A Gaussian process emulator for simulating ice sheet-climate interactions on a multi-million year timescale: CLISEMv1.0

Jonas Van Breedam[1], Philippe Huybrechts[1], Michel Crucifix[2]

[1]Earth System Science & Departement Geografie, Vrije Universiteit Brussel, Brussels, Belgium
5 [2]Georges Lemaître Centre for Earth and Climate Research (TECLIM), Earth and Life Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

Correspondence to: Jonas Van Breedam (jonas.van.breedam@vub.be)

Abstract. On multi-million year timescales, fully coupled ice sheet - climate simulations are hampered by computational limitations, even at coarser resolutions and when using asynchronous coupling schemes. In this study, a novel coupling method
10 CLISEMv1.0 (CLimate-Ice Sheet EMulator version 1.0) is presented where a Gaussian process emulator is applied to the climate model HadSM3, coupled to the ice sheet model AISMPALEO. The temperature and precipitation fields from HadSM3 are emulated to feed the mass balance model in AISMPALEO. The sensitivity of the evolution of the ice sheet over time is tested with respect to the number of predefined ice sheet geometries the emulator is calibrated on. Additionally, the model performance is evaluated to the formulation of the ice sheet parameter (being either ice sheet volume, either ice sheet area, or
15 both) and to the coupling time. Sensitivity experiments are conducted to explore the uncertainty introduced by the emulator. Also different lapse rate adjustments are used between the relatively coarse climate model and the much finer ice sheet model topography. It is shown that the ice sheet evolution over a million-year timescale is strongly sensitive to the definition of the ice sheet parameter and to the number of predefined ice sheet geometries. With the new coupling procedure, we provide a computationally efficient framework for simulating ice sheet-climate interactions on a multi-million year timescale that allows
20 for a large number of sensitivity tests.

1 Introduction

Earth System Models provide the state-of-the-art for quantifying feedbacks between the different components of the climate system on a decadal to centennial timescale (Eyring et al., 2016). On millennial to multi-millennial timescales, Earth System Models of Intermediate Complexity are used to explore the feedbacks in the climate system between the ice sheets, the atmosphere and the ocean (Eby et al., 2013; Van Breedam et al., 2020). Those fully coupled models, even at coarser resolution, are computationally very expensive and other techniques have been proposed to simulate ice sheet-climate interactions on a (multi) million-year timescale.

The basic asynchronous method, also called the direct asynchronous method is a simple and straightforward coupling technique to address the different response times between ice sheets and the atmosphere (Pollard, 2010). The direct asynchronous method consists in running a climate model, typically a General Circulation Model (GCM), for a few decades until steady state and then use the relevant climatic information over the polar regions as input to an ice sheet model. The ice sheet model is typically run for a few thousand years before the climatic information is updated (e.g. DeConto and Pollard, 2003; Gasson et al., 2016). This procedure is repeated for the entire time span of interest. The indirect asynchronous coupling, also called the matrix method or the GCM lookup table (Pollard, 2010) is based on predefined, idealised GCM snapshots that span the possible forcing during the entire simulation period. The matrix look-up table is more sophisticated, with the creation of a matrix of climate model output from the extremes of the forcing and some intermediate climate states in between (Ladant et al., 2014; Stap et al., 2017; Berends et al., 2018). The coupling procedure linearly interpolates the climatic fields from precursor climate model runs.

An alternative approach is to consider a Gaussian process emulator. A Gaussian process emulator is a statistical model that fits a Gaussian process to data in order to link input with output fields of a model, generally referred to as the simulator (Andrianakis and Challenor, 2012). Emulators have been used for a number of applications in climate science, for instance as a tool to predict the future climate evolution (Levermann et al., 2020) or sea-level rise as a result of land ice melting (Edwards et al., 2021), based on large ensembles of simulations with each different model input parameters. It is also a useful technique to couple different components of the climate system that would require large computational resources such as an atmosphere-ocean coupling (Tran et al., 2019). An emulator has been used to assess the sensitivity of the climate during the Pleistocene (Araya-Melo et al., 2015) and the late Pliocene (Lord et al., 2017). In these simulations, the ice sheets are static and defined by different ice sheet geometries. So far, an emulator has not been used to study the climate system including dynamic ice sheets.

Here a Gaussian process emulator is presented that is calibrated on the climatic output from the climate model HadSM3 to force an ice sheet model in order to predict the climate over Antarctica during the late Eocene. The late Eocene to Eocene-

Oligocene Transition is chosen because of the large contrast in continental glaciation and large variations in climate forcing
such as CO_2 concentrations at this time (Pagani et al., 2011; Zhang et al., 2013). The ice sheet model runs continuously over a
multi-million year period and passes the ice sheet geometry information (referred to as the ice sheet parameter) to the emulator
(statistical representation of the climate model). The emulator calculates the climatic variables based on the prescribed external
forcing (carbon dioxide concentration and orbital parameters) and the actual ice sheet parameter and returns temperature and
precipitation data to the ice sheet model. This coupling procedure is novel, but various implementation choices may influence
the result: the approach for lapse rate adjustment, the coupling time between ice sheet and climate, and the definition of
emulator input variables. Also, the number of GCM experiments on which the emulator is tuned might have an influence on
the predicted climate. The key questions to be answered can be summarized as follows:

1. The ice sheet parameter is defined by a number that represents the influence of the ice sheet in the climate system
   (Araya-Melo et al. 2015, Lord et al., 2017). The ice sheet mainly influences the local climate by its distinct albedo,
   by its height and by its freshwater input into the ocean (not taken into account in this study). Therefore, it is not trivial
   how the ice sheet parameter should be defined in a single number. Is ice volume a proper way to define the ice sheet
   parameter? Does ice area represent the climatic changes better? Or is it best to calibrate the emulator based on both
   ice volume and ice area?

2. The emulator needs a number of input ice sheet geometries to simulate the climate for a range in orbital parameters
   and CO_2 values for the given ice sheet geometry. How many ice sheet geometries and climate model experiments are
   needed? Does the spacing between different ice sheet geometries influence the model performance?

3. The lapse rate adjustment between a coarse resolution climate model and the high-resolution ice sheet model is usually
   applied by a constant value for the moist adiabatic lapse rate over the domain. Common values are 5˚C per km (Ladant
   et al., 2014), 6.5˚C per km (Löfverström et al., 2015), 7˚C per km (Thompson and Pollard, 1997) or 8˚C per km
   (Berends et al., 2018). The lapse rate above ice-covered regions is found to be 4.9 ˚C per km, smaller than the typical
   values for the moist adiabatic lapse rate (Gardner et al., 2009). Moreover, the near surface lapse rate varies spatially
   and temporarily between diurnal and seasonal cycles as opposed to the free adiabatic lapse rate that has a rather
   constant value (Marshall et al., 2006). What is the influence of using a different lapse rate on the ice sheet evolution?

4. With asynchronously coupled climate-ice sheet model runs one generally updates the climatic information every
   several thousand years only, given the long response time of the ice sheets and the computational limits. However,
   the choice of the coupling time might have an influence on the ice sheet evolution over time. What is the optimal
   coupling time to have a realistic, yet efficient model running ?

5. What is the uncertainty introduced by the emulator and what is its influence on the coupled ice sheet-climate
   simulations?

2 Model description

In this section, the new coupling method CLISEMv1.0 is described together with the climate model HadSM3 and the ice sheet model AISMPALEO. CLISEMv1.0 is calibrated on climatic output from HadSM3 and provides the forcing fields (monthly temperature and precipitation) for the ice sheet model. The ice sheet model AISMPALEO returns the ice sheet volume and/or area to CLISEMv1.0.

