**Late Pliocene lakes and soils: A data – model comparison for the analysis of climate feedbacks in a warmer world**

Matthew J. Pound and Julia Tindall (on behalf of all authors)

**We wish to thank the reviewers for their constructive comments. Below we have collated the three reviews into a single document and responded (in bold print) to each comment. For reference we have used the four digit "C" code before each reviewers comment so that it can be related back to the original review.**

**Abstract**

(C1648) P3176, line 12: ": : :seasonal increases in [precipitation] the Northern Hemisphere: : :" – this is vague. Do you see this over all of the Northern Hemisphere? Figure 4 does not show this.

**We thank the reviewer for highlighting this inconsistency. In the revised manuscript we shall be more specific.**

1. **Introduction**

**1.1 Background**

**1.2 The importance of soils and lakes in palaeoclimate studies**

(C1616) Page 3178, Line 1: and also Contoux et al., 2013, which focuses on the Megalake Chad in the mid-Pliocene. By the way, the reference to Sepulchre et al.

2008 is wrong, it is Sepulchre et al. 2009 who investigates the feedback of the lake on the mid Holocene climate. Sepulchre et al 2008 investigates the water balance of the Chad basin under mid Holocene conditions, which is different.

**We will include the Contoux et al., 2013 reference in the introduction and correct the Sepulchre reference**

(C1616) Page 3178, line 21: ‘As lakes and soils have had significant regional impacts on mid Holocene precipitation (e.g. Krinner et al., 2012), it stands to reason that similar affects could be seen in the Late Pliocene.’ First, Krinner et al., 2012 do not change the soils, so this citation is only appropriate for lakes, you should try to find another citation for soils. Moreover, the ‘have had significant regional impacts’ is probably a little too affirmative in light of recent results (Contoux et al., 2013) and of previous results (Sepulchre et al., 2009, Brostrom et al., 1998) which depict only a minor effect of lakes. ‘could have had’ seems more appropriate, to my opinion.

**This section of the introduction will be rephrased to more accurately reflect the differences in impacts found by different studies.**

(C1648) P3179, line 1-2: Please include a sentence of why you want to include these new boundary conditions (too improve the model-data mismatch) and what you are going to compare your results with.

**We have added an additional sentence**

## 2. Methods

### 2.1. Construction of the lakes and soils database

(C1616) On the database side, details on lake extent calculations (dry and wet) and discussion on spatial and temporal uncertainty, especially for large lakes (megalakes) are needed. This database is going to be used as a reference, so the assumptions you make to calculate your extents needs to be documented. It is also important that you assess the uncertainties or the unknowns, both spatial and temporal, especially for all megalakes (Chad, Fazzan, Makgadikgadi, Eyre, and Northwestern USA lakes), not only the Zaire megalake.

(C1635) The information in the database (suppl. data) is very sketchy and additional explanations need to be documented and added: - how were the lake extents calculated? - is it possible to give (spatial and temporal) errors/uncertainties for the used calculations? - For Mega-lakes (with probably the highest influences on the results) it would be useful to add short comments on how the original references were interpreted; so that the authors calculations and interpretations are reproducible; also for further and steady completion and contribution by the scientific community to this data base The column "surface area" lacks settling the rounding differences to the relevant digit. This would then also account for data checking as these numbers should not be produced simply by unchecked calculations.

**We thank both reviewers for bringing this to our attention. We acknowledge and apologise for the lack of detail in the supplementary database. This will be rectified in a revised manuscript with an additional three sections (see the text below) in the supplementary information describing the methodology and uncertainties in the lakes and soils data and a supplementary table with more information.**

**Construction of the lakes and soils database and uncertainties associated with Late Pliocene lakes and soils.**

**Late Pliocene lake and soil data was compiled from published literature (Supplementary table 1). This supplementary section will seek to explain the full databasing methodology, outline possible uncertainties and document specific uncertainties associated with the Late Pliocene mega-lakes.**

**Database construction methodology**

**Adding a new site to the database first involves producing a latitude and longitude of the locality. This does not present a problem even with older literature. Further to this, it is considered that continental drift from the Late Pliocene to the present day is minimal (Salzmann et al., 2008) and we therefore do not Palaeo-rotate localities.**

**In the Lakes and Soils database the absolute maximum age range has been recorded, taking into account all published errors and the published dating technique has also been recorded (Supplementary Table 1). By providing the method used by the original authors to date the site, it provides a qualitative means of assessing the accuracy of the dating. For example it is more likely that a radiometrically dated locality will have a more tightly constrained age than one dated through stratigraphic correlation. However, we would strongly advise caution about simply using this information from Supplementary Table 1, to discount or select data without reading the original publications. In an effort to provide maps suitable for future time-slice focussed palaeoclimate modelling and in the absence of the dating accuracy required for such maps, we have provided maps designed for a warm-humid (wet-lakes) and a cold-arid (dry-lakes) climate regime.**

