Author’s response to reviewer comments:

“Towards an advanced atmospheric chemistry-enabled ESM with dynamic land surface processes: Part I – Linking LPJ-GUESS (v4.0) with EMAC modelling system (v2.53)”

by

Matthew Forrest et al.

We thank the reviewers giving time to review the manuscript and for their insightful comments which will significantly improve the manuscript. Here we reproduce the reviewers’ comments in full and address them in turn. The reviewers’ comments are in black, our responses are in light blue. We include proposed alterations to the manuscript to address the reviewers concerns in green. Page and line numbers refer to the original manuscript.

Reviewer 1

Review of the article "Towards an advanced atmospheric chemistry-enabled ESM with dynamic land surface processes: Part I - Linking LPJ-GUESS (v4.0) with EMAC modelling system (v2.53) by Forrest et al.

General comments

The article describes a one-way coupling of the vegetation model LPJ-GUESS with the ECHAM5 atmospheric model that is implemented in the modelling system EMAC. Several aspects of the resulting simulated vegetation are displayed and evaluated to be in good agreement with observations. Additionally, it is pointed out that this is a first important step on the way to build an Earth System Model (ESM) including both atmospheric chemistry as well as dynamic vegetation.

I appreciate this initiative as this ESM will be a very helpful tool to approach many important scientific questions like those listed in the abstract and the introduction. In view of the large effort it takes to construct such an ESM it is also appropriate to report already the first development steps to the modelling community. Also the text is well written. There are only two aspects, which (in my opinion) should be improved in the manuscript before publication.

First, it should be clear from the title and the abstract that no detailed plan to construct an ESM nor any results based on an ESM are presented and that the only content of the article is the coupling of LPJ-GUESS to EMAC and the evaluation of the resulting vegetation. The reader is confused by the structure of the abstract. At the beginning of the abstract ESMs are explained, then it’s mentioned that the coupling of LPJ-GUESS to EMAC is presented, then the development of ESMs is motivated that include dynamic vegetation and atmospheric chemistry, to finish with a sentence that simulated vegetation patterns are in agreement with observations. It would be much more straightforward to describe the contents of the paper first and then to motivate this work or the other way around, but not to mix both aspects in the abstract.

Very good point, the abstract structure should be simplified and perhaps we over-sold the ESM aspect of the work. We propose a new and simpler title:

Including vegetation dynamics in an atmospheric chemistry-enabled GCM: Linking LPJ-GUESS (v4.0) with EMAC modelling system (v2.53)

And we have reformulated the abstract (also taking into account other comments) as follows:

"Central to the development of Earth System Models (ESMs) has been the coupling of previously separate model types, such as ocean, atmospheric and vegetation models, to provide interactive feedbacks between the system components. A modelling framework which combines a detailed representation of these components, including vegetation and other land surface processes, enables the study of land-atmosphere feedbacks under global change. This includes the methane cycle and lifetime and the atmospheric chemistry of reduced carbon; fire effects and feedbacks; future nitrogen deposition rates and fertilisation scenarios; ozone damage to plants; and the contribution of biogenic volatile organic compounds to aerosol load and, via cloud condensation nuclei activation, to cloud formation (e.g., precipitation cycles). Here we present the initial steps of coupling LPJ-GUESS, a dynamic global vegetation model, to the atmospheric chemistry enabled atmosphere-ocean general circulation model EMAC. The LPJ-GUESS framework includes a comparatively detailed individual based model of vegetation dynamics, a crop and managed-land scheme, a nitrogen cycle and a choice of fire models; and hence represents many important terrestrial biosphere processes and provides a wide range of prognostic trace gas emissions from vegetation, soil and fire. When development is complete, these trace gas emissions will form key inputs to the state-of-art atmospheric chemistry representations in EMAC allowing for bi-directional chemical interactions of the surface with the atmosphere. Then the full model will become a powerful tool for investigating land-atmosphere interactions. Initial results show that the one-way, on-line coupling from EMAC to LPJ-GUESS gives a reasonable description of the global potential natural vegetation distribution and reproduces the broad patterns of biomass, tree cover and canopy height when compared to remote sensing datasets. Based on this first evaluation, we conclude that the coupled model provides a suitable means to simulate dynamic vegetation processes into EMAC."

Second, the description of the coupling is incomplete in some aspects. LPJ-GUESS has a daily time step. This should be mentioned in section 2.2 and not only in the appendix. As the atmosphere model resolves the daily cycle, I guess, EMAC is building daily averages at the end of the simulation day and then passing it to LPJ-GUESS. Please describe this. What does it mean in terms of photosynthesis and stomatal conductance? These variables have a strong daily cycle, depend on each other, and also depend on the daily cycle in atmospheric conditions, but they are calculated by LPJ-GUESS on a daily basis.

Yes, we agree that this should be described, and yes, EMAC builds daily averages which are passed to LPJ-GUESS. We propose to add the following text as new paragraph after line 20 on page 4.

"Photosynthesis, respiration and hydrological processes operate on a daily time step and require daily temperature, precipitation and incident short wave radiation. However, monthly climate data may be provided, in which case the model interpolates daily values from the monthly values. In these circumstances, the number of precipitation days in the monthly periods may also be provided to disaggregate total precipitation into distinct rain events. In the case of unmanaged natural vegetation (as simulated here), vegetation dynamics (such as establishment and mortality), disturbance, turnover of plant tissues and turnover between litter pools, and allocation of carbon and nitrogen to plant organs all occur on an annual basis."

What variables are passed from EMAC to LPJ-GUESS? Precipitation, snow, solar insolation (split in visible and NIR?), wind?, temperature (surface temperature, 2m temperature, temperature of the lowest atmosphere level?), atmospheric humidity?, etc. A list of variables could easily be added to section 2.3 and would give the reader much more insight, what coupling of LPJ-GUESS to an atmosphere model means.

Yes, we should explicitly list the variables and also describe how they are calculated (daily averages). We propose to replace the paragraph from line 10 to 17 on page 5 with the following paragraphs:

"To provide appropriate climate forcing for LPJ-GUESS, EMAC calculates the daily mean 2 m temperature, daily mean net downwards shortwave radiation and the total daily precipitation at the end of the simulation day and provided it to LPJ-GUESS. Atmospheric CO_2 concentration and nitrogen deposition are also provided on a daily basis from EMAC to LPJ-GUESS. Thus the LPJ-GUESS land-surface state is forced completely by the EMAC atmospheric state.

In turn, LPJ-GUESS provides fractional vegetation cover, leaf area index, daily net primary productivity and average height of each PFT to EMAC. Parameterisations for determining albedo and roughness length are implemented in EMAC, however they are not enabled in the simulations presented here. Thus, we are demonstrating only a one-way coupling where the land surface state does not affect the atmospheric state. The boundary conditions for the atmospheric model (in particular the surface energy and water fluxes) come from the pre-existing land surface representation.

The overall strategy is to tighten the coupling between LPJ-GUESS and EMAC in well-defined, consecutive steps and to assess the effects of one model on the other in a step-wise and logical manner. Here we report the effect of EMAC climate on LPJ-GUESS. The next step is to enable the albedo and roughness length schemes and to use the vegetation and forest fractions (which are used in the standard land surface scheme to determine the hydrological fluxes) to form a bidirectional coupling of interactive vegetation and climate. The work is underway (Tost et al. 2018) and will be presented in a future publication.

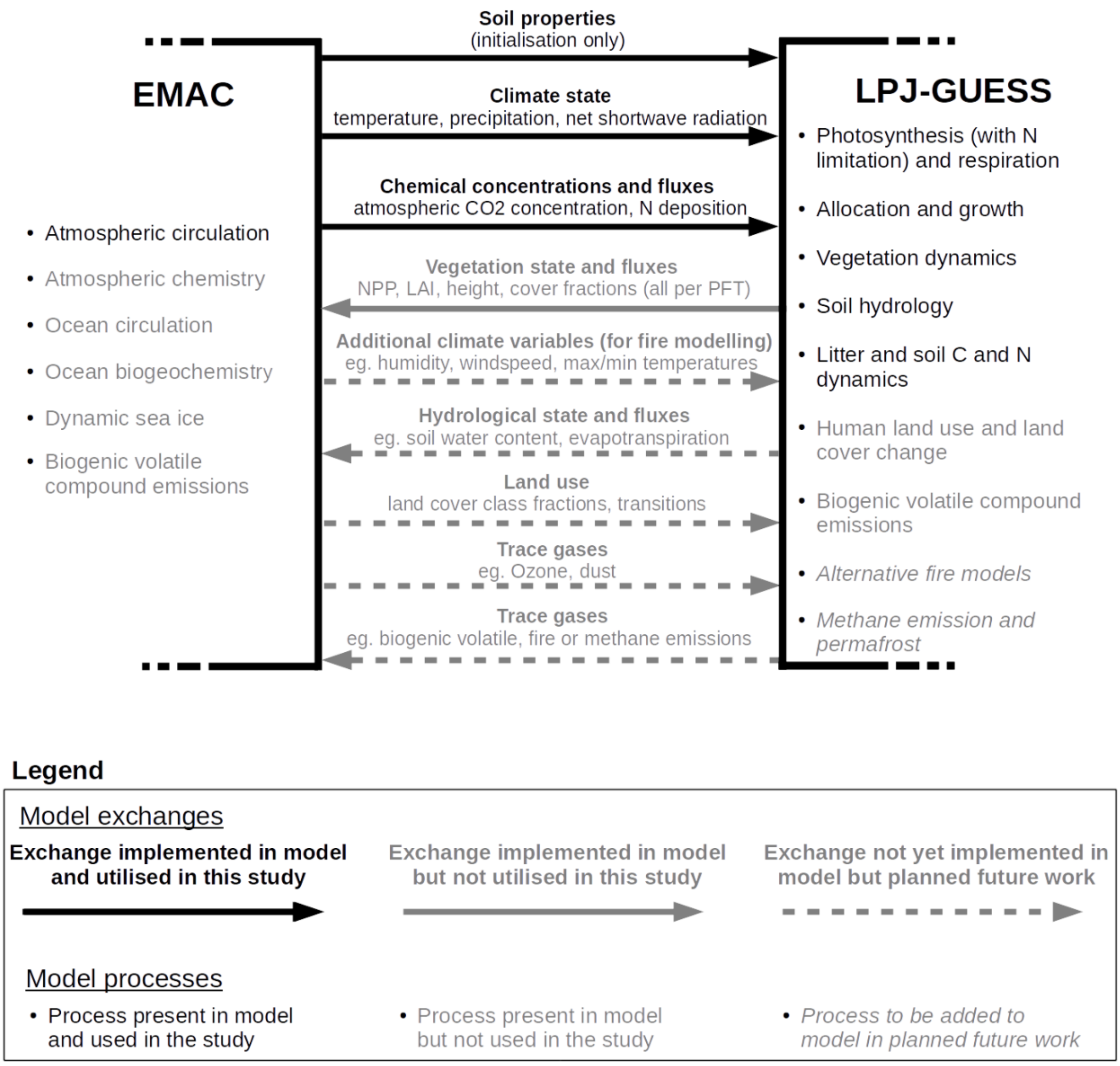
Future planned development steps are to enable land use and agriculture in LPJ-GUESS within EMAC, to include a more process-based representation of fire and include the relevant emissions, to fully replace the soil-vegetation part of the hydrological cycle in EMAC with that in LPJ-GUESS and to use LPJ-GUESS to close the land surface energy balance. When completed, these developments will extend the EMAC model into a full Earth system model including atmosphere (ECHAM5) with comprehensive chemistry (see Jöckel et al., 2010), vegetation and land surface processes (LPJ-GUESS) and an ocean component (MPIOM) (see Pozzer et al., 2011) with ocean biogeochemistry (see Kern, 2013). "

New reference:

Tost, H., Forrest, M., and Hickler, T. (2018) Interactive vegetation influences on climatological meteorological fields and trace gas emissions vol. 20, p. 12047, http://adsabs.harvard.edu/abs/2018EGUGA..2012047T.

