Author Comment to 'Review' in RC1

1. Explain the use of the *.md5 files

*The *.md5 files contain a md5 checksum which can be used to check whether changes have been made to the respective *.nc files. The md5 checksum of the .nc files should thus be identical to the ones in the *.md5 files. Such checks can be done with several freely available online tools. We will clarify this by explaining it on the dataset webpage of the data repository (the NIRD Research Data Archive).*

2. Comment on the novelty of the dataset

The reviewer is of course right that the PMIP model results underlying our dataset are publicly available, and that everyone could go to the PMIP archives and repeat our efforts. The assertion that this could be done "very straightforwardly", however, underestimates the amount of work put into the compilation of this dataset considerably. There has been a long series of decisions and considerations involved, starting from compiling data availability across variables and models, to scientific questions of how to treat differences in topography and land-sea mask, and how to re-reference the temperature and humidity fields. The authors expect that the availability of i) a direct visual impression of PI-LGM PMIP3 anomalies and model spread, ii) the clear description of the concept (f.e., CDO/NCO procedure and the idea of anomaly addition to original forcing) and iii) the freely available download of the relevant variables in one standard format will be valuable for future LGM modeling attempts. Indeed, the authors have already received interest from several colleagues about the dataset. Moreover, users inexperienced with PMIP/CMIP/CDO/NCO now have a clear presentation of this opportunity in LGM modeling, which increases the likelihood of modeling attempts by such users. Also, models for which a fully coupled LGM run is not (yet) available, or models that are inherently ocean-only, now have a relatively accessible way to simulate the LGM without needing to prepare such forcing datasets themselves.

3. How does one estimate NY (intra-annual) forcing typical of G-IG anomaly?

The construction of a "normal year" forcing (NYF) is indeed not trivial. Large and Yeager (2004; http://opensky.ucar.edu/islandora/object/technotes:434) describe their procedure for the CORE NYF in detail. However, our objective is not to construct a LGM normal year forcing in a strict sense, but to provide LGM-PI anomaly fields that can be used in combination with the existing CORE (and other) forcing fields. If added to the CORE NYF one obtains a LGM "normal year" forcing under the assumption that –beyond changes in the monthly mean state– the temporal and spatial statistics of the forcing fields are the same under LGM conditions. We see that this assumption was not made explicit in our manuscript, and we propose to add text to clarify that we focus on CORE forcing formats, and that the use of an anomaly dataset implies certain assumptions.

4. How should we handle the range of inter-model spreads?

The authors address how to use the model spread at the end of the manuscript: '..., all mean anomalies show a distinct spatial pattern that we expect to be indicative of the LGM-PI changes. […]. For modelling purposes, the inter-model disagreement of PMIP3 provides the user with leeway to adjust the amplitude of the forcing (within the model spread). Such adjustments can improve model-proxy data agreements, such as described for salinity in Sect. 3.7.' Moreover, such adjustments, or 'tuning', of an ocean-only model could inform fully coupled models by revealing model sensitivities. A good example of this is SSS, for which the PI-LGM anomaly is likely too large in the North Atlantic for most PMIP models, pointing to a potential limitation in hydrological cycling of these models (Sect. 3.7).' We acknowledge that we do not know how to objectively handle model spread beyond using it to "justify" model tuning – but think that a discussion of PMIP3 model spread is beyond the scope of the article (p. 6, l. 24-25). The user will be provided with the model spread for each variable in order to improve the usability of the model spreads (see also comment 7).

5. How can we include information from available observations?

We believe that for most atmospheric variables (e.g., wind, humidity, radiation), global coverage through LGM proxy data will be never achieved. For model forcing we thus have to rely on model simulations of the variables of interest. Comparison of model results to proxy data that are available (such as estimates of AMOC strength, sea ice extent or productivity for example) is a useful tool for model evaluation (Braconnot et al., 2012). For some atmospheric variables however (e.g., air temperature) one could use proxy data to guide corrections to the model mean anomaly (in addition to the model spread). We leave it to the individual modeling groups to adjust the mean anomalies if considered necessary/needed for their purposes. We will extend the text with information on how the authors think observational/proxy estimates can best be used.

6. Why are some fields interpolated to 6-hourly and others to daily fields?

Our intention is to provide the data in CORE forcing format (p. 3, l. 15-16), which could be either NY or interannually varying CORE forcing. For the variables specific humidity, wind components and air temperature CORE forcing is based on the NCEP-reanalysis, which has a standard time resolution of 6 hours (Large and Yeager, 2004). For radiation fluxes, daily is the highest time resolution in the CORE forcing (Large and Yeager, 2004), and therefore used in our dataset. Similarly, for precipitation and SSS, we conform to the CORE forcing standard (monthly time resolution; Large and Yeager, 2004). We will makes this clearer in the manuscript by extending the caption of Table 1, elaborating in lines 15-17 on page 3 and stressing our focus on the CORE format earlier in the manuscript/abstract.