**To make the soils database internally consistent we opted to use the USDA soil classification guide (Soil Survey Staff, 1999). This is a simple classification scheme and each of the soil orders (with the exception of Inceptisols and Entisols) can be associated with specific vegetation types (which facilitated producing a global distribution map using the Salzmann et al. (2008) data – model hybrid vegetation reconstruction as an additional data source) (Soil Survey Staff, 1999). The USDA soil classification scheme is a hierarchical classification with 12 soil orders, each of which can be further classified into suborders, groups, subgroups, families and series (Soil Survey Staff, 1999). The majority of paleosol localities were already reported in the USDA soil classification scheme, whereas a minority could be easily assigned to a soil order based on the original author's identification or through their soil horizon descriptions.**

**Late Pliocene lakes required a size and surface area of the water body to allow a map suitable for palaeoclimate modelling to be produced. By either using numbers provided in the published text (e.g. Dodson and Ramrath, 2001), extracted from scaled diagrams of the palaeoenvironment (e.g. Tiercelin, 1986), calculated from scaled maps of reconstructed lake extent (e.g. Drake et al., 2008), calculated from scaled geological outcrop maps (e.g. Yeniyol, 2012) or through a combination of these. It is worth noting at this stage that if a lake extent could not be confidently calculated from the published literature it was not included in the database. An example of this would be the lake deposits reported from around the palaeo-Yukon River in north-west North America (Matthews et al., 2003; Pound et al., In Prep.). To facilitate mapping Late Pliocene mega-lakes, a north, south, east and west latitude – longitude point was recorded, to give a geographical indication of the lakes extent. This enabled the size and location to be more accurately translated into the gridded maps for the modelling study. As well as the size of the lakes, published details such as chemistry, the location of inflows and outflows, type of lake (e.g. evaporitic, mesotrophic etc.) and any evidence for orbital controls and specific events (see supplementary table 1).**

**Developing a dry-lakes scenario map required additional information to be taken into account. This ranged from utilising the published distribution of geological outcrops showing a shift from lacustrine to fluvial or sub-aerial sediments within the Late Pliocene (e.g. Schuster et al., 2009), the distribution of evaporite deposits (e.g. Sáez et al., 1999) or the original author's palaeoenvironmental interpretation (e.g. Salama, 1987). The original author's interpretation of a lakes response to climates provided the most rigorous data for generating the dry-lakes scenario. This either meant that the original author clearly stated that the lake was reduced to a certain size, or that it would have been absent. In the absence of a statement from the original author on how a lake would have responded to a drier climate we used the geological evidence presented in the paper. For example, if a Late Pliocene geological outcrop containing one or multiple layers of fluvial**

**or sub-aerial sediments was within the reconstructed boundaries of a lake then at some point in the Late Pliocene the lake must have been smaller. These were then used as marker points to recalculate the lake dimensions and hence the surface area. This incorporates an element of uncertainty and with continued work more accurate information may be published on how many of these lakes responded to changes in Pliocene climate.**

**Each lake in the database represents a different sedimentary basin, the vast majority of which have no direct link to any of the other basins. As has been previously discussed in Peters and O'Brien (2001), our current geological understanding of these lakes means we cannot conclusively state that any of these lakes definitely co-occurred with any other.**

**Specific uncertainties of the Late Pliocene mega-lakes**

**In the manuscript we highlighted the specific uncertainties surrounding Mega-lake Zaire as it has the greatest uncertainty, in the following supplementary section we shall detail the uncertainties of the other mega-lakes.**

(C1616) 1/ You need to detail how the lake surfaces were calculated, and for both scenarios (wet and dry). I did not check the extent of all lakes, but of the two lakes I checked, I cannot reproduce the extents you find, and that's a problem: I calculated the extent of lake MegaFazzan, based on the Upper Pliocene extension (Figure 11 from Drake et al., 2008): the lake covers approximately on grid cell of the figure, i.e. roughly 27000 km2 . Drake et al. 2008 mention 135000 km2 in the text during more humid periods of the Pleistocene. Hence I don't understand how you find the 152856 km2 mentioned in the Supplementary material. What are your assumptions in order to find this extent? Moreover I think there is a problem with the longitude of the center of Fazzan (27 E, line 59 of Suppl.)

**We calculated a lake extent for Mega-Fazzan using the scale bar on the map, this gave a length x width of 396 x 386. The lake covers more than one grid cell, which in Figure 11 of Drake et al. (2008). It is also worth noting that the grid cells in Drake et al. (2008) are 2° latitude x 2° longitude in size, thus one grid cell would be 90132 km$^2$, not the 27000 km$^2$ area calculated by the reviewer. With regard to the longitude of Fazzan the reviewer is correct, this has been corrected.**