2.1 Climate model HadSM3

The climate model HadSM3 (Williams et al., 2001) is an atmosphere-slab ocean General Circulation Model (GCM). It has a resolution of 2.5˚ in latitude and 3.75˚ in longitude, with 19 levels in the vertical (Gordon et al., 2000). The MOSES-1 scheme is chosen as the land surface scheme (Cox et al., 1999) with a tundra-like albedo on the Antarctic continent where no ice is present and an albedo for snow where ice is present. Sea surface temperatures are reconstructed based on a best estimate from Evans et al. (2017) for the late Eocene in order to calibrate the corrective heat fluxes from the slab ocean model. These corrective heat fluxes represent the seasonal deep water exchange and horizontal heat transport that is present in the real ocean. The oceanic heat fluxes are exchanged between the atmosphere and the slab ocean model in the mixed-layer, which is 50 m thick in our simulations. This way, realistic sea surface temperatures are simulated for the different climate model simulations.

The model is chosen because of its good performance over the Antarctic ice sheet for the present-day and the Last Glacial Maximum (Connolley and Bracegirdle, 2007; Maris et al., 2012). Moreover, HadSM3 is a computational efficient climate model that allows performing a large number of experiments (Valdes et al., 2017). The paleogeographic reconstruction for the simulations is based on the method presented in Baatsen et al. (2016) and makes use of the Gplates software (Müller et al., 2018). The paleogeographic reconstruction represents the continental configuration at 39 Ma to be representative for the late Eocene. As a result, the Antarctic continent has a slightly different position compared to the present-day. The bedrock topography for Antarctica is derived from the Wilson et al. (2012) maximum bedrock elevation reconstruction (Figure 1). Simulated temperatures for a warm orbital configuration (maximum austral summer insolation) and a cold orbital configuration (minimum austral summer insolation) are shown in Fig. 2a and 2b, respectively.

2.2 Antarctic ice sheet model AISMPALEO

The Antarctic ice sheet model AISMPALEO is a three-dimensional thermomechanical ice sheet/ice shelf model (Huybrechts and de Wolde, 1999) used for simulating ice sheet dynamics during periods in the past (Goelzer et al., 2016a, 2016b) and the future (Seroussi et al., 2020, Van Breedam et al., 2020). The Shallow Ice Approximation is used to calculate the grounded ice flow, which is a result of internal deformation and basal sliding where the pressure melting point is reached. The model comprises a component taking into account the solid Earth response due to ice loading. This component consists of a rigid

elastic plate on top of a viscous asthenosphere to allow for deviations from local isostatic loading. The surface mass balance is computed using the Positive Degree Day (PDD) method (Janssens and Huybrechts, 2000), where the yearly sum of daily average temperatures above 0 ˚C is used to determine the melt potential. The standard deviation of the mean daily temperature is 4.2 ˚C, representing random weather fluctuations and the daily cycle. The difference in snow and ice albedo in the ice sheet model is taken into account by using a PDD factor for snow melting of 0.003 m ice equivalent (i.e.) per ˚C per day and a PDD factor for ice melting of 0.008 m i.e. per ˚C per day. Monthly mean temperature and precipitation are used from HadSM3 to drive the PDD model. The rain limit is chosen at 1 ˚C and determines whether precipitation falls as snow or as rain. Meltwater retention allows runoff to be retarded and/or to eventually refreeze in the snowpack. Ice shelf formation is included and calculated using the Shallow Shelf Approximation. Ice shelves start to form when the grounding line reaches the coast and the influx of ice from the continent exceeds the ablation (surface ablation and basal melting). A constant basal melt rate of 1 m per year is used in all the simulations. The ice sheet model is run at a resolution of 40 km to allow for long integrations.

2.3 CLISEMv1.0: set-up and calibration

The Gaussian Process (GP) Principal Component Analysis (PCA) emulator (Wilkinson, 2010; Bounceur et al., 2015; Lord et al., 2017) used in this study is a statistical representation of the climate model HadSM3 (the simulator). The emulator is calibrated on a relatively small number of climate model runs and aims to predict the climate for any combination of climatic forcing of the original climate model runs. To allow for reliable predictions, the initial climate model runs need to fill the entire multi-dimensional input space. In our case, the multi-dimensional input space consists of the orbital parameters (eccentricity e, obliquity ε and longitude of perihelion ϖ), the carbon dioxide concentration forcing, and the ice sheet size and extent defined by the ice sheet parameter. The theoretical background has already been discussed in detail in previous papers (Araya-Melo et al., 2015; Bounceur et al., 2015; Lord et al., 2017) and here the focus is on the implementation of the dynamic ice sheet component.

It is recommended to have at least 10 experiments per input parameter (Loeppky et al., 2009). Therefore, the recommended minimum number of experiments with our five input parameters (three orbital parameters, CO_2 concentration and ice sheet parameter) would be 50. Since the atmosphere-slab ocean model is time-efficient, it is chosen to run 100 climate model runs with 5 variable forcing parameters. The ice sheets have a very distinct climatic imprint compared to the orbital parameters and the CO_2 level, which all result in smooth climatic fields. Because of the large difference in albedo between ice and tundra at the edge of the ice sheet, the climatic imprint of a certain ice sheet geometry has a sharp boundary. The number of ice sheet geometries taken into the model design of the emulator might therefore have a large impact on the performance of the emulator. To test the impact of the ice sheet parameter, four different emulators are constructed based on a different number of predefined ice sheet geometries or based on a different spread of the ice sheet geometries. The different emulators are named according to the number of ice sheet geometries in the model design, being 8, 12 and 20 for respectively EMULATOR_8, EMULATOR_12a, EMULATOR_12b and EMULATOR_20.

Except for the number of predefined ice sheet geometries, also the spread of the different prescribed geometries is varying between the different emulators, depending on whether the ice sheet parameter is defined by ice area or by ice volume. EMULATOR_8, EMULATOR_12a and EMULATOR_20 have a good spread between the different ice sheet geometries in terms of ice volume and ice area. EMULATOR_12b is well defined for ice volume, but poorly defined by ice sheet area as there are several experiments with the same ice sheet area yet different ice sheet geometry (see Table A1 in Appendix for the experimental parameter values). This way, the influence of the spread in ice sheet volume/ice area on the emulated climate is investigated. The spacing of the different ice sheet geometries is expected to be crucial for medium sized ice sheets, because they constitute a transition zone towards a fully glaciated continent. EMULATOR_20 has the smallest spacing of ice sheet geometries around the crucial medium sized ice sheets, separated at the minimum distance that corresponds to the resolution of the climate model. The maximum ice sheet geometry in the model design for EMULATOR_12 is smaller than the maximum ice sheet geometry in the model design for EMULATOR_20 and EMULATOR_8. The objective designing EMULATOR_12 is to evaluate to what extent the emulator can still be used in an extrapolation regime beyond the largest ice sheet geometry.

A model design with 100 GCM experiments is constructed where each experiment has a different combination of orbital parameters, CO_2 concentration values and ice sheet geometry (see Table A1 in Appendix). Insolation values are well approximated as a linear combination of the eccentricity and longitude of perihelion (Loutre, 1993) and therefore, the terms $e\sin\varpi$ and $e\cos\varpi$ in combination with the obliquity ε are used for the orbital parameter variation in the model design. The range of orbital parameters is taken from Laskar et al. (2004) for the period 40 Ma to 33 Ma. The eccentricity has a maximum value during the period between 40 Ma and 33 Ma of 0.063 and the obliquity is sampled in the range of 22 - 24.5°. The CO_2 interval ranges from ~~1150~~ 550 ppmv to ~~550~~ 1150 ppmv, roughly equivalent to ~~4xCO₂~~ 2xCO₂ to ~~2xCO₂~~ 4xCO₂. The ice sheet parameter consists of 8, 12 or 20 predefined ice sheet geometries. They are constructed based on preliminary steady-state ice sheet model runs for a range of different climatic forcings (for EMULATOR_12a and EMULATOR_12b) or from ice sheet geometry snap shots during the build-up of a continental scale ice sheet (for EMULATOR_8 and EMULATOR_20). The final ice sheet geometries are chosen to provide a range from an almost ice-free Antarctic continent up to a fully glaciated continent. Tundra is present between the ice sheet margin and the coast. The parameter combinations are constructed using a Latin hypercube design where the minimum Euclidean distance between two parameter combinations is maximised (Figure 5). With this model design, 100 GCM runs are performed until the climate (atmosphere and slab ocean) is in steady-state with the forcing (40 years).

