gmdd-5-917-2012 – Final response to the comments of anonymous referees #1 and #2

Dear editors and reviewers,

we would like to thank anonymous referees #1 and #2 very much for their careful and constructive reviews. Following we address all comments that appeared during the review of the discussion paper. The comments by the referees are given in blue color, our response is formatted black. We have considered all comments for the preparation of a revised version of the manuscript.

<u>Response to comments from anonymous referee #1:</u>

"COSMOS – please put the model version number in the title."
We agree that the first version of the manuscript lacked a precise reference of the COSMOS model version. Yet, we believe that putting the model version number (COSMOS-landveg r2413, 2009) into the title will cause it to be quite lengthy and cumbersome. We therefore rather would like to refer to the version number <u>both</u> in the abstract and the model description. This change has been implemented in the revised version of the manuscript.

"In the abstract – 'warmer and wetter' is a bit general – say 'warmer and wetter in the global mean'"
In the revised manuscript, we adjusted the sentence from "The mid-Pliocene as simulated with our COSMOS-setup and PRISM boundary conditions is both warmer and wetter than the PI." to "The mid-Pliocene as simulated with our COSMOS-setup and PRISM boundary conditions is both warmer and wetter <u>in the global mean</u> than the PI."

"Pg 920, line 13. 'Our version' of COSMOS sounds a bit disconcerting. How does it vary from the 'standard' version. Does it have a version number or reference? You cite Roeckner et al for the ECHAM5 atmosphere, but here you are using it at a different resolution to that described there? Please be clear throughout the model description which aspects are 'unique' to your version of ECHAM/COSMOS, and how they vary from the standard version. Ideally provide a reference which described your exact reference."
Unfortunately, we cannot give a citation that describes and discusses the exact COSMOS setup that we used for the PlioMIP experiments. PlioMIP is based on a very special experimental design that involves prescription of a reconstructed paleo-vegetation – an experimental design that we, and apparently also other COSMOS-users, do not usually consider.
Our setup is comparable to studies referenced in the discussion paper, but an exact match of the experimental procedure cannot be found:
We have been using COSMOS in T31/L19 and GR30/L40 resolution. This is somehow a standard global low-resolution setup as it is used for the preparation of various studies of the climate system (e.g. the publications cited in the discussion-manuscript: Brovkin et al., 2009; Fischer and Jungclaus, 2010; Varma et al., 2012; Wei et al., 2012; Wei and Lohmann, 2012). Yet, in none of these publications JSBACH is used with the dynamical part of the vegetation module being switched off, as it is the case for our PlioMIP simulations of control experiment 1 and the mid-Pliocene simulations of experiment 1 and 2. The standard setup at our work group as well relies on the use of the dynamical part of the vegetation module (e.g. Wei et al., 2012), but within the PlioMIP framework it was decided that all models shall use the same vegetation cover. Therefore, we needed to switch off the dynamic vegetation

module in order to follow the PlioMIP experimental guidelines that stipulate the prescription of a fixed paleo-vegetation.

Other publications, focusing on more fundamental aspects of the performance of the COSMOS model (or its components), that could be used as a benchmark with respect to our results, also differ in one or more details of the methodology. This might e.g. be the use of an AMIP-style setup without a coupled ocean model (e.g. Roeckner et al., 2004; Roeckner et al., 2006; Hagemann et al., 2006; Wild and Roeckner, 2006), or the use of a different model resolution of atmosphere and ocean (e.g. Jungclaus et al., 2006; Wild and Roeckner, 2006; Raddatz et al., 2007).

In order to emphasize the peculiarity of our model setup we adjusted the respective part of the section "Model description" (pg. 920, line 13 and following) as follows:

"Our version of COSMOS includes the ECHAM5 atmosphere model <u>in T31-resolution with 19 levels</u>, the MPI-OM ocean model <u>in GR30 resolution with 40 levels</u>, and the land-vegetation model JSBACH. <u>Our setup is identical to the COSMOS-1.2.0 release, that has been developed in the Millennium project (Jungclaus et al., 2010), but additionally includes a dynamical vegetation module (Brovkin et al., 2009). In this version, COSMOS has been used for the preparation of various publications (Brovkin et al., 2009; Fischer and Jungclaus, 2010; Varma et al, 2012; Wei et al., 2012; Wei and Lohmann, 2012), but in our experiment 1 and the mid-Pliocene simulation of experiment 2 the dynamic vegetation module has been switched off in order to be consistent with the PlioMIP protocol.</u>"

Additionally, we added at the end of subsection 2.1 (pg. 922, line 15 and following) a reference to publications that describe an ECHAM5-setup with identical resolution: "<u>This setup of the ECHAM5 model has been used by various authors (Brovkin et al., 2009; Fischer and Jungclaus, 2010; Varma et al., 2011; Wei et al., 2012; Wei and Lohmann, 2012).</u>"

"P922, line 5. What is the reasoning behind using a solar constant of 1365W/m2 in atmosphere-only, and 1367 in the coupled model? You address this later in the sensitivity studies, but I just wondered why they are different values. "

There is no specific reasoning for choosing a slightly different solar constant in the coupled and standalone version of our model setup – the solar constant is simply a hard-coded parameter in the ECHAM5-model that we cannot adjust. The solar constant is automatically set to 1365 Wm^{-2} if ECHAM5 is run standalone, and to 1367 Wm^{-2} if the atmosphere is coupled to an ocean model. We presume that the solar constant is used as a tuning parameter in the different setups and keep the standard values. In the revised manuscript this has been emphasized by adjusting the respective passage in subsection "Experimental methodology" of the discussion (pg. 938, line 16 and following):

"Since all other paleoclimate simulations conducted <u>in COSMOS</u> rely on these standard ECHAM5 parameters, <u>and since the values of the solar constant are not adjustable (but hard-coded) parameters of ECHAM5,</u> we also preserved them in the PlioMIP simulations."

