al., 2013).

Here, we explore if there is further independent evidence for a late 1980s regime shift in the land carbon sink through analyzing carbon fluxes from biospheric models of various complexity and observational constraints (e.g. satellite-based vegetation activity).

Our emphasis is on global pattern of net primary production (NPP) since this key carbon flux is known to be a robust driver of carbon sink variability (Luyssaert et al., 2007; Zhao and Running, 2010). A particular focus is in identifying which land regions may have contributed to the potential shift and what underlying mechanisms may have caused it. Specifically, we analyze data-driven NPP data based on an established satellite-constrained biogeochemical model as well as process-based NPP data from nine terrestrial biosphere models that participated in a recent model intercomparison project 'trends and drivers of the regional-scale sources and sinks of carbon dioxide (TRENDY)' (Section 2).

2 Methods

2.1 Data and models

We analyze temporal patterns in various metrics of the terrestrial carbon cycle based on three independent data sources. First, we analyze data-driven NPP fields based on simulations with the satellite-constrained biogeochemical Carnegie-Ames-Stanford-Approach (CASA) model (van der Werf et al., 2006) for the period of available satellite vegetation data 1982-2011 (Zhu et al., 2013). This updated and extensively validated model runs at a 0.5º spatial resolution on a monthly time step. NPP is a measure of the amount of carbon fixed by plants during photosynthesis and accumulated as biomass. The CASA model is conceptually relative simple and a number of potentially important processes and mechanisms, such as related to nutrients (e.g., carbon and nitrogen) are not considered explicitly (van der Werf et al., 2006). However, factors that influence vegetation productivity may be indirectly captured through the satellite-based fraction of available photosynthetically active radiation absorbed by plants (fAPAR), a key driver in the NPP CASA light use efficiency parameterization (van der Werf et al., 2006). This is demonstrated by two recent studies showing that trends in satellite-based vegetation cover are consistent with expectations of growth enhancement via the CO_2 fertilization effect (Donohue et al., 2013; Los, 2013). Yet there is also new evidence suggesting that in dense forested ecosystems (tropical rainforests) fAPAR may not be fully responsive to the CO_2 fertilization effect (Forkel et al. 2015). Temporally varying driver data used for the CASA simulations include in addition to satellite-based fAPAR (fAPAR3g) (Zhu et al., 2013) also land temperature (CRU TS3.21) (Harris et al., 2014) as well as precipitation and surface solar radiation. We pay particular attention to uncertainties in these observational-based datasets and corresponding effects on NPP estimates. While land surface temperature data are considered to be relatively robust, a substantial limitation is that presently only one consistent satellite fAPAR dataset exists that covers the last 3 decades (Zhu et al., 2013) which we consider a minimum record length for meaningful change point analysis. Nonetheless, to account at least partially for observational uncertainties we evaluated different data sources for precipitation and surface solar radiation which are known to have substantial uncertainties (Wild, 2009; Greve et al., 2014) (see Methods in the Supplement). Based on a final selection of driver datasets (1 fAPAR, 1 temperature, 3 precipitation and 3 solar radiation; see Table S4 in the Supplement) we performed multiple CASA simulations with all possible input combinations (a total of 9 simulations) and analyzed the NPP ensemble mean with our change point methodology. We used the spread in these simulations as a measure of 'observational-based' uncertainty. In a complementary analysis, we also analyzed the heterotrophic respiration (R_h) and net ecosystem production (NEP; estimated as NPP – R_h) ensemble means from these sets of CASA simulations. To ensure that the carbon pools are at steady state, the CASA model was spun up for 250 years using a driver climatology based on the study period 1982-2011.

Second, we analyze a GCB for the period 1959-2011 and consisting of CO_2 emissions from fossil fuel burning and cement production as well as net LUC emissions, atmospheric CO_2 growth rates and oceanic uptake (Le Quéré et al., 2013). Uncertainties in these budget terms are also provided and utilized to estimate uncertainties in net land uptake and the residual land sink (through the sum of squared errors).

Third, we analyze process-based NPP, R_h and NEP data based on ensembles of nine single terrestrial biosphere models that participated in the recent TRENDY model intercomparison project (Sitch et al., 2015). Compared to CASA, the TRENDY models are substantially more complex and also run at significantly shorter time steps to resolve the diurnal cycle needed when coupled within Earth system climate models (Sitch et al., 2015). An important distinction is that in the TRENDY models vegetation characteristics (e.g. fAPAR) are simulated prognostically (unlike in CASA where such information is inferred from satellite observations). In the TRENDY experiments (Sitch et al., 2015) the models were driven with observed climate and atmospheric CO_2 data (S2 experiments) as well as with observed atmospheric CO_2 data only (S1) and in order to isolate the variability due to climate, the difference between these two experiments (S2 – S1) was taken. We analyze the NPP, R_h and NEP ensemble means (based on anomalies) from the nine participating models (Community Land Model 4CN, Hyland, Lund-Potsdam-Jena (LPJ), LPJ-GUESS, ORCHIDEE, Sheffield-DGVM, TRIFFID and VEGAS; for model details see Sitch et al., 2015) and used the spread among them as a measure of 'model-based' uncertainty in our change point framework. Based on a recent comprehensive evaluation against observations, it was found that most of the TRENDY models are capable of simulating the short- and long-term first-order dynamics of the terrestrial carbon cycle (Piao et al., 2013).

2.2 Statistical methodology

We apply a consistent change point methodology on the various metrics of the terrestrial carbon cycle to identify pattern of regime shifts (characterized as abrupt, substantial and sustained changes) and to contrast them to pattern showing either no or more gradual changes. We thus determine in a first step the statistical model that best fits the time series under investigation based on three options: 1) a constant mean, 2) a shift in the mean and 3) a linear trend. While there are numerous alternative statistical models (e.g. shifting trends as seen in satellite vegetation data at local to regional scales (Piao et al., 2011)), our choice of these three models is based on our primary objective to identify large-scale pattern in global and continental carbon fluxes that would be consistent with the recently observed regime shift in the land carbon sink (Sarmiento et al., 2010; Beaulieu et al., 2012).

In the 'shift in the mean' model, the shift is located through a change point detection algorithm that includes discrimination against a trend and the background autocorrelation (red noise) by considering all positions in a time series as a potential change point from 5 to n-5, with n being the record length (Beaulieu et al. 2012b). In a previous study, we found that by restricting the search for change points in this manner detection of spurious shifts at the beginning or end of a series can be avoided (Beaulieu et al. 2012a). In the change point method applied here, we also further developed the Beaulieu et al. (2012b) methodology to account for known explicit uncertainties in the time series under investigation. One important limitation here is that this statistical change point model cannot distinguish between a rather drastic shift (e.g., change from one year to the next) and a more smooth shift over the span of several years. Adding additional parameters could in principle provide more information on the nature of the shift (e.g., smooth versus abrupt) but this would also make the model prone to overfitting given the rather short time series in this study.