The design matrix of input data D (n x p) has 100 rows (number of experiments) and 5 ($e\sin\varpi$, $e\cos\varpi$, ε, CO_2, ice volume or ice area) or 6 columns ($e\sin\varpi$, $e\cos\varpi$, ε, CO_2, ice volume, ice area). Each simulation performed by the climate model is characterized by a row of matrix D, called the input vector x_i. The climatic output (temperature and precipitation) where HadSM3 is a function of the input vector is called $f(x_i)$ from all 100 experiments is saved in the matrix Y. Each column of Y

contains the output for one experiment. Here, the matrix Y only contains climatic output data on the ice sheet model grid (201x201 grid points) because our interest is the climate evolution over Antarctica.

The climatic output is modelled as a Gaussian process, defined by a mean function m(x) and a covariance function V(x,x'). The prior mean function is defined by a linear combination of a set of linear regression functions (Eq. 1), where β is a vector of regression coefficients corresponding to the mean function and h(x) is a vector of known regression functions of the inputs. The covariance function consists of the correlation function and the scaling value σ^2 (Eq. 2). The squared exponential correlation function is chosen with the inclusion of the so called nugget ν and correlation length δ hyperparameters (eq. 3). The nugget was originally meant to deal with sampling variability of the simulator output, but even when this sampling variability is small, it is effective to prevent numerical instability and to compensate for inadequate correlation priors (Andrianakis and Challenor, 2012; Araya-Melo et al., 2015). The correlation length δ can be understood as a measure of how quickly the model output is changing as a function of the input (Wilkinson, 2010). The larger the distance between two input vectors, the quicker the correlation goes to zero.

$$m(x) = h(x)^T \beta, \tag{1}$$

$$V(x,x') = \sigma^2 [c(x,x')], \tag{2}$$

$$c(x,x') = exp\left\{ \sum_{i=1}^{p} \left(\frac{d_i}{\delta_i} \right)^2 \right\} + \nu I, \tag{3}$$

The formalism followed here is Bayesian and the prior mean and prior covariance functions (Eq. 1-3) are updated (Eq. 4-5) based on the climate model output data. All values of β are a priori equally probable and we assume a vague conjugate prior (β, σ^2) that is proportional to σ^2. The posterior distribution of the model data (temperature and precipitation) is a Student's t distribution with n-q degrees of freedom (which is close to Gaussian) with a mean m*(x) and covariance V*(x,x'):

$$m^*(x) = h(x)^T \hat{\beta} + t(x) A^{-1} (y - H\hat{\beta}), \tag{4}$$

$$V^*(x,x') = \frac{(y-H\hat{\beta})^T A^{-1}(y-H\hat{\beta})}{n-q-2} [(c(x,x') + \nu I) - t(x)^T A^{-1} t(x') + P(x)(H^T A^{-1} H)^{-1} P(x')^T], \tag{5}$$

A PCA is performed on the climatic output and between 17 and 20 components are kept before calibrating the emulator (see model code for PC, length scales and nugget values). The Gaussian process model with the posterior means and variances is applied to each principal component. Once the PCA emulator is calibrated (by optimizing the nugget and length scales), the scores of each principal components can be estimated for arbitrary input values, with an associated covariance that effectively measures the prediction uncertainty.

As described before, four different emulators are constructed (EMULATOR_8, EMULATOR_12a, EMULATOR_12b and EMULATOR_20) and each emulator is calibrated with either ice volume, ice area or with both ice volume and ice area as the ice sheet parameter. This gives a total of 4 x 3 = 12 distinct calibrated emulators to simulate the ice sheet evolution. Calibration of an emulator is achieved by adjusting the length scales δ, the nugget ν (uncertainty band) and the number of Principal Components (PCs) in order to minimize the root mean square error between the simulated and predicted climatic fields in leave-one-out experiments. It is assumed that the sampling error introduced by model variability is almost negligible because we are using a slab-ocean climate model version where the internal climate variability is small and the climate states quickly converge to a mean value. Therefore, the adopted nugget is chosen to be 0.001 (small non-zero uncertainty band around the data) for the temperature and 0.01 for the precipitation emulation. The emulator calibration is done for precipitation and temperature data for each month. During the calibration process, the number of PCs is varied between 5 and 25. It is chosen to keep 20 PCs to explain the observed variation in temperature and 17 PCs to explain the variation in precipitation, since these numbers gave the best emulator performance (as explained further on).

R's Nelder-Mead optimization function (Nelder and Mead, 1965) is used to maximise the likelihood of the emulator (Kennedy and O'Hagan, 2000; Lord et al., 2017). We obtain a low length scale for the ice sheet parameter (between 0.02 and 0.5). When looking at the summer temperature for all GCM runs, there is a smooth, almost linear dependency between the ice sheet parameter and the simulated summer temperatures (Figure 6) and therefore we increased the length scale for the ice sheet parameter. The optimization function is used to get the correlation lengths for the orbital parameters and the CO_2 forcing right and the correlation length for the ice sheet parameter was manually chosen to be 1.2 for all emulators.

245

The emulator performance is determined by the variance of the emulator and the reliability of the emulator. The variance is a measure of the uncertainty of the mean predictions of the emulator. The reliability of the emulator determines how well the emulator is calibrated (i.e. how well the emulator estimates its own uncertainty). Ideally, the emulator is well-calibrated and the uncertainty is low. The calibration is investigated by leave-one-out experiments where the simulated temperature is predicted based on the calibrated emulator while leaving out one experiment at a time. The results are visualised in Fig. 7 where the number of grid points that is predicted within 1 (grey) to 4 (red) standard deviations from the simulated temperature is given for each of the GCM runs for EMULATOR_20. Overall, all emulators perform well, since > 68 % of the grid points are predicted within 1 standard deviation. The calibration based on ice volume and ice area separately shows a very similar reliability and uncertainty. Even though the variance for the emulators calibrated both on ice area and on ice volume is lower, it has a lower reliability because it struggles to capture the output dependency on both variables simultaneously.

The spatial difference between the simulated and the predicted climatic fields is visualised in Fig. 8. January temperatures are shown for experiment *xaemdk* that performs poorly and for experiment *xaemdv* that performs quite well. These experiments are run with the second smallest and smallest ice sheet geometry, respectively. The predicted temperatures are warm-biased

260 up to 8°C for the tundra region and cold-biased up to 10 °C for the ice-covered region for experiment *xaemdk*. The bias is much smaller for experiment *xaemdv* with errors less than 2 °C over most of the Antarctic continent. Other experiments that perform bad such as *xaemdg* and *xaembb* have a very high eccentricity, high obliquity and summer during aphelion or perihelion. They have the most extreme insolation values and lay at the edge of the experimental design, which may explain why the emulator does a more poor job at predicting the simulated temperatures. .

