Biogeosciences Discuss., 12, C8535–C8548, 2015
www.biogeosciences-discuss.net/12/C8535/2015/

Biogeosciences
Discussions

Interactive comment on "Climate-driven shifts in continental net primary production implicated as a driver of a recent abrupt increase in the land carbon sink" *by* W. Buermann et al.

W. Buermann et al.

w.buermann@leeds.ac.uk

Received and published: 16 December 2015

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. 1). 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.

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 CO_2 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 CO_2 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 CO_2 fertilisation effect was already quite conservative (see Section 2.1).

p. 13771: How is the CO_2 fertilization effect on photosynthesis considered in CASA? Only through the FAPAR forcing dataset or is there an additional module that accounts for CO_2 fertilization? FAPAR might be not sensitive enough to the CO_2 fertilization effect especially in forest ecosystems that have upper FAPAR values. Based on modeling experiments it has been shown that the CO_2 fertilization effect contributes only minor to changes in FAPAR (Forkel et al., 2015). Therefore it might be possible to underestimate the CO_2 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 Biology, 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 Cycles 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.

Interactive comment on Biogeosciences Discuss., 12, 13767, 2015.

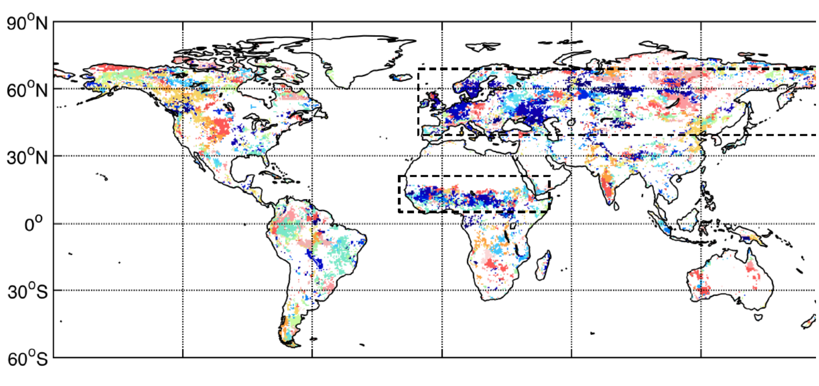

Fig. 1 **Spatial pattern of abrupt shifts in data-driven NPP**. Map
timing of abrupt shifts in data-driven (CASA) annual NPP for the sa
period 1982-2011. Here results are shown using a 5-model approa
disussion). All shifts shown are statistically significant ($P < 0.05$) ba
Monte Carlo simulations that take into account explicit uncertaintie