7. While model spread is plotted in the figures, I could not locate this as a variable in the data.

We understand this can be a valuable additional variable for the potential user, and will add it to the dataset for all variables named 'variablename_spread'.

8. There may be some grid-scale interpolation artifacts along coastlines (Fig. 1)

The most pronounced changes in humidity occur along the coastlines. This is not an artifact of the time or vertical interpolation, but a consequence of the different lgm land-sea mask as compared to the piControl land-sea mask. As some ocean becomes land along most coastlines, a local reduction occurs there in specific humidity.

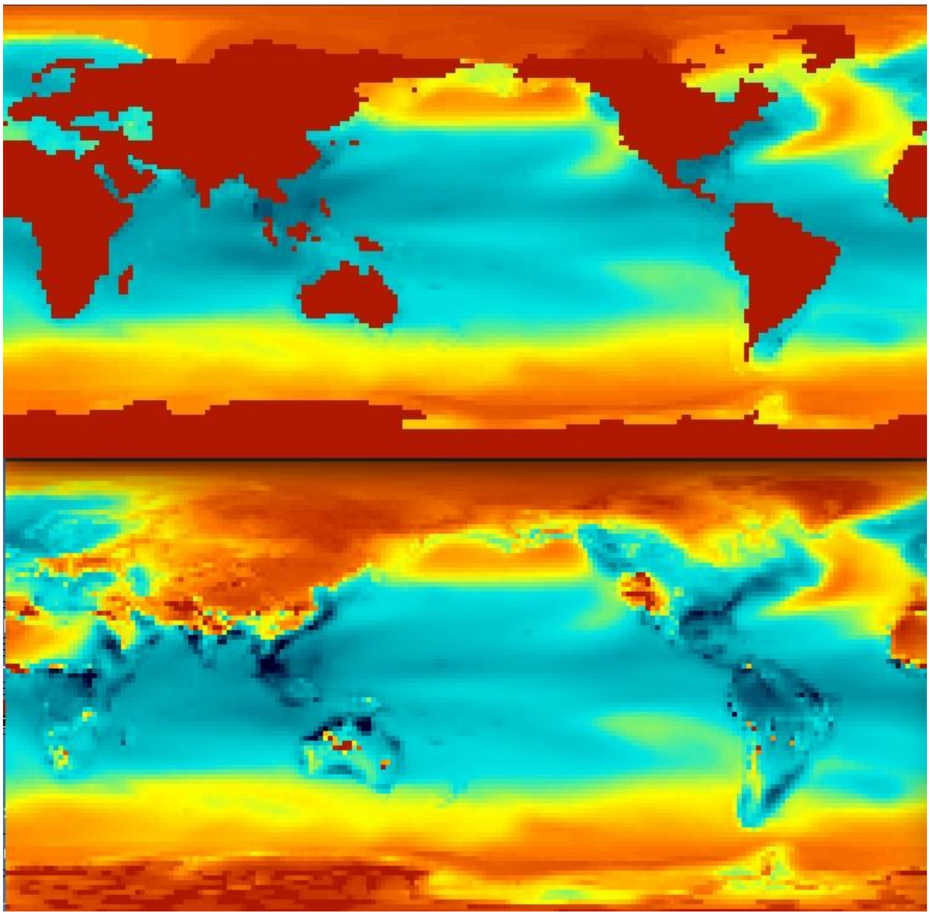
We realize that this is suboptimal, although it is only clearly visible for specific humidity and not relevant for groups that would apply an lgm land-sea mask. However, we want the files to be of use for modeling groups that want to use a pre-industrial land-sea mask as well (e.g. for idealized experiments). To remedy this inconsistency we propose the following approach:

1) Mask the multi-model mean anomaly as we present it now by the maximum lgm land area across all models,

2) Extrapolate variables over land using a distance-weighted average,

3) Mask the data with a decreased present-day land mask (decreased such that we insure that any pre-industrial land-sea mask is covered by the anomaly for all its ocean grid cells).

The area affected by land-sea mask changes will thus be filled with the extrapolated model mean anomaly. For humidity, an example of the result is given in the figure below (new anomaly above, old approach below. Note that we added an extra model – see comment 3 in RC2):



Coastal effects under the area affected by a changing land-sea mask are thus removed (for all variables). We note that the extrapolatation over the large Arctic continental shelves (which are land during the lgm model runs, and sea during the piControl model runs) leads to artificial structures in the specific humidity anomaly field (top figure). These structures are only relevant if a user applies a present-day land-sea mask, and - we expect - will not cause any problems as the anomaly gradients are similar to those in other regions (f.e. the North Pacific). In addition, no topography correction would be needed for surface temperature, as the field would only be based on open-ocean model data.

We tested this approach in NorESM-OC, and see no problems with the initialization caused by the masking.

9. The salinity fields look strange and unlike the figure in the manuscript (Fig. 2) (unless I'm doing something wrong to access the file; I used ncread('Salinity_anomaly_1deg.nc','sos',[1,1,4],[360,180,1]); in MATLAB).