2.4 Coupling procedure between AISMPALEO and the emulator

265

Due to computational limitations, the coupling procedure between ice sheets and climate on a multi-million year timescale is always asynchronous. However, the emulator-ice sheet coupling leaves the possibility for a very short coupling time because once the emulator is calibrated, it can be run in standalone modus (i.e. without the need to use the simulator). In these simulations, the ice sheet model is initialized from an ice-free state. After a predefined time (1000 years in the standard experiments), the ice sheet model passes the ice sheet parameter (the actual ice volume, ice area or both ice volume and ice area) to the emulator and the emulator provides temperatures and precipitation as a function of the orbital parameters, CO_2 forcing and the ice sheet parameter. In our simulations, temperature and precipitation are interpolated to the ice sheet model grid using a bilinear interpolation scheme, in order to have smooth climatic fields. In the standard experiments, a constant lapse rate correction of 5 °C per km is applied between HadSM3 and AISMPALEO. The climatic information is lapse rate corrected for the nearest input ice sheet geometry in terms of ice volume or ice area.

270

275

3 Sensitivity of the ice sheet evolution to the model set-up

In this section, the sensitivity of the ice sheet evolution is tested for the different emulators. The influence of the different number of ice sheet geometries in the model design is investigated in combination with how the ice sheet parameter is defined. In addition the sensitivity of the ice sheet evolution is tested to the coupling time and to the lapse rate adjustment between the coarse climate model and the much finer ice sheet model. The performance of the four different emulators is assessed. The ice sheet evolution is forced over a 3 million year time period with the real orbital forcing from 38 Ma to 35 Ma (Laskar et al., 2004) and CO_2 scenarios assuming a linear decrease in concentrations from around 980 ppmv to 720 ppmv.

280

3.1 Sensitivity to the definition of the ice sheet parameter

The ice sheet parameter represents the shape and area of the ice sheets. In previous studies (Araya-Melo et al., 2015; Lord et al., 2017), it is defined by indexing the different ice sheet geometries that have been used to generate the simulation outputs. . Yet, there are several other options on how to define the ice sheet parameter. Here, it is proposed to define the ice sheet parameter by quantifying the ice sheet volume, the ice sheet area or by a vector combining both aspects. This is done for the four different experiment designs. The ice sheet area and ice sheet volume are good parameters to define the ice sheet's influence on climate because the first parameter affects the local albedo and the latter has an influence on the elevation and

285

hence on the temperatures through adiabatic cooling. In case the ice sheet parameter is defined by both the ice sheet volume and the ice sheet area, both variables are calculated after each iteration of the ice sheet model.

Figure 9a shows the ice sheet evolution for the four emulators calibrated on ice volume. EMULATOR_12a, EMULATOR_12b and EMULATOR_20 show the transition towards a continental scale ice sheet in a very narrow CO_2 interval of 845 to 875 ppmv. On the other hand, EMULATOR_8 does not seem to show any sensitivity to the CO_2 forcing during the 3 million year-long simulation (and also not on a longer timescale). The reason is that the prescribed ice sheets are separated too much in the initial climate model runs. Because of the large difference in albedo between ice and tundra, the prescribed ice sheets create a sharp boundary at the ice sheet margin that is visible in the temperature field. If insufficient prescribed ice sheet geometries are used, the ice sheet does not grow enough to see the temperature regime obtained when HadSM3 is run with the next input ice sheet geometry. Consequently, the emulated temperatures remain too high at the ice sheet margin. It appears that the threshold on the number of needed ice sheet geometries is somewhere between 8 and 12. Using 20 input ice sheet geometries does not change the model performance much.

Another option is to calibrate the emulator based on the ice area, which is more directly linked to the albedo. The glaciation threshold for EMULATOR_12a and EMULATOR_20 shows a very similar sensitivity to CO_2 of about 860 ppmv (Figure 9b). EMULATOR_8 grows immediately to a medium sized ice sheet and cannot grow further towards a continental scale ice sheet for the reasons already quoted. EMULATOR_12b was poorly defined on ice area (several ice sheet geometries had a similar area, but different geometry) and the ice sheet grows immediately towards a continental scale. Therefore, in addition to having sufficient ice sheet geometries, a good spacing of the ice sheet parameter is another requirement to use an emulator for coupled ice sheet-climate simulations.

The simulated temperatures during the first 100 kyr of the simulations are visualised during a strong insolation maximum after 47,000 years and a strong insolation minimum after 60,000 years (Figure 10) for all four emulators calibrated on ice volume. The corresponding ice sheet sizes are given in Figure A5 (appendix). The temperature patterns are very similar, especially above the tundra regions. The main difference in simulated temperatures between the different emulators is caused by the size of the ice sheet, which in turn is determined by the number and spacing of ice sheets used for the calibration.

When the coupling is based on both the ice volume and the ice area, the transition to a fully glaciated climate gives very distinct results for each of the emulators (Figure 9c). EMULATOR_12a shows the transition to a continental scale glaciation for a similar CO_2 threshold than for the emulators tuned on ice volume or ice area of around 890 ppmv. EMULATOR_12b shows the transition to a fully glaciated continent immediately when the simulations start, because of the poor definition on ice area. Remarkably, the transition to a fully glaciated continent for EMULATOR_20 happens for a much lower CO_2 threshold of 765 ppmv. However, the reliability of EMULATOR_20 calibrated both on ice area and ice volume is lower than the reliability of

the emulator on either ice volume or ice area (see section 2.3), even though the variance is smaller when additional information

325 on the ice sheet parameter is added. The poor emulation originates from calibrating the emulator based on 6 variables, while actually only 5 input forcing parameters are reasonably independent. The additional information on the ice sheet parameter is strongly correlated in most cases, but not everywhere because the spread between ice volume and ice area is not equal. This is visualised in Fig. 11 where three different schematic ice sheet geometries are shown with their respective ice volume and ice area (normalized). For the second ice sheet geometry, the ice area increases with 0.8 units, while the ice volume increases with

330 0.2 units. Oppositely, the next ice sheet geometry is defined by an ice volume increase of 0.8 units and an ice area increase of 0.2 units. In EMULATOR_20, the smallest prescribed ice sheet geometries have a significant increase in ice area while the ice volume increase is relatively smaller. On the other hand, the largest prescribed ice sheet geometries have a much larger increase in ice volume than ice area (Figure 4). For EMULATOR_12a, the relative increase in ice area and ice volume is more equal and therefore, the reliability and the performance of the emulator is better. EMULATOR_8 grows to a fully glaciated continent

335 for a CO_2 threshold around 810 ppmv. As mentioned earlier, the lack of sufficient ice sheet geometries is also visible here with complete growth and decline close to the glaciation threshold.

The difference in ice sheet area and ice sheet volume evolution during the build-up of an ice sheet is further visualised in Fig. 12. When the ice sheet starts growing, both the ice sheet area and the ice volume increase. However, the timing between ice

340 sheet volume growth and ice sheet area growth are not fully synchronous. When the climate is cool during an austral summer insolation minimum, the area where the mass balance is positive increases and first the area where ice is accumulating increases. It takes more time for the ice volume to adjust and by the time the ice volume reaches a maximum, the ice area starts decreasing again during an austral summer insolation maximum. This can be seen in Fig. 12c where the first snapshot of the ice sheet geometry shows a relatively large area with ice, but ice volume at that moment is negligible. The ice area will start

345 decreasing towards an insolation maximum, while the ice volume is still growing in response to the large area with a positive mass balance. The delay of the ice volume response compared to ice area is even more clear when an ice sheet is growing from bare bedrock to a continental scale ice sheet(Figure 12a and 12b).

These remarks suggest a possible improvement that would consist in a better experiment design with ice sheets spanning a 2-

350 D space more optimally. On the other hand, the difference in relative magnitude of ice volume and ice area during the build-up of an ice sheet also suggests that it is not easy to create an optimal set of variables for the model design.

