7th of May, 2020

Dear Editor Kirsten Elger,

Thank you for your time to handle our manuscript. We have written a response to the comments of reviewer #3 below. Additionally, we included a general reply to some of the main points raised by the three reviewers (where reviewers #1 and #2 have reviewed the first version of the manuscript, and reviewer #3 has reviewed the second version of the manuscript) to further elaborate on our choices made for the third version of the manuscript. To address the comments of the different reviewers, we included a new section on the limitations and robustness of the dataset at the end of the manuscript, as well as a new figure (A1) and smaller changes throughout the text. There are revisions that we potentially could make to the data set (including basin wide river runoff anomalies and omitting the re-referencing procedure). Since we already have two data set versions archived with a doi, we refrained from preparing a new data set version at this time as discussed with you, and would like to await the further outcome of the review process.

Yours sincerely,
Anne Morée and Jörg Schwinger

General reply to some of the main points raised by reviewers

Since some time has passed since our first submission, we would like to start this response letter with some general considerations, taking into account the most important points raised by the three reviewers of our manuscript. We believe that there has been some misconception of what we intend to provide with our data set and what is needed and is practical in forced ocean modelling.

1) We do not aim to provide an LGM normal year forcing (as available data do not allow this). It is our *intention* to provide large scale, LGM-PI anomaly fields derived from PMIP3 models. This is not (only) due to data limitations, but, equally important, due to the fact that individual model results are highly uncertain, particularly at smaller spatial scales. Calculating estimates of large scale anomalies as a model mean is an established practise in climate research. One major point raised by the reviewers is that a model mean will be inconsistent with atmospheric dynamics (i.e. the underlying physical equations of motion/state will not be satisfied by the resulting forcing fields). This is correct, but not particularly relevant in the context of forced ocean modelling. The CORE forcing itself is constructed based on the NCEP-R1 reanalysis, to which either observation based corrections are applied (temperature, winds) or, for some variables (radiation and precipitation fields), the reanalysis data is replaced by observational estimates. No efforts are made to preserve dynamical consistency. This CORE forcing data set is widely used (lately in the CMIP6 endorsed OMIP model intercomparison), and we do not see a reason why something that is acceptable for CMIP6 should not be acceptable for our data set. Also, the use of model mean fields as boundary condition in modelling is not unheard of. E.g., HappiMIP (Mitchell et al., 2017) uses multi model average SSTs as forcing. Also outside of the paleoclimate and MIP community multi-model CMIP output is used as forcing (e.g., Chowdhury and Behera, 2019). Last, forcing

specifically with PMIP3 anomalies (only for fewer variables) is practiced in the state of the art studies by for example Muglia et al. (2015/2018) and Khatiwala et al. (2019).

The use of individual model anomalies (or absolute values) would be of little practical value, we believe (although we could easily provide these). First, this is an issue of computational resources. Very long integration times are necessary to run a forced ocean model into equilibrium with LGM forcing, particularly for biogeochemistry in general and carbon isotopes in particular. Using five model integrations based on five individual PMIP model forcings is not really an option. Even more important though, the ocean model's salinity restoring scheme has to be tuned (as described in our manuscript) for each individual forcing to produce a reasonable large scale ocean circulation. We do not believe that any ocean modelling group would want to apply this time consuming procedure for 5 or even more individual forcings.

2) Re-referencing of 2m to 10m quantities: This is a procedure that is applied to NCEP 2m temperature and specific humidity when the NCAR reanalysis data is processed for the CORE forcing. We have applied the same procedure to PMIP3 model output for the first versions of our data set. Reviewers have criticized this as a) inconsistent with the estimated boundary layer stability and the algorithms used in the models from which the input quantities are taken, and b) that using monthly averages as an input for this procedure is introducing errors. Both are valid points, and we have done some further analysis on point b), which is presented in detail below (see our specific response to reviewer #3, comment 1). We summarize the results here: Over the open ocean re-referencing only has a small effect (the difference between 2m and 10m temperature is less than 0.1 K for the vast majority of ocean grid points, Fig. 1 columns 1 and 3). The effect is larger over sea ice (typically around 1 to 1.5 degree) and land. The error made by using monthly mean inputs for the re-referencing is also small over the open ocean, but it can be substantial over sea-ice (and over land) (Fig. 1 column 2). Still, these errors are small compared to the uncertainty of the underlying model ensemble, which is larger than 8 K almost everywhere north and south of 60° (ensemble range of PMIP3 models, Fig. 1 last column). Results for specific humidity are discussed below (see our specific response to reviewer #3, comment 1).

Given this analysis, we conclude that it would be more robust to skip the re-referencing step altogether. For our data set we are taking the difference between two temperatures, so the effect of omitting the re-referencing is virtually zero over the open ocean (<0.1 K for the vast majority of open ocean grid cells; Fig 1 column 3). Larger differences occur over sea ice, particularly over the central Arctic Ocean. Here, the 2m temperature anomaly is up to 4 K larger (colder) than the anomaly based on 10m values. In practise however, these differences at very high latitudes will have little influence on model simulations forced by our data set, since the ocean is anyway insulated by thick sea-ice from the atmosphere. Also, the PMIP model ensemble range exceeds 20 K over the central Arctic Ocean.

Given the lack of time resolved input data, the unavoidable error made by using time-average input to re-referencing, and given the small impact on the actual anomaly fields (relative to the uncertainties), we propose to skip the re-referencing step in a revised version of our data set, and discuss the impact of this in our revised manuscript (in a new section, Sect. 4).

3) Using 3d model output

It has also been brought up that we could use more models if we would use 3D model output. In our revised data set, we have improved the surface salinity anomaly estimate by using 3D output (increasing number of available models from 2 to 5). It is also true that we could use 7 instead of 5 models for atmospheric conditions if using 3D output. However, we would need to calculate 10m (or 2m) temperature and specific humidity from the original 3D fields. Given our analysis on the errors made by re-referencing using monthly mean fields, this is not an option. We would like to stress that the calculated model mean anomalies are already quite robust with 5 models, i.e. the addition of one model from 4 to 5 models did not change the results significantly, as visible from the difference between version 1 and 2 of our dataset.

4) Freshwater budget

Our data set was criticized for not including river runoff flux anomalies, leaving out one component of the water cycle. This is correct, and we would be able to amend this in a third version of our data set. We propose providing basin wide total and fractional anomalies, which could be used to scale the pre-industrial river fluxes by modelling groups. We note that providing gridded anomalies for river runoff is not an option because of different land-sea masks and river mouth locations in the different models.

Author Comment to Review #3

We thank reviewer #3 for his/her time to provide constructive feedback on the version of our manuscript from 10th of September 2019. Our response to the three comments is provided below. Specifically, we propose to add data on basin scale river-runoff anomalies to our data set and to omit the re-referencing procedure, based on an analysis presented below. We also include a new section (Sect. 4) on the robustness and limitations of the dataset and a new Figure (A1) to address the reviewer's concerns.

Yours sincerely,
Anne Morée and Jörg Schwinger

1. **Computation of values @10m (t10m and q10m)**

The values at 2m height and the surface values are used to rescale the variables at 10m height. In atmospheric models, t@2m is not a prognostic variable. It is diagnosed with an iterative procedure (as in Large and Yeager, 2004). Input are surface values (temperature tas, pressure psl), and values at the first model level, generally between 10 to 100m height (temperature temp[k=1], wind u[k=1], v[k=1] and humidity q[k=1]). This computation of t@2m is an estimation, know to be very non precise, and not fully physically based. At least in the model it uses, or may use, the exact stability computed in the model, and the full high frequency outputs. Reapplying this procedure to recompute t@10m is prone to give large errors, especially because the estimated stability could be different from the one actually used in the model. As the temperature of the first level is available in the CMIP5 database, it would be better to directly compute t@10m from tas and temp[k=1]. At least, the authors should check on one model than applying the procedure twice gives the same result than the direct computation.

- Which wind u, v and humidity q do you use for the computation? It seems to be u@10m, v@10m, and q@2m, which for most model are also diagnostic variables computed with the same algorithm. This introduces an additional source of error.
- The procedure is applied on monthly mean. As the iterative procedure is highly non linear, is it justified?

Author response to comment 1
We have analysed the error made by using monthly means as an input to the re-referencing (the 2nd bullet point in the comment above). For this exercise we took the original NCAR reanalysis output (6-hourly time resolution), formed monthly means and applied the re-referencing. We compared this output with the correct re-referencing (i.e., 6-hourly output re-referenced and then averaged). We call the difference between the two the "re-referencing error" in the following text. We note that we can only derive this error for the pre-industrial state due to data availability, while our data set relies on the difference between LGM and PI quantities. If the re-referencing would have the same effect (and error) under an LGM and PI state, we could omit this procedure for creating our data set. Below we show that (over the ocean), the effect of re-referencing is small, except for sea ice covered regions. Consequently, anomalies based on 2m or 10m quantities are

virtually identical, again except for ice covered regions. Results of these analyses are shown in Fig. 1 for temperature, and Fig. 2 for specific humidity.

For temperature, re-referencing is mainly important over sea ice (disregarding the land, which is not relevant for our ocean forcing data set), where T(10m) is typically 1 to 1.5 K warmer than T(2m) (Fig 1, column 1). For the majority of open ocean grid cells the effect of re-referencing is smaller than 0.1 K. The re-referencing error is also small (<0.1 K) over the open ocean, but can be substantial over sea ice (Fig. 1 column 2).

For specific humidity the role of sea ice is less pronounced. The effect of re-referencing is largest over the low latitude ocean (between 40°S and 40°N) due to high absolute humidity values (Fig. 2 column 1). The re-referencing error is largest in the subtropics and again over sea ice (Fig. 2 column 2). Particularly at high latitudes the re-referencing error can be larger than the effect from the re-referencing itself (albeit for small absolute values).