The paper shows results for a one-way coupling. No information from LPJ-GUESS are used in the calculations within EMAC during the simulation. That means, that theatmosphere model still needs the old land surface representation (in particular for the surface energy balance calculation). Please mention this. Generally, I think, a diagram illustrating the data flow between EMAC and LPJ-GUESS (also in terms of output) would be very helpful.

Yes, the reviewer is right, information from LPJ-GUESS is not used by EMAC. We explicitly state this in the revised text for section 2.3 as described in the answer to the previous point. Further follow the suggestion that a diagram would be useful and propose to include we include the following figure:



"**Proposed New Main Text Figure 1**: The main processes and exchanges in the coupled model framework. Processes/exchanges with normal black text/black solid arrows are included in the framework and used in the simulations presented here; processes/exchanges with normal grey text/grey solid arrows are included in the framework but not used in the simulations presented here; and processes/exchanges with italic grey text/grey dotted arrows are not included in the framework but planned in future work. All exchanges happen on a daily basis, except for soil properties which happen only during the initialisation phase."

And we propose to refer to the diagram with the following text at the end of section 2.3 (overview of coupling implementation):

"The process and exchanges currently included in the modelling framework, as well as planned future additions, are shown in Fig **Proposed New Main Text Figure 1."**

Specific comments

Please omit "advanced" in the title. It is not explained in the article in what manner the atmosphere chemistry model is advanced.

Done (see revised title above).

Replace in the title "land surface processes" by "vegetation". The article is only concerned with vegetation. Many other dynamic land surface processes that are relevant in Earth System modelling (as lakes/wetlands, permafrost, erosion, hydrological discharge) are not mentioned.

Done (see revised title above).

page 1 line 10: please skip "fully". I don't believe that really everything in your surface description is computed porgnostically and nothing is prescribed (e.g. hydrological soil properties etc.).

Done (see revised abstract above).

page 5 line 10: What is the chemical input from EMAC to the vegetation model? Constant CO2, constant N deposition? Perhaps it's better to specify it as the atmospheric chemistry model in EMAC is not used for this study.

Done, our revised text for section 2.3 explicitly discussed the chemical input (CO_2 and N dep) and states that they are provided daily. We propose to insert the following text at line 25 on page 5 to make it clear that atmospheric chemistry was not fully enabled:

"Note that for reasons of computational burden, the atmospheric chemistry calculations of which EMAC is capable were not activated in these simulations."

page 5 line 12: What is SMIL?

SMIL stands for 'Submodel interface layer', a component of MESSy. The reviewer was correct to flag it, this rather technical jargon has no place in the main text and has been removed during the reformulation of section 2.3.

page 6 line 2: Why do you kill the vegetation in the spin-up run? What does this mean?

LPJ-GUESS starts from 'bare-ground' i.e. no vegetation and no soil C or N pools. But to have vegetation in an N limited model one requires plant available N, it is a bit of a chicken-and-egg situation. To overcome this, the model is run without N limitation (so that vegetation can grow) but with N deposition (to allow N pools to accumulate) for 100 years. This builds up reasonable N pools, but the vegetation has not been limited by N during its development and so is not what we want. Therefore, we 'kill' the vegetation by removing it and putting its C and N into the soil pools and start the vegetation again, this time with N limitation and pre-existing N pools.

To explain this, we propose to change the sentence which runs from line 1 to 3 of page 6 with the following fuller and hopefully clearer explanation:

"Here we followed the standard LPJ-GUESS procedure of starting with 'bare ground', ie. no vegetation and no C or N in the soil and litter pools, and running for approximately 500 years to allow the vegetation to reach equilibrium. Having no plant available N present in the soil at the start of the simulation would inhibit and distort vegetation growth if N limitation was enabled. To overcome this, we follow the standard protocol and run LPJ-GUESS for 100 years without N limitation but with normal N deposition to build up the N pools. After 100 years there is sufficient N in the pools, but the vegetation is inconsistent with the desired state as it has been growing without N limitation. Therefore, the vegetation is removed (and the C and N put into the litter pools), and the vegetation is allowed to regrow, this time with N limitation enabled, for a further 400 years.

page 6 line 26: It would be nice to have some more information about NME scores.

Yes, this a good point, particularly as the meaning of NME scores are inverted compared to r^2 values, ie. an r^2 value of zero means no correlation or explanatory power, but an NME of zero means perfect agreement between model and observation. We propose to insert the following text at page 6 line 27.

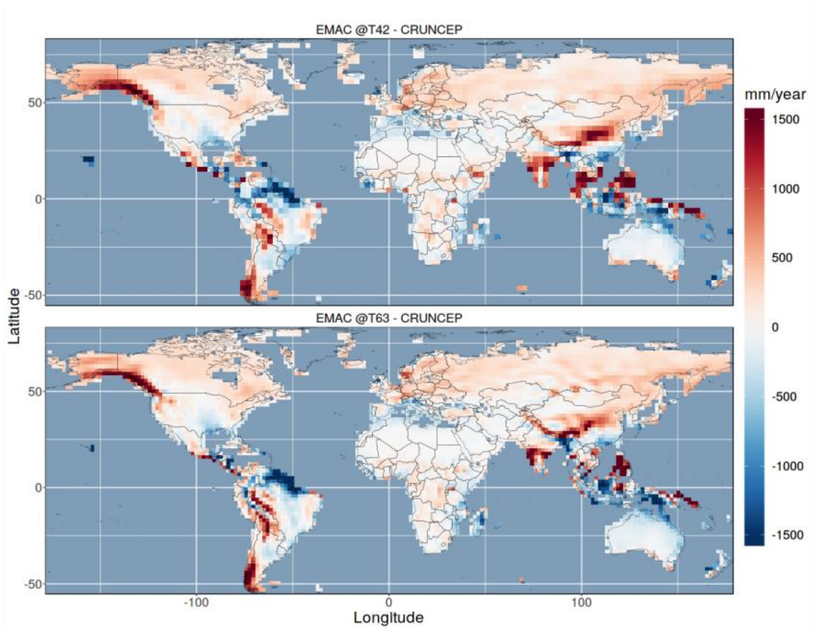
"It should be noted that the NME is rather different from a coefficient of correlation or a coefficient of determination. It does not attempt to derive a correlation but instead sums the differences between the model and the observation. It can be thought of as quantifying the deviation from the one-to-one line of perfect data-model agreement, rather than the deviation from a line of best fit. This means that is a rather direct and unforgiving metric, since every deviation of the model from the data is penalised (uncertainty is not included) and there is no possibility for the line of best fit to move to compensate for systematic biases. It also means the values are interpreted in the opposite direction to a correlation coefficient; an NME score of zero implies perfect agreement between observation and model, whereas an r^2 of zero would imply no correlation between the two. By the normalisation implicit in the method, using the mean value of the observations in place of the model gives an NME of unity."

page 7 line 12: Here it is speculated that a bias in vegetation is caused by a precipitation bias in EMAC. Please mention, how large this precipitation bias in EMAC is (with respect to the bias corrected CRUNCEP data).

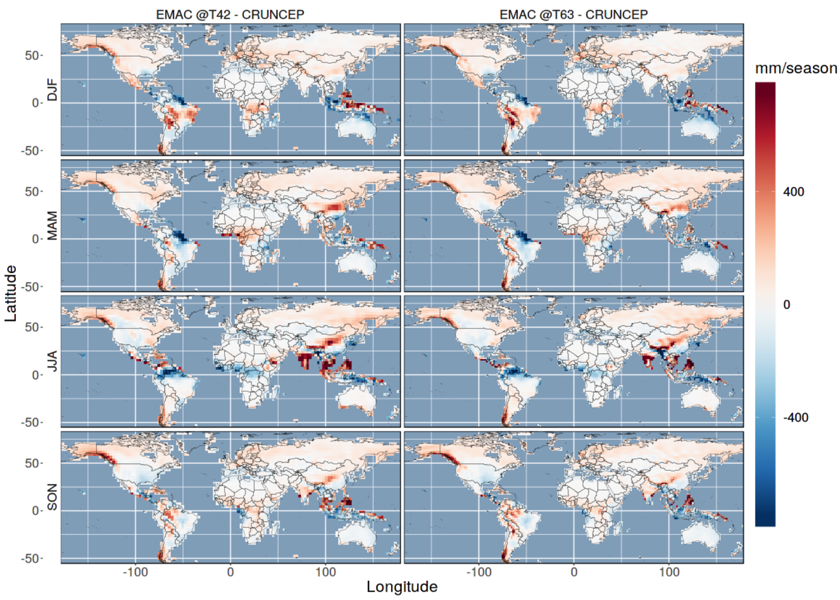
To answer this and other comments by both reviewers, we propose to include an appendix with plots showing the bias between the EMAC produced climate and the CRUNCEP corrected climate. We would like to stress that this is not intended to be a thorough investigation of the biases in EMAC, but rather supporting information to enable a more concrete and less speculative interpretation of our results. In particular, the bias in the net shortwave radiation quantifies the difference in radiation available to the plants using the relevant albedo values, not the gross flux.

Furthermore, we now propose to include results from an 'offline' LPJ-GUESS simulation driven by the CRUNCEP data (but using the same code and settings as the EMAC simulations) following the recommendation of reviewer 2. This also enables a better attribution data-model disagreement in the EMAC simulations. These simulations are referred to as 'CRUNCEP' in the following proposed text.