"P923. As for ECHAM, please do the same for MPI-OM. Is there a model description paper which shows e.g. climatologies for your particular setup, e.g. your particular resolution, position of poles etc. If not, please be very clear how your version differs from that in the referenced papers. "

We added a sentence that lists publications that are based on the same MPI-OM setup (but not the same COSMOS setup) at the end of subsection 2.3 (pg. 924, line 3 and following):

"This setup of the MPI-OM model has been the basis for various publications (Brovkin et al., 2009; Fischer and Jungclaus, 2010; Varma et al., 2011; Wei et al., 2012; Wei and Lohmann, 2012)."

For clarification we added another modification to subsection 2.4 (pg. 924, line 15 and following). The respective passage reads now:

"Other paleoclimatological studies that employed a COSMOS setup comparable to the one used for preparing the PlioMIP simulations are documented for the Holocene (Fischer and Jungclaus, 2010; Wei et al., 2012; Wei and Lohmann, 2012; Varma et al., 2011), Last Glacial Maximum (Zhang et al., 2012), and the Miocene (Knorr et al., 2011). <u>The setups described in these publications are not identical to the one used in this study, in particular none of them uses JSBACH with the dynamic vegetation module being switched off.</u>"

On the Max Planck Institute for Meteorology's ECHAM5 model-description-website the reference for the respective data sets is given as "from ECMWF, Reading, England". It seems that there is no proper citation in any of the COSMOS publications that we are aware of. To make the origin of the fields more clear, the respective sentence (pg. 925, line 8 following) has been adjusted to:

"Global datasets of soil wetness and the contribution of orography to surface roughness length originate from <u>input files for global forecast models developed at the ECMWF (e.g., White, 2003).</u>"

We tested the influence of an increased solar constant on global average SAT and precipitation of the mid-Pliocene simulation of experiment 1 in a sensitivity study using a solar constant of 1367 Wm^{-2} as in the coupled setup. This sensitivity study is needed since the solar constant in ECHAM5 is a hard-coded (and fixed) parameter that differs between a standalone and a coupled setup. We did not want to alter these fixed settings in our experiments, since we presume that a meaningful model tuning is the reason for the chosen solar constants.

The sensitivity simulation has been integrated for 50 model years, the average global temperature climatology has been calculated by averaging over the last 30 years. The SAT anomaly between the (standard) mid-Pliocene simulation of experiment 1 and the modified (sensitivity) simulation based on the increased solar constant is small for the global average (~0.004 K). Due to the design of experiment 1 the SAT over the ocean is rather constant as a result of the prescription of SST. Therefore, an increased solar constant cannot cause pronounced changes of SAT over the ocean (see Fig. 1). Over land, changes in SAT are evident, but the distribution is quite patchy and positive and negative anomalies largely cancel out. As a result, the global mean SAT is nearly unchanged (as stated in our manuscript). The reason for the compensating effect of warming and cooling in the temporal spatial average is related to minor changes in land-sea contrast, but the details have not been examined. The important outcome of this sensitivity study is, that our PlioMIP mid-Pliocene simulation of experiment 1 with a solar constant of 1365 Wm^{-2} is comparable to a standalone simulation with a solar constant of 1367 Wm^{-2} if one considers global average temperature and precipitation. Therefore, the presence of a slightly modified solar constant in experiment 1 is negligible for a general climatology as presented in Table 2 of the manuscript, and our mid-Pliocene simulations of experiment 1 and 2 are comparable as

well.

It may be that in the discussion manuscript this information has not been presented clearly enough. We applied the following modifications in order to address this problem (pg. 926, line 6 following):

"In experiment 1, the solar constant is set to 1365 Wm^{-2}, in experiment 2 its value is 1367 Wm^{-2}. The difference in the forcing does not have an appreciable impact on the global average climatology of experiment 1 (for details refer to the discussion of the experimental methodology further down)."

