Response to the first anonymous reviewer

We would like to thank this reviewer for their constructive comments, which have helped us to improve our manuscript. Here we address each question the reviewer raised, one by one.

> The Biogeosciences Discussion paper of Hang et al. reports biophysical constraints over terrestrial primary production. Using a simple model of light-use efficiency and photosynthesis along with latitudinally distributed PAR, fAPAR, elevation etc. they suggest the primary constraint on terrestrial productivity is sparse vegetation cover imposed by water limitation.
>
> Overall I like the simplicity of the approach and analysis. The application of successive constraints on primary production from GPP=ϕ0 ·a·PARtoa to GPP=ϕ ·a·PAR·fAPAR· ci−Γ/ci+2Γ∗ helped elucidate how and why we observe its spatial distribution. Figures 1 and 3 are fascinating. To me this paper provides a quantitative framework for understanding what we've known for [possibly] decades regarding biophysical constraints over terrestrial productivity. The addition of remotely sensed data and the conclusion that potential primary production is most limited by sparse vegetation cover due to water limitation appeared more novel to me. The discussion and implications of the paper fall far short of what I would hope for a paper published in Biogeosciences. Below are my thoughts on why I think the paper needs pretty substantial revision. I hope they are helpful to the authors.
>
> 1. I found the paper through the results section interesting and informative. The Discussion section was, however, remarkably uninformative. It did not deepen fundamental understanding of biophysical ecology or place the current results within the historical context of the field. This was disappointing because the authors seem to have interesting and unique results that should lead to substantial advancements.

This suggestion to improve the Discussion is much appreciated. In response, we have extended the sub-section "Implications for modelling strategy" by three paragraphs. After a brief retrospective on the history of LUE modelling, where we cite a number of additional references, we point out the key difference distinguishing our conceptual model from other LUE models: namely that our model is built on an explicit theoretical basis (the co-limitation hypothesis and the least-cost hypothesis, both independently supported) and testable assumptions (no differences in the controls of LUE among C3 plants, and soil effects on GPP acting via fAPAR rather than LUE). This provides at least two substantial benefits. First, it leads to a simple, traceable model in which *no* parameter is tuned to the observations that the model is supposed to predict. Second, it incorporates the CO_2 effect into a LUE model in a natural way, in contrast with other LUE models where the CO_2 effect either has to be imposed using supplementary equations, or is lacking.

I. On a similar note, the introduction begins with the idea that climate-CO2 feedbacks in global models are poorly constrained, as indeed they are. Section 4.2 discusses some of the current limitations to predicting fAPAR within this context, but there is no meaningful discussion of how the results presented in this study enable better model constraint. Is this possible? Why and how?

This comment also has been addressed by means of a more thorough discussion under "Implications for modelling strategy". We indicate that the conceptual model proposed here provides a first step along the way to achieving a 'next-generation model' with a stronger theoretical and empirical basis. Moreover, we have added (a) a discussion in some depth (citing some highly relevant papers published during the last two years) of how nutrient limitation might be better represented in models and (b) a simple exploratory data analysis on the controls of fAPAR. We have also emphasized the importance of predicting fAPAR as a research goal for next-generation DGVMs.

2. The discussion section considers nutrient limitation on several occasions. It seems largely to argue nutrients are at best a second-order effect, yet the model itself has no nutrients in it [apart from CO2]. Thus it does not seem that much can be made of nutrient limitation in the present study—it was never designed to do so. Rather on this point the discussion is framed around straw men. Would it not be far more interesting to discuss the present results in the context of where and how interactions with nutrients are likely to be manifested?

We thank the reviewer for pointing out this weakness in the manuscript. We have been able to sharpen our thinking on nutrient limitation significantly, while improving the way it is treated in the text. A new sub-section, entitled "[CO_2] and nutrient supply effects", has now been included in the Discussion. Based on a number of recent, independent analyses of observations, we argue that the way forward for the inclusion of nutrient limitations (and potential interactions between nutrient supplies and CO_2 fertilization) in primary production modelling involves making a clear distinction between effects on GPP, NPP and biomass growth, with probably the most important effects of site fertility being manifested through the allocation of NPP – to different plant compartments, and to export to the rhizosphere *versus* biomass growth. We note that this approach is different from the current paradigm of nutrient limitation in DGVMs.

Within this context, three points for your consideration:

I. Huston and Wolverton focus on NPP and ANPP rather than GPP, which is the focus of the present study, so it is not entirely surprising that interpretations regarding [any] controls over productivity differ. I found the third paragraph of the discussion neither insightful nor informative [and borderline disrespectful], so I would recommend deleting it.

Indeed, Huston and Wolverton were focusing on NPP, and moreover we now consider that this distinction is very important – a point that we had not fully appreciated earlier. Therefore, we deleted this paragraph as the reviewer suggested.