We propose to include the following bias climate plots:

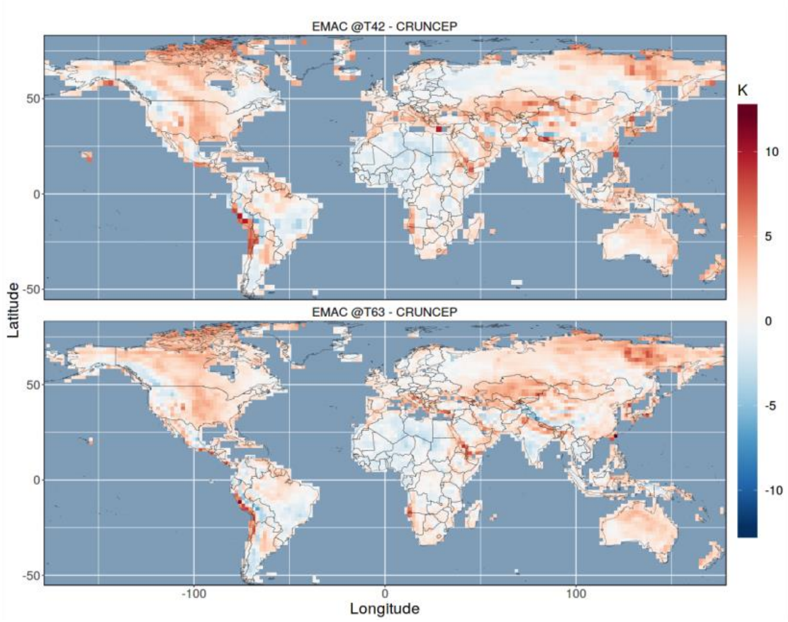


a)

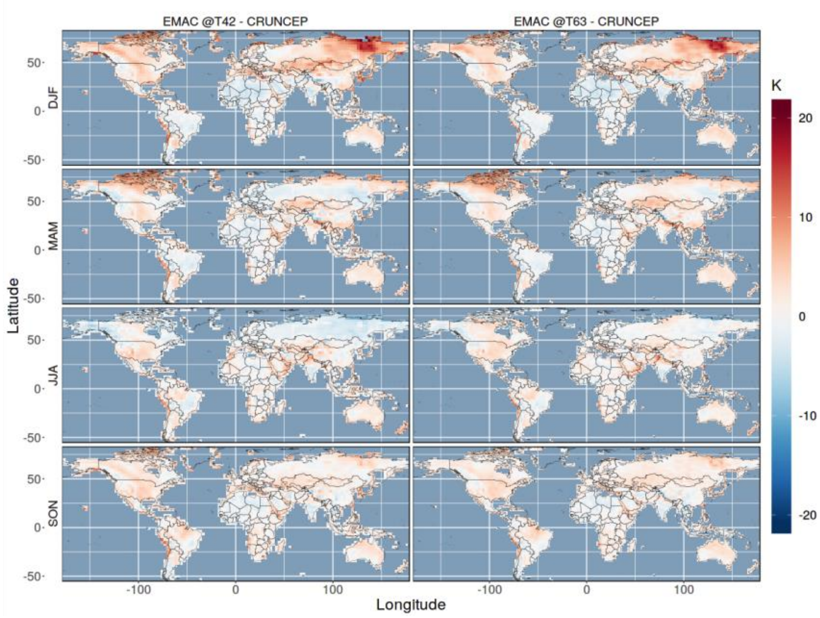


b)

New Appendix Figure 1. The a) mean annual precipitation bias and b) mean seasonal precipitation bias between the observed CRUNCEP (1981-2010) and EMAC simulations (last 50 years of simulation). Note that ensure visibility of relatively low precipitations biases, the plotted values are capped at 750 mm/season and 1500 mm/year in the seasonal and annual plots, respectively.

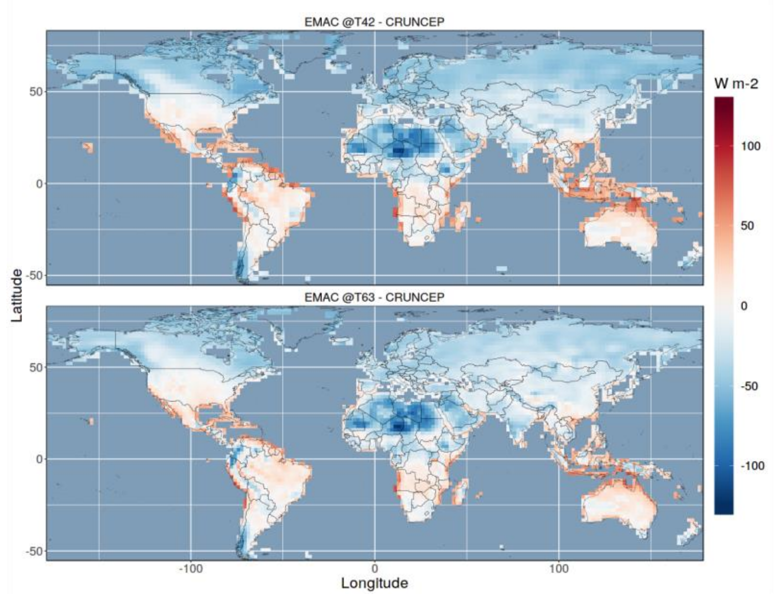


a)

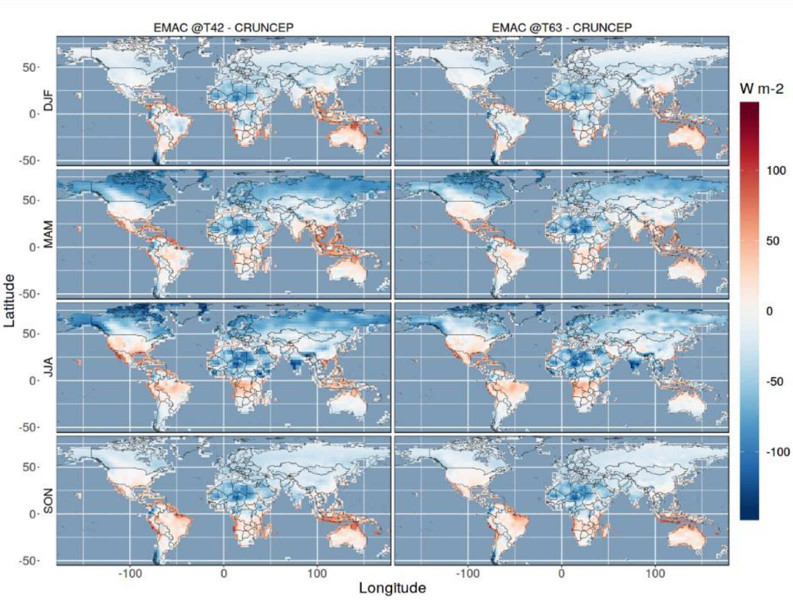


b)

Proposed New Appendix Figure 2. The a) mean annual temperature bias and b) mean seasonal temperature bias between the observed CRUNCEP (1981-2010) and EMAC simulations (last 50 years of simulation).



a)



b)

Proposed New Appendix Figure 3. The a) mean annual net shortwave radiation bias and b) mean seasonal net shortwave radiation bias between the observed CRUNCEP (1981-2010) and EMAC simulations (last 50 years of simulation). Note that these plots compare shows the radiation available for vegetation in the CRUNCEP and EMAC forced LPJ-GUESS simulations, and consequently the gross shortwave radiation has been adjusted by different albedo values. The CRUNCEP radiation has been adjusted using the standard LPJ value of 0.17 applied (temporally and spatially invariant), and the EMAC radiation has been adjusted by the spatially and temporally varying albedo values in the land surface scheme.

We propose to include the following text to introduce these plots as a new final paragraph of section 4 (Model evaluation).

"Whilst it is not within the scope of this work to evaluate the biases of climate state produced by EMAC, knowledge of these biases is very useful for disentangling the causes of model-data disagreement in the simulated vegetation. To this end, we include bias plots of seasonal and annual biases in surface temperature, precipitation and net (plant-available) short wave radiation of the EMAC T42 and T63 climate with respect to the CRUNCEP bias-corrected, re-analysis climate dataset (Appendix B)."

To answer the particular point, we propose to replace the paragraph on page 7 from line 8-18 with the following:

"The simulations reproduce the global patterns of vegetation cover well (Fig. 1), although some regional discrepancies are visible. The most obvious mismatch between all the simulations and the reconstructed megabiomes is the underestimation of the abundance of vegetation in the high northern latitudes. The tendency is relatively small in the *CRUNCEP* simulation, but larger for the EMAC simulations. The higher resolution *T63* simulation is better than the *T42*, as it shows a greater tundra and boreal forest extent. Therefore, this mismatch can be attributed to a high-latitude growing season low temperature and low plant available radiation bias in the EMAC climate (**Proposed new appendix figures 2 and 3**) at low resolution, which is somewhat mitigated at higher resolution due to a better representation of the synoptic scale systems in *T63* (Roeckner et al., 2006).

The extent of the temperate forest of the east coasts of the USA and China was underestimated in the simulations using the EMAC climate, which is not seen in the *CRUNCEP* simulations. Inspection of the climate bias plots (**Proposed new appendix figures 1-3**) shows that this can be attributed to a negative bias in precipitation in the southern areas, and negative bias in the plant available radiation in the northern area of coastal China. Large negative precipitation biases reduce the extent of the tropical forest in Indonesia and Papua New Guinea, and in the north-east coast of South America (not seen in the *CRUNCEP* simulations). In central Africa and the interior of Brazil, the extent of the tropical forests is also much reduced compared to *CRUNCEP*. However, in this case the reasons are not immediately clear from examination of the bias plots, although some seasonal biases in precipitation and plant available radiation are not clearly apparent. The extent of the Savanna and Dry Woodlands vegetation type is also under-simulated in Australia and eastern Africa as a result of a negative precipitation bias, although this extent is also not well represented in the *CRUNCEP* simulation. The extent of the Sahara Desert is also underestimated in the EMAC and *CRUNCEP* simulations, as some areas grow sufficient grass cover to be classified as short grasslands. As this is present in the *CRUNCEP* simulation, it is not related to a climate biases, but rather the over-simulation of grasses in very arid regions by LPJ-GUESS.

Given that the *T42* and *T63* results are broadly similar, for brevity of the presentation the subsequent *T42* results will be omitted from the subsequent spatial benchmarks, although the *T42* summary metric results will be tabulated and discussed."