The adjusted passage in the discussion of the experimental methodology reads now (pg. 938, line 16 and following):

"Since all other paleoclimate simulations conducted in COSMOS rely on these standard ECHAM5 parameters, and since the values of the solar constant are not adjustable (but hard-coded) parameters of ECHAM5, we also preserved them in the PlioMIP simulations. Our results of experiment 2 might therefore be based on a slightly different solar constant than those contributed by other PlioMIP groups, and the comparability of our experiments 1 and 2 might be flawed. We address the question of comparability by investigating the influence of a change in the solar constant on annual global average SAT (T) and precipitation (P). For this purpose we set up a modified standalone atmosphere mid-Pliocene simulation with an increased solar constant (1367 Wm^{-2}). All other settings are identical to our mid-Pliocene simulation of experiment 1. The anomaly of SAT between the simulation with high and low solar constant is negligible over the ocean. This is understandable since SST is prescribed in the standalone setup. Over land the change in SAT is not vanishing, but the time-averaged anomalies of global mean ΔT and ΔP introduced by the change in the solar constant are negligible – the values of T and P of both mid-Pliocene simulations are identical if rounded to the second decimal place (not shown). We therefore conclude, that the general global average climatology of experiment 1 is not influenced appreciably by a small modification of the solar constant by 2 Wm^{-2}, that therefore our climatology of PlioMIP experiment 1 with 1365 Wm^{-2} is comparable to one that we would have retrieved with a solar constant of 1367 Wm^{-2}, and that our experiments 1 and 2 are comparable."
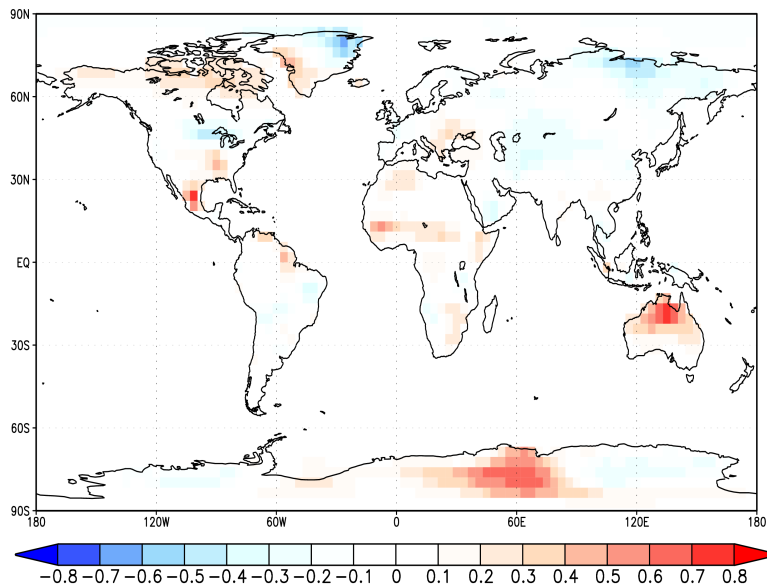
Fig. 1: Anomaly of SAT in K between the Pliocene simulation of experiment 1 and a similar simulation with increased solar constant (1367 W/m^{-2}). Shown is the time average of simulation years 820 to 849. As a result of the experimental methodology, temperature changes over the ocean are suppressed. Over land, warming and cooling largely cancel out, the global average temperature anomaly is ~0.004 K.

We addressed this remark by adding a more detailed explanation of our approach of adjusting the land-sea-mask to mid-Pliocene conditions. The respective passage of the manuscript (pg. 926, lines 20 and following) reads now:

"The setup of the mid-Pliocene simulations of experiments 1 and 2 is based on the preferred mid-Pliocene data set. <u>Our modelling approach deviates from the protocol in that we include major changes in the land-sea-mask of the ocean model (i.e. a closure of the Hudson Bay, and</u> adjustments in the West Antarctic), but neglect minor changes in the coast line related to sea level change. Therefore, the land-sea-mask of the PI control simulations is preserved for <u>the mid-Pliocene,</u> with the exception of a closure of the Hudson Bay, and adjustments at the western Antarctic <u>continent</u>."

We did not apply an explicit scaling that ensures that the sub-gridscale mid-Pliocene topography had the same magnitude as the PI sub-gridscale parameters. In general there are two possibilities of treating paleo-sub-gridscale topographic parameters: As a first approach one could just preserve the present day set of parameters and completely ignore the reconstructed changes in topography. This methodology is not consistent since changes in the elevation do not reflect in the sub-gridscale parameterization – but the high resolution information from present day (which is not available for topographic reconstructions) would still be available in the paleo-simulation. As a second approach one could recalculate the sub-gridscale parameterization from the reconstructed topography. This approach is consistent, since the changes in topography reflect in the sub-gridscale parameterization – but of course the so computed paleo-parameters will not be based on a highly variable topography as it is the case for the control simulation.