Since we are actually interested in the difference of two temperatures for our dataset, we analyze next the effect of taking the 2m temperature (and specific humidity) anomaly without any re-referencing compared to taking the 10m anomalies. This is shown in Figs. 1 and 2, 3rd column. Consistent with our analysis above, for both temperature and specific humidity, the difference is generally small (<0.1 K and <0.05 g/kg, respectively) over the open ocean. For specific humidity, we find differences that are of the same order of magnitude as the re-referencing error over sea ice. For temperature the difference is up to 4 K over the central Arctic Ocean. This is most probably caused by a more stable boundary layer and thus a larger effect of re-referencing under LGM compared to PI conditions. Also the re-referencing error could be smaller if the LGM boundary layer over sea ice is more stable (i.e., less alteration to an unstable state justifies usage of monthly means better), but this is difficult to quantify with available data.

We come to the conclusion that it might be better to omit the re-referencing step. The re-referencing of monthly mean values comes with a significant error that will be different between PI and LGM states for very stable boundary layers (over sea ice). The omission of re-referencing has no impact over open ocean regions. For Arctic temperatures, there seems to be a systematic difference between 2m and 10m temperature anomalies of up to 4 K. We note that this difference will not significantly influence simulation results, since at these latitudes the ocean is covered by thick sea ice in the LGM anyway. Also, we note that the large PMIP3 model spread at high latitudes (Fig 1 and 2, last column) might serve to justify this decision.

We therefore propose to omit the re-referencing step for a revised version of our data set. We have amended our manuscript to discuss the assumptions and limitations of our data set, and we have included a discussion of omitting the re-referencing step (new Sect. 4).

2. Actual interest of mean values

A procedure to force ocean models at the LGM is undoubtedly very useful. But I am very skeptical about using an average of different CMIP model outputs. The inter model spread is very large, as mentioned in the paper. Atmospheric circulation pattern differs, and the internal coherence of a

mean dataset is doubtful. It seems important that a user may evaluate the uncertainty coming from the forcing. And I can't imagine a way to build a variety of forcings from this dataset. The model spread is not relevant for this, as large values may locally and temporally come from different models. Perturbing the dataset with a fraction of the spread will generate incoherent patterns. From the above rationale, the use of an inter model average is far from being an obvious protocol of LGM experiments. The dataset should probably include:

- Anomaly of individual models. For model that has a small ensemble, mean of ensemble could be provided if the intra ensemble spread is small (but how to define "small" ?).
- Absolute values for individual model, as applying anomalies to a given dataset is inconsistent. One may want to try the absolute values as forcing data.

Author response to comment 2
We provide individual model anomalies for each variable in a new figure A1 to visualize the difference between the model anomalies for each variable in the dataset as to inform the reader with more detail than just the model spread in the dataset. As described in our general response, we would also be able to provide such individual model anomalies and/or absolute fields in a revised version of our data set. As outlined however (see above, point 1), we believe that the individual model fields would be of little practical value. This approach is limited by computational resources and trade-offs between integration length, ensemble size, and the need of tuning for each individual forcing. We believe that the use of an ensemble mean anomaly forcing with typically long paleo-simulation integration times is a valid and useful application. Regarding the missing "internal coherence of a mean dataset", we refer to our general response above (point 1).

3. Water budget closure

The dataset provides precipitation anomalies. Evaporation will be computed from the CORE formula. To close the water budget, ocean modeller are missing the input from river and land ice melting. River input seems to be available for only 4 of the 5 model used in this study. Anyway, if the individual model data are given in the dataset (as suggested above), friver could be made available for some of them. This assume that a general interpolation procedure can be designed for all models, which maybe difficult because the variety of solution for river runoff in each model.

Author response to comment 3
We propose to add information on river runoff anomalies to a revised version of our manuscript. Since a gridded anomaly field would not be practical due to differences in the land-sea mask and river mouth locations between the models (as noted by the reviewer), we propose to calculate river runoff-anomalies (absolute values and fractional change) on basin scale (North/South Atlantic, North/South Pacific, and Indian Ocean). These anomalies could then be used by modelling groups to scale their pre-industrial river runoff.
We note that forced ocean models will inevitably have an imbalance between freshwater sources and sinks (there is no regulating feedback in such a model setup). For longer integrations (several hundred years or longer) such models usually implement a balancing of freshwater fluxes to avoid long term salinity drift. Such balancing can be accomplished e.g. by increasing/decreasing the prescribed precipitation fluxes based on diagnosed imbalances.

Last, we noticed that the wind anomaly was not updated in our revision of Fig. 5, which we corrected now.

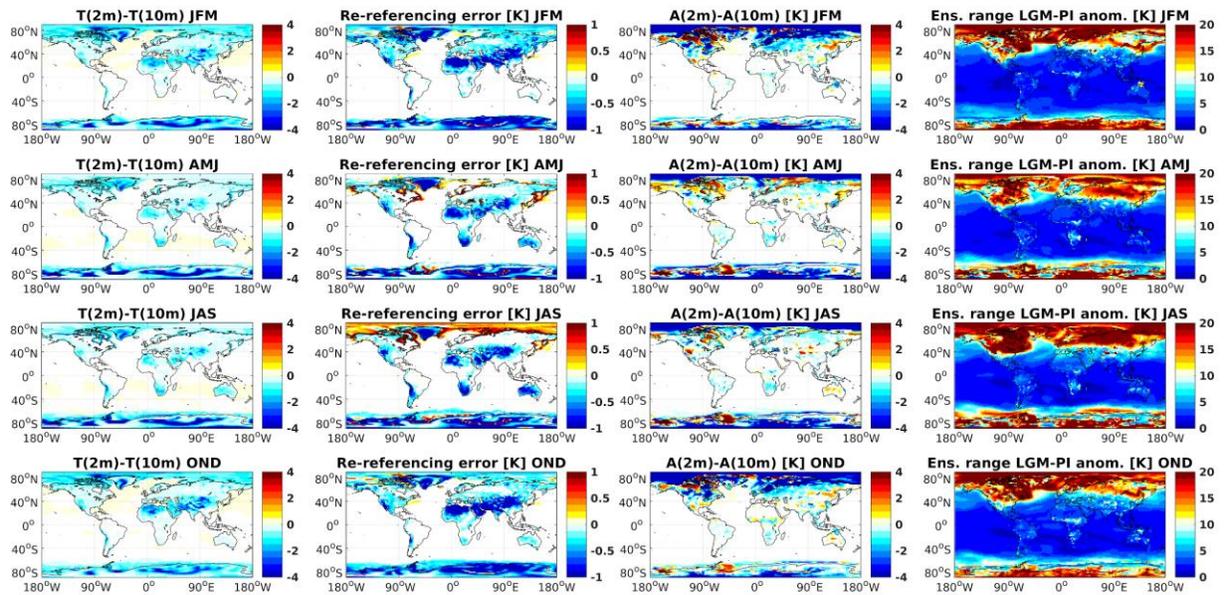Yours sincerely,
Anne Morée and Jörg Schwinger

Figure 1: (1st column) climatological difference between T(2m) and (after re-referencing using 6-hourly input data) T(10m); (2nd column) re-referencing error defined as the difference between T'(10m), which is calculated using monthly mean input data (i.e. as done in our dataset), and T(10m); (3rd column) difference between the LGM-PI anomaly based on T(2m) and T(10m), where 10m temperatures for LGM and PI have been re-referenced using climatological monthly mean PMIP3 output; (4th column) uncertainty estimate (ensemble range of the 5 PMIP3 models as provided in our dataset) of our LGM-PI temperature anomaly field. Masked grid cells in the first three columns values smaller than 0.1 K. Climatologies of the 1st and 2nd columns are calculated over 30 years of NCEP-R1 data (1980-2009).
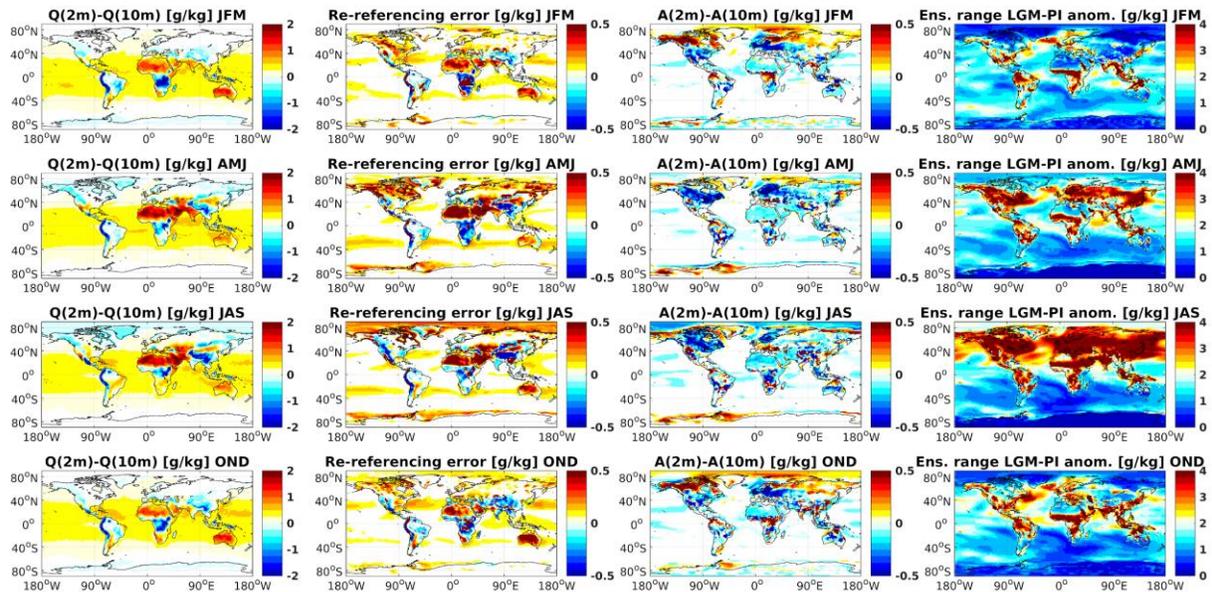
Figure 2: as Fig. 1 but for specific humidity. Masked grid cells in the first three columns indicate values smaller than 0.05 g/kg.