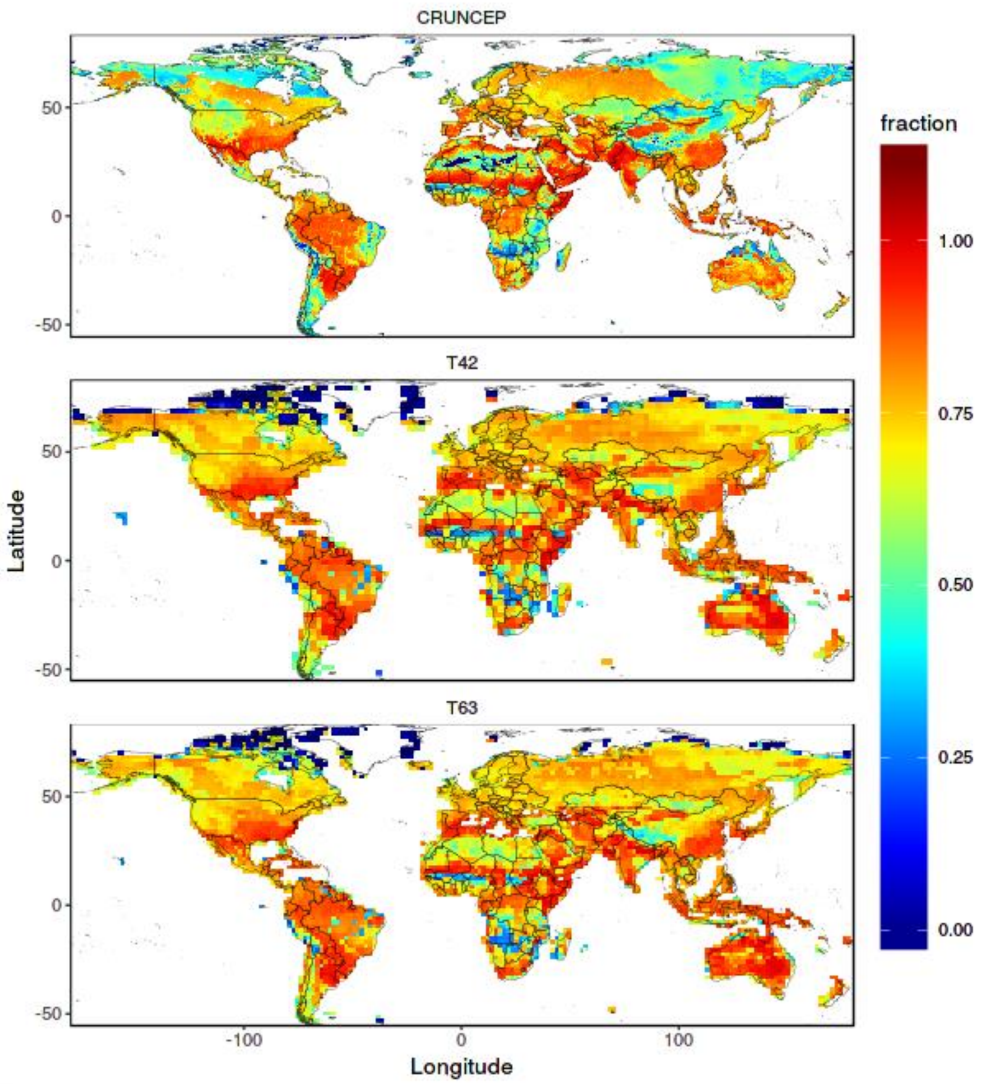
page 7 line 30: Here it is speculated that imperfect process representation in LPJ-GUESS is responsible for a bias in tree cover over North Africa. Please, mention again, if there is also a precipitation bias of EMAC in this region.

It can be seen in the plots above that the precipitation bias in this region is not large. Furthermore, in the new plot showing the tree cover bias in the CRUNCEP simulation, this bias is also observed. We therefore propose the change the last sentence on page 7 to read:

"A final source of disagreement is due to the inevitable imperfect process representations in LPJ-GUESS. This is exemplified by the over-estimation of tree cover north of the central African tropical forest in both the EMAC and *CRUNCEP* forced simulations."

page 9 line 2: Again speculation. Here about N limitation in the tropics causing a bias in biomass. Do you have a nitrogen limitation factor? Is it possible to prove the N limitation in your model output?

The newly included CRUNCEP simulations shows that LPJ-GUESS with observed climate underestimates tropical biomass, but to a much smaller extent. Furthermore, is it possible to determine the N limitation from our model output. To answer this point, and the point of reviewer two, we propose a new figure to be put in an appendix. We propose the following figure, which show the nitrogen limitation as the ratio of the N limited photosynthetic rate ($V_{c,max}$) to the non-N limited photosynthetic rate (weighted across the vegetation in a gridcell by leaf biomass).



Proposed New Appendix Figure 4. The ratio of nitrogen limited to nitrogen non-limited photosynthetic rates (weighted by leaf biomass) in all simulation runs.

This figure shows that in fact N limitation does not appear to be any more limiting in the tropics in the EMAC simulations than in the CRUNCEP simulations. We propose to amend the speculative (and incorrect) sentence on page 9 line 2 to read:

"This underestimation is seen to a lesser extent in the *CRUNCEP* simulation. The different degrees of underestimation of tropical biomass by the *CRUNCEP* and EMAC *T63* simulations cannot be explained by the different nitrogen deposition forcing data (**Proposed New Appendix Figure 4**), however negative seasonal biases in plant available radiation and precipitation are visible in these regions (**Proposed New Appendix Figure 1,3**)."

page 13 line 19: I'm very sceptical about a reduction of climate biases by vegetation dynamics. In most cases/regions there seems to be a positive feedback between climate and vegetation. That means, that climate biases will be enhanced by switching on vegetation dynamics.

We understand the reviewer's point, but we argue that a bias can be reduced by a positive feedback, in the case that bias was caused by the underestimation of a particular process or interaction. For example, a low precipitation bias could be reduced by a positive feedback on the process of precipitation cycling by vegetation.

None-the-less, we accept that our wording was perhaps naïve and over-optimistic, and propose to change the text on page 13, lines 18-20 to use more neutral language and to read:

"Furthermore, using dynamically simulated land surface boundary conditions (in this case from LPJ-GUESS) in a bidirectionally coupled model will alter the atmospheric state and therefore the climate biases. This will be the subject of future studies."

Technical corrections

page 1 line 9/10: "Then it" instead of "At this point, the full model" to shorten the abstract

Done.

page 2 line 32: "a fully" instead of "an fully"

Done.

page 3 line 4: "resulting damage" instead of "resulting to damage"

Done.

page 5 line 11: "not affect the climate" instead of "not effect the climate"

Done.

page 7 line 27" "are the climate" instead of "is the climate"

Done.

page 8 line 9: "to ensure" instead of "to be ensure"

Done.

page 11 line 12: "crown area of trees at 50 m2 in LPJ-GUESS" instead of "crown area of trees LPJ-GUESS is 50 m2"

Done.

page 11 line 28: "affect" instead of "effect"

Done.

page 13 caption of Tab.1: "the vegetation simulated" instead of "the vegetated simulated"

Done.

page 13 line 10: "that in this" instead of "that this"

Done.

page 13 line 10: What is the meaning of PNV?

Potential Natural Vegetation. This acronym is defined on page 6, line 12. But we also now clarify the term in the conclusions (page 13 line 10) be saying:

"PNV (potential natural vegetation with no human impacts)"

page 13 line 13: "utilising the recently" instead of "utilising a the recently"

Done.

page 13 line 25: "bidirectionally" instead of "bi-bidirectionally"

Done.

page 14 line 28: "Creation of three" instead of "Creation of a three"

Done.

page 16 line 23: "biomass due to" instead of "biomass to"

Done.

page 16 line 27: "was defined by the following classes in the Globcover 2009 dataset" instead of "was defined as classes"

Done.

Reviewer 2

First, apologies for this last-minute review; I appreciate that it doesn't allow much time for online discussion, but other commitments prevented an earlier response. Referee #1 has made a number of very good points and I agree with all, although I have bigger difficulties with many aspects of this paper. I here only give additional comments to those made by Ref #1.

Major comments

1. As far as I can tell, the authors are running a non-bias corrected GCM together with a version of LPJ-GUESS that uses pre-industrial nitrogen levels, and only considers 'natural' vegetation (i.e. something is non-existent across large parts of the globe). It is mentioned that the GCM has

temperature and precipitation biases (though little information is given), and of course it also has biases in other climate variables. The abstract concludes that 'initial results show that the one-way, on-line coupling from EMAC to LPJ-GUESS gives a good description of the global vegetation patterns ...'.

If a vegetation model which predicts artificial vegetation is fed with wrong data and anyway gives a good description of global vegetation patterns, isn't something seriously wrong? Or is the comparison just not very discerning?

Yes, we now realise that we over-stated our results and were also unclear in the intention of our comparison. Our intention was not to show the we have a perfect description of the land surface, but rather to show that the implemented coupling works as expected and that forcing LPJ-GUESS with EMAC-produced climate gives sensible continental to global scale patterns in some key vegetation indicators for which global observations are available. Similarly, the inclusion of NME scores was not meant to say that we had sufficient agreement (or better agreement than other models), but instead to quantify the effects of changing resolution and accounting for human land use.

To rectify this, we propose the following changes. The sentence in the abstract flagged by the reviewer would read:

"Initial results show that the one-way, on-line coupling from EMAC to LPJ-GUESS gives a reasonable description of the global potential natural vegetation distribution and the reproduces the broad patterns of biomass, tree cover and canopy height when compared to remote sensing datasets."

We hope this clarifies that for the direct comparison (expert-reconstructed potential natural vegetation) we have reasonable results and that for other comparison (which, as you point, are not apples-to-apples because of the lack of human land use) we are satisfied with the broad patterns.

We also propose to include the following text at line 27 on page 6.

"These summary metrics are not meant to demonstrate a particular level agreement better than some arbitrary threshold. Nor are they mean for strict evaluation or comparison to other models, since here we evaluating only the first milestone of model development and so the model is known to be incomplete (particularly with regard to human land use) and no tuning has been performed. They are included to quantitively evaluate the differences between models runs at difference resolutions and for assessing the effect of human land use via a land cover correction factor (see below)."

Furthermore, we are aware of land use issue and attempt to quantify it by including a land use correction in Table 1 (described in Appendix C). In doing so, we attempted to show that if account for human land use (albeit in a very simple way which assumes that dominant effect of human land use is deforestation) then the agreement between the model and observations improves.

2. The paper states that a human land use and agricultural framework is included in LPJ-GUESS, but not enabled in this study. I cannot understand this. The authors are all from Europe and all the vegetation and land-cover they can see is affected by humans.

The reviewer makes a good point. As mentioned above, we are aware of the land use issue and attempt to quantify it with the land use correction. The changes to the text described above attempt to acknowledge that the model is incomplete in this regard. We absolutely agree that including land use and agriculture would be essential for scientific results derived from the model for the present day, future or recent past (paleo applications would of course be different). However, for the proof-of-concept (or perhaps better said 'proof-of-functioning') demonstration of the model coupling presented here, we don't believe that this is necessary.

3. Similarly, N deposition rates are set for the decade 1850-1859, and seem to be kept constant. Given that nitrogen is a key nutrient, that values have changed enormously since the 1850s, and that LPJ-GUESS can account for this, why proceed with such an artificial assumption?

This was an unfortunate oversight on our part that we only noticed after the model runs were complete. We regret this. It was not a bug in the code exactly, just a mis-specification in the settings file. But similar to the land use point above, we don't believe that the use of pre-industrial N deposition data invalidates our proof-of-functioning, as the results are as expected. Of course, for future studies (including the evaluation of the two-way coupling) we will ensure that temporally-correct N deposition data is used.