For our paleo-simulations we chose to follow the second approach. In the first version of the manuscript the lack of a high-resolution paleo-topography has already been mentioned (see pg. 927, line 22 and following). To illustrate the potential influence of a low-resolution topography on the sub-gridscale parameterization, we add here a comparison of all fields that influence the sub-gridscale orography parameterization for PI and mid-Pliocene conditions (Fig. 2 and 3). Since the number of fields is large, in the revised version of the manuscript we propose to add only one figure, depicting the differences in orographic peaks elevation, as an example.
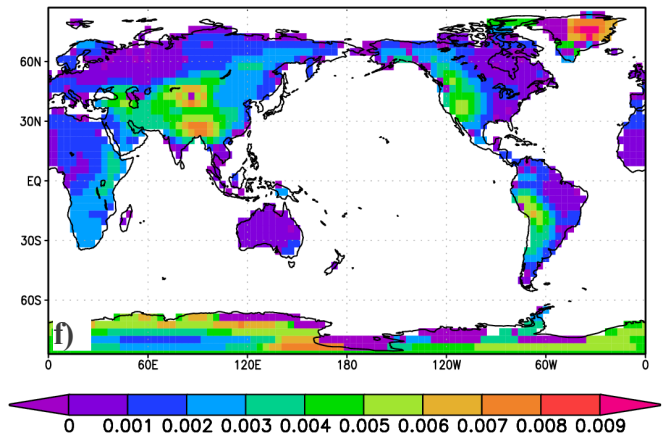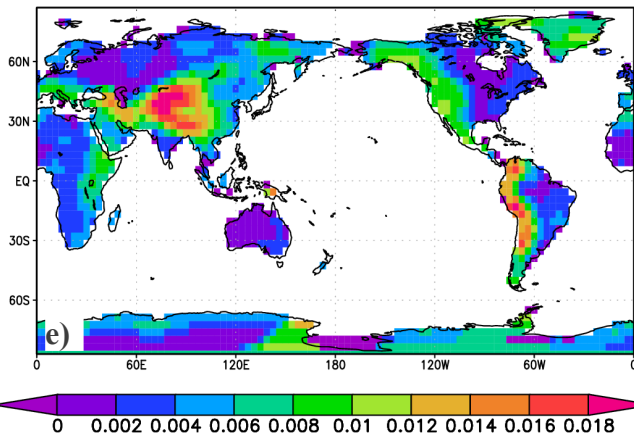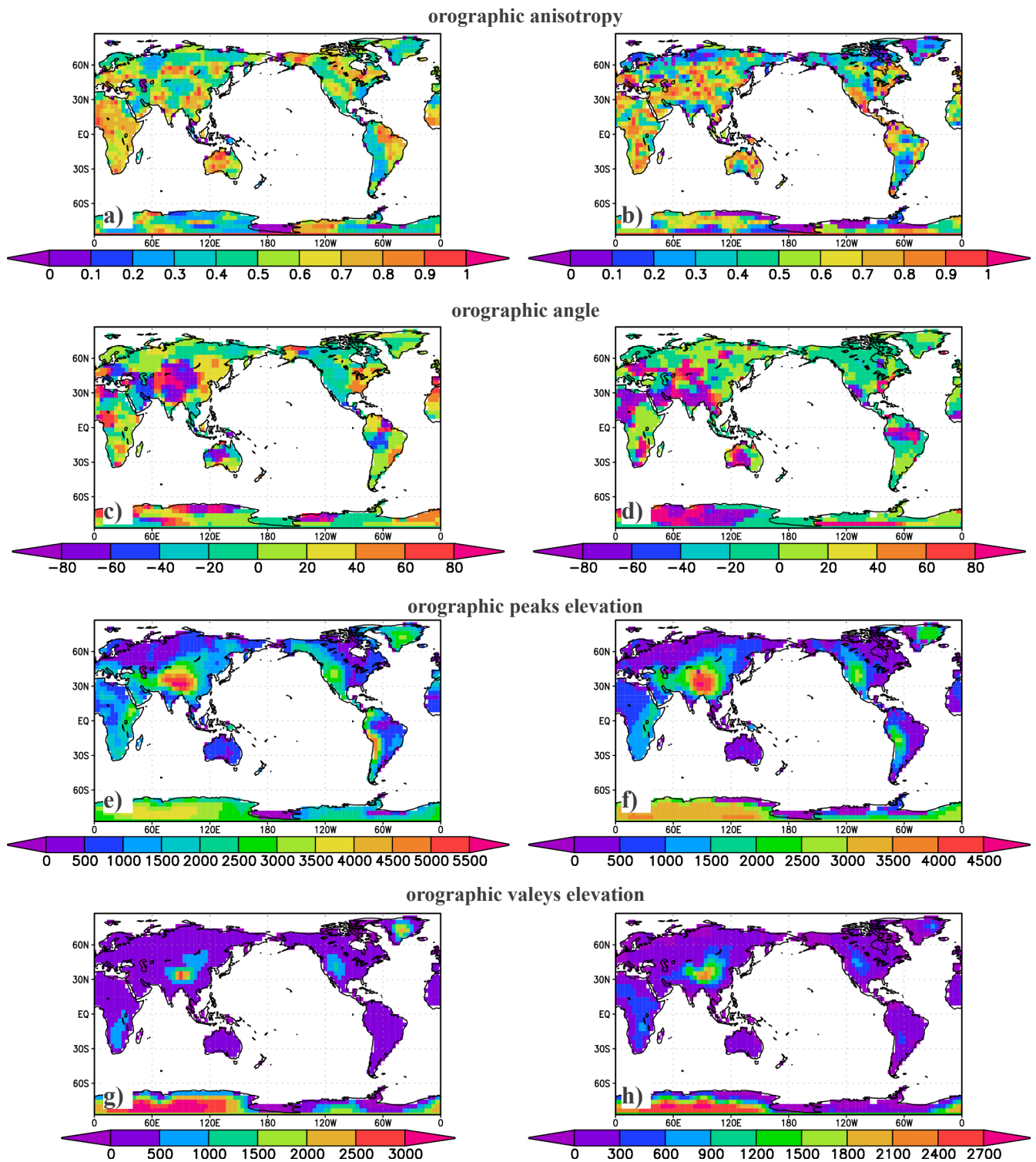
mean orography
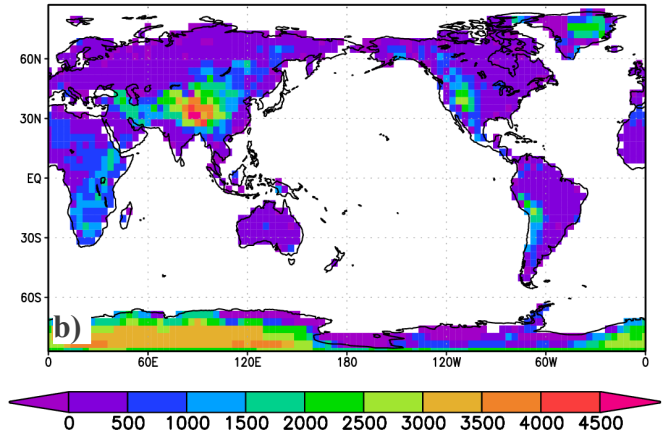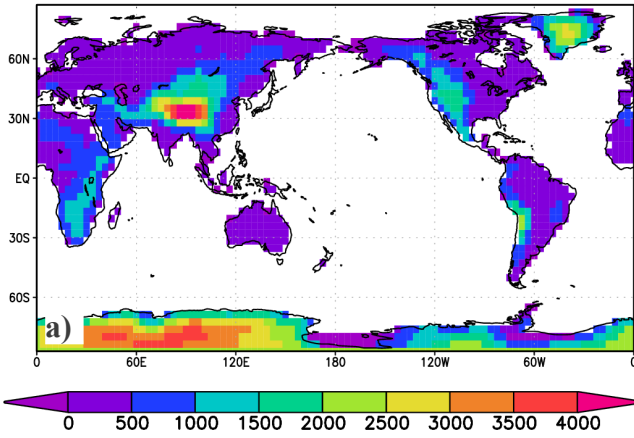
standard deviation of orography
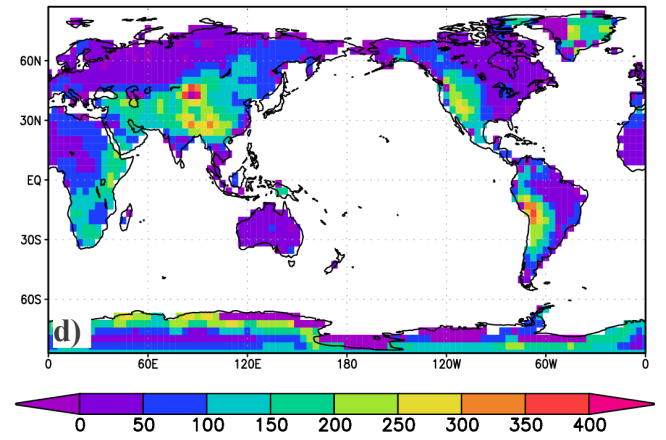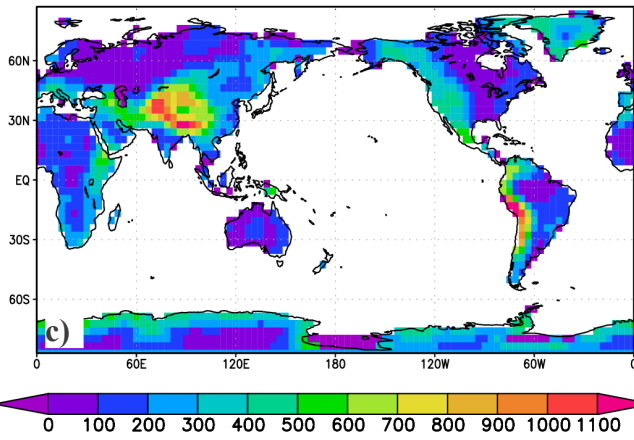
orographic slope