(C1616) For lake MegaChad, you give 446760 km2 for wet scenario and 296378 km2 for dry scenario, and again, there is absolutely no way to know what are your assumptions in order to find these extents. Moreover, you cite Schuster et al. 2009, who mention a surface of 'more than 350000 km2 ' for a maximum level of 325 m asl during humid periods, and Otero et al. 2010, who mention frequent connections between the Chadian and Niger provinces (via the outflowing of MegaChad in the Benue river) during the Pliocene based on results from Otero et al., 2009, PPP. Some information on lake surface is given in Ghienne et al. 2002, who give a relation between lake level, volume and extent in table 1, computed from the TOPO6 dataset. They find an extent of 448000 km2 for a level of 321 m asl. Nevertheless, Leblanc et al. 2006 using another topographic dataset (SRTM 30), suggest an area of 340400 km2 for a maximum lake level of 325 m asl (see Leblanc et al. 2006, text and figure 6). So, there is still uncertainty, not on maximum level, but on the maximum extent. This should be included in the discussion of uncertainties, and you should explain how you find 446760 km2 for the wet scenario. For the dry scenario, what did you assume for the level? Schuster, 2002 describe alternating dry to wet episodes during the Pliocene in the Northern subbasin. I don't think one can have more information than that, so how did you do? I would probably have taken something like 290 m asl (i.e. a 'big Chad lake', with the Bahr el Ghazal valley being inundated, although there is no way to know if that's a good assumption), which gives an extent of 140000 km2 with Leblanc's curve, and 180000 km2 with Ghienne's curve. In short: what are your assumptions on extents? What process lies between the surfaces published in the literature and the extents you give? How do you calculate wet and dry extents?

**As the reviewer highlights (with a focus on Lake Chad) there is considerable uncertainty surrounding the reconstruction of lakes during the Late Pliocene. In response to a previous comment we have attempted to provide additional supplementary information outlining the uncertainties surrounding the mega-lakes. In specific response to this comment: we calculated our extent of mega-lake Chad using the map, descriptions and geological information published in Schuster et al. (2009) and Otero et al. (2010). For our wet-lake scenario we calculated the extent based on figure 1 in Schuster et al. (2009) and as our calculated spatial extent lies within the previously calculated estimates, although towards the upper estimate of Ghienne et al. (2002), we decided to use that. However, we acknowledge that we should provide the reader with some scale of the uncertainty and we feel this has been provided in our additional sections to be added to the supplementary information in our revised manuscript. To generate the extent of mega-lake Chad for our dry-lake scenario we used the geological column of Koro-Toro (Schuster et al., 2009), dated in Otero et al. (2010) as an approximate edge of the lake. An assumption based on the presence of fossil roots thought to represent plants growing in a seasonally wet and dry climate and the variability of the geology, from clay rich sandstone to lacustrine sediments, when compared to the massive pelites and diatomites further up the geological column.**

### 2.2. Preparing the data for inclusion in a climate model

(C1648) P3180, line 18: Remove “due to”

**Done**

## 2.3. Uncertainties in reconstructing soils and lakes from geological data

(C1648) Inceptisols and Entisols are stated as not included in the soil synthesis. Firstly, it is important to include a description of what these actually represent – Climate of the Past has a wide readership that may not be familiar with such terminology! What might the effect of omitting these types of soils be considering they cover a large proportion of the Earth today? What are the properties that might change the climate (e.g. albedo values)?

**For a brief decription of what an entisol and inceptisol represent see Page 3181 Lines 12-13. In addition to this and in response to the second part of this comment we would add to this section: It is difficult to assess the likely impacts on albedo and texture, had we been able to determine the Late Pliocene distribution of Inceptisols and Entisols, as these soil orders have limited pedogenic development they are more intimately tied to their parent material than the other soil orders. This could mean that Inceptisols and Entisols could have any combination of albedo and texture depending on Late Pliocene surface geology.**

(C1648) The discussion merits a section on using the Pliocene biome reconstruction by Salzmann et al. (2008) to determine soil type in regions where there is no data. It is my understanding that this reconstruction was created using data from BIOME4. Therefore, your results from BIOME4 have inherently been influenced by the fact that the boundary conditions are no full independent of this model.

**The Late Pliocene biome reconstruction from Salzmann et al. (2008) is a hybrid map combining 240 palaeobotanical data sites with a “best-fit to data” BIOME4 output (forced by HadAM3 predicted Late Pliocene climate), which were merged using expert knowledge. Although this means that limited regions of the biome reconstruction do rely more on model**

**predictions than real data, the overall product is primarily based on an exhaustive database of Late Pliocene plant fossil localities. Developing a global Late Pliocene soil map with only 54 paleosol localities required either extensive interpolation (with all the possible errors that may have come with that), or the use of another dataset (the hybrid biome reconstruction) and the knowledge that most soil orders (with the exception of Inceptisols and Entisols) are related to particular vegetation types.**

### 2.4 Modelling

(C1616) You need to describe, even briefly, how lake surface is treated in your model. Are there several layers? How is the temperature of the lake calculated? Has this model (MOSES2.2/TRIFFID with lakes) been used for other studies (paleo or present) and have the results been confronted to data?

**We have already mentioned that lakes are added to the model by increasing the amount of surface water, while reducing other surface types as appropriate. In essence this is a very simple way of representing a lake. It has only 1 layer (the surface) and is not dynamically changed throughout the model run. Large amounts of precipitation will not make the lake larger and large amounts of evaporation will not make the lake smaller (this is already mentioned in the text).**

(C1616) Page 3183, lines 5-10: 'the initial vegetation pattern for the control run was prescribed from PRISM3D' is that true just for the control run or also for the other simulations? Please clarify.