To illustrate that using pre-industrial N deposition does not invalidate out results we propose to include a new figure to quantify the degree of N limitation. This figure shows the ratio of the N limited photosynthetic rate to the photosynthetic rate with limitation. We have introduced the figure in answer to a comment by reviewer one (please see **Proposed New Appendix Figure 4** above) and we propose to include this figure as an appendix plot in the revised manuscript.

To answer the reviewer's current point, this figure shows that the patterns of nitrogen limitation are broadly similar with both transient N deposition (CRUNCEP) and pre-industrial N deposition (T42, T63). In fact, in some regions the T63 and T42 simulations actually show less N limitation than the CRUNCEP simulations, showing that climatological factors (such as other constraints on growth and different N mineralisation rates due to different temperatures) have a larger effect than N deposition rates. Therefore, we can conclude that whilst nitrogen is critical for understanding and describing ecosystems, N deposition is not so critical that with only pre-industrial levels one would get wildly different or inaccurate results.

We intend to include the following test to briefly discuss this point and introduce the N limitation figure:

"The degree of nitrogen limitation experienced by the vegetation in the simulations (quantified by the ratio of nitrogen limited to non-nitrogen limited photosythetic rates) is shown in Fig **Proposed New Appendix Figure 4**, which shows that pre-industrial nitrogen deposition does not result in additional nitrogen limitation. However, in future work, temporally-appropriate nitrogen deposition data will be used."

4. Much of the paper is vague about biases in EMAC and their importance. The authors explain (p5, L21) that 'it is expected that such biases will be reduced at higher spatial resolutions', but no evidence or quantification is provided. This is a serious weakness of the paper, and surprising as I most model groups know the biases of their GCMs pretty well these days.

Yes, and we apologise for the vague and sloppy language. We propose to provide bias plots of seasonal temperature, precipitation and net shortwave radiation in an appendix. Please see our response to reviewer one for the bias plots (**New Appendix Figures 1-3**) and accompanying changes to the text.

5. On p6 we read that the simulations correspond loosely to the last couple of decades, in order to gain some insight into biases that may be present when LPJ-GUESS is forced by EMAC climate. But LPJ-GUESS runs over centuries, so how can results from a 20 year simulation (which used constant SST) give much insight into anything?

We apologise for the poor description here (as the reviewer also mentions in point 6.) To clarify here for the reviewers: throughout the simulation period (for all 500 years), LPJ-GUESS was driven by the climate variables that were being dynamically simulated by EMAC. As the reviewer points out, LPJ-GUESS is normally run for centuries and the procedure was no different here. The reference to the 'last couple of decades' was simply referring to the fact that we have SSTs from 1998, CO_2 concentration of 367 ppm (which also correspond to approximately 1998) and we are comparing to satellite data from which the observations were taken in the 2000s.

In order to clear up the confusion we propose the following changes. Addition of a sentence on page 5, line 27 which reads:

"Throughout both the simulations, LPJ-GUESS was driven exclusively by climate variables from EMAC, at no point were external climate datasets used."

Also, modification the sentence starting on page 6 line 21 to read:

"Instead, our goal with this evaluation is to perform steady state simulations where the climate and CO_2 forcing are constant and correspond approximately to conditions in the recent past. Thus, after 500 years of simulation, we can compare the equilibrium vegetation to satellite products based on observations in the early 2000s. Whilst we can't expect perfect agreement since (among other reasons) this is not a full transient simulation, the simulations should be sufficient to check if the model coupling is working as intended, and to gain some insight into biases that may be present when LPJ-GUESS is forced using EMAC climate."

6. Actually, the LPJ-GUESS setup as given on p6, L1-5, is confusing. Here we read about a 400 year run after the vegetation has been killed off, and with nitrogen limitation accounted for. Does this mean the authors used the N-deposition of the 1850s across some period from 1600 to 2000? I think human populations have increased by more than a factor of 10 over this period, and N-deposition should reflect this to some extent.

Yes, we realise that the description of the setup was incomplete. We have proposed revised text in answer to a comment from reviewer one which hopefully makes the details and logic of the spin-up clearer.

However, this alone does not answer all of the concern raised here by the reviewer. To answer the point about N-deposition, yes, it was constant at 1850s levels through the simulation. To explain this, we propose to extend the sentence on page 5 line 31 to read:

"Nitrogen deposition rates were prescribed using data from Lamarque et al. (2013) for the decade 1850-1859 throughout the *T42* and *T63* simulations."

We would also like to emphasise that the simulations presented here are intended to be proof-of-implementation, not transient simulations with the intention to reconstruct a particular time period with great accuracy. We hope that second piece of revised text to the reviewer's point 5 will make this clear in the manuscript. Viewed in this light, the 400 years of spin-up don't correspond to calendar years 1600-2000, but rather they are a simulation period required by the model to allow the modelled vegetation to come into equilibrium with the climate. In such a context, the exact amount of N deposition is not critical, but rather it is important that there is N deposition occurring with reasonable spatial patterns.

And which meteorology was used for the 100+400years of simulation. Was this the constant SST, non-bias corrected EMAC, or was it CRU? When trying to interpret e.g. Fig 1 or indeed all results I really missed this information.

Yes, we apologise that this important information was unclear. Throughout the 100+400 years constant SST, non-bias corrected EMAC meteorology was used. We have added an explicit statement of this fact as described in our answer to the review's point 5.

7. The fair evaluation of LPJ-GUESS would have been in its 'offline' mode, driven by CRU data. These results should also have been presented in Figs. 1-3, so we see how much influence 20 years of EMAC has on the simulations.

We apologise once more for the inexact language and see how the text may have misled the reader into believing that we intended to evaluate stand-alone LPJ-GUESS. This is categorically not our intention, although we can see how the confusion can have occurred with following text at line 10 page 6:

“As LPJ-GUESS has been evaluated in detail in previous work, it is beyond the scope of this work to perform a full model evaluation and propose improvements for the dynamic vegetation model. Instead we performed basic evaluation using …”

This was poorly worded and would lead the reader to believe that we are evaluating standalone LPJ-GUESS. In fact, we want to stress that LPJ-GUESS has already been evaluated extensively and we want to evaluate the functioning of the coupled model system. To clear this up we propose the following text to replace the two sentences starting at page 6 line 10:

“As LPJ-GUESS has been evaluated in detail in previous work, for example, net primary production (e.g. Zaehle et al., 2005; Hickler et al., 2006), modelled potential natural vegetation (Hickler et al., 2006; Smith et al., 2014), stand-scale and continental-scale evapotranspiration (AET) and runoff (Gerten et al., 2004), vegetation greening trends in high northern latitudes (Lucht et al., 2002) and the African Sahel (Hickler et al., 2005), stand-scale leaf area index (LAI) and gross primary productivity (GPP; Arneth et al., 2007), forest stand structure and development (Smith et al., 2001, 2014; Hickler et al., 2004), global net ecosystem exchange (NEE) variability (Ahlström et al., 2012, 2015) and CO 2 fertilisation experiments (e.g. Hickler et al., 2008; Zaehle et al., 2014; Medlyn et al., 2015), it is beyond the scope of this work to perform a full model evaluation and propose improvements for the dynamic vegetation model. Instead we evaluated the coupled model setup to consider how LPJ-GUESS responds when EMAC climate is used as the forcing data and to investigate any biases in the vegetation produced. For this we used an expert-derived Potential Natural Vegetation (PNV) map and using remotely-sensed data sets of tree cover (Dimiceli et al., 2015), canopy height (Simard et al., 2011) and biomass (Avitabile et al., 2016; Thurner et al., 2014)”

New references:

Ahlström, A., G. Schurgers, A. Arneth, and B. Smith (2012) Robustness and Uncertainty in Terrestrial Ecosystem Carbon Response to CMIP5 Climate Change Projections. Environmental Research Letters 7(4): 044008.

Ahlström, Anders, Michael R. Raupach, Guy Schurgers, et al. 2015The Dominant Role of Semi-Arid Ecosystems in the Trend and Variability of the Land CO2 Sink. Science 348(6237): 895–899.

Arneth, Almut, Paul A. Miller, Marko Scholze, et al. (2007) CO2 Inhibition of Global Terrestrial Isoprene Emissions: Potential Implications for Atmospheric Chemistry. Geophysical Research Letters 34(18): L18813.

Hickler, Thomas, Benjamin Smith, Martin T. Sykes, et al. (2004) Using a Generalized Vegetation Model to Simulate Vegetation Dynamics in Northeastern Usa. Ecology 85(2): 519–530.

Hickler, Thomas, Lars Eklundh, Jonathan W. Seaquist, et al. (2005) Precipitation Controls Sahel Greening Trend. Geophysical Research Letters 32(21).

Hickler, Thomas, I. Colin Prentice, Benjamin Smith, Martin T. Sykes, and Sönke Zaehle (2006) Implementing Plant Hydraulic Architecture within the LPJ Dynamic Global Vegetation Model. Global Ecology and Biogeography 15(6): 567–577.

Hickler, Thomas, Benjamin Smith, I. Colin Prentice, et al. (2008) CO 2 Fertilization in Temperate FACE Experiments Not Representative of Boreal and Tropical Forests. Global Change Biology 14(7): 1531–1542.

Lucht, Wolfgang, I. Colin Prentice, Ranga B. Myneni, et al. (2002) Climatic Control of the High-Latitude Vegetation Greening Trend and Pinatubo Effect. Science 296(5573): 1687–1689.

Medlyn, Belinda E., Sönke Zaehle, Martin G. De Kauwe, et al. (2015) Using Ecosystem Experiments to Improve Vegetation Models. Nature Climate Change 5(6): 528–534.

Zaehle, S., S. Sitch, B. Smith, and F. Hatterman (2005) Effects of Parameter Uncertainties on the Modeling of Terrestrial Biosphere Dynamics. Global Biogeochemical Cycles 19(3): GB3020.

Zaehle, Sönke, Belinda E. Medlyn, Martin G. De Kauwe, et al. (2014) Evaluation of 11 Terrestrial Carbon–Nitrogen Cycle Models against Observations from Two Temperate Free-Air CO2 Enrichment Studies. New Phytologist 202(3): 803–822.