The most likely model among the three statistical models fitted is determined based on the Schwarz Information Criterion (SIC), which compares their likelihoods with a penalty for the number of parameters fitted. If the 'shift in the mean' model seems the most likely, we calculate in a second step the direction and magnitude of the shifts (subtracting means prior and after the shift) and the corresponding *P*-value by integrating the full uncertainty of the data using Monte Carlo simulations. To perform the Monte Carlo simulation, we draw 1000 normally distributed synthetic series having the same statistical properties as the time series of interest. A new feature is that the series are simulated with uncertainty additivity: the squared variance of each data point is added to the overall time series squared variance and the square root of this sum provides the synthetic series variance. This therefore takes into account the explicit uncertainties in the various time series under investigation. The change point method is applied to the synthetic time series and a SIC difference between the model with a shift in the mean and no shift is calculated for each time series. This provides a distribution for the SIC difference under the hypothesis of no shift in the mean. The *P*-value is the estimated probability to find a SIC difference at least as extreme as the one observed, under the hypothesis of no change. This methodology assumes that the errors of the model are independent and normally distributed with a constant variance. We test normality of the residuals using the Lilliefors test, the independence is verified using the Durbin-Watson test and the constant variance is verified using a Fisher test. All tests are available and performed using MATLAB. If independence is not respected, we generate synthetic series with the same first-order autocorrelation as observed in the respective time series residuals.

3 Results

3.1 Shifts in data-driven NPP

Applying our change point methodology (Section 2.2) on data-driven global NPP fields reveals a marked spatial clustering of abrupt and sustained increases in NPP across northern Eurasia and northern Africa in the late 1980s (Fig. 1). At more regional levels, the impact of severe disturbance events such as the mountain pine beetle outbreak in the late 1990s in the temperate and boreal forests of western North America (Kurz et al., 2008) is also disclosed (via rapid and sustained decreases in NPP). A similar analysis without constraining to only statistically significant results at the grid point level implies that the coherent pattern of abrupt and sustained NPP shifts across northern Eurasia and northern Africa are spatially even more extensive (Fig. S1 in the Supplement).

This point is further illustrated when we apply our change point framework on data-driven NPP time series representative of large land regions and highlights the important role of the northern extratropics (magnitude of NPP shift: ~0.7 PgC yr^{-1}), and the northern Eurasian continent (~0.5 PgC yr^{-1}) in particular, in the regime shift in carbon uptake by terrestrial plants in the late 1980s (Table 1 and Fig. S2 in the Supplement). While the northern African region also exhibits a robust albeit smaller increase in data-driven NPP (~0.2 PgC yr^{-1}) in the late 1980s, no corresponding NPP shifts are discernable for tropical/southern and global land areas (Table 1). It is well known that factors like ENSO (van der Werf et al., 2004) and volcanic aerosols (Lucht et al., 2002) have a large influence on variability in the carbon balance of terrestrial ecosystems (particularly at interannual time scales) and these phenomena may have also played a role in the prominent late 1980s NPP regime shifts. An analysis that explicitly accounts for such effects however suggests that these two factors are not the causes of the shift, but indicates that a shift in data-driven NPP in the late 1980s emerges also for tropical/southern and global land regions albeit with limited statistical significance (Table 1). In a complementary analysis, we also re-assessed the robustness of the earlier reported abrupt shift in the 'residual' global land carbon sink in the late 1980s (Sarmiento et al., 2010; Beaulieu et al., 2012a) by analyzing a GCB for the period 1959-2011 (Le Quéré et al., 2013) with our change point methodology that also accounts for explicit uncertainties in the individual budget terms (Section 2.2). Results confirm the presence of a regime shift in the global 'residual' land sink (~1-1.3 PgC yr^{-1} depending on statistical treatment) in the late 1980s (Table 1 and Fig. S3 in the Supplement). Taken together, the good agreement in the timing of regime shifts in the global 'residual' land carbon sink and continental data-driven NPP may imply that the latter is a significant driver of the increased terrestrial carbon uptake in the late 1980s.

3.2 Drivers of the late-1980s shift in data-driven NPP

In order to unravel the mechanisms leading to the continental shifts in data-driven NPP, we focus on the two target regions of northern Eurasia and northern Africa that predominantly contributed to this late 1980s shift (see Fig. 1a). A factorial analysis for specific seasons shows that the northern Eurasian continent experienced a marked increase in spring temperatures and spring satellite vegetation activity (fAPAR) in the late 1980s that together drove a substantial increase in spring NPP (Fig. 2a). This relatively sudden springtime warming was also associated with a marked earlier spring onset (~5 days; see Fig. S4 in the Supplement) and the enhanced productivity in the early part of the growing season appears to have also benefited plant productivity in subsequent summers (Fig. 2a and Fig. S4 in the Supplement). Increased plant productivity in both of these seasons contributed predominantly to the pronounced and sustained increases in annual NPP in the late 1980s (Fig. 2a). Over northern Africa including the dry Sahel, marked increases in data-driven NPP during wet and dry seasons that are driven by both increases in rainfall as well as satellite fAPAR triggered a pronounced increase in annual NPP in the late 1980s (Fig. 2b). A closer inspection shows that in the period after this shift rainfall increased specifically during the later portion of the rainy season, which effectively lengthened the more productive growing season (Fig. S4 in the Supplement).