3.2 Sensitivity to the coupling time between AISMPALEO and the emulator

The coupling time in the first set of experiments is 1000 years. The sensitivity of the ice sheet evolution to the coupling time is tested by applying five different coupling times, ranging from 10 years to 2000 years. The smallest time step is in the same

355 order of magnitude as the time step used in the ice sheet model. This way, it can be regarded as an example of a direct coupling between the climatic component and the ice sheet model. Fig.13 shows the ice sheet evolution during one precession cycle for

the five coupling times. The climatic information for the largest time step of 2000 years is updated eleven times during this interval and the ice sheet evolution clearly responds stepwise. Another observation is that the higher the coupling time step, the more delayed the ice sheet responds to the forcing and the lower the amplitude of the ice sheet volume. Decreasing the coupling time step results in a smoother ice volume evolution. The differences between a coupling time of 500 years, 250 years and 10 years are becoming very small, suggesting that the solution converges. To make a compromise between model efficiency and model accuracy, we opted for performing the multi-million year sensitivity simulations with a coupling time step of 500 years as a lower limit and 2000 years as an upper limit.

360

365 The coupling time is doubled and halved to respectively 2000 and 500 years to test its influence on the ice sheet evolution for EMULATOR_12a and EMULATOR_12b calibrated with ice volume as the ice sheet parameter (Figure 14). Halving the coupling time step increases the computational time with about 40 %, while doubling the coupling time decreases the computational time with the same percentage. When the coupling time decreases, the ice sheet volume has a larger amplitude and is slightly more sensitive to changes in insolation.

370

For EMULATOR_12a, the glaciation threshold is more sensitive to CO_2 changes when the coupling time is decreased. The difference in glaciation threshold between a coupling time step of 500 years and 1000 years is negligible, but the difference with a coupling time step of 2000 years is about 30 ppmv. The continental-scale glaciation for EMULATOR_12b with a coupling time step of 1000 years occurs for lower CO_2 values than for a coupling time step of 500 and 2000 years due to the complex interaction between ice sheet response and forcing. The shorter the coupling time, the more sensitive the ice sheet is to the forcing. For a coupling time step of 1000 years, the ice sheet grows more than for a coupling time step of 2000 years, but does not decline as much as for a coupling time step of 500 years and therefore grows quicker to the fully glaciated state.

375

Comparing the glaciation threshold with respect to the CO_2 forcing between the different coupling time steps, there is a decreasing difference between EMULATOR_12a and EMULATOR_12b for a decreasing coupling time step. The difference in glaciation threshold is about 40 ppmv for a coupling time step of 2000 years, 30 ppmv for a coupling time step of 1000 years and 10 ppmv for a coupling time step of 500 years, suggesting that the solution converges for decreasing coupling time steps.

380

3.3 Sensitivity to the lapse rate adjustment between HadSM3 and AISMPALEO

385 The lapse rate is the change in temperature with elevation and its value is highly dependent on the moisture in the air. Since the climate model is relatively coarse compared to the ice sheet model, we need to apply a lapse rate correction to the elevation difference between the climate model and ice sheet model. The lapse rate is, in reality, both temporarily and spatially variable, and here we test different model choices. One experiment includes temporally variations in the lapse rate correction, and another experiment includes both spatial and temporal variations. The temporally variable lapse rate is calculated as the average

near-surface lapse rate for all grid points on the Antarctic continent for each month (Figure 15 and Figure A1). The spatially variable lapse rate is included by calculating the local near-surface lapse rate (dT/dZ) simulated by HadSM3 for the 4 adjacent ice covered grid points for each month. The average monthly lapse rate varies roughly between the wet adiabatic lapse rate during summer (December-January) and the dry adiabatic lapse rate in winter (June-July). These values are higher than the constant lapse rate that was applied in the standard experiments and therefore, the lapse rate corrected temperatures are lower for a growing ice sheet. The resulting ice sheet evolution shows a more stepwise change towards full glaciation (Figure 16). The larger temperature difference between the climate model and the ice sheet model makes it harder for the ice sheet to grow further until a threshold is reached and a large area is cold enough for snow accumulation.

4 ~~Bayesian sensitivity~~Uncertainty analysis with EMULATOR_ ~~100b~~12b calibrated ~~with~~ on ice volume

The additional value of the use of an emulator for coupled ice sheet-climate simulations is that the mean climate predictions come with the estimate of its variance and that two different predictions have a covariance. Here, the uncertainties caused by the emulator variance are explored in order to sample climate trajectories. The covariance between output points given by the emulator is used to update the mean and variance of a climate prediction at a given iteration (e.g. time iteration i) of the ice sheet model, given the climate used at the previous point iteration (i-1). It is then possible to sample the updated distribution. This provides a climate sample at iteration i and the procedure continues to obtain climate samples at iteration i+1 and so forth. The process yields a sample climate trajectory. Strictly speaking, the emulated climate at iteration i is correlated with the outputs at iteration i-1, i-2, etc. and all of them should be used to update the mean and variance at iteration i. However, since the Gaussian process emulator has an exponential decaying correlation function which is short-ranged (in contrast to a power-law), it is expected that the covariance structure of emulated climate trajectories that is associated to the emulator variance is effectively captured by the first-order autocorrelation.

Now that the general principle is explained, more mathematical details about the procedure are provided with specific attention to the fact that a PCA emulator is used. At any time step i, the current temperature is estimated on the basis of principal components. To this end, the emulator mean is computed for each PC score, given the orbital parameters, the CO_2 concentration and the ice level at the current time step $m(x_i)$ and at the previous time step $m(x_{i-1})$, along with the computed covariances associated with these elements, denoted $V(x_i,x_i)$, $V(x_{i-1},x_i)$ and $V(x_{i-1},x_{i-1})$, The mean and covariance at time step i are then updated given the scores for the corresponding principal component T_{i-1} which was effectively applied at time step i-1. The PC score at time step i which will finally be applied to the ice sheet model is drawn from this distribution. The procedure is repeated for each PC score. The temperature field reconstructed from these PC scores is further perturbed by a random field with variance equal to the residual variance not captured by the principal components. This approach provides us with a random draw of the temperature field consistent with the information provided by the PCA emulator (Eq. 7-9).

$$m(x_i) \vee T_{i-1} = m(x_i) + \frac{V(x_i, x_{i-1})}{V(x_{i-1}, x_{i-1})}(T_{i-1} - m(x_{i-1})),$$ (7)

$$\sigma^2(x_i) \vee T_{i-1} = V(x_i, x_i)\left(1 - \left(\frac{V(x_i, x_{i-1})^2}{V(x_i, x_i)V(x_{i-1}, x_{i-1})}\right)\right),$$ (8)

$$T_i \sim N\left(m(x_i) \vee T_{i-1}, \sigma^2(x_i) \vee T_{i-1}\right),$$ (9)

425

The uncertainty of the emulator is explored by performing 50 Monte Carlo simulations including the variance on EMULATOR_12b calibrated with ice volume. Since temperature and precipitation are emulated separately for each month, a choice had to be made on which parameter is most decisive for ice sheet growth. It appears that summer temperature has a main control on the evolution of the ice sheet over time. Therefore, the uncertainties in the emulated January temperature are

430 explored and emulated precipitation in a similar manner as the previous experiments. The temperatures of the other months are reconstructed based on the annual cycle of the 100 input experiments. First, the mean temperature for each month and for each input ice sheet geometry is calculated. Then, the annual cycle with respect to the January temperatures is calculated for each input ice sheet geometry (Figure 17). The mean difference between each month and the emulated January temperature is applied to calculate the temperature of the other months. By applying a constant temperature anomaly compared to the January

435 temperature, the temperature of the ice-free regions in January is overestimated for the June temperatures when these regions are snow-covered. This has a minor influence on the results because ice melt does not occur during the austral winter months anyway. So, the emulator is almost identical to EMULATOR_12b, except that covariances are used to sample trajectories around the mean. If the variances are assumed to be zero at all time steps, the mean trajectory already presented in Fig. 9a is approximated (slight difference due to the application of the annual cycle to the January temperature).