**The initial vegetation pattern for the control run was prescribed from PRISM3D data, and vegetation was dynamically altered by the model throughout the initial 500 year spin up. The simulations with prescribed lakes and soils were run with the all boundary conditions (including vegetation) from the end of the 500 year spin up; so they were started from**

**with a vegetation pattern predicted by MOSES2/TRIFFID after 500 years. This will be added to a revised manuscript.**

(C1616) It is necessary to show the soils map used in the control experiment, for comparison, otherwise the reader does not know where the changes are located. It would also be useful (if not necessary) to provide a map of soil albedo changes and soil texture changes (see Specific comment Page 3184, lines 4-6) for the reader to understand where the changes in temp/precip come from.

**There is no one soils map in the PRISM3 dataset. We have therefore included an anomaly map of the albedo (the most influential parameter) in the manuscript and all other parameters in the supplementary information.**

(C1635) How were the other geological boundary conditions (f.e. elevation model; coast lines; what about the Black Sea?) etc. ) for the modelling experiments defined? Even if these boundary conditions are described in Dowsett et al. 2010 some of the major points should be described in this paper to avoid questions or unclarities concerning e.g. tectonics, sea level and other mechanisms.

**This will be included in a revised manuscript**

(C1648) The model description set-up is quite limited. Was TRIFFID run in dynamic or equilibrium mode? This could make a difference to the spin-up of the model in terms of vegetation type. Trees for example take more than a thousand years to equilibrate typically and therefore equilibrium mode is normally required due to computational expense. More details are required, especially if other groups wish to reproduce your results using the boundary conditions you have produced. Please also include information about the albedo of the lakes.

**TRIFFID was run in equilibrium mode for the first 50 years of the control run. After this TRIFFID was run in dynamic mode throughout. This is clearly**

**sufficient to spin up the vegetation of the control run. The climate of the lakes and soils runs is similar to the control run and it appears the vegetation has reached equilibrium at the end of the 350 years of this run – as it is not changing. This will be included in a revised manuscript.**

**The albedo of water is 0.06. (This compares to albedo's of:**

**Broadleaf trees =0.1; Needleleaf trees=0.1, c3/c4/shrub=0.2, ice=0.75.) This looks like the changing albedo of the lakes will have only a minor effect on the climate.**

(C1635) Page 3183, line 16: surface type "urban" – why is this type used in a Late Pliocene simulation? Please explain or change and assign the urban surface type according to the fraction of the other surface types of the grid.

**Urban is one of the surface types that is allowable under MOSES2.1/TRIFFID and was included in the model description for completeness. This is not used in the Pliocene run as there was no urban land fraction in the Pliocene. The text will be changed to make this clear.**

(C1648) P3182: Please put a clarification of why you do not include the dry-lake scenario in your modelling. It would have been interesting to see the sensitivity of the climate to this uncertainty.

**We will explain in the manuscript that the changes between the dry-lakes experiment and the wet-lakes experiment were minimal and supply the climate results in the supplementary information**.

(C1635) Page 3183, line 26: Table 1: add the albedo numbers to an additional column in Table.

**This is already in the text – but can be added to the table. Light soils have albedo of 0.35, medium soils have albedo of 0.17 and dark soils have albedo of 0.11**

(C1616) Page 3184, lines 24-25: "The BIOME4 model was driven from the average annual climate data obtained from the last 30 yr of each HadCM3 experiment'. So you did not use the anomaly procedure to force BIOME4? This seems inconsistent with previous work from Salzmann et al (2008) and Pound et al. (2011). Moreover, you do not precise which CO2 level you used to force the BIOME4 model, and which resolution you use
**Yes we did use the anomaly method. Sorry this was not clear in the text. We used 405ppm CO2 (consistent with the HadCM3 simulation). The resolution is the same as HadCM3 3.75deg * 2.5deg. The manuscript has been changed to make this clear.**

## 2. Results

**All the following comments refer to the results section. Most of the comments relate to a need to include quantitative information when describing the results or discussing mechanisms that have operated to change the climate. We have actioned these two key points in our results section and thank the reviewers for the suggestions that have improved the manuscript.**

(C1616) You need to talk about all the large features seen on the figures, especially changes seen over the oceans. For example, in the precipitation response to changes in soil and soil+ lakes, there are some strong modifications of the ITCZ, with +/- dipoles suggesting a shift of the ITCZ, particularly in the Pacific Ocean. There is also warming around Greenland and seasonal cooling in Antarctica. These features are not discussed in the text, do you think they are robust, or do you think it is an artifact? I doubt that the lakes or even soils can have such an impact away from the changes, and this is not what you claim in the paper. Whatever the reasons for these changes, you cannot omit to talk about it. My feeling is that these changes (as well as those over the Amazon basin maybe) could be related to internal model variability. You have to check if a longer climatological mean (i.e. 50 years or even 100 years, since you have 350 years of integration this should not be a problem) would reduce the differences which are seen away from the zones of soils and lake changes.