3.3 Shifts in process-based NPP

The exploited biogeochemical model (CASA) for data-driven NPP simulations has a relatively simple structure and provides an integrated view (via satellite fAPAR) of the many interacting factors that influence NPP variability. Further, data-driven NPP estimates are also influenced by observational uncertainties in both satellite (e.g., volcanic aerosols, cloud cover, signal saturation) and key climate driver data that are only partially accounted for in our data-driven NPP simulations (see Section 2.1). We thus explored if process-based terrestrial biosphere models driven by climate and atmospheric CO_2 observations also show evidence of a marked shift in NPP in the late 1980s. Results based on the TRENDY ensembles (Section 2) show that for the satellite period (~ last 3 decades) a NPP shift in the late 1980s emerges as a prominent feature, but only in experiments that capture variability due to climate exclusively (Fig. 3, Fig. S5 and Table S1 in the Supplement). Regional attributions associated with the shift are similar as in the case for data-driven NPP, but differences in NPP sensitivities to climate (inferred from differences in the magnitude of the shifts) are evident (Table 1 and Table S1 in the Supplement). For example, the magnitude of the late 1980s NPP shift in Northern Eurasia based on TRENDY is only about half the size of the corresponding shift in data-driven NPP (Table 1 and Table S1 in the Supplement). One reason for this marked difference may be that seasonal carry-over effects in NPP are largely absent in the TRENDY models; e.g. the late 1980s Northern Eurasian annual NPP shift in TRENDY is solely due to spring contributions (coincident with the time of the climate forcing) whereas for data-driven NPP it is comprised of about equal contributions from spring and summer (Figs 2a and 3a). It should be noted that such seasonal carry-over effects due to an earlier spring onset are indeed observed at multiple eddy covariance flux sites across temperate and boreal ecosystems (Richardson et al., 2010) and that the phenology response (at seasonal and interannual time scales) in data-driven approaches is generally considered more robust (Raczka et al., 2013).

The TRENDY simulations are not restricted to the satellite period allowing us to assess whether the identified late 1980s NPP shifts also emerge as dominant pattern when the study period is extended to the last 5 decades (to be consistent with the time frame of the GCB). Results show that the late 1980s shift over northern Eurasia is a stable pattern. For northern Africa, however, an even more prominent shift is identified in the late 1960s (Fig. S5 and Table S1 in the Supplement). This may suggest that this region by itself is not important enough to influence the global land sink (since there is no evidence for a corresponding shift in the global residual land sink; see Table 1). Further, in TRENDY experiments in which atmospheric CO_2 and climate drivers are varied, the shift appears to be masked by an increasing trend in NPP associated largely with the CO_2 fertilization effect (Fig. 3, Fig. S6 and Table S1 in the Supplement). In fact, the high NPP sensitivity to changes in atmospheric CO_2 concentrations in many of the current generation of terrestrial biosphere models (Arora et al., 2013) and the potential role of nutrient limitations (Zaehle, 2013) and/or climate feedbacks (Smith et al., 2015) in mitigating this sensitivity is presently a subject of intense research.

3.4 Shifts in R_h and NEP

In how far the identified climate-driven regime shifts in NPP in the late 1980s translate into a sustained carbon sink (consistent with the shift seen in the residual land carbon sink from the GCB; Table 1) depends in part on associated responses in key carbon loss fluxes such as R_h which (apart from its dependence on substrate supply from NPP) often depends on climatic factors in a similar fashion as NPP (Lucht et al., 2002). A limitation, however, is that currently no data-driven analog for R_h estimation exists and one has to revert to alternative methods including more uncertain process-based simulations. Nevertheless to estimate the degree at which shifts in NPP may be potentially offset by corresponding shifts in R_h we apply our change point framework also on the R_h as well as NEP fluxes from the CASA and TRENDY simulations (see Section 2). Results show that for the large land regions of interest regime shifts in NPP are repeatedly accompanied by substantial shifts in R_h (Table 1, Tables S1-S2 and Figs S2, S5-S6 in the Supplement) often with 1-2 year lags (seen most clearly when ENSO and volcanic influences are accounted for). A consequence is that corresponding shifts in NEP are often less robust or not detectable (Table 1 and Table S3 in the Supplement). For example, in the case of the two focal regions northern Eurasia and northern Africa, the late 1980s shifts in data-driven NPP are offset by corresponding shifts in R_h at levels of 64-66% and 80-90% (depending on statistical treatment), respectively (Table 1). As stated, the estimated shifts in the R_h fluxes are more uncertain and may represent more upper bound estimates as a new study suggests that carbon models have a tendency to transfer carbon too quickly through the plant-soil systems because of severe biases in simulated soil carbon and/or too high R_h sensitivities to climate (Carvalhais et al., 2014).

At global scale, a shift in R_h can also be estimated as the residual between a NPP shift and a corresponding shift in the residual land sink based on the GCB (Anderegg et al., 2015). Using methods that include such residual calculations as well as our direct model estimates (see Methods in the Supplement), we estimate a shift in global NPP in the late 1980s of 1.14 ± 0.34 PgC yr^{-1} and a corresponding shift in R_h of 0.36 ± 0.48 PgC yr^{-1}. Our best estimate for an associated global shift in NEP is 0.63 ± 0.30 PgC yr^{-1} that amounts to roughly 60% of the magnitude of the late 1980s shift in the residual land sink from the GCB (1.12 ± 0.14 PgC yr^{-1}; based on the three estimates shown in Table 1).

4 Discussion

Our findings provide independent evidence from a biospheric modeling perspective for the abrupt strengthening of the 'residual' land carbon sink in the late 1980s (Sarmiento et al. 2010) and suggest that the underlying driver is a shift in global NPP in response to coordinated large-scale climate shifts. However, the late 1980s climate perturbations may also substantially influence fire regimes, but the paucity of data on burned area and related carbon emissions extending back to the early 80s severely limits estimating corresponding impacts. For northern Eurasia (which is responsible for the largest contribution to the late 1980s regime shift in data-driven NPP), however, it is not anticipated that the observed profound spring warming and greening (inferred through fAPAR) in the late 1980s may have lead to substantial changes in fire emissions since the fire activity peaks later in the season (van der Werf et al., 2006). For northern Africa, changes in fire regimes associated with the late 1980s shift towards wetter conditions may have a substantial influence on net carbon balance albeit with uncertain direction since a shift towards wetter conditions may increase (more fuel load) or reduce (shortening the dry season) fire emissions (Andela and van der Werf, 2014). Models that can potentially quantify this influence are still in their early phase of development. While much uncertainty (specifically pertaining to magnitude) remains in estimating the contribution of climate-driven changes in the major land carbon fluxes to the late 1980s regime shift in the land carbon sink, our regional NPP attributions are consistent with a reported decrease in the interhemispheric gradient in atmospheric CO_2 in the 1990s relative to the 1980s that is attributed to an increase in the northern carbon sink (Wang et al., 2013).

Other factors not related to climate may have also played a role in the late 1980s regime shift of the land carbon sink. A potential large contribution in this regard may be from land-use and land cover changes across northern Eurasia through agricultural abandonment and rapid changes in forest management in the aftermath of the late 1980s post-Soviet collapse. While such processes are accounted for in net LUC emission estimates compiled in the GCB (and therefore included in our analysis; see Table 1) corresponding effects may not be fully captured due to a lack of robust data especially in the period prior the Soviet collapse (Achard et al., 2006). However, at least in the case of agricultural abandonment newly available estimates (Schierhorn et al., 2013) of associated carbon sinks for the post-Soviet period 1990-2009 suggest a minor contribution (~0.03 PgC yr^{-1}).