Fig. 2: Overview of parameters of the sub-gridscale orography parameterization for PI (left) and mid-Pliocene (right); **a)** and **b)**: mean orography / m, **c)** and **d)**: standard deviation of orography / m, **e)** and **f)**: orographic slope / degrees.

Fig. 3: Overview of parameters of the sub-gridscale orography parameterization for PI (left) and mid-Pliocene (right); **a)** and **b)**: orographic anisotropy / degrees, **c)** and **d)**: orographic angle / degrees, **e)** and **f)**: orographic peaks elevation / m, **g)** and **h)**: orographic valleys elevation / m.

We find that our method of generating the paleo-sea-ice distribution is valid, since it is based on a very simple physical principle that holds in particular for a simplified version of the real world (which our climate model is). Our method of generating the mid-Pliocene sea-ice forcing for experiment 1 assumes that any warming above the freezing point at a particular grid-cell will remove the contained frozen water. In standalone atmosphere models this is a consistent approximation.

We following shortly explain the idea behind this method for the sea-ice reconstruction: Imagine a glass of water that contains ice cubes where the temperature of the glass is fixed through an external bath. If one measures the temperature of the water while the ice is melting it will become evident that the temperature of the water does not start to rise before all the ice has been molten (if one assumes that the water is well mixed). As long as there is still ice available, the heat that is entering the glass from the outside is used break up the crystal structure of the water molecules. Our atmosphere model consists of rather coarse grid-cells. These are – per definition – well mixed, since they represent the smallest available spatial structure. We therefore assume that any warming of the sea surface at a particular grid-cell (as it is evident from the proxy records) cannot happen before all the sea-ice at this grid-cell has been fully removed. Furthermore, the methodology of using a yearly recurring monthly climatology of sea-ice (as it is the case in the PlioMIP simulations of experiment 1) implies that the thermodynamics at a specific grid-cell is in equilibrium (since no long term changes of the sea-ice cover occur). Therefore, the term "well mixed" does not only refer to the spatial but also to the temporal scale.

We refrain from adding such an extensive explanation to the final manuscript since this topic has been already addressed in the discussion of our manuscript on page 940. There, we also mention the equivalence of our mid-Pliocene sea-ice distribution to one that we would have retrieved directly from the paleo-SST field. An energetically more consistent setup can be obtained only in the coupled version (experiment 2).

Our approach of generating the mid-Pliocene vegetation forcing for JSBACH is a translation of the modern observation of vegetation distribution, that is given within the PlioMIP framework, to a modern vegetation as simulated with COSMOS. As discussed in the manuscript, our approach leads to a smoothed vegetation field, since the calculation of the fractional amount of a JSBACH PFT that corresponds to a specific biome is performed via averages over larger areas. As suggested by anonymous referee #1, we prepared a sensitivity study in which we investigate the influence of the lack of vegetation variability on the global climatology of a modern climate. We conducted the PI control simulation of experiment 1 again with a total integration time of 50 model years, but this time did not use as a vegetation forcing the long-term averaged PI-vegetation distribution that COSMOS simulates,

but the modern observation on which we applied the average vegetation properties that have been generated with the mapping procedure. The vegetation distribution that has been generated by this method (Fig. 4) is comparable to the one of the PI simulation of experiment 1, but is as expected smoothed due to the large-scale averaging.

The results from this sensitivity study show that the global average SAT is not strongly affected: The change in the vegetation forcing causes a global average warming of only ~0.02 °C. Yet, there are local anomalies of SAT especially over Asia and North America that are noteworthy (cf. Fig. 5). We propose to present this information in an appendix to the final version of the manuscript. Within the manuscript itself we would like to only state that the alternative implementation of fixed PI vegetation does not appreciably affect the global average climatology.
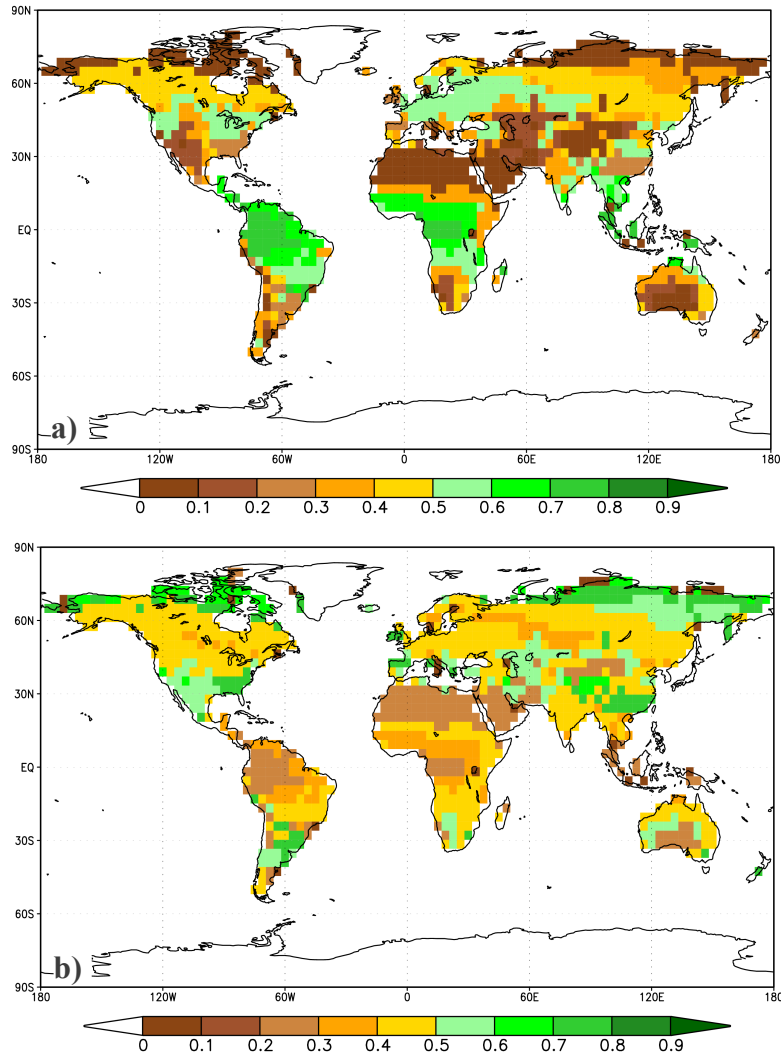
Fig. 4: Fraction of **a)** forest and **b)** grass cover as for the PI simulation of experiment 1 (cf. the discussion manuscript for details), but based on the modern observation on which the average JSBACH representation of biomes has been applied.
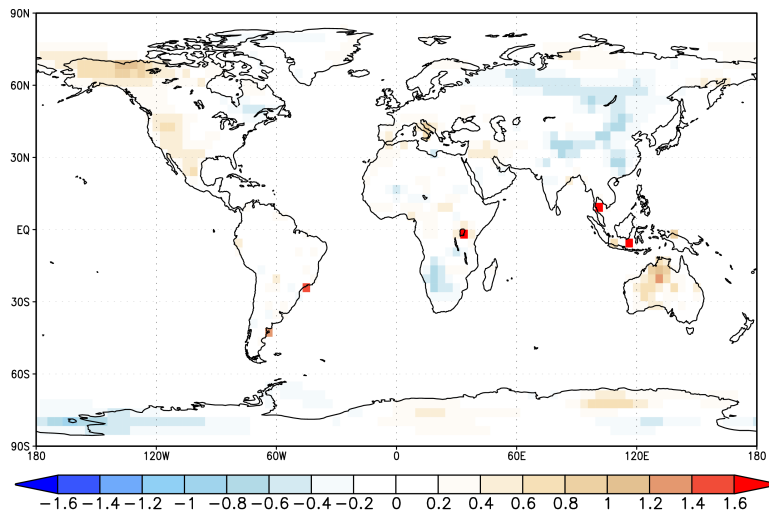
Fig. 5: Anomaly of SAT in K between a sensitivity study that is forced with modern observations that have been translated to the JSBACH vegetation representation, and the PI simulation of experiment 1 (that is based on vegetation directly retrieved from a PI simulation with COSMOS). Shown is the time average of simulation years 820 to 849. The global average SAT anomaly is ~0.02 K.