**We have looked in more detail and agree that a 30 year climatological mean is not sufficient for this study and have increased the climatological means to 100 years. The original reason for using the final 30 years of the simulation only was to ensure that they dynamic vegetation in TRIFFID was fully spun up and providing the correct feedbacks to the climate. However, it appears that this is a minor consideration compared to the internal model variability that appears in a number of regions on 30 year timescales and provides misleading results on the effects of lakes and soils.**

**We now average over years 250-350 of the simulation. The ITCZ changes now do not occur in some of the simulations. This is because this feature is due to internal model variability in the ITCZ that can happen over large timescales (note however that the ITCZ signals we saw in the previous figures were not due to a shift but were all overprinted on the maximum rainfall (which gives only about a 10% change in rainfall over a very small area), and so were not as large as they appeared from the original anomaly plots. Since, whether or not the signals appear is dependent on the averaging period used, they are extremely unlikely to be attributed to Pliocene soils and this will be mentioned in the new version of our paper.**

**Amazon – reduction of precipitation in soils run. The signals still there and are there in every averaging period I have looked at; along with the reduction in precipitation (which has dropped by about 10%) we also see a reduction in evaporation, and a reduction in soil moisture at the deepest**

**soil level. It appears that the changes over the Amazon are robust and due to the differing boundary conditions that our model received due to the changes in soil parameters.**

**We have carefully looked at all the soil parameters that we changed between standard HadCM3 soils and Pliocene soils and the only parameter change to affect the Amazon was the Clapp-Hornberger beta exponent. Which was increased from between 6 and 8.5 (depending on region) to 9 over the whole Amazon with the new soils**

**Greenland – soil changes in DJF do appear robust. (Slight warming near the poles is not robust as it was not present in a different averaging period). Greenland – lake changes have totally disappeared with the new averaging period, however we now see a cooling of between 0.5deg and 1.0deg over a small region of the N. Atlantic at around 45N. This is extremely unlikely to be due to any of the lake parameters and is instead likely due to long term intrinsic model variability. Although we have presented only results in this paper that are highly significant, the nature of significance testing along with the fact that different years of a model run are not independent (but instead are autocorrelated) it is unavoidable that we will see features in the model results that are not dependent on the forcings used.**

(C1616) You need to relate the changes in precip/temp to the changes in boundary conditions AND explain the mechanisms behind it (otherwise you cannot state that you do an 'analysis of climate feedbacks'). For example, you say 'lakes in Australia have created a small reduction of desert in this region' (page 3190, lines 12-13). However on figs 4 and 5, there are no changes in temperature over Australia, and a decrease in precipitation in Pliocene Wet lakes experiment. What mechanism is at stake then? Is it a difference in cloud cover? Minimum temperature? In summary, I think it will be easier for you and for the readers to understand the underlying mechanisms if you provide more figures, in particular albedo and roughness changes for soils, but also for lakes (a lake surface is generally darker and flatter, and this can change temperature and winds, and hence modify pressure and precipitation).

(C1648) Although the changes in vegetation/precipitation/temperature are described in this paper when Pliocene lakes and soils are included there is very little discussion of the mechanisms that are actually operating. Why is the precipitation greater in certain regions when lakes are included? You need to explain this in terms of the atmospheric-land surface dynamics that are occurring. Quite a few inferences are made where the results could be examined in more depth i.e. albedo changes can be plotted along with energy balance maps over the regions of interest.

**Precipitation is likely greater in regions when lakes are included due to the recycling of lake water. (ie water will evaporate from the lake, which will increase humidity/ form clouds and increase precipitation near the lake).**

(C1648) A separate plot to show non-linearity when wet lakes and soils are included together compared with when they are separately included is needed. It may also be useful to perform a statistical analysis to determine whether the presence of realistic soils and lakes is more important in different regions.

**We feel a non-linearity plot is no longer needed as the new climate plots show the linearity clearly. It would be difficult and confusing to do a statistical analysis to determine whether the presence of realistic soils and lakes is more important in different regions. This is because the studies we are doing are not idealised, we are making very different changes in different regions, and there are very different local considerations. So for example, if a soil type has changed between a region, the climate response is a combination of many factors including soil albedo changes, changes due to other soil parameters (see supplementary information), the base climate, the base vegetation coverage and atmospheric circulation in the region and other non-local effects. The best way to show how changes in soils correspond to climate is perhaps the simplest, and is what we have attempted in this version of the paper. So for example we can by comparing figures showing boundary condition changes (e.g. soil albedo) with figures**

**3 and 4. We see that there is a clear relationship between soil albedo and temperature, but sometimes other processes are important (eg. The Amazon). At a first glance it would be easier to do a statistical analysis for lakes, but again we have included different sized lakes at every location so we would not be able to say with certainty whether differences at different locations were to do with the size of the lakes, the locality, the underlying climate/vegetation etc. Although we can qualitatively discuss which lakes appear most important to the Pliocene climate and in which season these lakes have maximum impact, to do a full statistical analysis would require a more idealised modelling study which is beyond the scope of this paper.**

(C1648) Although the authors describe the changes in vegetation for their different sensitivity experiments it is essential to include some quantitative results of the changes. For example instead of just stating shrubs expanded, include the percentage increase in shrub in a particular region. It is very difficult to see in the current figure.
Although the authors describe the changes in vegetation for their different sensitivity experiments it is essential to include some quantitative results of the changes. For example instead of just stating shrubs expanded, include the percentage increase in shrub in a particular region. It is very difficult to see in the current figure.