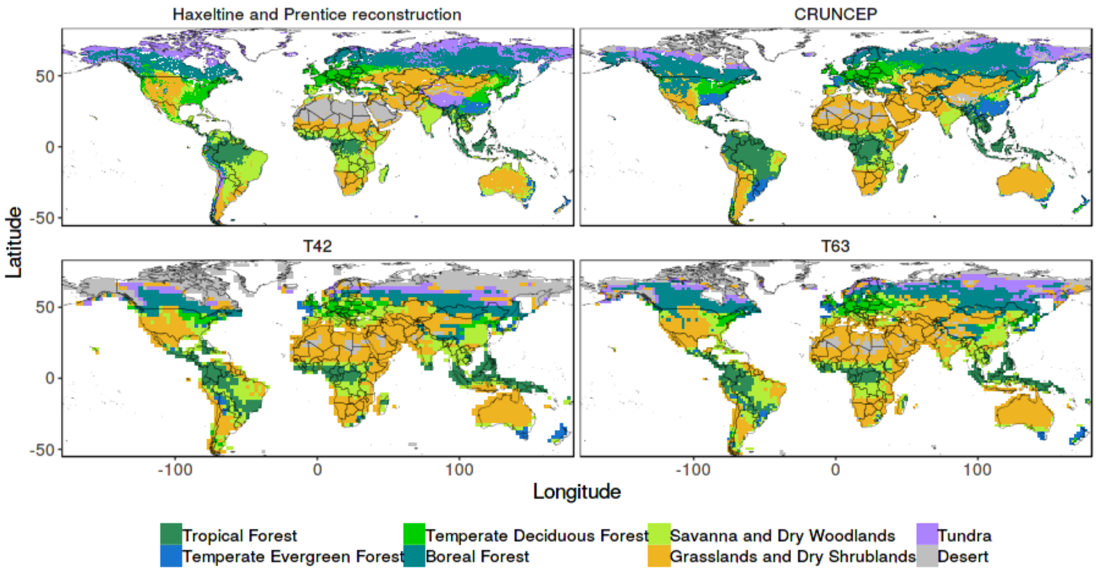
However, we feel that the reviewer makes an excellent point, there is no reason not show standalone LPJ-GUESS driven by CRUNCEP data (and the other standard input data and settings) and doing so would be illuminating. So, in addition, to the above clarification, we have performed such a simulation and propose to include it in the manuscript. To describe this simulation, we propose to replace the lines 6-7 on page 6 with the following text:

To aid the interpretation of the EMAC simulations, we also performed an 'offline' LPJ-GUESS simulation using observed climate data from the CRUNCEP bias-corrected, re-analysis dataset (Wei et al 2014). The simulation was performed using exactly the same code and parameter settings as the EMAC *T42* and *T63* simulations, but code was compiled as a stand-alone model.   The spin-up procedure, CO_2 concentration and nitrogen deposition follow Smith et al. (2014). Note that this is full a transient simulation, after the 500 year spin-up, a further 113 years were simulated using CRUNCEP data. The plots presented here show model output averaged over the years 1981-2010 and are referred to simply as the *CRUNCEP* simulation.

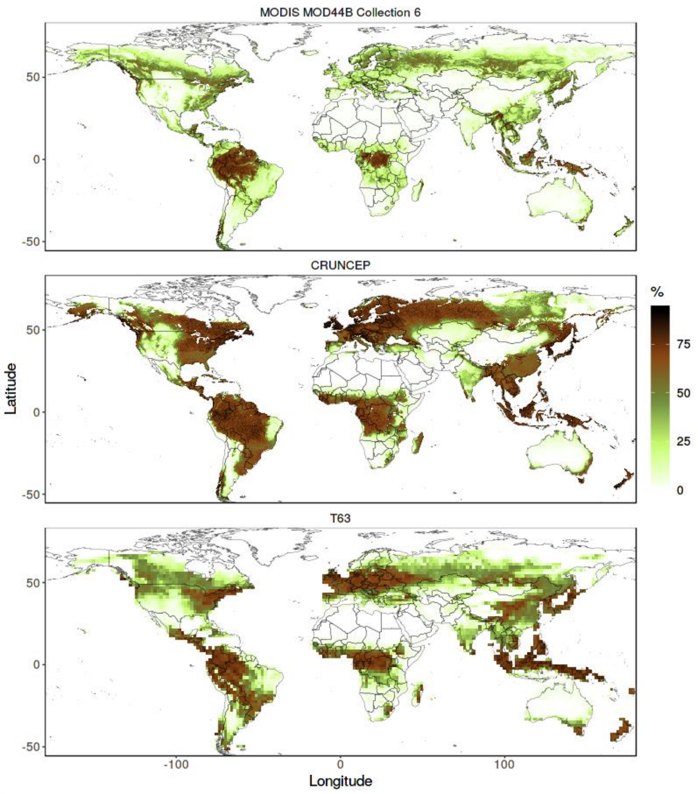
New citation:

Wei, Y., Liu, S., Huntzinger, D. N., Michalak, A. M., Viovy, N., Post, W. M., Schwalm, C. R., Schaefer, K., Jacobson, A. R., Lu, C., Tian, H., Ricciuto, D. M., Cook, R. B., Mao, J., and Shi, X.: The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data, Geoscientific Model Development, 7, 2875–2893, https://doi.org/https://doi.org/10.5194/gmd-7-2875-2014

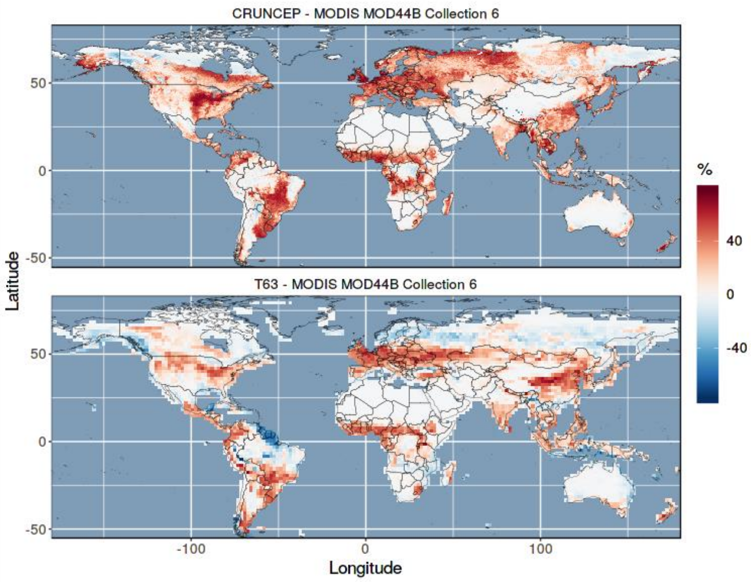
We then propose to include the *CRUNCEP* simulation in Figures 1-4, as follows:



Proposed Modified Figure 1: Distribution of PNV megabiomes simulated by LPJ-GUESS within EMAC *T42* and *T63*) and with observed climate data (*CRUNCEP*) compared to an expert-derived PNV map.

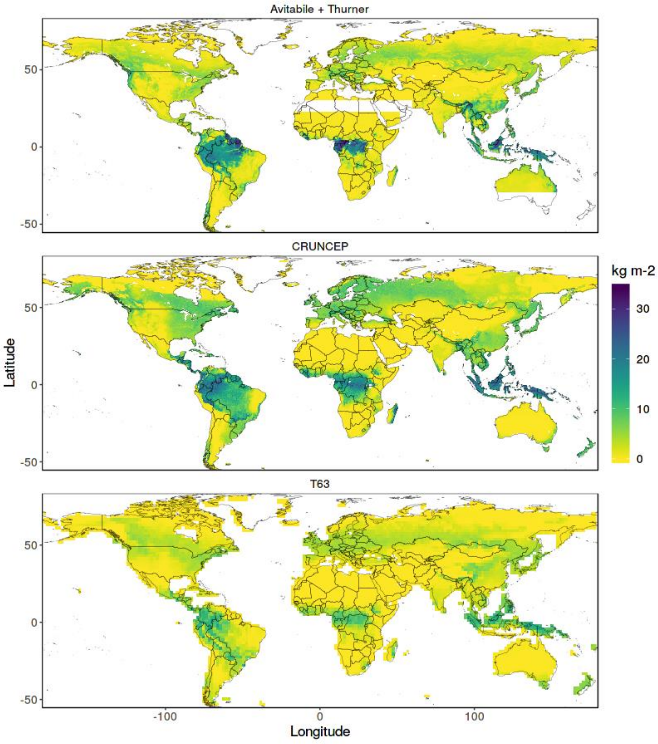


a)

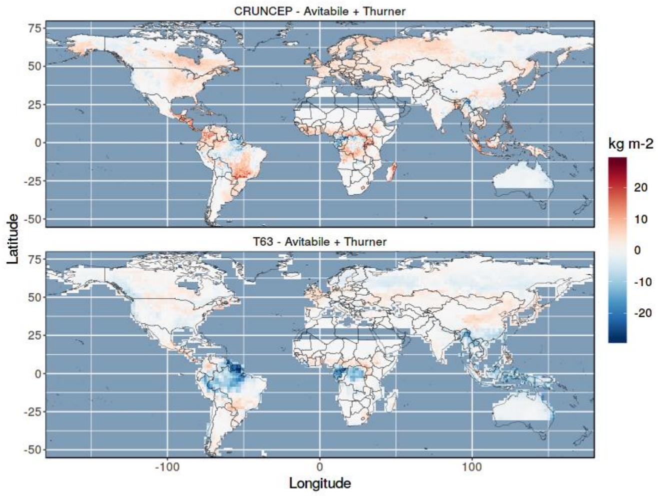


b)

Proposed Modified Figure 2: Comparison of tree cover from the *T63* (EMAC climate) and *CRUNCEP* (observed climate) simulations with observed tree cover from Dimiceli et al. (2015) a) absolute values and b) the difference between simulations minus observations.

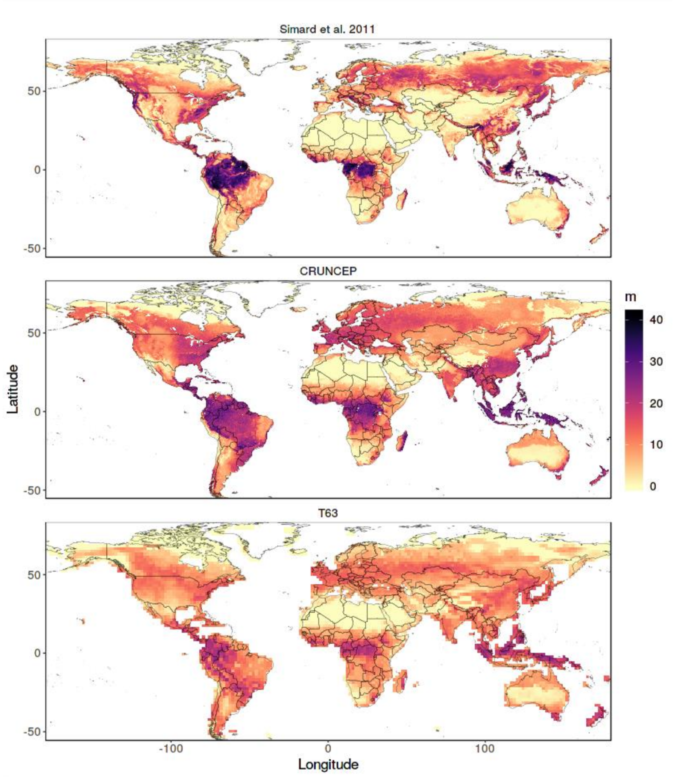


a)