The salinity field is the only field provided on a 1x1 degree grid. The field was provided for all grid-cells (extrapolated over land) such that it can be used with any land-ocean mask, as described in the text (Sect. 3.7). As this causes confusion, we propose to follow the same approach as described in our answer to comment 8 – thus providing the data on a decreased pre-industrial land-sea mask.

10. Please say what NCO is

NetCDF Operators (NCO) is a NetCDF toolkit (http://nco.sourceforge.net/#Definition) that can be used besides or in addition to CDO. We will clarify this in the manuscript text.

11. There are several thorny issues associated with forcing a model with multi-model means. For one, the fields are no longer dynamically consistent. An implication is that there could be strangely conflicting contributions e.g. to surface salinity from relaxation and precipitation. Second, have these models all been run to equilibrium? Third, computing ensemble means tends to damp uncorrelated variability between members, which reduces the variance of forcing fields. Is there a way to correct for this and generate a "normal year"?

i-a) Dynamic consistency: The CORE forcing itself is not dynamically consistent, since it is a blend of different sources (NCEP reanalysis, satellite and surface observations). This is apparently not seen as a major problem in the ocean modelling community (the CORE forcing is widely used, e.g. recently in CMIP6-OMIP). We are not aware either of any study investigating such dynamical inconsistencies in the context of ocean modeling. From this standpoint, we believe that the inconsistencies that are introduced by using a multi-model mean for our anomaly forcing are not a major issue. We will comment on this topic and how it can be prevented for certain variables (see also i-b) in the text.

i-b) Salinity relaxation and precipitation: Freshwater balance and salinity relaxation need special attention for multi-centennial forced ocean model simulations. Using the CORE forcing, precipitation is prescribed but evaporation and SSS evolve freely such that imbalances can and will develop. Groups *applying the CORE forcing (and our anomaly forcing) will be able to deal with such complications (e.g. enforce freshwater balance globally by adjusting precipitation, balance salinity relaxation to conserve total salinity). In NorESM-OC, this is for example done by enforcing global freshwater-flux balance (by adjusting precipitation with a global correction factor). Also, salinity relaxation is applied such that the global net flux of salt into the surface ocean is zero. Again, these issues occur already when the original CORE forcing is used and are not specific to our anomaly forcing. We will clarify this in the text.*

ii) Equilibrium: CMIP/PMIP models have been run to equilibrium as much as feasibly possible considering the high costs of computing. The authors comment on this in lines 1-6, p. 2. The output of the different CMIP/PMIP model experiments is considered a reasonable estimate of the past global climate state – which is required for our use. We expect that issues with persistent model drift more likely would arise for interior/deep ocean model fields. Nevertheless, we do acknowledge, that part of the inter-model spread may be explained by differences in the extent to which the respective PMIP model has approached equilibrium. As we comment on equilibration already in l. 1-6 on p. 2, we think no further action is required.

iii) Damped uncorrelated variability: We agree with the reviewer that this is an issue with our multi model approach that is not mentioned in our manuscript. The preferred method of using an ensemble of N fully coupled models to force a standalone ocean model would be to create an ensemble of N forcings and run the standalone model for each of the forcings separately. Of course, the drawback of this approach is the N-fold increase in computational resources (cpu and storage). We believe that for the sake of achieving long integrations of the LGM ocean state, it is justifiable to use a multi model mean anomaly. In addition, the anomalies are generally small (< 10 %) as compared to the forcing field. We see that it would be good to mention these considerations explicitly in our manuscript, and we will do so in a new version.

In general, we acknowledge that it is important to mention the sources of error coming from the use of multi-model means and will extend the text with comments on the main issues.

12. Are effects of evaporation included in the precipitation file?

The variable 'pr' presented in the dataset represents the total amount of precipitation (liquid and solid phases, and from all types of clouds – both large-scale and convective)). Evaporation is thus not included in the precipitation fluxes, and should be calculated by the model itself based on evaporative forcing (through f.e. temperature and humidity). We will improve clarity on this in the section on precipitation.

13. It would be helpful to provide a river runoff file.

The authors found that the CMIP river runoff variable (friver, 'Water Flux into Sea Water From Rivers (kg m-2 s-1)') is only available for CNRM-CM5, IPSL-CM5A-LR, MIROC-ESM and MRI-CGCM3 for the lgm. Besides that, the differences in land-sea masks of these models are problematic as river outlets may end up off-coast or on land when applied in a forced ocean model. Averaging over such 'point-sources' will create a little meaningful product in our opinion, as the river outlets vary *considerably between models. In NorESM-OC, we route the preindustrial river runoff to the nearest ocean grid-cell, but such solutions are very model dependent, and we expect that modelling groups devise a suitable solution for their specific model for the treatment of runoff. We will comment on river runoff in the precipitation section in order to address the relevance of river routing in LGM setups.*

14. Sect. 3.6 line 10 'due to' changes?

'due to' is indeed missing here – we will add it to the sentence. Thank you for noting this mistake.