**We do not refer to individual PFTs in the text as we focus on the output from the BIOME4 model. This produces biomes based on the dominant PFT in an individual grid cell. We will ensure that the text is clear on this and we will add red boxes to the figure to highlight the regions we discuss in the text.**

### 3.1 Late Pliocene soils

### 3.2 Late Pliocene lakes

### 3.3 Impact of soils and lakes on simulating Late Pliocene climate and vegetation

(C1616) Section 3.3 needs to be re-written, changes in temp/precip quantified and related to boundary condition changes and explained.

(C1616) On the modelling side, I do not think the paper does an 'analysis of climate feedbacks' as suggested by the title. Such an analysis implies to detail the mechanisms underlying temperature, precipitation and biome changes. This is too rarely done in the text, and there are no figures to illustrate those potential mechanisms. Moreover, the changes seen in the temperature and precipitation are not quantified nor correctly described, and some features are simply omitted in the text (especially changes occurring above the oceans).

(C1635) Some of the large differences between the control run and the presented data were not described and/or discussed: - changes on the ocean (f.e. mid-Pacific MAP /Pliocene soils: how can the soil data create these large changes over the ocean?) - some of the differences between the experiments are only marginally described: is it possible to quantify instead of using expressions such as "a small increase", "modest increase" etc

(C1648) Although the authors state quite clearly there has been an improvement in the regional data-model comparison for the Pliocene when realistic soils and lakes are included there is no quantitative evidence of this in the manuscript. Could the authors compare their temperature and precipitation patterns with available late Pliocene data and show statistically that there is closer match when realistic soils and lakes are included compared with the control?
**We have temperature and precipitation estimates for a limited number of palaeobotanical sites for the Late Pliocene. For the regions identified in this study as showing changes, the number of numerical proxy data sites are too low to do meaningful statistics.**

(C1635) Page 3188, line 13: please specify "0.1% confidence level" and/or give example.

**Discussed in the text**

(C1616) Page 3188, line 14 : biome instead of 'BIOME'
**Done**

### 3.3.1 PRISM3 + soils

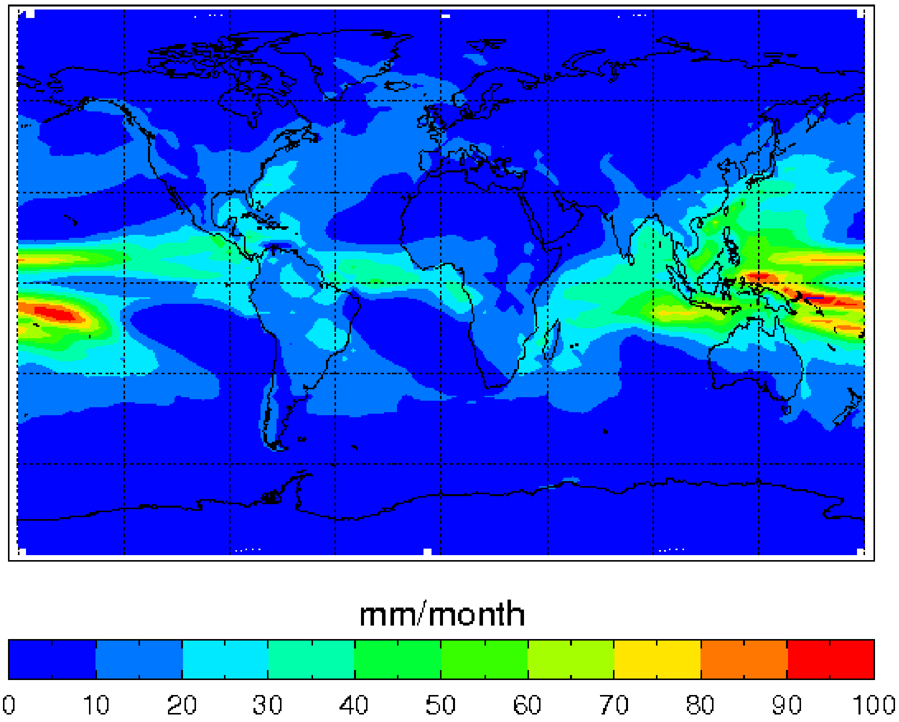
(C1616) Page 3188, paragraph 3.3.1 : about temperatures: There is also a cooling in Northern Red Sea, as well as the extreme East of the Arabian Peninsula on the annual mean. You don't talk about the warm oceanic anomaly surrounding Greenland in annual and DJF means. This feature has to be explained, and it seems to be robust because it is reproduced in the Soils + Wet lakes experiment. Moreover, all these features have to be related to the corresponding changes in soils. If you don't relate the temp and pre-cip changes to soils changes, there is no point in explaining these differences. There are also temperatures differences over Antarctica in JJA, which are not mentioned and not explained. About precipitation: 'a small increase', 'a reduction': please quantify. Actually, the biggest changes in precipitation are over South America, as you mention, but also over the ocean, especially tropical Pacific and Southern Indian ocean. Once again, these features are not mentioned and not explained. In particular, the tropical Pacific anomaly seems robust because it is reproduced in the Soils + Wet lakes experiment.