"P931, line 10. co2 – most importantly, 405ppmv is the pliomip standard, and used by other groups in pliomip. This is probably really the reason you used that value! "

This remark has been addressed in the revised version of the manuscript. Indeed, the uncertainty in pCO_2 seems to be high and different values are documented in the literature (e.g., Kürschner et al., 1996; Raymo et al., 1996).

"P932. the TOA inbalances – in the atmosphere-only run, this is not really an 'energy gain' but an energy inbalance – the model does not gain energy at this rate, it effectively also throws the energy away because the SSTs are fixed. In the atmosphere-ocean case, this is actually an energy gain (assuming the energy budget of the model is closed). Not sure what you mean by 'inherrent feature of the model forcing'. "

We addressed this comment in the revised version of the manuscript (pg. 932, lines 18 and following). The adjusted text reads: "In experiment 2 there are residual net energy inputs into the climate system of slightly more than 1.5 Wm^{-2} for PI control and mid-Pliocene. In experiment 1, the climate in the mid-Pliocene simulation appears to be slightly farther from radiative equilibrium with a persisting net energy <u>imbalance</u> of 3.5 Wm^{-2}, while the PI control simulation shows a small TOA net energy <u>imbalance</u> of ~0.4 Wm^{-2}. Since the atmosphere is well-equilibrated during the last 30 y of experiment 1 (Fig. 5a), net energy fluxes are unlikely to be caused by remaining energy buffers in the model climate (like the deep ocean in experiment 2), but are presumably <u>caused by inconsistencies in the prescribed SST field.</u>"

Typos:

We have corrected all the mentioned typos in the manuscript.

Additional References:

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: The atmospheric general circulation model ECHAM5. PART II: Sensitivity of Simulated Climate to Horizontal and Vertical Resolution, Report 354, Max-Planck-Institut fuer Meteorologie, Hamburg, 2004.

White, P. W.: IFS Documentation, Part IV: Physical processes (CY25R1), available at: `http://www.ecmwf.int/research/ifsdocs/pdf_files/Physics.pdf` (access: 21 June 2012), 2003.

Response to comments from anonymous referee #2:

"In section 3.2 (Mid-Pliocene simulations), you describe how your have adjusted the land ocean distribution in the atmosphere and ocean components in respect to the contemporary setup. Here in particular you explain the changes around the Antarctic continent. Do you flatten the land surface to zero and digging out the ocean to 500 m depth? How is the exchange of momentum, heat, and flux water at these points implemented? Please clarify it. "

The land surface is flattened to 0 m height and the ocean is dug out. This cannot be done completely consistent for the differing grids of the ocean and the atmosphere model. The changes in the ocean's land-sea-mask are rather minor, since only few grid-cells have been altered, and those altered are close to a grid pole and therefore small. Interpolation of the altered ocean's land-sea-mask to the coarse atmosphere grid (T31) leads to a complete loss of this change, i.e. in the interpolated version the dug out ocean is again replaced by land. This of course leads to minor inconsistencies in the exchange of momentum, heat and water flux (which is performed by the OASIS3 coupler). These inconsistencies are generally present at gateway regions in the COSMOS setup if a non-fractional land-sea-mask is being used. In our opinion there is no way of correcting this problem: Either one would run the simulation with a surplus of land at this region (as it has been done in this study), or one would run the simulation with (a manually introduced) surplus of ocean. None of these two methods is correct, therefore we used the atmosphere's land-sea-mask that is generated by interpolation from the ocean grid (which favors land in the given case, and is closer to the optimal solution). We have added a comment on this drawback of the model in the discussion of the experimental methodology, section 5.1 (pg. 939, line 14 and following):

   "The small removal of land in the ocean model at the West Antarctic cannot be adequately reproduced on the non-fractional land-sea-mask of the atmosphere model. Therefore, at this location minor inconsistencies in the exchange of momentum, heat and freshwater may arise. This is a general problem of models with a non-fractional land-sea-mask."