A remarkable finding is that two key climatic constraints on plant growth (temperature and precipitation) have shifted in the late 1980s in a way as to facilitate an abrupt and sustained increase in continental-scale terrestrial NPP. This bears the question if there is an underlying link that would explain why these large-scale climate pattern varied nearly synchronously. The Arctic Oscillation (AO) is the most important climate mode in the northern extratropics (Thompson and Wallace, 1998) and also a prominent mode in coupled global (Los et al., 2001) and hemispheric (Buermann et al., 2003) climate and satellite vegetation greenness data. Consistent with these results, we find that over the satellite period 1982-2011 the winter AO is tightly correlated with northern Eurasian spring temperatures (r=0.60, P<0.001) and spring fAPAR (r=0.40, P=0.03), respectively (Fig. 4). In the late 1980s, the AO together with its regional manifestation the North Atlantic Oscillation (NAO) (Hurrell, 1995) underwent an extreme shift into their respective positive phases, thereby moving North Atlantic winter storm tracks northward and enabling advection of mild maritime air deep into the northern Eurasian land mass (Thompson and Wallace, 1998). Our results show that the northern Eurasian biomes responded rapidly to the associated substantial spring warming as evidenced through synchronous increases in satellite-based vegetation activity (Fig. 4a). Change point analysis on these drivers of NPP also confirms the existence of this prominent late-1980s shift (showing significant shifts in winter (JFM) AO (1989, p=0.07), Northern Eurasian spring temperature (1989, p<0.001) and Northern Eurasian spring fAPAR (1990, p<0.001), respectively). In the aftermath of this shift, however, spring temperatures and vegetation activity stayed at elevated levels (causing a sustained impact on plant carbon uptake; see Fig. 2a) while the AO/NAO exhibited a negative trend, a fact that may be explained by a more gradual warming response to greenhouse gas forcing that is superimposed on the more oscillatory influence of the AO/NAO.

Northern African wet season rainfall pattern are strongly influenced by Atlantic sea surface temperature (SST) variability (Hoerling et al., 2006). In this regard, the warming of the North Atlantic relative to the South Atlantic that resumed in the late 1980s to mid 1990s caused a northward displacement of the Atlantic intertropical convergence zone (ITCZ) and increased rainfall rates across northern Africa, which led to a recovery from earlier severe drought conditions (Hoerling et al., 2006). This increased moisture supply also led to rapid increases in satellite fAPAR (Fig. 4b). An open question is to what extent AO/NAO and Atlantic SST forcings may have interacted (Xie and Carton, 2004) in the wake of the apparent coordinated regional climate shifts over northern Eurasia and northern Africa in the late 1980s. It is well established that ENSO (van der Werf et al., 2004) and volcanic eruptions (Lucht et al., 2002) have a dominant influence on the terrestrial carbon cycle at interannual time scales and much of recent research has focused on associated links (Cox et al., 2013; Wang et al., 2014). Our findings here may suggest that North Atlantic climate variability and corresponding impacts on adjacent vast land masses may be more important in regards to abrupt, substantial and more sustained shifts in the terrestrial carbon cycle.

5 Conclusions

Our results point to a mechanism whereby North Atlantic climate variability modulates the global terrestrial carbon cycle. New research suggests that a large portion of the variability in the North Atlantic may be externally forced by anthropogenic aerosols (Booth et al., 2012) and the pronounced warming trend in the Arctic regions, known as Arctic amplification (Cohen et al., 2014). Arctic amplification specifically is thought to intensify under climate change (Deser et al., 2010) and this may drive the AO/NAO more into their respective negative phases (Cohen et al., 2014) which, based on our results, would substantially reduce carbon uptake by terrestrial plants and weaken the land carbon sink. This illustrates the pressing need for improved knowledge of North Atlantic climate variability and associated forcing mechanisms in order to more credibly project the evolution of the land carbon sink and carbon cycle climate feedbacks under climate change.

References

Achard, F., et al. Areas of rapid forest cover change in boreal Eurasia. For. Ecol. Manag., 237, 322–334, 2006.

Andela, N., & van der Werf, G. R. Recent trends in African fires driven by cropland expansion and El Nino to La Nina transition. Nature CC, 4, 791–795, 2014.

Anderegg, W. R. L. et al. Tropical nighttime warming as a dominant driver of variability in the terrestrial carbon sink. Proc. Nat. Acad. USA, doi: 10.1073/pnas.1521479112, 2015.

Arora, V. K., et al. Carbon-concentration and carbon-climate feedbacks in CMIP5 earth system models, J. Clim. 26, 5289-5314, 2013.

Beaulieu, C., Sarmiento, J. L., Mikaloff Fletcher, S., Chen, J. & Medvigy, D. Identification and characterization of abrupt changes in the land uptake of carbon. Glob. Biogeochem. Cycl., 26, GB1007, 2012a.

Beaulieu, C., Chen, J. & Sarmiento, J. L. Change-point analysis as a tool to detect abrupt climate variations. [in special issue: Climate predictions: the influence of nonlinearity and randomness] Phil. Trans. R. Soc. A, 370, 1228-1249, 2012b.

Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T. & Bellouin, N. Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. Nature, 484, 228-232, 2012.

Buermann, W., et al. Interannual covariability in Northern Hemisphere air temperatures and greenness associated with El Niño-Southern Oscillation and the Arctic Oscillation. J. Geophys. Res., 108, doi:10.1029/2002JD002630, 2003.

Carvalhais, N., et al. Global covariation of carbon turnover times with climate in terrestrial ecosystems, Nature, 514, 213–217, 2014.

Ciais, P., et al. in IPCC Climate Change 2013: The Physical Science Basis. (eds Stocker, T. F., et al.) 465–570 (Cambridge University Press, Cambridge).

Cohen, J., et al. Recent Arctic amplification and extreme mid-latitude weather. Nature Geosci., 7, 627–637, 2014.

Cox, P. M., et al. Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. Nature, 494, 341-344, 2013.

Denman, K. L., et al. in IPCC Climate Change 2007: The Physical Science Basis (eds Solomon, S., et al.) 499–587 (Cambridge University Press, Cambridge).

Deser, C., Tomas, R., Alexander, M. & Lawrence, D. The seasonal atmospheric response to projected Arctic sea ice loss in the late twenty-first century. J. Clim., 23, 333–351, 2010.