References

Chowdhury, P., and Behera, M. R.: Evaluation of CMIP5 and CORDEX derived wave climate in Indian Ocean, Clim Dyn, 52, 4463-4482, 10.1007/s00382-018-4391-0, 2019.

Khatiwala, S., Schmittner, A., and Muglia, J.: Air-sea disequilibrium enhances ocean carbon storage during glacial periods, Science Advances, 5, eaaw4981, 10.1126/sciadv.aaw4981, 2019.

Mitchell, D., AchutaRao, K., Allen, M., Bethke, I., Beyerle, U., Ciavarella, A., Forster, P. M., Fuglestvedt, J., Gillett, N., Haustein, K., Ingram, W., Iversen, T., Kharin, V., Klingaman, N., Massey, N., Fischer, E., Schleussner, C. F., Scinocca, J., Seland, Ø., Shiogama, H., Shuckburgh, E., Sparrow, S., Stone, D., Uhe, P., Wallom, D., Wehner, M., and Zaaboul, R.: Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design, Geosci. Model Dev., 10, 571-583, 10.5194/gmd-10-571-2017, 2017.

Muglia, J., and Schmittner, A.: Glacial Atlantic overturning increased by wind stress in climate models, Geophysical Research Letters, 42, 9862-9868, citeulike-article-id:13822889, doi: 10.1002/2015gl064583, 2015.

Muglia, J., Skinner, L. C., and Schmittner, A.: Weak overturning circulation and high Southern Ocean nutrient utilization maximized glacial ocean carbon, Earth and Planetary Science Letters, 496, 47-56, https://doi.org/10.1016/j.epsl.2018.05.038, 2018.

A Last Glacial Maximum forcing dataset for ocean modelling

Anne L. Morée[1], Jörg Schwinger[2]

[1]Geophysical Institute, University of Bergen and Bjerknes Centre for Climate Research, Bergen, 5007, Norway

[2]NORCE Climate, Bjerknes Centre for Climate Research, 5007 Bergen, Norway

5   *Correspondence to*: Anne L. Morée (anne.moree@uib.no)

Abstract. Model simulations of the Last Glacial Maximum (LGM, ~21 000 years before present) can aid the interpretation of proxy records, help to gain an improved mechanistic understanding of the LGM climate system and are valuable for the evaluation of model performance in a different climate state. Ocean-ice only model configurations forced by prescribed atmospheric data (referred to as "forced ocean models") drastically reduce the

10   computational cost of paleoclimate modelling as compared to fully coupled model frameworks. While feedbacks between the atmosphere and ocean-sea-ice compartments of the Earth system are not present in such model configurations, many scientific questions can be addressed with models of this type. The data presented here are derived from fully coupled paleoclimate simulations of the Palaeoclimate Modelling Intercomparison Project phase 3 (PMIP3). The data are publicly accessible at the NIRD Research Data Archive at

15   https://doi.org/10.11582/2019.00019 (Morée and Schwinger, 2019a). They consist of 2-D anomaly forcing fields suitable for use in ocean models that employ a bulk forcing approach and are optimized for use with CORE forcing fields. The data include specific humidity, downwelling longwave and shortwave radiation, precipitation, wind (v and u components), temperature and sea surface salinity (SSS). All fields are provided as climatological mean anomalies between LGM and pre-industrial times. These anomaly data can therefore be added to any pre-industrial

20   ocean forcing data set in order to obtain forcing fields representative of LGM conditions as simulated by PMIP3 models. These forcing data provide a means to simulate the LGM in a computationally efficient way, while still taking advantage of the complexity of fully coupled model set-ups. Furthermore, the dataset can be easily updated to reflect results from upcoming and future paleo model intercomparison activities.

1 Introduction

25   The LGM (~21 kya) is of interest to the climate research community because of the relative abundance of proxy data, and because it is the most recent profoundly different climatic state of our planet. For these reasons, the LGM is extensively studied in modelling frameworks (Menviel et al., 2017; Brady et al., 2012; Otto-Bliesner et al., 2007). Model simulations of the past ocean ~~can not~~cannot only provide a method to gain a mechanistic understanding of marine proxy records, they can also inform us about model performance in a different climatic

30   state of the Earth system (Braconnot et al., 2012). Typical state-of-the-art tools to simulate the (past) Earth system are climate or Earth system models as, for example, used in the Coupled Model Intercomparison Project phase 5 (CMIP5; Taylor et al~~. (~~., 2011~~)~~). Besides simulating our present climate, these CMIP5 models are also used to simulate past climate states (such as the LGM) in the Palaeoclimate Modelling Intercomparison Project 3 (PMIP3). However, the computational costs and run-time of such fully coupled model frameworks are a major obstacle for

1

their application to palaeoclimate modelling. Palaeoclimate modelling optimally requires long (thousands to ten thousands of years) simulations in order to provide the necessary time for relevant processes to emerge (e.g. $CaCO_3$ compensation) (Braconnot et al., 2007). Complex fully coupled models can typically not be run into full equilibrium (which requires hundreds to thousands of years of integration) due to computational costs (Eyring et al., 2016). Therefore, the PMIP3 models exhibit model drift (especially in the deep ocean, e.g. Marzocchi and Jansen, (2017)). The 3rd phase of the PMIP project (~~PMIP3;~~ Braconnot et al., (2012)) limits global mean sea-surface temperature drift to under 0.05 K per century and requires the Atlantic Meridional Overturning Circulation to be stable (Kageyama et al., 2018).

~~The use of PMIP output as ocean forcing is an accepted practice in ocean modelling (e.g., Muglia and Schmittner (2015)).~~ We refer to a "forced ocean model" as a model of the ocean-sea-ice-atmosphere system in which the atmosphere is represented by prescribed 2-D forcing fields. It can be used whenever ocean-atmosphere feedbacks are of minor importance and has the advantage of reducing the computational costs – making longer or more model runs feasible. The use of PMIP output as~~in~~ ~~ocean~~ forced ocean modelling~~ing~~ is a~~n accepted~~ practice ~~in ocean modelling~~ (e.g., Muglia and Schmittner, (2015); Khatiwala et al., 2019). We present 2-D (surface) anomaly fields of CMIP5/PMIP3 experiments 'lgm' (representing the Last Glacial Maximum state of the Earth system) minus 'piControl' (representing the pre-industrial state) calculated from monthly climatological PMIP3 output. The PMIP3 output is the result of global boundary conditions and forcings (such as insolation and ice sheet cover) applied in the fully coupled PMIP3 models (Braconnot et al., 2012). Our dataset (Morée and Schwinger, 2019a) is a unique compilation of existing data, processed and reformatted such that it can be readily applied in a forced ocean model framework that uses a bulk forcing approach similar to Large and Yeager (2004). Since this approach has been popularized through coordinated model intercomparison activities (Griffies et al., 2009), a majority of forced ocean models today uses this approach. The 2-D anomaly fields presented here can be added to the pre-industrial forcing of a forced ocean model in order to obtain an atmospheric forcing representative of the LGM. The data are climatological mean anomalies, and as such suitable for equilibrium LGM 'time-slice' modelling of the ocean. The description of the procedure followed to make this dataset (Sect. 3) should support any extension of the dataset with additional (PMIP-derived) variables if needed. The PMIP4 guidelines (Kageyama et al., 2017) can support users in designing a specific model set-up, for example regarding the land-sea mask, trace gas concentrations, river runoff or other conditions and forcing one would want to apply to a model. In Sect. 2, a general description of the dataset and data sources is provided alongside ~~with~~ an overview of the variables (Table 1).

2 General description of the dataset

The data presented in this article are 2-D anomaly fields of the LGM versus pre-industrial state ~~(experiment 'lgm' minus experiment 'piControl')~~ based on ~~the~~ PMIP3 (Braconnot et al., 2012). These anomaly fields can be used as atmospheric LGM forcing fields for ocean-only model set-ups when added to pre-industrial forcing fields (as done by e.g. Muglia and Schmittner, 2015; Khatiwala et al., 2019), and are optimized for use in combination with Coordinated Ocean-ice Reference Experiments (CORE) forcing fields (Griffies et al., 2009). The use of an anomaly forcing implies the assumption that no changes in temporal or spatial variability occurred between the lgm and piControl states beyond changes in the mean. The basis of the anomaly~~is~~ data is monthly climatological

PMIP3 output. Any variables presented on sub-monthly time resolution are therefore time-interpolated. ~~Since this is a limitation of the available data, we have to assume that any sub-monthly variability (e.g. the diurnal cycle) is preserved from the preindustrial climate state to the LGM state.~~ The anomalies are calculated as the mean of the difference between monthly climatologies of the 'lgm' and 'piControl' PMIP3 model runs. In cases where

5   modelling groups provided more than one ensemble member, we included only the first member in our calculations. ~~Even though PMIP3 simulations have limitations and a large inter-model spread, PMIP3 is the state of the art for modelling of past climates at present (Braconnot et al., 2012; Braconnot and Kageyama, 2015). Furthermore, no global proxy-based reconstructions of the variables presented here are available to provide a proxy-based LGM forcing dataset. Using mean coupled model output as forcing is thus considered the best~~

10   ~~available option for use in forced ocean models.~~ The data is the mean anomaly of five PMIP3 models (CNRM-CM5, IPSL-CM5A-LR, GISS-E2-R, MIROC-ESM and MRI-CGCM3: Table 2), as only these models provide output for all variables ~~at or close to the desired atmospheric height~~. A discussion on the limitations of our dataset is provided in Sect. 4.