**We have now quantified the changes between the lakes/soils runs and the control runs.**

(C1648) P3188, line 25-: Have you looked at the standard deviation of the precipitation differences over South America to demonstrate model variability?

**We have looked at the standard deviation of the precipitation (plot below) and there is a small amount of standard deviation over the Amazon region.**

**As we discuss in the text (Section 3.3.1) we believe the change in MAP over the Amazon to be robust.**




### 3.3.2 PRISM3 + wet-lakes scenario

(C1616) Page 3189, line 8: ‘a modest increase’: please quantify, and on the color scale, it’s the darkest blue, so it doesn’t seem so modest.

**We have modified the results in response to all the reviewers’ comments requesting numerical values for changes in climatology and mechanisms that have changed the climate.**

### 3.3.3 PRISM3 + soils + wet-lakes scenario

## 4 Discussion

(C1616) Notably, you compare your results with Krinner et al., 2012, which focused on the Mid-Holocene, but not to Coe and Bonan 1997, Bröstrom et al. 1998, Sepulchre et al., 2009, which also investigated lake Chad feedbacks in the mid-Holocene, neither to Burrough et al. 2009, which investigated megalake Makgadikgadi feedbacks during the LGM, or to Contoux et al. 2013, which is the first study to focus on the lake feedbacks in the Pliocene, although only for megalake Chad (see Specific comments).You could also note that Sepulchre et al., 2009, Krinner et al., 2012 and Contoux et al., 2013 find a decrease of precipitation above the Megalake Chad, which is different from your results. You could also discuss the reasons for this discrepancy (note that Contoux et al. 2013 found that response away from Megalake Chad was dependant on boundary layer parameterization, but that the drying response over Megachad was robust)

**We have added more information to the discussion, including the Contoux et al. (2013) reference recommended by the reviewer.**

(C1616) Page 3191, line 9: 'the global distribution .. is significantly different from present day' especially for soils, this cannot be assessed on figure 2. Please include Control run soils map.

**Due to copyright we are unable to provide modern distribution maps for soils and lakes. Further to this modern soils and lakes have been heavily influenced by humans and may no longer represent a "natural" distribution and or extent. We have however included references where an interested reader can find the maps.**

## 5 Conclusions

(C1648) It would be beneficial to include a statement about future avenues of research such as using a surface water scheme model (e.g. HYDRA) coupled to a climate model to understand the implications of a two-way feedback on Pliocene climate.

**We agree and our final sentence in the conclusions states that lake and soil feedbacks should be further explored in future palaeoclimate modelling.**

(C1616) If the future of this database is to be used in climate models, it would be useful to know if one can get the netcdf files for lake percentage and soil coverage, and from where one can get them
**The boundary conditions will be available on the USGS PRISM4 website. However, the recent shutdown of the U.S. federal government has prevented us from generating a web link.**

**Figures:**

(C1616)Fig 1: you should also precise that triangles are soil data and circles lake data. – I **This will be included in the figure caption**

(C1616)Fig 2: the map of soils used for the control simulation is necessary, to assess the differences between Pliocene control and Pliocene + soils simulations.

**There is no one soils map in the PRISM3 dataset. We have therefore included an anomaly map of the albedo (the most influential parameter) in the manuscript and all other parameters in the supplementary information.**

(C1616)Fig 3 and 4: the caption of these figures should be something like this "Top : mean annual : : : Middle : same for boreal winter (December to February). Bottom: same for boreal summer (June to August)." Moreover I think these figures are really too small (I have to zoom up to 600% on my screen, and the colors are quite blurry), and there are no marks on the maps. The scales are quite unusual too.

**The figure captions will be modified and we will endeavour to make the panels as large as possible within the constraints of an A4 page**

(C1616)Fig 5: Once again I think the panels are too small. The key for the biomes cannot be read unless zooming at +500%. Suggestion: the changes are small between each panel. It would be easier to locate them if you could find a way to plot only the grid cells where the biome is different than the Pliocene control

**Once again we will endeavour to make the individual panels and key as large as possible, within the constraints of an A4 page. Although we cannot produce maps just showing the grid cells that change biome, as this would remove other important data (such as the relation of biomes to one another). We will highlight the regions we discuss in the text to make it easier for the reader to locate themselves.**

(C1635) Figure 2: Please add a map of the present day conditions and a map of the used control run (standard Prism3 control) Figure 3, caption: change to: The differences of mean annual surface temperature for the soils and lakes experiments from the standard Prism3 control. Figure 4, caption: change to: The differences of mean annual precipitation for the soils and lakes experiments from the standard Prism3 control.