Donohue, R. J., Roderick, M. L., McVicar, T. R. & Farquhar, G. D. CO2 fertilisation has increased maximum foliage cover across the globe's warm, arid environments. Geophys. Res. Lett., 40, 3031–3035, 2013.

Enfield, D. B., Mestas-Nunez, A. M. & Trimble, P. J. The Atlantic Multidecadal Oscillation and its relationship to rainfall and river flows in the continental U.S. Geophys. Res. Lett., 28, 2077-2080, 2001.

Forkel, M., Migliavacca, M., Thonicke, K., Reichstein, M., Schaphoff, S., Weber, U. and Carvalhais, N.: Codominant water control on global interannual variability and trends in land surface phenology and greenness, Glob. Change Biol., 21, 3414–3435, 2015.

Greve, P., et al. Global assessment of trends in wetting and drying over land. Nature Geosci., doi:10.1038/ngeo2247, 2014.

Hoerling, M. J., Hurrell, J., Eischeid, J. & Phillips, A. Detection and attribution of 20th century Northern and Southern African rainfall change. J. Clim., 19, 3989–4008, 2006.

Harris, I., Jones, P. D., Osborna, T. J. & Lister, D. H. Updated high-resolution grids of monthly climate observations – the CRU TS3.10 dataset. Int. J. Climatol., 34, 623–642, 2014.

Hurrell, J. W. Decadal trends in the North Atlantic oscillation regional temperatures and precipitation. Science, 269, 676–679, 1995.

IPCC. Climate Change 2013: The Physical Science Basis: Summary for Policymakers. (eds Stocker, T. F. et al.) (Cambridge University Press, Cambridge).

Kurz, W. A., et al. Mountain pine beetle and forest carbon feedback to climate change. Nature, 452, 987–990, 2008.

Le Quéré , C., et al. Trends in the sources and sinks of carbon dioxide, Nature Geosci., 2, 831–836, 2009.

Le Quéré, C., et al. The global carbon budget 1959–2011. Earth Syst. Sci. Data, 5, 1107–1157, 2013.

Los, S. O., et al. Global interannual variations in sea surface temperature and land surface vegetation, air temperature, and precipitation, J. Clim., 14, 1535–1550, 2001.

Los, S. (2013). Analysis of trends in fused AVHRR and MODIS NDVI data for 1982-2006: Indication for a CO2fertilization effect in global vegetation. Glob. Biogeochem. Cycl., 27, 318-330, 2013.

Lucht, W., et al. Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. Science, 296, 1687-1688, 2002.

Luyssaert, S., et al. CO2 balance of boreal, temperate, and tropical forests derived from a global database. Glob. Change Biol., 13, 2509–2537, 2007.

Piao, S., et al. Changes in satellite-derived vegetation growth trend in temperate and boreal Eurasia from 1982 to 2006. Glob. Change Biol., 17, 3228–3239, 2011.

Piao, S., et al. Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO2 trends, Glob. Change Biol., 19, 2117–2132, 2013.

Raczka, B. M., et al. Evaluation of continental carbon cycle simulations with North American flux tower observations, Ecol. Monographs, 83, 531–556, 2013.

Richardson, A. D., Influence of spring and autumn phenological transitions on forest ecosystem productivity, Phil. Trans. R. Soc. B, 365, 3227-3246, 2010.

Sarmiento, J. L. et al. Trends and regional distributions of land and ocean sinks. Biogeosciences, 7, 2351–2367, 2010.

Schierhorn, F., et al. Post-Soviet cropland abandonment and carbon sequestration in European Russia, Ukraine, and Belarus. Glob. Biogeochem. Cycl., 27, 1175-1185, 2013.

Sitch, S. et al. Recent trends and drivers of regional sources and sinks of carbon dioxide. Biogeosciences, 12, 653–679, 2015.

Smith, W. K. et al. Large divergence of satellite and Earth system model estimates of global terrestrial CO2 fertilization, Nature Clim. Change, doi:10.1038/nclimate2879.

Thompson, D. W. J. & Wallace, J. M. The Arctic oscillation signature in the wintertime geopotential height and temperature fields. Geophys. Res. Lett., 25, 1297–1300, 1998.

van der Werf, G. R., et al. Continental-scale partitioning of fire emissions during the 1997 to 2001 El Nino/La Nina period. Science, 303, 73-76, 2004.

van der Werf, G. R., et al. Interannual variability in global biomass burning emissions from 1997 to 2004. Atm. Chem. Phys., 6, 3423–3441, 2006.

Wang, Y., Li, M. & Shen, L. Accelerating carbon uptake in the Northern Hemisphere: evidence from the interhemispheric difference of atmospheric CO2 concentrations. Tellus B, 65, 20334, 2013.

Wang, X., et al. A two-fold increase of carbon cycle sensitivity to tropical temperature variations. Nature, 506, 212-215, 2014.

Wild, M. Global dimming and brightening: A review. J. Geophys. Res., 114, D00D16, 2009.

Xie, S. P. & Carton, J. A. Tropical Atlantic variability: Patterns, mechanisms, and impacts. Earth's Climate, 121-142, 2004.

Zaehle, S. Terrestrial nitrogen−carbon cycle interactions at the global scale. Phil. Trans. R. Soc. B, 368, 20130125, 2013.

Zhao, M. & Running, S. W. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009, Science, 329, 940–943, 2010.

Zhu, Z., et al. Global data sets of vegetation Leaf Area Index (LAI) 3g and Fraction Of Photosynthetically Active Radiation (FPAR) 3g derived from Global Inventory

Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the period 1981 to 2011. Rem. Sens., 5, 927–948, 2013.

Acknowledgements

We gratefully acknowledge support for this study from the National Aeronautics and Space Administration Carbon Cycle Science Program (grant NNX11AD45G). We also thank Emanuel Gloor, Stephen Sitch, and John Chiang for constructive comments that improved the manuscript. Finally, we thank the TRENDY modelling group for making their data available.

Author Contributions

W.B., C.B., B.P. and G.J.C. designed the analyses. W.B., C.B. and B.P. conducted the analyses. All authors contributed to the writing of the manuscript.

Competing financial interests

The authors declare no competing financial interests.