II. The discussion of forest FACE +23% NPP stimulation at eCO2 [Norby et al. 2005] ignores (a) follow on papers showing nutrients limit primary production at eCO2 in forest FACE [e.g., Norby et al. 2011] and (b) the broad diversity of long-term responses to eCO2 controlled by nutrients in other studies/ecosystem types. Granted there is a 1-sentence nod to nutrient limitation of CO2 fertilization in the following section. This however does not satisfactorily describe a concept nor its relevance to the interpretation of the data presented here.

We appreciate these detailed comments. The interactions of nutrient availability and the CO_2 response of NPP, or more precisely of biomass growth, constitute a complex subject with a large (and frequently confusing) literature. It would be far beyond the scope of the present manuscript to attempt a review of the subject. (Such a review is planned, perhaps a year from now.) Nonetheless, to avoid the impression of oversimplification, we have included in our revised Discussion reference to the influential paper by Norby et al. (2011) and also to some other papers (Lee et al., 2011, Vicca et al. 2012, Aoki et al. 2012, and Fernández-Martínez et al. 2014) that deal in different ways with the nutrient dependence of primary production and/or the CO_2 fertilization effect.

III. The discussion of modeled vs flux-tower GPP in the 7th paragraph suggests the model's over-estimate cannot be related to nutrients on the basis there is no overestimate in the tropical biome—i.e., black symbols w/GPP>2500 g C m^{-2} a^{-1}, presumably [Figure 2]. Does this contention follow logically? There is good evidence to suggest tropical forest productivity is light limited whereas temperate and boreal forest productivity is nutrient limited. Perhaps the authors are correct, but again this point would benefit tremendously from deeper analysis.

On reflection, we agree with the reviewer that our argument was not watertight. On the other hand we are unconvinced of the logic behind Nemani et al.'s estimation of limiting factors (we already hinted at this in the Introduction). There is evidence for light limitation of primary production in tropical forests, and for N limitation in particular in temperate and boreal forests. But the suggestion has also been made repeatedly that tropical forest growth is limited by P and/or base cation availability, and more generally, there is ample experimental evidence that growth limitation by one factor does not (in defiance of Liebig's so-called 'law') preclude a response to others. This is another large subject, which we have opted to avoid, because at the moment we have not marshalled enough evidence for the deeper analysis that we would like to carry out.

> 3. The discussion of elevation effects was very interesting. Why not a similar approach to the remainder of the discussion section?

See above for a summary of the revisions we have made to the Discussion, which mean that it is now more balanced, covering a broader range of topics in comparable depth.

Response to the second anonymous reviewer

> *I don't think this should be published as a research paper. However, it may have some value as a teaching tool. The paper makes a good start by noting the large uncertainties in the photosynthesis parameterizations used in the CMIP-5 inter comparison, and they note the variety of approaches used to parameterize productivity in existing models. However, the analysis that follows does little to address any of these uncertainties; their estimate of global GPP is an extreme outlier (nearly 2x the mean); the structure of model itself is not innovative (similar to CASA and SDBM), and no new result is presented to demonstrate how existing models could be improved. The sequential approach of narrowing-in, starting from the top of atmosphere solar radiation to consider, atmospheric attenuation, elevation, satellite determined FPAR and temperature limits is implicit in other light use efficiency models. It would be helpful if there were some discussion of the relative merits of this approach relative to other alternatives. While I would agree with the observation (last line of the abstract) that water is the most important factor limiting global productivity, there is no analysis to support this. I also agree with the conclusion that Earth system models should make greater use of satellite data, but again, this is basically a statement of opinion.*

This reviewer has missed the key point that our model is derived from first principles, represented in an extremely simple way – and yet, it is able to reproduce key features of the geographic and seasonal patterns of GPP, suggesting that a great deal of the complexity in current DGVMs could (usefully) be dispensed with. A related point was made in the paper previously published in *Biogeosciences* by Smith et al. (2013). Our MS is complementary to that paper as its focus was on demonstrating traceability of land carbon-cycle parameterizations from observations; our focus is on the theoretical underpinnings of models specifically for the primary production component of the carbon cycle.

In our revision of the MS we have taken some trouble to spell out the ways in which our model is an advance, and we have suggested directions for further development. The main improvements are described in our responses to the first anonymous reviewer.

We respond now to the second reviewer's specific comments, one by one.

1. *the analysis that follows does little to address any of these uncertainties*

Our revision of the MS states more explicitly the following points which were implicit, and apparently not articulated clearly enough, in the first version. (1) As there are large disagreements among DGVMs, their fundamental assumptions cannot all be correct; hence the need to re-examine those assumptions, which can best be done by starting with a simple model developed from a sound and explicit theoretical basis. (2) Current DGVMs rely on many poorly known and implicitly tunable parameters, but the tuning is rarely done transparently or optimally; hence the value of a simpler model (as we present) in which the number of parameters is minimal, and no parameter is tuned to the observations that will be used for evaluation. (3) The complexity of current DGVMs make it generally impossible to track down the sources of differences among models, suggesting the need for a fresh start, building on the huge advances in data availability since most of the current "state-of-the-art" models were originally designed.