The variables are i) near-surface specific humidity ~~at 10 meters~~, ii) downwelling longwave radiation, iii)

15   downwelling shortwave radiation, iv) precipitation, v) wind (v and u components), vi) near surface temperature ~~at 10 meters~~, and vii) sea surface salinity (SSS) (Table 1). The SSS anomaly field can be used to ~~apply~~ adjust SSS restoring in LGM simulations.

All variables (Sect. 3.1-7) of the monthly climatological PMIP3 output have been regridded (Table 3, #1), averaged (Table 3, #2), and differenced (Table 3, #3) to calculate the anomaly fields. Additional procedures for each variable

20   are provided in the respective part of Sect. 3, together with a figure of each variable's yearly mean anomaly and model spread. Alongside the lgm-piControl anomaly for each variable, the model spread across all five models is made available. ~~For further comparison, we show t~~The individual model anomalies for each of the variables are presented in ~~(~~Fig. A1~~)~~. The~~is~~ inter-model disagreement is described for each variable in Sect. 3, and could for example be used to guide adjustments of the amplitude of the forcing anomaly for model tuning purposes.

25   Additionally, proxy-based reconstructions are available for some of the variables (e.g., temperature), which can constrain potential adjustments to the forcing anomaly fields. We leave it to the individual modelling groups to make such adjustments to their forcing fields for their specific application.

All operations were performed with NetCDF toolkits CDO version 1.9.3 (Schulzweida, 2019) or NCO version 4.6.9. The main functions used are documented in Table 3, and referred to in the text at the first occurrence. The

30   atmospheric anomaly data are on a Gaussian grid, with 192×94 (lon×lat) grid-points. The SSS fields is on a regular 360×180 (lon×lat) grid. Regridding any of the files to a different model grid should be straightforward (e.g., Table 3, #1), as it was ensured that all files contain the information needed for re-gridding. The variables, grid and time resolution are chosen to be compatible with the CORE forcing fields (Large and Yeager, 2004), which have been extensively used in the ocean modelling community (e.g. Griffies et al.~~,~~. ~~(~~2009~~;~~; Schwinger et al.~~,~~. ~~(~~2016~~)~~). We

35   anticipate that the variables selected here should be useful in different model set-ups as well. We intend to provide a data set that is flexible with respect to the use of different land-ocean masks in different models. Therefore, we account for changes in sea-level (i.e. a larger land area in the LGM), which can affect variables in coastline areas, by applying the following masking procedure: i) masking the multi-model mean anomaly with the maximum lgm land mask across all models, then ii) extrapolating the variable over land using a distance-weighted average (Table

40   3, #4), and iii) finally masking the data with a present-day land mask (based on the World Ocean Atlas 2013 1°

resolution land mask), but with the ocean extended in a 1.5 degrees radius over land. This choice ensures that the anomaly forcing data can be used with any pre-industrial land-sea mask. Through following this procedure, the grid points affected by land-sea mask changes are thus filled with the extrapolated model mean anomaly from the LGM coastaln ocean. In the case of NorESM-OC (Schwinger et al., 2016), the atmospheric anomaly fields were

5   added to its CORE normal-year forcing fields (Large and Yeager, 2004) to obtain an LGM normal-year forcing, under the assumption of an unchanged spatial and temporal variability for the respective variable. Note that the addition of the anomaly fields to the user's own model forcing could lead to physically unrealistic/not-meaningful results for some variables (such as negative precipitation or radiation). This must be corrected for by capping off sub-zero values (Table 3, #5) after addition of the anomaly.

10   3 The variables

3.1 Specific humidity anomaly

~~The monthly climatology of near-surface specific humidity is provided at 2 meters height in PMIP3. The bulk forcing method of Large and Yeager (2004) requires specific humidity (and temperature) at the same height as the wind forcing (10 meters). Therefore, specific humidity was re-referenced to 10 meters height for each of the four~~

15   ~~models following the procedure detailed in Large and Yeager (2004). The re-referencing required the use of wind (u and v components), sea level pressure (CMIP variable 'psl') and skin temperature ('ts', representing sea surface temperature over the open ocean), which were taken from the respective 'piControl' and 'lgm' CMIP5/PMIP3 output for each model.~~ The mean anomaly of near-surface specific humidity over the ~~four~~ five models was time interpolated (Table 3, #6) to a 6-hour time resolution from the monthly climatological PMIP3 output. The annual

20   mean lgm-piControl anomaly field (Fig. 1) shows a global decrease in specific humidity, as expected from decreased air temperatures (Sect. 3.6). The anomaly is most pronounced around the equator, where we see a decrease of 2-3 kg kg^{-1}, while the anomaly is near-zero towards both poles. The model spread of the anomaly shows a disagreement between the PMIP3 models generally in the order of 1-2 kg kg^{-1}, without any strong spatial pattern (Fig. 1).

25   3.2 Downwelling longwave radiation anomaly

The anomaly for surface downwelling longwave radiation is time-interpolated (Table 3, #6) to a daily time resolution. The annual mean anomaly field (Fig. 2) shows globally decreased downwelling longwave radiation in the 'lgm' experiment as compared to the 'piControl' experiment, in the order of 10-30 W m^{-2} over most of the ocean due to a generally cooler atmosphere (Sect. 3.6). The largest anomalies lie close to the northern ice sheets,

30   with up to -90 W m^{-2} lower radiation in the 'lgm' experiment than in the 'piControl' experiment. Ice is likely also the main contributor to the high (60-90 W m^{-2}) inter-model spread in North Atlantic and Southern Oceans. The remainder of the ocean exhibits a better agreement, with inter-model spreads generally below 20 W m^{-2} (Fig. 2).

3.3 Downwelling shortwave radiation anomaly

The surface downwelling shortwave radiation anomaly field is time-interpolated (Table 3, #6) to daily fields as

35   done for downwelling longwave radiation. The annual mean anomaly is especially pronounced around the

Laurentide and Scandinavian ice sheets, where strong positive anomalies of over ~30 W m^{-2} exist (Fig. 3). Globally, the annual mean downwelling shortwave radiation anomaly generally falls in a range of -15 to +15 W m^{-2} over the ocean. The anomaly field shows negative anomalies as well positive ones in an alternating spatial pattern approximately symmetrically around the equator in the Pacific basin. The inter-model spread is largest in the North Atlantic region and along the equator (Fig. 3). Due to the large model disagreement of up to 50 W m^{-2} for this variable (Fig. 3), the inter-model spread and mean anomaly are of similar magnitude although a consistent pattern is present in the anomaly field.

3.4 Precipitation anomaly

The anomaly presented here is the lgm-piControl precipitation anomaly at the air-sea interface and includes both the liquid and solid phases from all types of clouds (both large-scale and convective), and excludes evaporation. The units were converted to mm day^{-1} to comply with the CORE forcing format (causing a deviation from the CF-1.6 convention). The resulting annual mean anomaly generally falls in the range of -2 to 2 mm day^{-1}, and is most pronounced along the equator (Fig. 4). The models show a mean increase in precipitation directly south of the equator in the Pacific basin, as well as in the Pacific subtropics off the western North-American coast. The North Atlantic also receives a mean positive precipitation anomaly, offsetting part of the positive salinity anomaly there, which is potentially relevant for the simulation of deep--water formation in this region (Sect. 3.7). Negative mean precipitation anomalies are most pronounced directly north of the equator and north of ~40° N in the Pacific basin as well as in the Atlantic Arctic. The inter-model spread is up to ~5 mm day^{-1} around the equator, likely due to the model disagreement about the sign and location of changes in the inter-tropical convergence zone (Fig. 4). Related to precipitation fluxes, river runoff fluxes also changed between the lgm and piControl model experiments. As land-sea masks and river routing and flux calculations are very model specific we cannot provide a gridded river run-off anomaly. Instead we provide mean absolute and relative large-scale river runoff changes integrated over ocean basins (North/South Atlantic, North/South Pacific, Indian Ocean). These anomalies can be used by modelling groups to scale pre-industrial river runoff.

and influenced by differences in land-sea masking, we recommend individual modelling groups to close their water budget based on their model specifics. Additionally, we propose groups to consult with the PMIP guidelines when doing so (Kageyama et al., 2017). expect modelling groups to find a suitable solution for their setup themselves, and recommend consulting the PMIP guidelines when doing so (Kageyama et al., 2017).

3.5 Wind anomalies, u and v components

Both for the u and v component of the wind speed, the lgm-piControl anomaly is time-interpolated to 6-hourly fields. The annual mean meridional wind velocity (v, southerly winds) anomaly shows a pronounced increase (~3-5 m s^{-1}) in southerly winds around the NW edge of the Laurentide ice sheet as well as over the NW edge of the Scandinavian ice sheet (Fig. 5). Alongside that, a pronounced decrease (~ 3-5 m s^{-1}) in southerly winds is simulated along the eastern North American coast and the Canadian archipelago. The open ocean anomalies are generally small (at most ±1 m s^{-1}). The inter-model spread has no pronounced pattern but is sizable, with ~ 1-5 m s^{-1} disagreement between the PMIP3 models. The mean zonal wind velocity (u, westerly winds) anomaly shows alternating negative and positive anomaly bands with an approximate ±2 m s^{-1} range (Fig. 6). This pattern is

5

stronger in the Northern Hemisphere north of ~45° N. The inter-model spread (~1-3 m s⁻¹) has little structure except for the ~4-5 m s⁻¹ disagreement in the Southern Ocean south of ~40° S, and the ~3-5 m s⁻¹ disagreement in the North Atlantic (Fig. 6).

3.6 Temperature anomaly

5   The near-surface atmospheric temperature ~~at 2 m height from PMIP3 is re-referenced to 10 meters (as done for specific humidity, Sect. 3.1), and~~ is time-interpolated to calculate the 6-hourly mean anomaly for temperature. The annual mean anomaly is most pronounced in the North Atlantic, where open ocean anomalies exceed 10 K. Elsewhere, the annual mean temperature ~~generally~~ anomaly is ~~around~~ ~ 2.5 K. There is a clear pattern in the model spread: The models show a large spread (>10 K) north of ~45° N, as well as south of ~40° S (5-10 K), likely due

10 to the disagreement about ice cover. At lower latitudes, ~~and over the ocean~~ the model spread is generally smaller (0-3 K) (Fig. 7).