"The mid-Pliocene 3dim ocean temperature distribution has been obtained by utilizing two interpolations; firs onto a coarser grid and second back to the finer target grid. Does this kind of smoothing introduce large errors? How large are the differences in the worst case? "

The chosen interpolation method (consisting of two different interpolations) is basically a necessary "artifact" of the curvilinear grid of the ocean model. A transformation script needs to be run in order to convert the regular grid of the reconstruction to the ocean grid. We stated this on pg. 928, line 18 and following, but would not too much emphasize to the errors that are introduced by our (and any other) interpolation method - the deep-ocean temperature data set (Dowsett et al., 2009) is used as an initial condition (rather than a forcing) that itself adjusts to the energy fluxes during the course of the simulation. For an interpolation of a quantity that has a conserved global inventory during a simulation (e.g. the global ocean salinity field) such a check of the introduced errors would be necessary, but for the deep-ocean temperature field those errors actually do not play a role.

"Page 927, Line 25-27: You talk about the distribution of glaciers, but I doubt that the model resolution is high enough to resolve any glacier. Should you name this distribution ice cape distribution or, in particular, ice sheet distribution? "

We have changed the term "glacier distribution" to "ice sheet distribution".

"Page 933, Line 1-3: I am with you that the prescribed SST might cause the size of the imbalance. Immediately it came to my mind, how does the latitudinal difference in the SST look like? Since you present this quantity later (Fig. 6), you might add here a comment about the coming analyzes. "

We added a remark at pg. 933, line 3 and following:

"For the PI simulation, this conclusion is supported by zonal average ocean SAT plotted versus latitudes (see Fig. 6b). If one assumes that the PI simulation of experiment 2 is close to radiative equilibrium, then it is evident that in many regions of the mid- and high-latitudes the prescribed PI SST of experiment 1 is too warm."

"Page 934, Line 24-25: The warming is partly driven by changes in the topography. How large are the height difference in Greenland for example and to what an extent is the detected warming driven by the height effect alone considering a lapse rate of -6.5 K/km? "

On pg. 934, line 24 and following, we have added to the manuscript a short note on the estimate of the contribution of topography to the general warming over Greenland at the example of a single ice-free grid-cell:

"Over Greenland the warming due to the topographic effect seems to be the major driver for mid-Pliocene warmth. For a typical location over Greenland, that is free of ice in the mid-Pliocene simulation of experiment 2, we find a warming of ~15 °C with respect to PI. The mid-Pliocene elevation is decreased by 1800 m. Therefore, at this location about 80% of the warming is caused by the reduction of the elevation if one assumes a lapse rate of 6.5 K/1000 m."

"Page 935, Line 1-4: What might cause the exceptional strong warming in the Weddell Sea? Do the convection sites change their location in the Southern Ocean? "

The warming seems to be linked to a strong reduction in the sea-ice compactness (cf. Fig. 8 and 11), therefore the warming might be caused by the change in the albedo of the ocean surface.

"Tab. 1: Since in the text on page 930, lines 19-21, you lump together the plant functional types (PFTS) 1-4 to a generalized forest type, and 5-8 to a generalized grass type, you might highlight these clustering by either add an additional horizontal line, extra space between these groups, or an additional column highlighting these groups. "

We have added now an additional column to Table 1. It emphasizes whether a PFT belongs to the forest or grass type.

"Tab 2: The main difference between experiment 1 and 2 is the coupling to an active ocean model. Since this table might be consulted frequently to quickly resolve this main difference, you might add a corresponding additional column or comment in the table caption. "

We have added comments ("A" or "AO") to column 1 of Table 2 to emphasize the differences.

"Fig. 1: In the Hudson Bay the difference between the setups is clearly identifiable, but in the southern Pacific offshore of Antarctica, the changes are a little bit hidden. Have you tried another projection to resolve this issue? "

We agree that the change of the land-sea-mask is not very obvious, but rather would like to keep the figure in its previous form, i.e. with a regular longitude-latitude projection (as it was asked for in the guidelines for the PlioMIP experimental description manuscripts).

We have replaced the respective gray contour lines by green contour lines that are more visible.

We have added a second axis for temperature anomalies at the right-hand side of Fig. 6 a-c. It is adjusted it in a way that the zero line is maintained.

We have added an additional anomaly axis at the right-hand side of Fig. 6d.

We have replaced the gray isolines that depict the density of sea-ice by green isolines that are more visible. We also have added information regarding the contour line interval to the figure caption.

We have added information regarding the contour line interval to the figure captions. We also have added a bottom topography that has been calculated from the zonal minimum sea floor elevation of each ocean basin.

Yes, it is. We have added a remark to the caption of Fig. 13.

Yes, due to an unnoticed bug in the plot script there was a rounding error for the contour line captions. This error has been fixed, and the contour line interval is now 2 Sv. The contour line interval has been mentioned in the caption of Fig. 14.