**Figure 2: Due to copyright we are unable to provide modern distribution maps for soils and lakes. Further to this modern soils and lakes have been heavily influenced by humans and may no longer represent a "natural" distribution and or extent. We have however included references where an interested reader can find the maps. We have also included the parameter anomaly maps in the supplementary materials. Figure 3: Caption is changed. Figure 4: Caption is changed.**

(C1648) P3191, line 9-10: It is important to show that the distribution of soils and lakes is different than today by including maps of present distribution.

**Due to copyright we are unable to provide modern distribution maps for soils and lakes. Further to this modern soils and lakes have been heavily influenced by humans and may no longer represent a "natural" distribution**

**and or extent. We have however included references where an interested reader can find the maps.**

(C1648) Figure 2: The colour scale for the lakes map is quite difficult to interpret when printed. Please make this clearer.

**We will modify this to make it easier to see when printed**

(C1648) Figure 3: The colour scale is saturated at the lower end when printed. Please modify. The figure caption needs to state that these are anomalies, whether they are significant and what the reference climatology to which the anomalies are calculated.

**Anomalies relative to the standard PRISM3 control run (with modern lake and soil distribution). Everything plotted is significant at the 0.1% confidence level. Colour scale has been modified.**

(C1648) Figure 4: Again when printed the colour scale is saturated at the upper and lower bounds. Please also state that these are anomaly plots.

**As above**

(C1648) Figure 5: It is very difficult to distinguish the differences between these sub-plots (see point 4 above). In particular, the synergy of Pliocene soils and lakes feedbacks is not clear compared with when only one of these land surface attributes is changed. I suggest alongside these plots also showing difference plots of Leaf Area Index. This way it will be easier to see where obvious changes in vegetation occur.

**We have focussed on biomes as these represent a means to relate model results to palaeobotanical data in a realistic way. Using biomes rather than the individual PFT distributions or individual characteristics (such as LAI) also widens the readership of a publication (please see the diversity of papers citing Salzmann *et al.*, 2008; Pound *et al.*, 2011).**

**References used in replying to reviewer comments**

Chang, H., Ao, H., An, Z., Fang, X., Song, Y., Qiang, X., 2012. Magnetostratigraphy of the Suerkuli Basin indicates Pliocene (3.2 Ma) activity of the middle Altyn Tagh Fault, northern Tibetan Plateau. Journal of Asian Earth Sciences 44, 169-175.

Dodson, J.R., Ramrath, A., 2001. An Upper Pliocene lacustrine environmental record from south-Western Australia - preliminary results. Palaeogeography, Palaeoclimatology, Palaeoecology 167, 309-320.

Drake, N.A., El-Hawat, A.S., Turner, P., Armitage, S.J., Salem, M.J., White, K.H., McLaren, S., 2008. Palaeohydrology of the Fazzan Basin and surrounding regions: The last 7 million years. Palaeogeography, Palaeoclimatology, Palaeoecology 263, 131-145.

Matthews Jr. JV, Westgate JA, Ovenden L, Carter LD, Fouch T, 2003. Stratigraphy, fossils, and age of sediments at the upper pit of the Lost Chicken gold mine: new information on the late Pliocene environment of east central Alaska. Quaternary Research 60, 9-18.

Pound, M.J., Haywood, A.M., Salzmann, U., Riding, J.B., Lunt, D.J., Hunter, S.J., 2011. A Tortonian (Late Miocene, 11.61-7.25 Ma) global vegetation reconstruction. Palaeogeography, Palaeoclimatology, Palaeoecology 300, 29-45.

Pound, M.J., Lowther, R.I., Peakall, J., Chapman, R.J., Salzmann, U. Palynology, palaeoenvironment and age of the gold bearing Late Pliocene White Channel Gravels from Bonanza Creek, Yukon Territory. Manuscript in preparation.

Sáez, A., Cabrera, L., Jensen, A., Chong, G., 1999. Late Neogene lacustrine record and palaeogeography in the Quillagua–Llamara basin, Central Andean fore-arc (northern Chile). Palaeogeography, Palaeoclimatology, Palaeoecology 151, 5-37.

Salama, R.B., 1987. The evolution of the River Nile. The buried saline rift lakes in Sudan - I. Bahr El Arab Rift, the Sudd buried saline lake. Journal of African Earth Sciences 6, 899-913.

Salzmann U, Haywood A.M., Lunt D.J., Valdes P.J., Hill D.J., 2008. A new global biome reconstruction and data-model comparison for the Middle Pliocene. Global Ecology and Biogeography 17, 432-447.

Schuster, M., Duringer, P., Ghienne, J.-F., Roquin, C., Sepulchre, P., Moussa, A., Lebatard, A.-E., Mackaye, H.T., Likius, A., Vignaud, P., Brunet, M., 2009. Chad Basin: Paleoenvironments of the Sahara since the Late Miocene. Comptes Rendus Geoscience 341, 603-611.

Tiercelin, J.J., 1986. The Pliocene Hadar Formation, Afar depression of Ethiopia. Geological Society, London, Special Publications 25, 221-240.

Yeniyol, M., 2012. Geology and mineralogy of a sepiolite-palygorskite occurrence from SW Eskisehir (Turkey). Clay Minerals 47, 93-104.