3.7 Sea surface salinity anomaly

Global mean salinity is initialized in PMIP3 models with a 1 psu higher salinity to account for the concentrating effect of the decrease in sea level (Kageyama et al., 2017). Sea surface salinity however, shows a more variable

15 annual mean lgm-piControl change due to changes in the global hydrological cycle (Fig. 8). The sea surface salinity anomaly is presented on a regular 1x1 grid for ease of use. The resulting annual mean SSS anomaly (Fig. 8) shows an increase in sea surface salinity (~1 psu) over the Southern Ocean south of ~55° S, as well as in the Arctic (>3 psu) and the Northern Indian Ocean (~1 psu). A ~2 psu anomaly is simulated in the Canadian Archipelago, the Labrador Sea and across the North Atlantic between what is now Canada and Europe (Fig. 8). Freshening is

20 simulated close to some continents, and is especially pronounced around Scandinavia (about -3 psu). Simulated ocean circulation can be very sensitive to fresh water forcing and thus SSS, especially in the North Atlantic (e.g. Rahmstorf, ~~(~~1996~~;)~~, Spence et al., ~~(~~2008~~)~~). Application of SSS restoring using the SSS anomaly field should therefore be done with caution and attention to its effects on the meridional overturning circulation. Tuning of the salinity anomaly in important deep-water formation regions of up to about ±1 psu, such as done by for example

25 Winguth et al. (1999), may be required to obtain a satisfactory circulation field in reasonable agreement with proxy data. Such adjustments fall well within the PMIP3 model spread (Fig. 8), and show the current limitations of fully coupled PMIP3 models to simulate the LGM hydrological cycle consistent with proxy records of ocean circulation.

4 Limitations of the dataset

The anomaly fields presented here are a model-based 'best-estimate' of the LGM anomaly relative to the pre-

30 industrial state. There are some important limitations to these data related to the temporal resolution, the use of model means, and the fact that we rely on modelling results only.

Proxy data with global coverage are unavailable for most of the variables needed to force stand-alone ocean models. We do not attempt to constrain the anomaly fields using the spatially limited information from available proxy data. Consequently, where PMIP3 models are in disagreement with proxy data, our dataset will be so, too.

35 The limitations (or uncertainty) of the PMIP3 simulations can be seen through the large inter-model spread, which

is provided with the anomaly data. This does not preclude the possibility that PMIP3 models collectively (i.e. such that the model spread is small) disagree with available proxy data. Nevertheless, PMIP3 is the state of the art for modelling of past climates at present (Braconnot et al., 2012; Braconnot and Kageyama, 2015).

By adding multi model mean anomalies to forcing fields, dynamical inconsistencies (e.g. between wind and temperature fields) will be created. This means that the resulting forcing fields do not strictly obey the equations of state/motion. A forcing data set would typically be dynamically consistent if the forcing would be the outcome of an atmospheric model or an advanced reanalysis. The advantage of using model mean fields is that large anomalies of individual models will be smoothed out where models disagree. We believe that currently a main challenge for paleo modelling activities is to achieve long enough integration times. Therefore, using a single forcing (as opposed to using multiple forcings from individual models) seems to be preferable. Regarding the dynamical inconsistencies, it is important to note that the CORE forcing itself is a mixture of reanalysis and observational data products and as such not dynamically consistent.

PMIP3 model output is publicly available only as monthly mean fields, which also results in some limitations for the anomaly forcing data set. First, although we interpolate the monthly mean anomaly fields to higher (e.g. 6-hourly) temporal resolution, we implicitly assume that any sub-monthly variability (e.g. the diurnal cycle) is preserved from the preindustrial climate state to the LGM state. Second, we are not able to accurately re-reference near-surface temperature and humidity to a different reference height. The CORE bulk forcing method of Large and Yeager (2004) actually requires near-surface specific humidity and temperature at the same height as the wind forcing (at 10 meters). Humidity and temperature are however provided at 2 meters height in PMIP3 (as in most atmospheric data products). A procedure to re-reference humidity and temperature from 2 to 10 meters (e.g. Large and Yeager 2004) requires input data in higher (sub-daily) time resolution in order to resolve different boundary layer stability regimes. However, for an anomaly forcing, the re-referencing only has an effect if it leads to different temperature/humidity increments under the PI and the LGM state. For the open ocean this is barely the case and taking a climatological anomaly of 2-meter-quantities and apply it at 10 meters height is unproblematic. Over sea ice, however, there could be a larger effect of the re-referencing (due to a significantly different atmospheric stability in the LGM state), especially regarding the temperature. Our analysis indicates that this is probably the case over the central Arctic Ocean (not shown). For all other regions, we estimate that the error made in applying the re-referencing approach on monthly climatological data resolved data, does not justify its application. In general, the error made by omitting the re-referencing is much smaller than the uncertainties of the anomalies (i.e. the model spread), particularly at high latitudes.

Regarding the robustness of the dataset, we observe that the inclusion of additional model data only leads to minor changes in the anomalies. An example of this is given by comparing version 1 (Morée and Schwinger, 2019b) and the current version 2 (Morée and Schwinger, 2019a) of this dataset, as the latter also includes the GISS-E2-R model for the calculation of the anomalies. Indeed, individual model anomalies (Fig. A1) show broad agreement, although the magnitude of the anomaly is less agreed on (as discussed in more detail for the individual variables in Sect. 3).

Despite the limitations described here, we believe that using the mean PMIP3 anomaly of coupled models as forcing is currently the best available option for use in stand-alone ocean models. For this purpose, our dataset provides lgm-piControl anomalies in standardized format for the most common variables used in ocean forcing.

45 Data availability

The data are publicly accessible at the NIRD Research Data Archive at https://doi.org/10.11582/2019.00019 (Morée and Schwinger, 2019a). The .md5 files contain an md5 checksum, which can be used to check whether changes have been made to the respective neNetCDF files.

65 Summary and Conclusions

The output of the fully coupled PMIP3 simulations of CNRM-CM5, IPSL-CM5A-LR, MIROC-ESM, and MRI-CGCM3 and GISS-E2-R is converted to anomaly datasets intended for use in forced ocean modelling of the LGM. All anomalies are calculated as the difference between the 'lgm' and 'piControl' PMIP3 experiments. In addition, all data are formatted in a way that further conversions (of for example units or the grid) can be applied in a straightforward way. The variables are provided in NetCDF format in separate files, and distributed by the NIRD Research Data Archive (Morée and Schwinger, 2019a). A climatological LGM forcing data set can be created for any forced ocean model by addition of the presented 2-D anomaly fields to the model's pre-industrial forcing. This approach enables the scientific community to simulate the LGM ocean state in a forced ocean model set-up. We expect that if additional forcing is needed for a specific model, the same approach as described above can be followed. This process is simplified by providing all main CDO and NCO commands used in creating the dataset (Table 3). All data represent a climatological year, i.e. one annual cycle per variable. The application of the data is thus suitable for 'time-slice' equilibrium simulations of the LGM, and optimised for use with the CORE forcing format (Large and Yeager, 2004).

The uncertainty of our anomaly forcing (approximated by the model spread of the PMIP3 models) is generally of similar magnitude as the multi-model annual mean. The attribution of the model spread to specific processes is beyond the scope of this article, but our results show that there is considerable uncertainty involved in the magnitude of the anomaly for all variables presented here. Nevertheless, all mean anomalies show a distinct spatial pattern that we expect to be indicative of the LGM-PI changes. Finally, there is no other way to reconstruct most of these variables than model simulations with state-of-the-art Earth system models such as those applied in the PMIP3 experiments. For modelling purposes, the inter-model disagreement of PMIP3 provides the user with leeway to adjust the amplitude of the forcing (guided by the size of the model spread, which is therefore provided alongside the variables in the dataset). Such adjustments can improve model-proxy data agreements, such as described for salinity in Sect. 3.7.
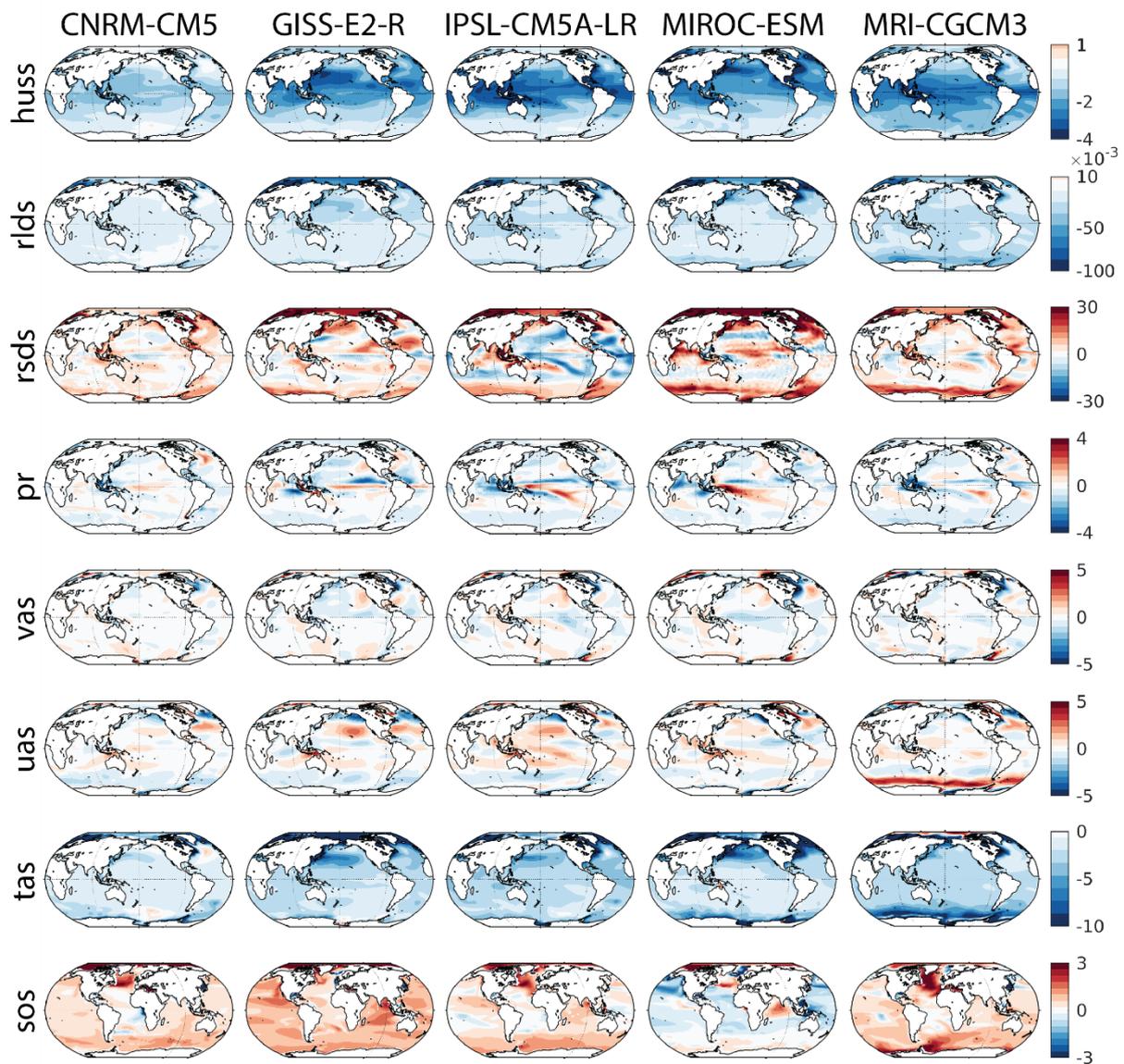
Appendix A.