Table 1. Timing and magnitude of abrupt changes in the terrestrial carbon cycle at global and continental scales. Timing of abrupt change (first data entry) as well as corresponding direction and magnitude (second data entry in units of PgC yr^{-1}) and *P*-values (in brackets; estimated through Monte Carlo simulations) are provided if a 'shift in the mean' model fits the respective time series best (see Section 2). The timing of a shift indicates the first year of a new regime. Shifts that are statistically significant ($P < 0.05$) are highlighted in bold. Additional tests were carried out for assessing the nature and robustness of the shifts including accounting for influences related to ENSO and volcanic eruptions (Covariates), and specifically removing the two years of largest impact of the strong Mt. Pinatubo volcanic eruption in the original time series (No Pinatubo). Plots of all time series analyzed are also provided (Figs S2-S3 in the Supplement).

Region	Original data	Covariates[f]	No Pinatubo[g]
Global Carbon Budget 1959-2011			
Residual land sink[e]	**1989, +1.03 (0.003)**	**1989, +1.28 (<0.001)**	**1989, +1.06 (0.003)**
Net land uptake[e]	**1989, +1.19 (0.004)**	**1989, +1.43 (<0.001)**	**1989, +1.23 (0.001)**
Data-driven (CASA) NPP 1982-2011			
Global	1995, +1.18 (0.239)[c]	1989, +1.12 (0.124)[c]	1989, +1.49 (0.084)[c]
Northern land (>30ºN)	**1988, +0.72 (0.010)**[c]	**1989, +0.62 (0.008)**[c]	**1988, +0.76 (0.003)**[c]
Tropic./south. land (<30ºN)	1995, +0.73 (0.266)[c]	1989, +0.50 (0.526)	1995, +0.70 (0.388)[c]
Northern Eurasia	**1988, +0.53 (<0.001)**[b]	**1989, +0.45 (0.001)**	**1988, +0.54 (<0.001)**[b]
Northern Africa	**1989, +0.20 (0.005)**	**1989, +0.17 (0.003)**	**1989, +0.21 (0.001)**
Process-based (CASA) R_h 1982-2011			
Global	**1996, +0.96 (0.001)**	**1990, +0.80 (0.028)**[c]	**1996, +0.94 (0.001)**
Northern land (>30ºN)	**1990, +0.44 (0.003)**[c]	**1990, +0.42 (<0.001)**[a,c]	**1990, +0.46 (<0.001)**[c]
Tropic./south. land (<30ºN)	**1996, +0.63 (0.003)**	1996, +0.49 (0.054)	**1996, +0.61 (0.002)**
Northern Eurasia	**1988, +0.35 (0.004)**[a,c]	**1990, +0.29 (0.004)**[b,c]	**1988, +0.37 (<0.001)**[a,c]
Northern Africa	**1988, +0.18 (<0.001)**	**1991, +0.14 (0.003)**[b]	**1988, +0.19 (<0.001)**
Process-based (CASA) NEP 1982-2011			
Global	1987[d]	1999, -0.68 (0.122)	1987[d]
Northern land (>30ºN)	1987, +0.30 (0.318)	1989[d]	1987, +0.31 (0.304)
Tropic./south. land (<30ºN)	1999[d]	1999, -0.56 (0.087)	1999[d]
Northern Eurasia	1988, +0.18 (0.061)	1989, +0.15 (0.154)[b,c]	1988, +0.18 (0.074)
Northern Africa	1992[d]	2003[d]	1992[d]

a. Not normally distributed (Lilliefors test, 5% critical level)

b. Variance not constant (F-test, 5% critical level)

c. Residuals not independent (Kruskal-Wallis, 5% critical level)

d. 'Linear trend' or 'constant mean' model fits data better than a 'shift in the mean' model

d. In the global carbon budget (Le Quéré et al., 2013), the net land uptake is estimated as the difference between global fossil fuel emissions and the sum of atmospheric CO_2 growth rate and oceanic uptake, while the residual land sink is the difference between net land uptake and LUC emissions

e. Variability related to ENSO and volcanoes were removed in the original time series through regressions against the multivariate ENSO index and stratospheric optical thickness after Beaulieu et al. (2012a)

f. The two Pinatubo years (1992, 1993) were removed in the original time series prior change point analysis

Figure captions

Figure 1. Spatial pattern of abrupt shifts in data-driven NPP. Maps show (a) timing and corresponding (b) direction and magnitude of abrupt shifts in data-driven (CASA) annual NPP for the satellite period 1982-2011. All robust NPP shifts shown here have passed the two key statistical criteria: (i) a 'shift in the mean' model fits the time series at each grid point best (evaluated trough the Schwarz Information Criterion) and (ii) the shift is also statistically significant ($P < 0.05$) based on Monte Carlo simulations that take into account explicit uncertainties (Section 2). Maps of shifts that passed only the first criteria are also provided (Fig. S1 in the Supplement). In (a), the focus regions northern Eurasia (10ºW-180ºE, 40ºN-70ºN) and northern Africa (20ºW-50ºE, 5ºN-20ºN) are outlined. The shifts are only assessed for the period 1987-2006, since for robust change point detection a minimum span of 5 years of data prior and after a shift is required (Section 2).

Figure 2. Temporal changes in continental data-driven NPP. Panels show annual and seasonal (CASA-based) NPP anomalies corresponding to the (a) northern Eurasian and (b) northern African focus regions (see Fig. 1a). All anomalies are relative to 1982-2011. Shaded contours represent 1σ uncertainties that account for biases in model driver data (Section 2). To understand which factors are mainly responsible for the identified shifts (see Table 1), we performed factorial NPP simulations in which only one model driver is varied whereas all others are kept constant (e.g., 'fAPAR only' corresponds to NPP simulations in which only fAPAR was varied, whereas temperature, precipitation and solar radiation data were kept at their climatological mean values). It should be noted that

satellite fAPAR, a proxy for vegetation cover, is often correlated with climate variables and this places limits on attributing a single model driver to changes in NPP. The first year of a new regime in annual NPP is outlined (thick dark red vertical lines; see Table 1) and the means in NPP anomalies prior and after the shift are also shown (dashed lines).

Figure 3. Temporal changes in continental process-based NPP based on nine terrestrial biosphere models. Panels show annual and seasonal NPP anomalies for the (a) northern Eurasian and (b) North African focus regions, based on ensembles of nine biosphere models that participated in the recent TRENDY model intercomparison study (Sitch et al. 2015). In the annual case, results for two sets of model simulations are shown: one in which climate and CO2 as model drivers were varied (S2) and another one that only takes into account the effect of climate variations (S2 – S1; see Section 2). All anomalies are relative to the 1982-2010 overlapping satellite era to facilitate comparisons with the data-driven NPP simulations (see Fig. 2). Mean ensembles were formed based on anomalies in the single TRENDY models to emphasize temporal changes in NPP and to suppress uncertainties arising from model differences in magnitudes. Shaded contours represent 1σ uncertainties corresponding to the spread in the single TRENDY models. The first year of a new regime in annual NPP is outlined (thick dark red vertical lines; Table S1 in the Supplement) and the means in NPP anomalies prior and after the shift are also shown (dashed lines). In panel (b), the last data point for the year 2010 is omitted since tropical rainfall input data for the TRENDY runs were erroneous for that year (S. Sitch, personal communication).