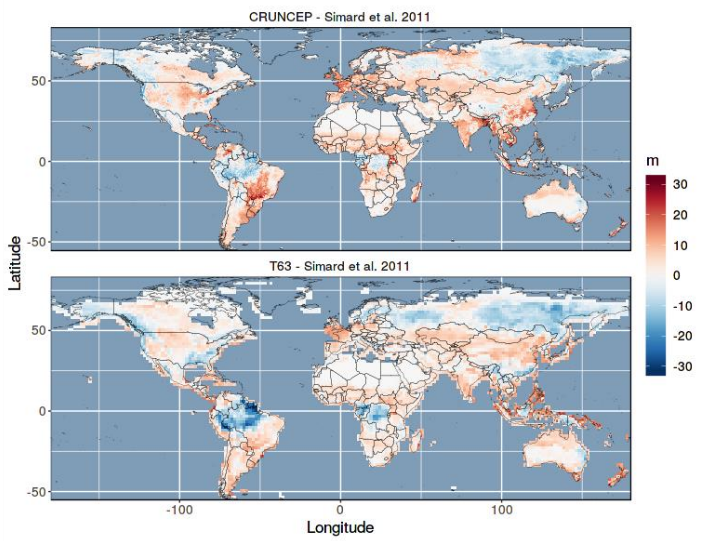


b)

Proposed Modified Figure 3: Comparison of biomass from the *T63* (EMAC climate) and *CRUNCEP* (observed climate) simulations with observed biomass from Avitabile et al. (2016) and Thurner et al. (2014), a) absolute values and b) the difference between simulations minus observation. Note that neither the Avitabile et al. (2016) nor the Thurner et al. (2014) biomass dataset provide biomass for a band across the Sahara, so no data are plotted there.



a)



b)

Proposed Modified Figure 4: Comparison of canopy height from the *T63* (EMAC climate) and *CRUNCEP* (observed climate) simulations with observed canopy height from Simard et al. (2011), a) absolute values and b) the difference between simulations minus observation.

We also propose to include the NME results for the *CRUNCEP* simulation in the results table, giving:

Dataset	without LUC			with LUC		
	T42	*T63*	*CRUNCEP*	*T42*	*T63*	*CRUNCEP*
Tree cover (Dimiceli et al., 2015)	0.94	0.85	1.1	0.81	0.69	0.62
Biomass (Avitabile et al., 2016; Thurner et al., 2014)	0.7	0.8	0.76	0.67	0.7	0.56
Canopy height (Simard et al., 2011)	n/a	n/a	n/a	0.96	0.81	0.77

Proposed Modified Table 1: NME scores for the vegetation produced by the *T42*, *T63* and *CRUNCEP* simulations compared to three remotely-sensed global datasets both with and without a LUC (Land Use Correction) where applicable. Note that lower scores imply better agreement between simulation and observation.

Although I appreciate that GMD is a place to report interim results, I am left with the feeling that this particular work is premature. I think the authors should run their model setup with the various anthropogenic impacts enabled (since they seem to have this capability), and they should properly account for GCM biases, before they compare with today's vegetation maps. Given that they seem to have all the model pieces in place, I cannot see why this wasn't done. And they need to compare LPJ-GUESS+EMAC with LPJ-GUESS+CRU in order to get a better sense of where discrepancies in vegetation cover and characteristics are coming from.

The reviewer makes several important points here which will we address in turn.

We are disappointed that the reviewer feels that this work is premature. We agree that we may have over-stated the degree to which model development is completed and will happily revise the text to further emphasize that the coupled-model presented here is only the first milestone on a development path which will lead to an ESM capable of addressing a broad range of research topics. However, the development work involved in combining these two modelling frameworks was a significant undertaking and included but was not limited to: factoring out the original land surface scheme code in EMAC, re-ordering the space and time loops in LPJ-GUESS , modifying the LPJ-GUESS code to allow restarts within the EMAC framework and at arbitrary points in time (not just the end of the year) whilst ensuring binary identical results between restarted and non-restarted simulations, and writing new input and return functions for LPJ-GUESS. As part of achieving this milestone, it also seemed prudent to perform a broad-strokes evaluation to check that the resulting model was working as expected, to determine if the resulting vegetation state forms a suitable basis for which to modify the land surface properties of EMAC and identify any prominent discrepancies in the vegetation state compared to observations which may inform future development work. Thus, we are making an LPG-GUESS – EMAC coupled system available that can be used and further developed by a larger community. We believe that this development work, combined the first evaluation, is sufficient for publication as a GMD 'Development and technical paper' manuscript (to use GMD's own classification).

To better describe our progress down a larger roadmap we propose the following changes.

On page 2 line 16 change "To construct an ESM with dynamic vegetation and fire," to:

"To take the first steps towards constructing an ESM with dynamic vegetation, anthropogenic influences and fire," The reformulated abstract, including closing sentence:

“Based on this first evaluation, we conclude that the coupled model provides a suitable means to simulate dynamic vegetation processes into EMAC.”

Furthermore, **Proposed New Figure 1** (see response to reviewer 1) and the accompanying text also attempts to make clear what has been done, and what is still to be done.

On the next point, we fully agree that anthropogenic impacts are critical when simulating and studying the Earth system, and including their impacts is essential for any scientific conclusions concerning the anthropocene. Of course, for paleo-applications, where anthropogenic impact is nil, the model would not require these inclusions. We fully intend to perform further development to integrate these processes into the EMAC modelling system via LPJ-GUESS. However, although the processes are included in LPJ-GUESS, enabling them with EMAC is non-trivial. The reasons are somewhat technical but in summary it is because the land use, land use transition and management data are provided to LPJ-GUESS via significantly different input streams than the other inputs, and passing this information through EMAC (and the associated re-gridding) requires significant amounts of further development. Performing a study with various anthropogenic impacts is undoubtedly an excellent and exciting idea but would require significant amounts of development and computing time (a full 500 year T63 simulation requires approximately 2 months on 144 CPU cores) and would amount to an entire study in itself. Such a study would go far beyond the scope of a ‘Development and technical paper’ manuscript presented here.

To answer the reviewer’s comment about ‘accounting for GCM biases’, we point out (and apologise again for not make this clear) that in the manuscript we aim for a simple evaluation of the coupled system, of which GCM biases are an implicit part. Correcting for GCM biases is not the goal, although understand the origins of such biases will inform future development work. We agree that such biases are a critical issue and should be discussed more rigorously, and to this end we propose to include LPJ-GUESS+CRU simulations as suggested (see above).

The whole manuscript also needs to be tightened up, with evidence offered in place of speculation. The example above, where impacts of resolution on a GCM were ’expected’ is a good example. GCM modellers should know and demonstrate such results, not rest upon guesswork.

We acknowledge that the speculation and imprecise language in the submitted draft is unacceptable. As well as including the climate bias plots and LPJ-GUESS+CRU plots suggested by the reviewers to allow discussion of quantified biases wherever possible, we also to propose to tighten up the manuscript as suggested by the reviewer by removing such speculation or supporting it with citations.

With regards to the example quoted above we propose to amend the speculative sentence on page 5 at line 21 to read:

“It is well-known that these biases are dependent on spatial resolution; see Roeckner et al. (2006) for a study of the biases at different resolutions in ECHAM5 (the GCM upon which EMAC was original based). Whilst it is not within the scope of this study to perform a detailed analysis of the biases in EMAC or their dependence on spatial resolution, the impact of the horizontal spatial resolution of the atmospheric simulation on the vegetation simulation is relevant.”

Other comments

p2, L11. It would be good to mention some of the key cycles that 'dynamic' vegetation models often lack too, e.g. not all have N-cycle, and few have P-cycles.

Yes, good point. We propose to amend the sentence to read:

"However, whilst all ESMs by definition have a carbon cycle, not all have truly dynamic vegetation or a nitrogen cycle, fewer still have prognostic representations of fire or a phosphorous cycle."

p2. Seems strange not to mention the EC-Earth ESM, which seems to have come much further in linking LPJ-GUESS inside an ESM model. Are there any links between the work described in this paper and the EC-Earth efforts? What are the similarities and differences in the approaches?

Yes, whilst we cited studies involving both the EC-Earth global ESM and RCA-GUESS regional ESM, these initiatives merit further discussion in this context. We propose replace the sentence at page 2 line 21 which currently reads:

"Furthermore, LPJ-GUESS has already been used with atmospheric models, both global (Weiss et al., 2014; Alessandri et al., 2017) and regional (Wramneby et al.,2010; Zhang et al., 2014) "

With the text:

"Furthermore, LPJ-GUESS has already been used both a global ESM, EC-Earth (Weiss et al., 2014; Alessandri et al., 2017), and a regional ESM, RCA-GUESS (Wramneby et al.,2010; Smith et al., 2011; Zhang et al., 2014). In both modelling systems, climate variables and daily soil moisture from the atmospheric component and its land surface scheme are aggregated over one simulation day and provided to LPJ-GUESS (Weiss et al., 2014; Smith et al., 2011). In the EC-Earth framework, LPJ-GUESS provides only time-varying leaf area index (LAI) to the atmospheric component which initially only affected physiological resistance (Weiss et al., 2014). This link was recently extended to include effects on albedo, surface roughness length, soil water exploitable by roots and snow shading by vegetation (Alessandri et al., 2017). The land surface scheme in RCA-GUESS splits the gridcell surface into two tiles, one of forest and one herbaceous vegetation, and LPJ-GUESS is used to dynamically adjust the LAI within those tiles and relative fractional coverage of needle-leaved and broad-leaved trees in the forest tile. These LAI and fractional cover values affect albedo, surface roughness and heat fluxes in the land surface scheme (Smith et al., 2011).

In the work reported here we have adopted a broadly similar approach with regards to forcing LPJ-GUESS with daily aggregated climate fields from the atmospheric model, although daily soil moisture values are calculated by LPJ-GUESS and not the land surface. In the model version described here, LPJ-GUESS return LAI, vegetation cover fractions, canopy heights and net primary production to EMAC, which allows dynamic calculation of transpiration (by using the vegetation cover provided by LPJ-GUESS as opposed to prescribed vegetation cover) and of albedo and surface roughness (using newly added parameterisations). However, this information is not (thus far) used by the land surface scheme. In other words, although there is two-way information flow and calculation of land surface properties, the model is only effectively coupled in one direction. Enabling the effect of LPJ-GUESS's dynamic vegetation on the atmosphere (via the land surface scheme) is still under development and will be reported in a future publication (for preliminary results see Tost et al., 2018). The integration of LPJ-GUESS into EMAC is independent of the development of the EC-Earth and RCA-GUESS, but we believe that there are possible synergies in terms of future model development. Furthermore, we consider this parallel development to be complementary in terms of scientific applications, in particular the representation of atmospheric chemistry processes in EMAC allows study of land-atmosphere interactions mediated by trace gas exchanges."

New references:

Smith, Benjamin, Patrick Samuelsson, Anna Wramneby, and Markku Rummukainen (2011) A Model of the Coupled Dynamics of Climate, Vegetation and Terrestrial Ecosystem Biogeochemistry for Regional Applications. Tellus A 63(1): 87–106.