Figure A1: Annual mean individual model anomalies for each of the variables (see Table 1) and models in the dataset. Units as in the remainder of this manuscript.

Author contributions. AM prepared, visualized and analysed the data and wrote the original draft of the manuscript. AM and JS together conceptualized the method and revised the manuscript. JS provided supervision throughout the study.

Competing interests. The authors declare that they have no conflict of interest.

References

Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J. Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M. F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S. L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the

10   Mid-Holocene and Last Glacial Maximum; Part 1: experiments and large-scale features, Clim. Past, 3, 261-277, 10.5194/cp-3-261-2007, 2007.
Braconnot, P., Harrison, S., Kageyama, M., Bartlein, P., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nature Climate Change, 2, 417-424, doi: 10.1038/nclimate1456, 2012.

15   Braconnot, P., and Kageyama, M.: Shortwave forcing and feedbacks in Last Glacial Maximum and Mid-Holocene PMIP3 simulations, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, 20140424, 10.1098/rsta.2014.0424, 2015.
Brady, E. C., Otto-Bliesner, B. L., Kay, J. E., and Rosenbloom, N.: Sensitivity to Glacial Forcing in the CCSM4, Journal of Climate, 26, 1901-1925, 10.1175/JCLI-D-11-00416.1, 2012.

20   Caubel, A., Denvil, S., Foujols, M. A., Marti, O., Dufresne, J.-L., Bopp, L., Cadule, P., Ethé, C., Idelkadi, A., Mancip, M., Masson, S., Mignot, J., Ionela, M., Balkanski, Y., Bekki, S., Bony, S., Braconnot, P., Brockman, P., Codron, F., Cozic, A., Cugnet, D., Fairhead, L., Fichefet, T., Flavoni, S., Guez, L., Guilyardi, E., Hourdin, F., Ghattas, J., Kageyama, M., Khodri, M., Labetoulle, S., Lefebvre, M.-P., Levy, C., Li, L., Lott, F., Madec, G., Marchand, M., Meurdesoif, Y., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Viovy, N., and Vuichard, N.:

25   IPSL-CM5A-LR model output prepared for CMIP5 piControl experiment, served by ESGF,WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.IPILpc, 2016.
Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J. P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P.,

30   Grandpeix, J. Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M. P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Clim Dyn, 40,

35   2123-2165, 10.1007/s00382-012-1636-1, 2013.
Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937-1958, 10.5194/gmd-9-1937-2016, 2016.
Griffies, S. M., Biastoch, A., Böning, C., Bryan, F., Danabasoglu, G., Chassignet, E. P., England, M. H., Gerdes,

40   R., Haak, H., Hallberg, R. W., Hazeleger, W., Jungclaus, J., Large, W. G., Madec, G., Pirani, A., Samuels, B. L., Scheinert, M., Gupta, A. S., Severijns, C. A., Simmons, H. L., Treguier, A. M., Winton, M., Yeager, S., and Yin, J.: Coordinated Ocean-ice Reference Experiments (COREs), Ocean Modelling, 26, 1-46, https://doi.org/10.1016/j.ocemod.2008.08.007, 2009.
JAMSTEC, AORI, and NIES: MIROC-ESM model output prepared for CMIP5 piControl, served by ESGF,

45   doi:10.1594/WDCC/CMIP5.MIMEpc 2015a.
JAMSTEC, AORI, and NIES: MIROC-ESM model output prepared for CMIP5 lgm, served by ESGF, doi:10.1594/WDCC/CMIP5.MIMElg, 2015b.
Kageyama, M., Denvil, S., Foujols, M. A., Caubel, A., Marti, O., Dufresne, J.-L., Bopp, L., Cadule, P., Ethé, C., Idelkadi, A., Mancip, M., Masson, S., Mignot, J., Ionela, M., Balkanski, Y., Bekki, S., Bony, S., Braconnot, P.,

50   Brockman, P., Codron, F., Cozic, A., Cugnet, D., Fairhead, L., Fichefet, T., Flavoni, S., Guez, L., Guilyardi, E., Hourdin, F., Ghattas, J., Khodri, M., Labetoulle, S., Lefebvre, M.-P., Levy, C., Li, L., Lott, F., Madec, G., Marchand, M., Meurdesoif, Y., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Viovy, N., and Vuichard, N.: IPSL-CM5A-LR model output prepared for CMIP5 lgm experiment, served by ESGF,WDCC at DKRZ doi:10.1594/WDCC/CMIP5.IPILlg, 2016.

Kageyama, M., Albani, S., Braconnot, P., Harrison, S. P., Hopcroft, P. O., Ivanovic, R. F., Lambert, F., Marti, O., Peltier, W. R., Peterschmitt, J. Y., Roche, D. M., Tarasov, L., Zhang, X., Brady, E. C., Haywood, A. M., LeGrande, A. N., Lunt, D. J., Mahowald, N. M., Mikolajewicz, U., Nisancioglu, K. H., Otto-Bliesner, B. L., Renssen, H., Tomas, R. A., Zhang, Q., Abe-Ouchi, A., Bartlein, P. J., Cao, J., Li, Q., Lohmann, G., Ohgaito, R., Shi, X., Volodin, E., Yoshida, K., Zhang, X., and Zheng, W.: The PMIP4 contribution to CMIP6 – Part 4: Scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum experiments and PMIP4 sensitivity experiments, Geosci. Model Dev., 10, 4035-4055, 10.5194/gmd-10-4035-2017, 2017.

Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J. H., Otto-Bliesner, B. L., Peterschmitt, J. Y., Abe-Ouchi, A., Albani, S., Bartlein, P. J., Brierley, C., Crucifix, M., Dolan, A., Fernandez-Donado, L., Fischer, H., Hopcroft, P. O., Ivanovic, R. F., Lambert, F., Lunt, D. J., Mahowald, N. M., Peltier, W. R., Phipps, S. J., Roche, D. M., Schmidt, G. A., Tarasov, L., Valdes, P. J., Zhang, Q., and Zhou, T.: The PMIP4 contribution to CMIP6 – Part 1: Overview and over-arching analysis plan, Geosci. Model Dev., 11, 1033-1057, 10.5194/gmd-11-1033-2018, 2018.

Khatiwala, S., Schmittner, A., and Muglia, J.: Air-sea disequilibrium enhances ocean carbon storage during glacial periods, Science Advances, 5, eaaw4981, 10.1126/sciadv.aaw4981, 2019.

Large, W. G., and Yeager, S. G.: Diurnal to decadal global forcing for ocean and sea-ice models: the data sets and flux climatologies, Tech. Note NCAR/TN-460+STR. National Center of Atmospheric Research, Boulder, Colorado, USA, http://opensky.ucar.edu/islandora/object/technotes:434, 2004.

Marzocchi, A., and Jansen, M. F.: Connecting Antarctic sea ice to deep-ocean circulation in modern and glacial climate simulations, Geophysical Research Letters, 44, 6286-6295, 10.1002/2017GL073936, 2017.

Menviel, L., Yu, J., Joos, F., Mouchet, A., Meissner, K. J., and England, M. H.: Poorly ventilated deep ocean at the Last Glacial Maximum inferred from carbon isotopes: A data-model comparison study, Paleoceanography, 32, 2-17, 10.1002/2016pa003024, 2017.

Morée, A., Schwinger, J.: PMIP3-based Last Glacial Maximum (LGM) pre-industrial (PI) anomaly fields for addition to PI ocean model forcing, version 2, Norstore, https://doi.org/10.11582/2019.00019, 2019a.

Morée, A., Schwinger, J.: Last Glacial Maximum minus pre-industrial anomaly fields for use in forced ocean modelling, based on PMIP3PMIP3 based Last Glacial Maximum (LGM) pre-industrial (PI) anomaly fields for addition to PI ocean model forcing, version 2, Norstore, https://doi.org/10.11582/2019.0001110.11582/2019.00019, 2019b.