Figure 4. Synchronous continental shifts in climate and satellite vegetation data and links to North Atlantic climate variability. In panel (a), temporal variations in spring (MAM) temperature and satellite-based vegetation activity (fAPAR) representative of the northern Eurasian target region are plotted alongside the winter (JFM) Arctic Oscillation (AO) and North Atlantic Oscillation (NAO) time series. In panel (b), temporal variations in wet season (May-Oct) precipitation and fAPAR for the northern African target region are plotted alongside the annual Atlantic Multidecadal Oscillation (AMO) time series. Plotted are both annual values (thin dotted lines) and a smoothed time series based on a seven-point binomial filter (thick lines). All time series are standardized anomalies relative to the satellite period 1982-2011. All climate indices time series are obtained from *www.esrl.noaa.gov*. Definitions of the AO and NAO are given in Thompson and Wallace (1998) and Hurrell (1995), respectively. The AMO is a detrended area-weighted average of North Altantic SSTs (0°-70°N) (Enfield et al., 2001).

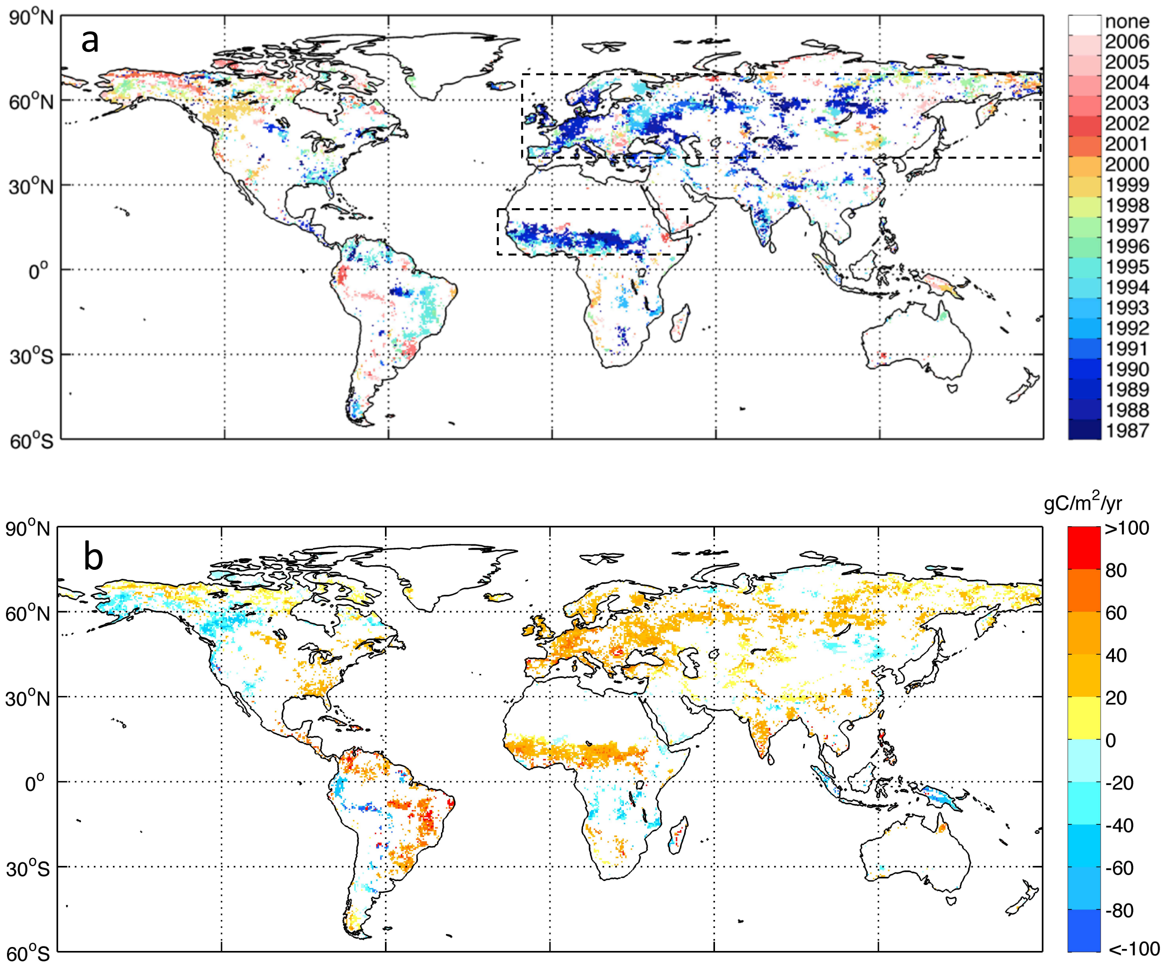
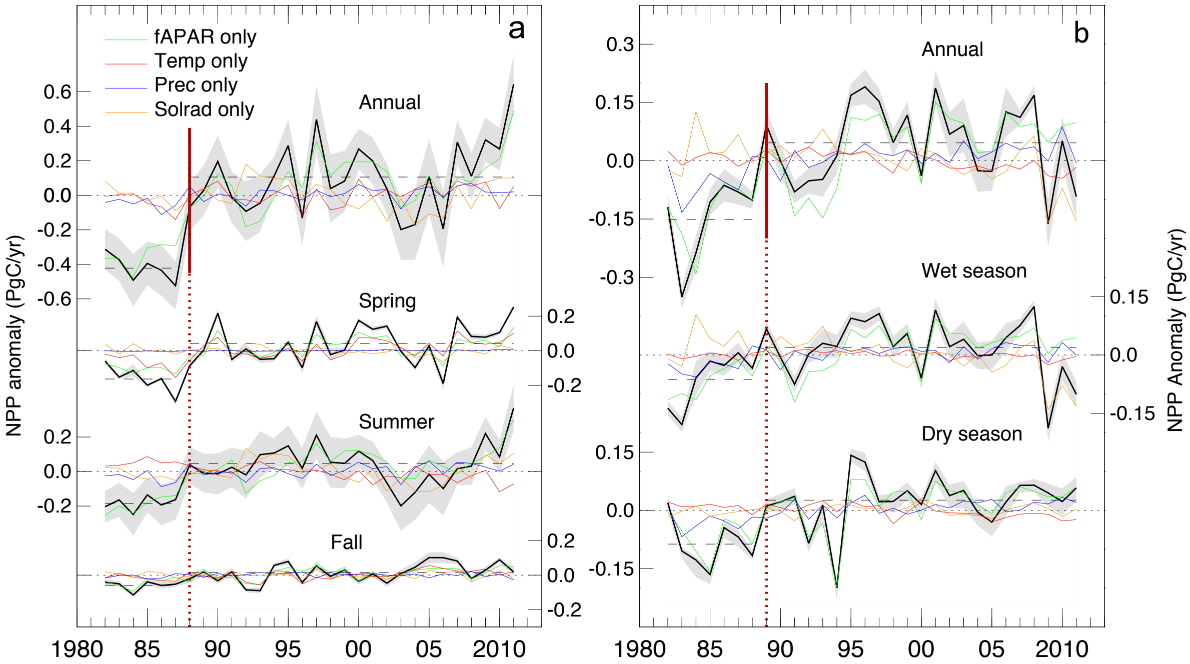
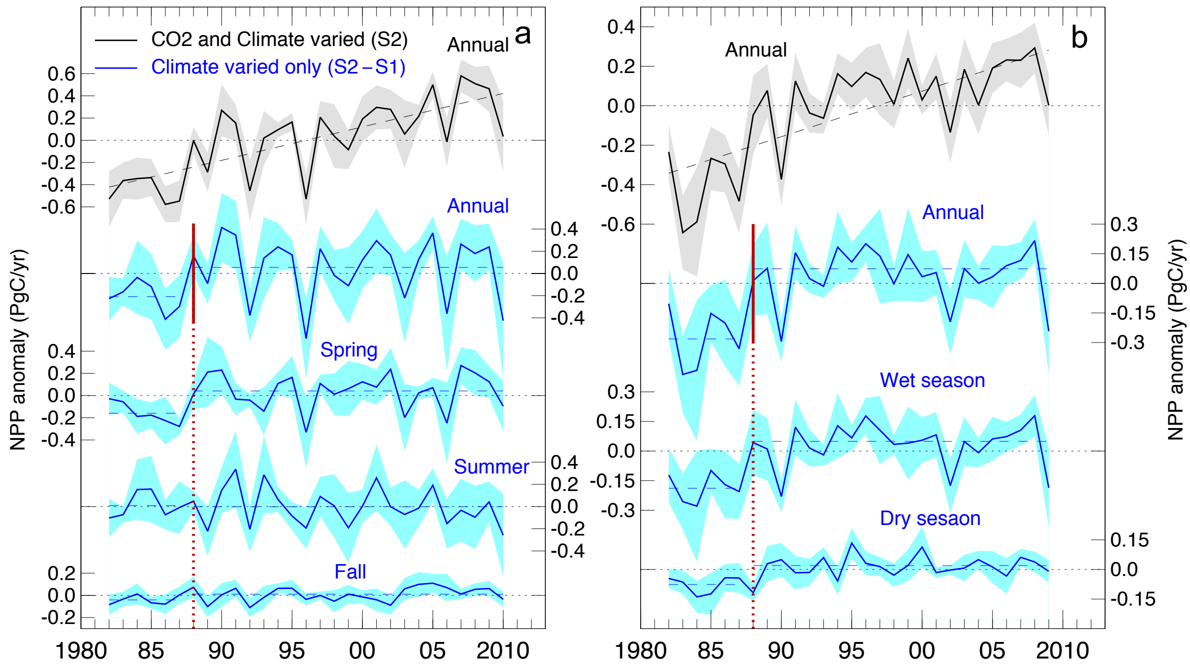