440

The resulting ice sheet evolution over time is shown in Fig. 18 for a coupling time of 500 years and a coupling time of 2000 years. The original simulations are shown by the blue curve and the approximation by emulating only January temperatures and applying a constant correction based on the annual cycle is represented by the green curve. Generally, the curves look very similar, but the simulations with emulating only the January temperatures underestimate slightly the ice sheet volume

445 compared to the original run and result in a glaciation threshold occurring for lower CO_2 values than for the original run. This is the results of applying a constant temperature correction over the entire continent with respect to the January temperature. The actual austral autumn/winter/spring temperatures are colder on the ice-free regions due to the effect of snowfall on the albedo.

450 The black curves represent 50 Monte Carlo simulations where the variance of the January temperature is added to the mean predictions. Both simulations, with a coupling time step of 500 years and 2000 years, have a similar variance. The overall uncertainty in the ice sheet evolution is also comparable for both coupling time steps with a difference between the lowest and highest glaciation threshold of about 25-50 ppmv for more than 95 % of the simulations. Striking is the asymmetry in glaciation

threshold for the experiments including the variance in comparison with the reference experiment (Figure 18). Most experiments including the variance predict the glaciation threshold to happen for lower CO_2 values than the reference experiment based on the mean temperature prediction. The reason is that the ice sheet in these simulations lose mass during an insolation maximum for a longer time and therefore does not manage to grow above the mean prediction. Fig. 18b also shows that the ice sheet is growing (peaks are higher) and melting (the ice sheet is not that stable) faster for a coupling time of 2000 years compared to a coupling time of 500 years. The main mechanism is the slow response time of the ice sheet that has more time to grow during an insolation minimum and more time to decay during an insolation maximum for a larger coupling time step.

5 Discussion

The aim of the new coupling technique CLISEMv1.0 is to create an efficient and accurate way to model ice sheet-climate interactions on time scales beyond what directly coupled models can achieve. In that sense, we build further on previous modelling attempts such as the asynchronous coupling method and the matrix method. The basic asynchronous method has the advantage that you do not need to have any prior information on the possible ice sheet geometries. A strong disadvantage is that this method does not allow for sensitivity experiments at a reasonable computing time since the whole chain of ice sheet model - GCM runs would have to be repeated. Nevertheless, this method has been very popular in paleoclimatic studies during all time periods in geological history where ice might have been present, from the Paleozoic ice houses (Horton et al., 2007; Lowry et al., 2014; Pohl et al., 2016) over the Eocene-Oligocene Transition (DeConto and Pollard, 2003), Miocene (Gasson et al., 2016) to the Quaternary ice ages (Charbit et al., 2002; Herrington and Poulsen, 2012).

With CLISEMv1.0, the forcing uncertainty is explored with preliminary GCM snapshots. The emulation of climate and precipitation is akin to Kriging or geospatial interpolation, but in 5 or 6-dimensional space with a large number of climate model runs. It is not the same as linear interpolation as the posterior mean includes a term that absorbs deviations from linearity. Crucially, the emulator comes with estimates of covariance, which measures the uncertainty introduced by using the emulator. Checking that this uncertainty is consistent with leave-one-out experiments is a key aspect of the emulator evaluation. A so-called "nugget" allows to introduce variability directly explained by the model inputs, such as model internal variability (Andrianakis and Challenor, 2012), but in our design this nugget is a small numerical value. For this reason, the use of a GP emulator is also not completely equal to interpolating the raw model output from the climate model. This is in contrast to the climate matrix method, that consists of a limited number of GCM runs for the end members in the forcing and linearly interpolates the climatologies based on the actual ice sheet geometry (Gasson et al., 2016; Stap et al., 2017; Berends et al., 2018).

485 Whether the asynchronous coupling method, the matrix look-up table or the GP emulator is used to simulate ice sheet - climate interactions on (multi)-million year timescales, a choice on the coupling time has to be made and that will affect the outcome. When using the matrix look-up table or the emulator, the result will also be dependent on the choice of the model design and on the number of GCM runs. For the GP emulator, additional choices need to be made about the length scale, nugget and the covariance function to use. The more complexity is added to the method, the more uncertainties might arise. On the other hand,

490 the GP emulator provides a posterior covariance which provides an objective criterion to verify that it is well calibrated and to evaluate the introduced uncertainties.

A common problem for the emulator and the matrix look-up table method, where the ice sheet parameter is defined by a single number, is that there is no control on the regions where ice starts to grow. The problem can be addressed by describing the ice

495 sheet location and geometry with a vector of several dimensions. In reverse, the definition of this vector and the experiment design have to be thought such as to provide a reasonably orthogonal experiment design, in order to avoid the issues experienced in this study by attempting to calibrate the emulator both on ice volume and ice area simultaneously. Optionally, ice sheets could be described with additional variables such as shape factors that are relatively independent on the other ice sheet parameters. We leave the suggestion of creating other ice sheet variables to improve the emulator performance for future

500 work.

To have a properly working emulator including dynamic ice sheets, it is crucial to have a good spacing between the input ice sheet geometries and sufficient different ice sheet geometries. Five different emulators showed to have a good set-up: EMULATOR_12a, EMULATOR_12b and EMULATOR_20 tuned on ice volume, and EMULATOR_12a and

505 EMULATOR_20 tuned on ice area (Figure 19). EMULATOR_8 included too little input ice sheet geometries to make the ice sheet grow to a continental scale ice sheet . EMULATOR_12b tuned on ice area did not produce reliable results because several input ice sheet areas had a different geometry, yet similar area. The glaciation threshold for each of these five emulators ranges between 845 ppmv and 875 ppmv. Overall, taking into account the radiative forcing of carbon dioxide, the differences in glaciation threshold are very small between the different ways of calibrating the emulators (see video supplement for the ice

510 sheet geometry evolution of EMULATOR_12a, EMULATOR_12b and EMULATOR_20 calibrated on ice volume).

It appears that the more prescribed ice sheet geometries are used in the emulator, the larger the amplitude of ice sheet growth during a cold orbital configuration. However, this has no impact on the CO_2 threshold to continental scale glaciation. The maximum ice sheet extent is also dependent on the largest predefined ice sheet geometry. EMULATOR_12a and

515 EMULATOR_12b had a smaller largest predefined ice sheet geometry than EMULATOR_20. The extrapolation of the emulated climatic variables lead to a climatic state that was too cold, allowing for a larger ice sheet than EMULATOR_20 that was calibrated on a larger range of prescribed ice sheet geometries.

Our simulations show that using a coupling time step of 500 years instead of 2000 years results in a quicker ice sheet response to the forcing. Such small coupling time steps are not common for multi-million year simulations with three-dimensional ice

520 sheet models coupled to a climate model. Gasson et al. (2016) used an asynchronous coupling method to simulate the ice sheet evolution during the Miocene with a coupling step of 2000 years, while Stap et al. (2017) used a coupling time step of 500 years for a one-dimensional ice sheet model forced by a climate model. Also the value for the lapse rate correction clearly has an influence on the ice sheet evolution over time and the emulator allows a number of sensitivity experiments. We have attempted to include a more realistic lapse rate that follows the seasonal and spatial variations. In HadSM3, the monthly

525 average lapse rate over the Antarctic ice sheet ranges between 7 ˚C per km in January (summer) to 10.5 ˚C per km in July (winter). The lapse rate over the Greenland ice sheet during the LGM had a similar range from ~5.5 ˚C per km during summer to 9.5 ˚C per km during winter (Erokhina et al., 2017). This near-surface lapse rate is influenced by atmospheric boundary processes, the surface type (snow, tundra) and the atmospheric circulation (Kageyama et al., 2005).