Muglia, J., and Schmittner, A.: Glacial Atlantic overturning increased by wind stress in climate models, Geophysical Research Letters, 42, 9862-9868, doi:10.1002/2015gl064583, 2015.

NASA-GISS: GISS-E2-R model output prepared for CMIP5 pre-industrial control, served by ESGF WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.GIGRpc, 2014a.

NASA-GISS: GISS-E2-R model output prepared for CMIP5 last glacial maximum, served by ESGF WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.GIGRlg, 2014b.

Otto-Bliesner, B. L., Hewitt, C. D., Marchitto, T. M., Brady, E., Abe-Ouchi, A., Crucifix, M., Murakami, S., and Weber, S. L.: Last Glacial Maximum ocean thermohaline circulation: PMIP2 model intercomparisons and data constraints, Geophysical Research Letters, 34, 10.1029/2007GL029475, 2007.

Rahmstorf, S.: On the freshwater forcing and transport of the Atlantic thermohaline circulation, Clim Dyn, 12, 799-811, 10.1007/s003820050144, 1996.

Schmidt, G. A., Kelley, M., Nazarenko, L., Ruedy, R., Russell, G. L., Aleinov, I., Bauer, M., Bauer, S. E., Bhat, M. K., Bleck, R., Canuto, V., Chen, Y.-H., Cheng, Y., Clune, T. L., Del Genio, A., de Fainchtein, R., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Lo, K. K., Matthews, E. E., Menon, S., Miller, R. L., Oinas, V., Oloso, A. O., Perlwitz, J. P., Puma, M. J., Putman, W. M., Rind, D., Romanou, A., Sato, M., Shindell, D. T., Sun, S., Syed, R. A., Tausnev, N., Tsigaridis, K., Unger, N., Voulgarakis, A., Yao, M.-S., and Zhang, J.: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive, Journal of Advances in Modeling Earth Systems, 6, 141-184, 10.1002/2013MS000265, 2014.

Schulzweida, U.: CDO User Guide (Version 1.9.6), Max Planck Institute for Meteorology, Bundesstraße 53, 20146 Hamburg, Germany, http://doi.org/10.5281/zenodo.2558193, 215 pp., 2019.

Schwinger, J., Goris, N., Tjiputra, J. F., Kriest, I., Bentsen, M., Bethke, I., Ilicak, M., Assmann, K. M., and Heinze, C.: Evaluation of NorESM-OC (versions 1 and 1.2), the ocean carbon-cycle stand-alone configuration of the Norwegian Earth System Model (NorESM1), Geosci. Model Dev., 9, 2589-2622, 10.5194/gmd-9-2589-2016, 2016.

Sénési, S., Richon, J., Franchistéguy, L., Tyteca, S., Moine, M.-P., Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Madec, G., Maisonnave, E., Planton, S., Saint-Martin, D., Szopa, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: CNRM-CM5 model output prepared for CMIP5 piControl, served by ESGF,WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.CEC5pc, 2014a.

Sénési, S., Richon, J., Franchistéguy, L., Tyteca, S., Moine, M.-P., Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J.,

Douville, H., Madec, G., Maisonnave, E., Planton, S., Saint-Martin, D., Szopa, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: CNRM-CM5 model output prepared for CMIP5 lgm, served by ESGF,WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.CEC5lg, 2014b.

Spence, J. P., Eby, M., and Weaver, A. J.: The Sensitivity of the Atlantic Meridional Overturning Circulation to Freshwater Forcing at Eddy-Permitting Resolutions, Journal of Climate, 21, 2697-2710, 10.1175/2007JCLI2103.1, 2008.

Sueyoshi, T., Ohgaito, R., Yamamoto, A., Chikamoto, M. O., Hajima, T., Okajima, H., Yoshimori, M., Abe, M., O'Ishi, R., Saito, F., Watanabe, S., Kawamiya, M., and Abe-Ouchi, A.: Set-up of the PMIP3 paleoclimate experiments conducted using an Earth system model, MIROC-ESM, Geosci. Model Dev., 6, 819-836, 10.5194/gmd-6-819-2013, 2013.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93, 485-498, 10.1175/BAMS-D-11-00094.1, 2011.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M. P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: description and basic evaluation, Clim Dyn, 40, 2091-2121, 10.1007/s00382-011-1259-y, 2013.

Winguth, A. M. E., Archer, D., Duplessy, J. C., Maier-Reimer, E., and Mikolajewicz, U.: Sensitivity of paleonutrient tracer distributions and deep-sea circulation to glacial boundary conditions, Paleoceanography, 14, 304-323, 10.1029/1999PA900002, 1999.

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Tanaka, T. Y., Shindo, E., Tsujino, H., Deushi, M., Mizuta, R., Yabu, S., Obata, A., Nakano, H., Koshiro, T., Ose, T., and Kitoh, A.: A New Global Climate Model of the Meteorological Research Institute: MRI-CGCM3;Model Description and Basic Performance, Journal of the Meteorological Society of Japan. Ser. II, 90A, 23-64, 10.2151/jmsj.2012-A02, 2012.

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Tanaka, T., Shindo, E., Tsujino, H., Deushi, M., Mizuta, R., Yabu, S., Obata, A., Nakano, H., Koshiro, T., Ose, T., and Kitoh, A.: MRI-CGCM3 model output prepared for CMIP5 piControl, served by ESGF,WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.MRMCpc, 2015a.

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., Tanaka, T., Shindo, E., Tsujino, H., Deushi, M., Mizuta, R., Yabu, S., Obata, A., Nakano, H., Koshiro, T., Ose, T., and Kitoh, A.: MRI-CGCM3 model output prepared for CMIP5 lgm, served by ESGF,WDCC at DKRZ, doi:10.1594/WDCC/CMIP5.MRMClg, 2015b.

Variable description	Units	Resolution (lon×lat), time	Variable name
Specific humidity	kg kg^{-1}	192×94, 1460	huss ~~10m~~
Downwelling longwave radiation	W m^{-2}	192×94, 365	rlds
Downwelling shortwave radiation	W m^{-2}	192×94, 365	rsds
Precipitation	mm day^{-1}	192×94, 12	pr
Wind (u and v components)	m s^{-1}	192×94, 1460	uas and vas
Temperature	K	192×94, 1460	tas ~~10m~~
Sea surface salinity	psu	360×180, 12	sos

Table 1: Summary of the data showing variable description, units, format (lon×lat, time), ~~Notes,~~and NetCDF variable name(s) ~~and the original PMIP3 variable name(s)~~. Formats follow CORE conventions (Large and Yeager, 2004). The wind component variables are provided in separate files (Morée and Schwinger, 2019a). In each NetCDF file (i.e., for each variable) the model spread is provided alongside the anomaly field named 'variablename_spread'.

5

Model name	Modelling group	Reference	Source data reference
CNRM-CM5	CNRM-CERFACS (France)	Voldoire et al. (2013)	piControl: Sénési et al. (2014a) lgm: Sénési et al. (2014b)
IPSL-CM5A-LR	IPSL (Institut Pierre Simon Laplace, France)	Dufresne et al. (2013)	piControl: Caubel et al. (2016) lgm: Kageyama et al. (2016)
MIROC-ESM	MIROC (JAMSTEC and NIES, Japan)	Sueyoshi et al. (2013)	piControl: JAMSTEC et al. (2015a) lgm: JAMSTEC et al. (2015b)
MRI-CGCM3	MRI (Meteorological Research Institute, Japan)	Yukimoto et al. (2012)	piControl: Yukimoto et al. (2015a) lgm: Yukimoto et al. (2015b)
GISS-E2-R	NASA/GISS (Goddard Institute for Space Studies, USA)	Schmidt et al. (2014)	piControl: NASA-GISS (2014a) lgm: NASA-GISS (2014b)

Table 2: PMIP3 models used in this study

#	CDO or NCO command
1	cdo remapbil,t62grid
2	cdo ensmean
3	cdo sub
4	cdo setmisstodis
5	ncap2
6	cdo inttime

Table 3: Package commands applied in this study. Detailed information on these commands can be found in the respective NCO and CDO documentation online. All operations were performed with either CDO version 1.9.3 (Schulzweida, 2019) or NCO version 4.6.9. The complete list of commands is available in the NetCDF files ~~as~~ under global attribute 'history'.

5

Humidity (kg kg^{-1})
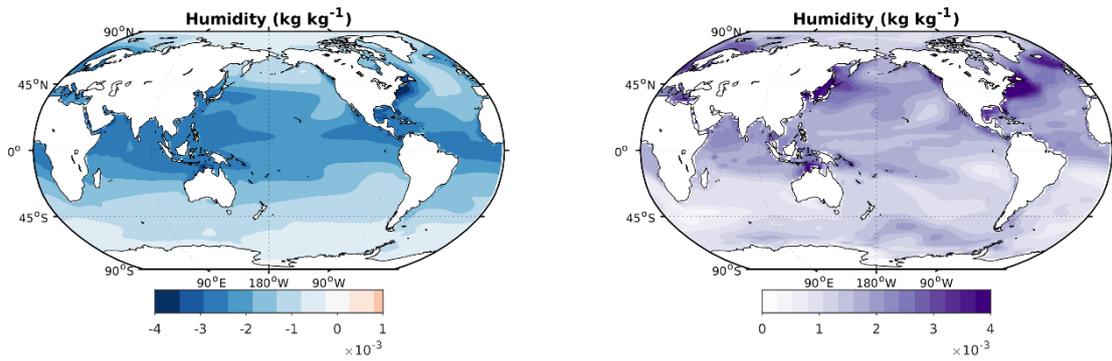
Humidity (kg kg^{-1})

Figure 1: Annual mean 10-meter height specific humidity lgm-piControl anomaly (left) and model spread (right) in kg kg^{-1}.