Tost, Holger, Matthew Forrest, and Thomas Hickler (2018) Interactive Vegetation Influences on Climatological Meteorological Fields and Trace Gas Emissions. In P. 12047. http://adsabs.harvard.edu/abs/2018EGUGA..2012047T, accessed November 13, 2018.

p3, L11 - refers to a 'companion' publication. As no real reference is given I assume they mean 'future' publication? In my experience these sometimes never appear (even if high-priority), and if one cannot already present an author list and title that can be cited I would re-phrase.

Yes, the reviewer makes a good point. Although the work on the bidirectional coupling for the 'companion publication' is well under way (see above response including reference to Tost et al., 2018), due to unfortunate personal circumstances this work has been delayed. We therefore propose to remove all references to a 'companion publication' and instead refer to a 'future publication' and to remove 'Part 1' from the title (see new title in response to reviewer 1).

p3, L27. The phrase 'tree-individual' sounds odd and is not helpful. Re-phrase.

In saying 'tree-individual' model we wished to inform the read that individual trees are distinguished in the model but individual grasses are not. We accept that this phrasing is not helpful and instead just say 'individual' here and in the revised abstract.

p4, L22. Why 'de facto'. Aren't all components of LPJ-GUESS or EMAC de facto components?!

We apologise for the unclear phrasing. We wanted to convey that GlobFIRM was first developed and used in the distinct but related LPJ-DGVM and can be considered to be an independent model embedded in LPJ-GUESS, but which is enabled by default in most LPJ-GUESS global simulations. We propose to instead simply state:

"The GlobFIRM fire model (Thonicke et al., 2001) is included in LPJ-GUESS."

p15, L15. Shouldn't you say 'This will extend' rather than 'This extends'? If the model is already a full ESM I don't see why you are reporting on the very limited and artificial setup you have here.

The reviewer is absolutely correct, we will follow the suggestion to say 'This will extend'.

p5, L26-27. Why keep a constant CO2 when feeding a vegetation model; the seasonal cycle is well known and documented. Any why 367 ppm? Values are over 400 ppm these days.

To answer the reviewer's questions: CO_2 was kept constant because we are producing a steady state, rather than a transient, simulation. The value of 367 ppm was chosen to be broadly consistent with the prescribed SST of 1998. For transient simulations CO_2 will of course not be kept constant. Values of over 400 ppm are not relevant in this study as we are not seeking to reproduce the last few years of climate. Finally, LPJ-GUESS normally takes a single global annual value of CO_2, and that is the approach repeated here. However, the reviewer is correct to point the seasonal cycle of CO_2 with amplitude of about 5 ppm. Since CO_2 is provided daily from EMAC to LPJ-GUESS, future simulations can include the seasonal cycle (and also spatial variation) for increased accuracy and its effect can be quantified.

We hope that the revised text included in our response to the reviewer's point 5 adequately explains that we are performing steady state simulations and motivates our choice of CO_2 value.

p6. Also, isn't LPJ-GUESS a stochastic model? If so, how was the randomness of the results accounted for?

The reviewer is correct to point out that LPJ-GUESS is stochastic (although the stochasticity is generating using a pseudo random number generator with a defined starting seed to give reproducibility between model runs). Randomness was handled in the standard fashion for LPJ-GUESS; by simulating multiple patches and averaging them. We propose to extend the sentence on page 5 lines 31-32 to read:

"… and for each gridcell 50 replicate patches were simulated and averaged to account for model stochasticity."

p6. Define and give references for CRUNCEP.

Yes, this is included in our proposed text describing the offline CRUNCEP simulation.

p7, L8. The authors conclude that 'the simulations reproduce the global patterns of vegetation cover well'. There are several point here. Firstly, the Sahara is largely missing, and that is rather a big deal. Secondly, I guess these patterns are mainly determined by the 100+400 years of simulation, rather than the last 20 years, but as noted above I don't understand which climate driver was behind these 500 years.

To answer the first point, yes, we were remiss not to discuss the Sahara in the original text. We propose to discuss the under-estimation of Sahara extent in the revised text as follows:

"The extent of the Sahara Desert is also underestimated in the EMAC and *CRUNCEP* simulations, as some areas grow sufficient grass cover to be classified as short grasslands. As this is present in the *CRUNCEP* simulation, it is not related to a climate biases, but rather the over-simulation of grasses in very arid regions by LPJ-GUESS."

To answer the second, the reviewer is correct that the patterns will have been determined in the main part by the full 500 years of simulation. As this was exclusively EMAC-produced climate variables (now made clear in our answer to point 5), we believe our interpretations and conclusions to be valid.

p7 and elsewhere. The authors sometimes say 'conservation remapping', sometimes 'conservative

We apologise for the typo and the lack of explanation. The correct term in 'conservative remapping' this has been corrected in the manuscript. We also removed the work 'bilinearly' which erroneously appeared on page 8 line 20.

To explain the method we propose to include a citation of the algorithm:

Jones, Philip W. (1999) First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. Monthly Weather Review 127(9): 2204–2210.

And a citation to the implementation:

CDO (2018) Climate Data Operators. http://www.mpimet.mpg.de/cdo.

p7, L27. Here it states 'The second source of disagreement is the climate biases in the EMAC derived climate, most obviously the underestimation of tree cover ...'. Usually one evaluates climate biases with reference to a temperature data set, not by looking at tree-lines. Again, I really miss any quantification of the EMAC errors going into this simulation, and without that I have no bases to judge the impact of LPJ-GUESS coupling.

We apologise for the unclear wording and the lack of quantification of EMAC errors. We propose to include plots of the bias in the EMAC climate to aid interpretation (see our answer to the reviewer's point 4).

Regarding the wording, we stress again that in this work it is was not our intention to comprehensively evaluate climate biases in EMAC (as the reviewer appears to have understood from our wording) but rather to examine the behaviour of LPJ-GUESS when forced with EMAC climate. In this case looking at tree lines is appropriate because we are assessing vegetation, not climate. We propose to change the sentence to:

"The second possible reason for discrepancies in modelled tree cover compared to observed tree cover is climate biases in the EMAC-produced climate. For example, this is apparent in the underestimation of tree cover on the north-east coast of South America, as already indicated in the biome plots (Fig. 1), which is clearly the result of a large negative precipitation in the region (see **New Appendix Figure 1**). "

p8, L2. The authors claim that biomass isn't directly relevant for land-atmosphere exchanges, but useful for evaluating DGVM performance. Well, canopy height is mentioned, and LAI (and hence BVOC and deposition parameters) could have been, but isn't biomass also one of the key outputs of ESMs? They are supposed to account for C-sequestration, NPP, etc. It is essential that an ESM can predict these outputs very well, but here they seem to be forgotten.

We apologise for the careless language. We fully agree that biomass is an important state variable and output in ESMs. Our statement was merely meant to reflect that from a technical perspective, total biomass (in terms of kgC/m^2) is not used directly by the land surface (and therefore not passed back to EMAC), whereas the *distribution* of biomass in terms of height (ie. canopy height) and area (ie. vegetative cover) are used directly in the land surface scheme. However, we realise that this was poorly worded and perhaps an unimportant point, and we propose to replace the sentence starting on page 8 line 2 with the following text to highlight to importance of biomass:

"Standing biomass is a key state variable in ESM and DGVMs as it is connected to productivity, carbon sequestration, evapotranspiration, vegetation cover, canopy height and other critical processes and variables which are relevant to vegetation functioning and land-atmosphere exchanges. As such, it is a useful quantity for evaluating DGVM/ESM performance."

p11. Canopy height is here evaluated, but N-availability is a key driver for this, and here the N-deposition component is from the 1850s.

Yes, we understand that N-availability is a driver of canopy height and other vegetation properties, and that using N-deposition from the 1850s will not perfectly reconstruct 20th Century vegetation, it was an unfortunate oversite on our part to use these values. However, the difference (in terms of N limitation) compared to the 'offline' CRUNCEP simulation with transient N deposition is not that large (see previous answers and **Proposed New Appendx Fig 4**). Therefore, we don't believe that this negates the point that we are trying to make with these simulations, which is that the coupled model functions as expected, reproduces the broad pattern of global vegetation and so the vegetation simulated provides a reasonable basis for an updated land surface parameterisation in EMAC.

p11, L25-29. Here again I am not sure what to make of the paper and the setup. For biomass inclusion of a land use correction makes the results worse, but the authors say that 'this is not a major concern .. as biomass does not directly affect the atmosphere'.

How can they say this? Biomass is directly linked to water flows, energy balances, LAI, BVOC emissions, deposition rates, canopy height, momentum exchange, vegetation extent and a host of related parameters. A failure to model biomass reflects a failure to model the vegetation.

This was a careless choice of words on our part, we simply wanted to point out that quantity 'biomass' (in terms of kgC/m^2) is not explicitly used by the land surface scheme (whereas quantities relative to biomass, such vegetation height and cover, are). We propose to replace the two sentences on page 11 line 27-29 with:

"This indicates the importance of including land use effects in a consistent and realistic way in the coupled model, and that improved simulation of biomass is also critical due its status as a key state variable in the land surface representation. In particular, the average global patch-destroying disturbance rate of 0.01 yr^{-1} could be re-evaluated and rather simplistic mortality and tissue turnover rates could be further developed in LPJ-GUESS."

p11. Again the authors suggest that LPJ-GUESS needs to be changed to perform better, but since offline simulations perform better (p7, L10), I would look to EMAC first. (I wonder if any co-authors from Lund, Potsdam or Jena would have made the same conclusions!)

We wholeheartedly agree with the reviewer that if we make statements about where LPJ-GUESS should be improved, it follows that we should support such statements with information about the performance of stand-alone LPJ-GUESS.

The inclusion of stand-alone LPJ-GUESS results makes is clear that the canopy height biases are independent of the forcing climate data used (see **Proposed Modified Figure 4**), so attempts to improve canopy height must therefore focus on the LPJ-GUESS representation of tree height, biomass and allometry.

p12. The text states that 'scores improve when moving to a higher resolution implies that .. leads to a tangible increase in model performance'. Again, I have trouble with the loose arguments. A change in GCM resolution will result in a change in GCM performance and biases. There is no need to 'imply'. The EMAC bias results should have been presented and analysed to establish problems with EMAC, and the offline LPJ-GUESS results should have been presented as the only true benchmark against which the linking can be assessed.

(This sentence was also rather circular by the way. Higher scores implies better performance? I though that that was definition of the score?)

Once again we apologise for the poor word choice and inexact meaning. We also realise that this sentence was not only circular but somewhat superfluous. We propose to replace the sentence with the following:

"While the result that the resolution of the atmospheric has such a significant effect on the vegetation is not surprising in itself, it does highlight that when considering the bidirectionally coupled model with dynamic (as opposed to prescribed) vegetation, thorough investigation must be made of the effect of the resolution of the atmospheric model on both model components, particularly considering feedbacks between them."

To answer the reviewers second point, we fully recognise this and now include both EMAC bias plots and offline LPJ-GUESS results as discussed above.