6 Conclusions

530 In this study, the computational efficient coupler CLISEMv1.0 that provides climatic fields for simulating ice sheet-climate interactions on a multi-million year has been described together with its sensitivity to the implementation and an uncertainty analysis. CLISEMv1.0 estimates the climate as a function of the orbital parameters, the CO_2 forcing and the ice sheet parameter where each forcing is defined by a single number. The ice sheet parameter is either defined by the ice sheet area, the ice sheet volume or by both the ice sheet area and ice sheet volume.

535

A set of different emulators was constructed to investigate the influence of the number of prescribed ice sheet geometries in the model design on the coupled ice sheet-climate simulations. The number of precursor ice sheet geometries has a large effect on the ice sheet sensitivity to CO_2 and orbital forcing, because of its large climatic imprint caused by the high albedo of ice. Also, the spread of the ice sheet geometries has been shown to have a significant impact on the performance of the coupled

540 ice sheet-climate simulations and has a larger effect than the definition of the ice sheet parameter. When there is an equal spread between the ice sheet area and the ice sheet volume of the input ice sheet geometries, the threshold to continental scale glaciation occurs in a very narrow CO_2 window of 860 ± 15 ppmv.

Once the emulator is well-calibrated, the emulator – ice sheet coupling method is very suitable to use for performing ice sheet-

545 climate simulations on a multi-million year timescale and to use for sensitivity tests. Here we tested the sensitivity of the ice sheet evolution to the coupling time and to the lapse rate adjustment. Our results indicated that the glaciation threshold to the CO_2 forcing converges for a decreasing time step. Also, shortening the coupling time slightly increases the sensitivity to CO_2 forcing. The shorter the coupling time, the larger the ice sheet grows during an insolation minimum and the more the ice sheets shrinks during an insolation maximum. This might have large consequences for paleoclimatic studies implementing

550 asynchronous coupling techniques, where the coupling time is usually on the order of several millennia. The elevation differences between coarse climate models and high resolution ice sheet models are usually corrected for by applying a constant lapse rate correction. The value of this lapse rate correction has a larger effect than the coupling time and we propose to take the real lapse rate correction that is observed in the climate model output.

555 The emulator-ice sheet coupling method is applied here for idealized CO_2 scenarios for the time period between 38 Ma and 35 Ma ago. In these simulations, temperature and precipitation are emulated to drive the mass balance of the ice sheet model. CLISEMv1.0 is a useful tool to investigate the ice sheet evolution during all major climatic transitions of the Cenozoic where the interaction between orbital parameters and CO_2 variations are thought to have played a significant role or even to investigate the existence of pre-Cenozoic glaciations throughout the Phanerozoic.

560 **Author contribution**

JVB designed the coupling method between the emulator and the ice sheet model. MC developed the Gaussian process emulator. PH developed the ice sheet model code. JVB wrote the manuscript with contributions from all co-authors.

Code and data availability

The code for the coupler CLISEMv1.0 between the climate and the ice sheet model is available on Zenodo:
565 https://doi.org/10.5281/zenodo.5245156. Also a video supplement showing the ice sheet evolution for the three best performing emulators calibrated on ice volume is available on Zenodo: https://doi.org/10.5281/zenodo.5242914. All data used in this paper is available upon request.

Competing interests

PH is topical editor of GMD.

570 **References**

Andrianakis, I. and Challenor, P.G.: The effect of the nugget on Gaussian process emulators of computer models, Computational Statistics and Data Analysis, 56, 4215-422, https://doi.org/10.1016/j.csda.2012.04.020, 2012.
Araya-Melo, P.A., Crucifix, M. and Bounceur N.: Global sensitivity analysis of the Indian monsoon during the Pleistocene, Climate of the Past, 11, 45-61, https://doi.org/10.5194/cp-11-45-2015, 2015.

575 Baatsen, M., van Hinsbergen, D. J. J., von der Heydt, A. S., Dijkstra, H. A., Sluijs, A., Abels, H. A., and Bijl, P. K.: Reconstructing geographical boundary conditions for palaeoclimate modelling during the Cenozoic, Clim. Past, 12, 1635–1644, https://doi.org/10.5194/cp-12-1635-2016, 2016.

Berends, C.J., de Boer, B. and van de Wal, R.S.W.: Application of HadCM3@Bristolv1.0 simulations of paleoclimate as forcing for an ice-sheet model, ANICE2.1: set-up and benchmark experiments, Geosci. Model Dev., 1, 4657-4675,
580 https://doi.org/10.5194/gmd-11-4657-2018, 2018.

Bounceur, N., Crucifix, M., and Wilkinson, R. D.: Global sensitivity analysis of the climate–vegetation system to astronomical forcing: an emulator-based approach, Earth Syst. Dynam., 6, 205–224, https://doi.org/10.5194/esd-6-205-2015, 2015.

Charbit, S., Ritz, C. and Ramstein, G.: Simulations of Northern Hemisphere ice-sheet retreat: sensitivity to physical mechanisms involved during the Last Deglaciation, Quat. Sci. Rev., 21, 243-265, https://doi.org/10.1016/S0277-
585 3791(01)00093-2, 2002.

Connolley, W.M. and Bracegirdle, T.J.: An Antarctic assessment of IPCC AR4 coupled models, Geophysical research Letters, 34, L22505, https://doi.org/10.1029/2007GL031648, 2007.

Cox, P.M., Betts, R.A., Bunton, C.B., Essery, R.L.H., Rowntree, P.R. and Smith. J.: The impact of new land surface physics on the GCM simulation of climate and climate sensitivity, Climate Dynamics, 15, 183-203,
590 https://doi.org/10.1007/s003820050276, 1999.

DeConto, R. M. and Pollard, D.: Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO_2, Nature, 421, 245-248, https://doi.org/10.1038/nature01290, 2003.

Eby, M., Weaver, A. J., Alexander, K., Zickfeld, K., Abe-Ouchi, A., Cimatoribus, A. A., Crespin, E., Drijfhout, S. S., Edwards, N. R., Eliseev, A. V., Feulner, G., Fichefet, T., Forest, C. E., Goosse, H., Holden, P. B., Joos, F., Kawamiya, M., Kicklighter,
595 D., Kienert, H., Matsumoto, K., Mokhov, I. I., Monier, E., Olsen, S. M., Pedersen, J. O. P., Perrette, M., Philippon-Berthier, G., Ridgwell, A., Schlosser, A., Schneider von Deimling, T., Shaffer, G., Smith, R. S., Spahni, R., Sokolov, A. P., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., Zeng, N., and Zhao, F.: Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity, Clim. Past, 9, 1111–1140, https://doi.org/10.5194/cp-9-1111-2013, 2013.

600 Edwards, T. L., Nowicki, S., Marzeion, B., Hock, R., Goelzer, H., Seroussi, H., Jourdain, N. C., Slater, D., Turner, F., Smith, C. J., McKenna, C. M., Simon, E., Abe-Ouchi, A., Gregory, J. M., Larour, E., Lipscomb, W. H., Payne, A. J., Shepher, A., AGosta, C., Alexander, P., Albrech, T., Anderson, B., Asay-Davis, X, Aschwanden, A., Barthel, A., Bliss, A., Calov, R, Chambers, C., Champollion, N., Choi, Y., Cullather, R., Cuzzone, J., Dumas, C., Felikson, D., Fettweis, X., Fujita, K., Galton-Fenzi, B. K., Gladstone, R., Golledge, N. R., Greve, R., Hattermann, T., Hoffman, M. J., Humbert, A., Huss, M., Huybrechts,
605 P., Immerzeel, W., Kleiner, T., Kraaijenbrink, P., Le clec'h, S., Lee, V., Leguy, G. R., Little, C. M., Lowry, D. P., Malles, J. H., Martin, D. F., Maussion, F., Morlighem, M., O'Neill, J. F., Nias, I., Pattyn, F., Pelle, T., Price, S., Quiquet, A., Radic, V., Reese, R., Rounce, D. R., Rückamp. M., Sakai, A., Courtney, S., Schlegel, N.-J., Shannon, S., Smith, R. S., Straneo, F., Sun,

S., Tarasov, L., Trusel, L. D., Van Breedam, J., Van de Wal, R., van den Broeke, M., Winkelmann, R., Zekollari, H., Zhao, C., Zhang, T. and Zwinger, T.: Projected land ice contributions to 21st century sea level rise, Nature, (accepted).