Longwave radiation (W m^{-2})

Longwave radiation (W m^{-2})
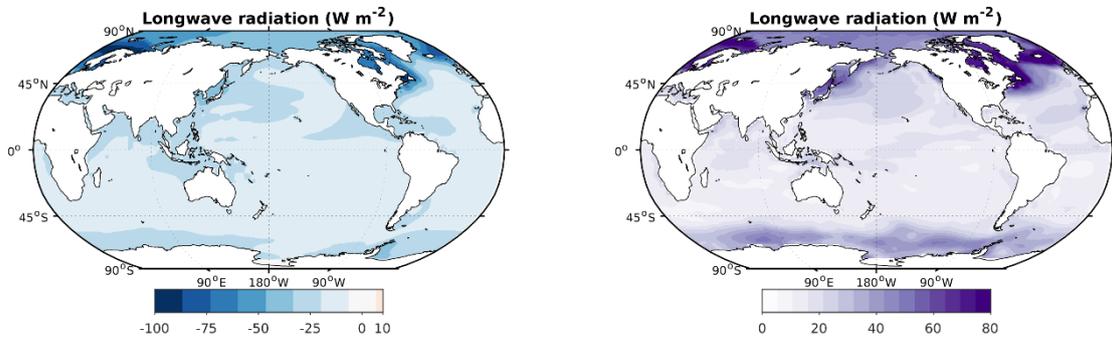
Figure 2: Annual mean downwelling longwave radiation lgm-piControl anomaly (left) and model spread (right) in W m^{-2}.
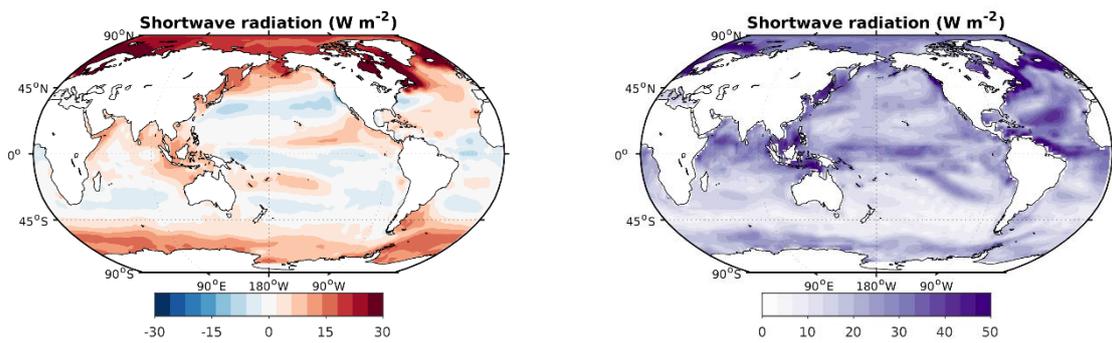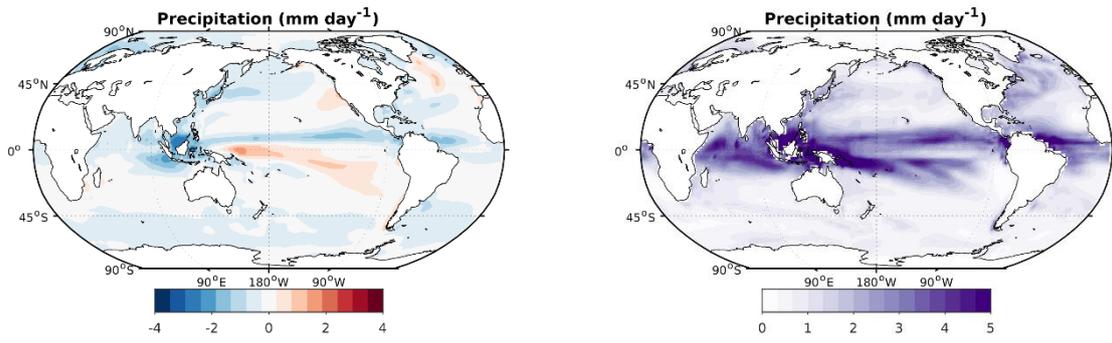
Shortwave radiation (W m^{-2})

Shortwave radiation (W m^{-2})

Figure 3: Annual mean downwelling shortwave radiation lgm-piControl anomaly (left) and model spread (right) in W m^{-2}.

Figure 4: Annual mean precipitation lgm-piControl anomaly (left) and model spread (right) in mm day^{-1}.
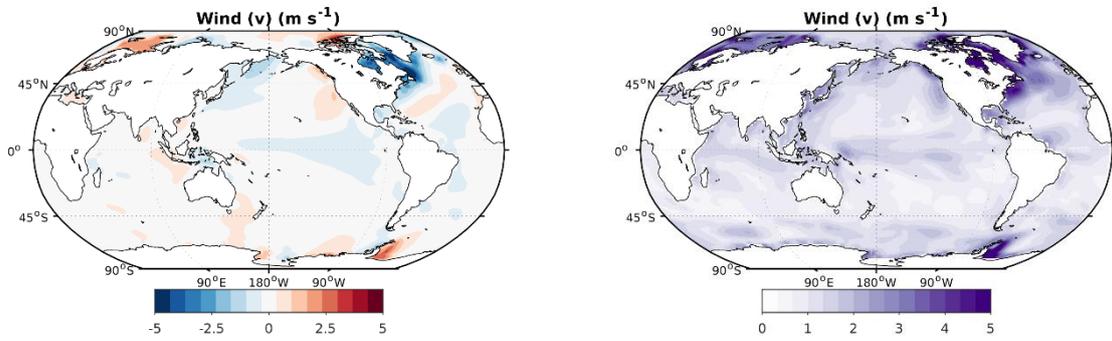

Figure 5: Annual mean meridional wind velocity lgm-piControl anomaly (left) and model spread (right) in m s^{-1}.
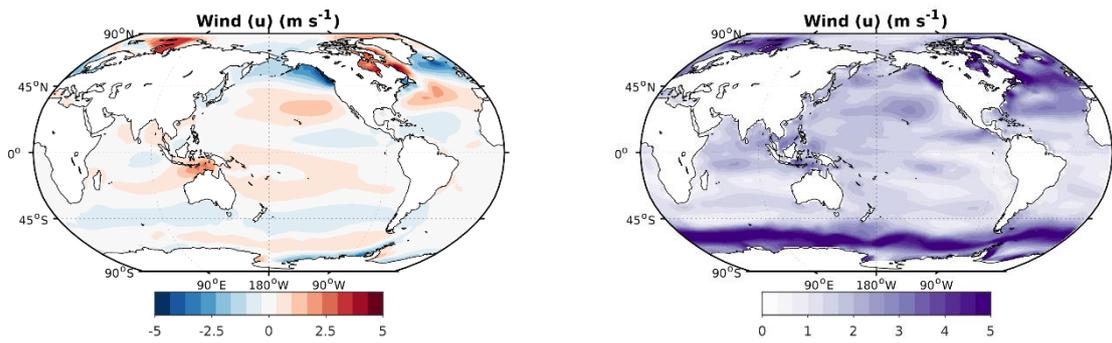

5

Figure 6: Annual mean zonal wind velocity lgm-piControl anomaly (left) and model spread (right) in m s^{-1}.
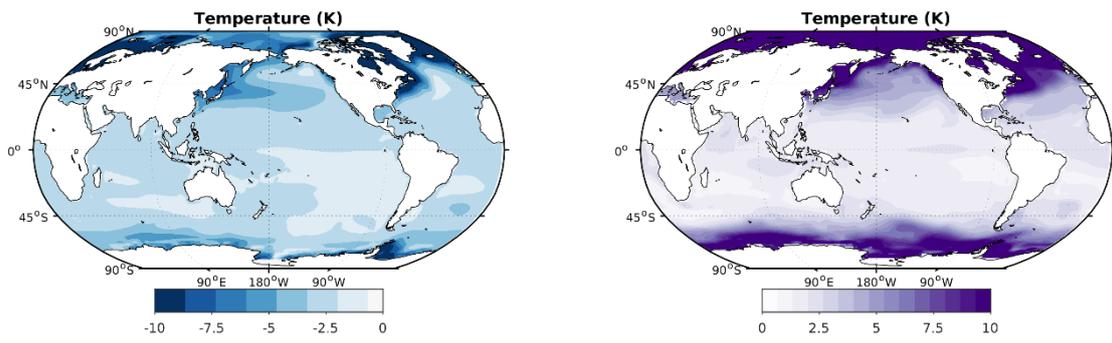
Figure 7: Annual mean 10-meter height temperature lgm-piControl anomaly (left) and model spread (right) in K.
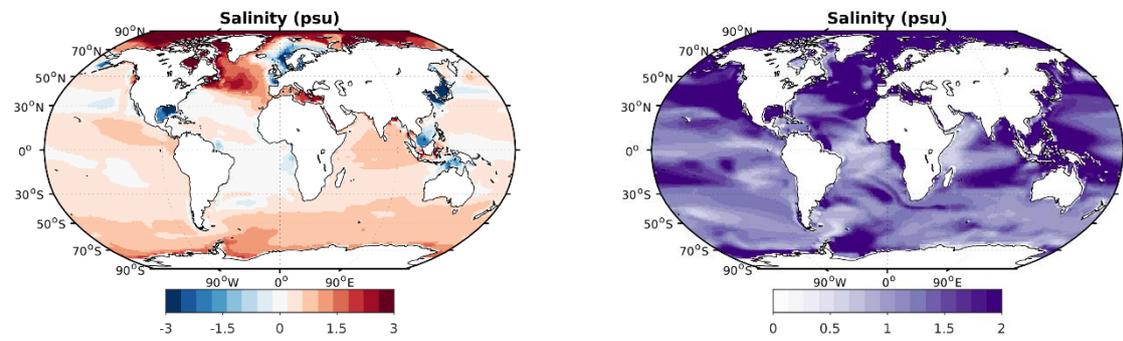

Figure 8: Annual mean sea surface salinity lgm-piControl anomaly (left) and model spread (right) in psu.