a
b
90°N
60°N
30°N
0°
30°S
60°S
none
2006
2005
2004
2003
2002
2001
2000
1999
1998
1997
1996
1995
1994
1993
1992
1991
1990
1989
1988
1987
gC/m²/yr
>100
80
60
40
20
0
-20
-40
-60
-80
<-100

Figure 1



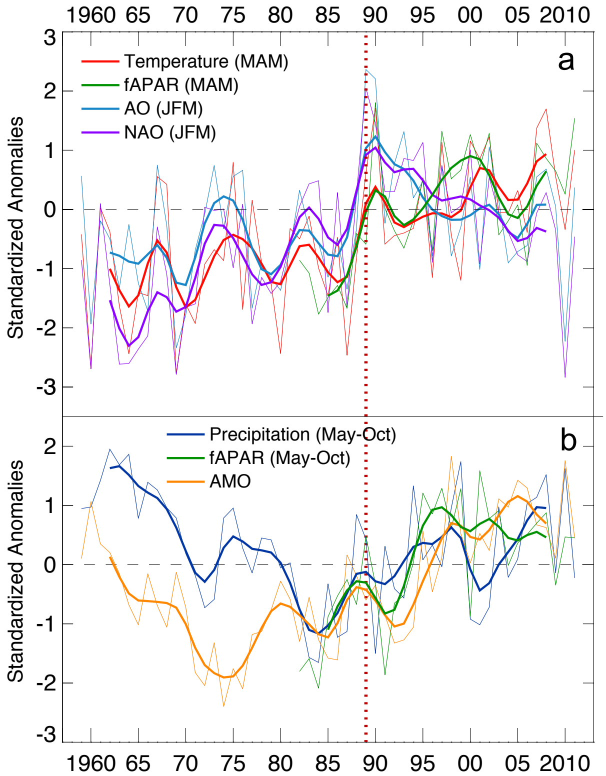
fAPAR only
Temp only
Prec only
Solrad only
a
Annual
Spring
Summer
Fall
NPP anomaly (PgC/yr)
b
Annual
Wet season
Dry season
NPP Anomaly (PgC/yr)
1980
85
90
95
00
05
2010

Figure 2



CO2 and Climate varied (S2)
Climate varied only (S2 – S1)
Annual
a
Annual
Spring
Summer
Fall
NPP anomaly (PgC/yr)
1980
85
90
95
00
05
2010
b
Annual
Annual
Wet season
Dry sesaon
NPP anomaly (PgC/yr)

Figure 3



a
Temperature (MAM)
fAPAR (MAM)
AO (JFM)
NAO (JFM)
b
Precipitation (May-Oct)
fAPAR (May-Oct)
AMO
Standardized Anomalies
1960 65 70 75 80 85 90 95 00 05 2010

Figure 4