2. *their estimate of global GPP is an extreme outlier (nearly 2x the mean)*

The generally accepted range for global GPP is about 110-130 Pg C a^{-1}. Unfortunately, the CMIP5 models' estimates range from about 100 to 210 Pg C a^{-1} with a mean close to 140 Pg C a^{-1}. So although our model estimate is presumably too high, it is only about 1.5 x the mean (and not outside the range) of current "state-of-the-art" models.

3. *the structure of model itself is not innovative (similar to CASA and SDBM)*

This is not quite true. In CASA and SDBM, a maximum light use efficiency (of NPP) is calibrated to observations, and somewhat arbitrary multipliers are used to reduce LUE as a function of temperature and drought. By contrast, our model is derived directly from Farquhar's photosynthesis model for C_3 plants as GPP = $\varphi_0\ I_{abs}\,(c_i - \Gamma^*)/(c_i + 2\Gamma^*)$, where the maximum LUE (φ_0) is based on independent measurements, and growing season temperature, evapotranspiration deficit and elevation modify LUE through their effects on c_i – which are innovatively predicted, based on a theory (and field evidence) first published in 2013.

4. *No new result is presented to demonstrate how existing models could be improved*

An expanded discussion on "implications for modeling strategy" in our revised manuscript now gives pointers as to how existing models could be improved. We argue that ecosystem models should be traceable to observations and explicit hypotheses. We suggest that better models could be achieved by incorporating improved understanding of fundamental plant processes, together with the greatly enhanced availability of large-scale observational data sets. Our model represents a first step towards this goal.

5. *The sequential approach of narrowing-in, starting from the top of atmosphere solar radiation to consider, atmospheric attenuation, elevation, satellite determined FPAR and temperature limits is implicit in other light use efficiency models. It would be helpful if there were some discussion of the relative merits of this approach relative to other alternatives.*

This sequential approach might be "implicit" in other LUE models in the sense that a similar demonstration could be done, in principle, with a different LUE model. However, (a) it has not been done, to our knowledge, (b) the results would be much less interpretable in the case of any model that adopts different parameter values for different biomes or PFTs, such as the MODIS GPP product, (c) the first reviewer evidently holds a different opinion, noting that "the application of successive constraints on primary production ... helped elucidate how and why we observe its spatial distribution" and described Figures 1 and 3 as "fascinating". As for alternatives, it is not obvious to us how else such an exercise could be done.

6. *While I would agree with the observation (last line of the abstract) that water is the most important factor limiting global productivity, there is no analysis to support this.*

The largest constraint on GPP occurs when fAPAR is introduced into the model. But it's true that we did not previously show any analysis to prove that the largest constraint on fAPAR is water availability.

Annual fAPAR is the remotely sensed equivalent of foliage projective cover (FPC). It has been known since Ray Specht's pioneering work in the 1970s (e.g. Specht, 1972) that water availability is the dominant control of evergreen FPC at a continental scale, so we did not think the point was controversial. However, we have now included an analysis of our fAPAR data set by sequential regessions. We show that nearly half of the spatial variability in fAPAR can be explained by the Cramer-Prentice α coefficient, a widely used measure of plant water availability (which estimates the ratio of actual to equilibrium evapotranspiration, similar to the ratio of actual to potential evapotranspiration, proposed by Specht to be proportional to FPC.) Mean growing-season temperature contributes about a further 10% and soil cation exchange capacity (a measure of fertility) an additional 1%.

7. *I also agree with the conclusion that Earth system models should make greater use of satellite data, but again, this is basically a statement of opinion.*

This statement of opinion is also probably not controversial but in any case it is not the main conclusion of this paper. Our main purpose was to analyse the controls on GPP, a topic of central importance for carbon cycle modelling yet one for which there is no consensus in the literature (as we discuss in our Introduction). A subsidiary purpose (which we have made more of in our revision) was to provide some pointers towards the design of a next-generation vegetation model. There is an element of subjective opinion in this too, inevitably, but we do not pretend otherwise.

References

Smith, M. J., Purves, D. W., Vanderwel, M. C., Lyutsarev, V., and Emmott, S.: The climate dependence of the terrestrial carbon cycle, including parameter and structural uncertainties, Biogeosciences, 10, 583-606, 2013.

Specht, R. L.: Water use by perennial, evergreen plant communities in Australia and Papua New Guinea, Australian Journal of Botany, 20, 273–299, 1972.