610   Erokhina, O., Rogozhina, I., Prange, M., Bakker, P., Bernales, J., Paul, A. and Schulz, M.: Dependence of slope lapse rate over the Greenland ice sheet on background climate, J. Glaciol., 63, 568-572, https://doi.org/10.1017/jog.2017.10, 2017.

Evans, D., Sagoo, N., Renema, W., Cotton, L.J., Müller, W., Todd, J.A., Saraswati, P.K., Stassen, P., Ziegler, M., Pearson, P.N., Valdes, P.J. and Affek, H.P.: Eocene greenhouse climate revealed by coupled clumped isotope-Mg/Ca thermometry, PNAS, 115, 1174-1179, https://doi.org/10.1073/pnas.1714744115, 2017.

615   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, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Gardner, A.S., Sharp, M., Koerner, R.M., Labine, C., Boon, S., Marshall, S.J., Burgess, D.O. and Lewis, D.: Near-Surface Temperature Lapse Rates over Arctic Glaciers and Their Implications for Temperature Downscaling, J. Clim., 22, 4281-4298,

620   https://doi.org/10.1175/2009JCLI2845.1, 2009.

Gasson, E., DeConto, R. M. and Pollard, D.: Dynamic Antarctic ice sheet during the early to mid-Miocene, PNAS, 113, 3459-3464, https://doi.org/10.1073/pnas.1516130113, 2016.

Goelzer, H., Huybrechts, P., Loutre, M.-F., and Fichefet, T.: Impact of ice sheet meltwater fluxes on the climate evolution at the onset of the Last Interglacial, Clim. Past, 12, 1721–1737, https://doi.org/10.5194/cp-12-1721-2016, 2016a.

625   Goelzer, H., Huybrechts, P., Loutre, M.-F., and Fichefet, T.: Last Interglacial climate and sea-level evolution from a coupled ice sheet–climate model, Clim. Past, 12, 2195–2213, https://doi.org/10.5194/cp-12-2195-2016, 2016b.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B. and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, Clim. Dyn., 16, 147-168, https://doi.org/10.1007/s003820050010, 2000.

630   Herrington, A. R. and Poulsen, C. J.: Terminating the Last Interglacial: The Role of Ice Sheet–Climate Feedbacks in a GCM Asynchronously Coupled to an Ice Sheet Model. J. Climate, 25, 1871–1882, https://doi.org/10.1175/JCLI-D-11-00218.1, 2012.

Horton, D. E., Poulsen, C.J. and Pollard, D.: Orbital and CO_2 forcing of late Paleozoic continental ice sheets, Geophys. Res. Lett., 34, L19708, https://doi.org/10.1029/2007GL031188, 2007.

635   Huybrechts, P. and De Wolde, J.: The Dynamic Response of the Greenland and Antarctic Ice Sheets to Multiple-Century Climatic Warming, J. Clim., 12, 2169–2188, https://doi.org/10.1175/1520-0442(1999)012<2169:tdrotg>2.0.co;2, 1999.

Janssens, I. and Huybrechts, P.: The treatment of meltwater retention in mass-balance parameterizations of the Greenland ice sheet, Ann. Glaciol, 31, 133-140, https://doi.org/10.3189/172756400781819941, 2000.

Kageyama, M., Harrison, S.P. and Abe-Ouchi, A.: The depression of tropical snowlines at the last glacial maximum: What

640   can we learn from climate model experiments? Quat. Int., 138-139, 202-219, https://doi.org/10.1016/j.quaint.2005.02.013, 2005.

Kennedy, M.C. and O'Hagan, A.: Predicting the Output from a Complex Computer Code when Fast Approximations are Available, Biometrika, 87, 1-13, https://doi.org/10.2307/2673557, 2000.

Ladant, J-B., Donnadieu, Y., Lefebvre, V. and Dumas, C.: The respective role of atmopsheric carbon dioxide and orbital parameters on ice sheet evolution at the Eocene-Oligocene transition, Paleoceanography, 29, 810-823, https://doi.org/10.1002/2013PA002593, 2014.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A.C.M. and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth, Astron. Astrophys., 428, 261-285, https://doi.org/10.1051/0004-6361:20041335, 2004.

Levermann, A., Winkelmann, R., Albrecht, T., Goelzer, H., Golledge, N. R., Greve, R., Huybrechts, P., Jordan, J., Leguy, G., Martin, D., Morlighem, M., Pattyn, F., Pollard, D., Quiquet, A., Rodehacke, C., Seroussi, H., Sutter, J., Zhang, T., Van Breedam, J., Calov, R., DeConto, R., Dumas, C., Garbe, J., Gudmundsson, G. H., Hoffman, M. J., Humbert, A., Kleiner, T., Lipscomb, W. H., Meinshausen, M., Ng, E., Nowicki, S. M. J., Perego, M., Price, S. F., Saito, F., Schlegel, N.-J., Sun, S., and van de Wal, R. S. W.: Projecting Antarctica's contribution to future sea level rise from basal ice shelf melt using linear response functions of 16 ice sheet models (LARMIP-2), Earth Syst. Dynam., 11, 35–76, https://doi.org/10.5194/esd-11-35-2020, 2020.

Loeppky, J.L., Sacks, J. and Welch, W.J.: Choosing the Sample Size of a Computer Experiment: A Practical Guide, Technometrics, 51, 366-376, https://doi.org/10.1198/TECH.2009.08040, 2009.

Löfverström, M., Liakka, J. and Kleman, J.: The North American Cordillera – An Impediment to Growing the Continent-Wide Laurentide Ice Sheet, J. Clim, 28, 9433-9450, https://doi.org/10.1175/JCLI-D-15-0044.1, 2015.

Lord, N.S., Crucifix, M., Lunt, D.J., Thorne, M.C., Bounceur, N., Dowsett, H., O'Brien, C.L. and Ridgwell, A.: Emulation of long-term changes in global climate: application to the late Pliocene and future, Climate of the Past, 13, 1539-1571, https://doi.org/10.5194/cp-13-1539-2017, 2017.

Lowry, D. P., C. J. Poulsen, D. E. Horton, T. H. Torsvik, and D. Pollard, Thresholds for Paleozoic ice sheet initiation, Geology, 42, 627 – 630, https://doi.org/10.1130/G35615.1, 2014.

Maris, M.N.A, de Boer, B. and Oerlemans, J.: A climate model intercomparison for the Antarctic region: present and past, Clim. Past. 8, 803-814, https://doi.org/10.5194/cp-8-803-2012, 2012.

Marshall, S.J., Sharp, M.J., Burgess, D.O. and Anslow, F.S.: Near-surface-temperature lapse rates on the Prince of Wales Icefield, Ellesmere Island, Canada: implications for regional downscaling of temperature, Int. J. Climatol., 27, 385-398, https://doi.org/10.1002/joc.1396, 2006.

Müller, R.D., Cannon, J., Qin, X., Watson, R.J., Gurnis, M., Williams, S., Pfaffelmoser, T., Seton, M., Russell, S.H.J. and Zahirovic, S.: GPlates: Building a Virtual Earth Through Deep Time, Geochemistry, Geophys. Geosystems, 19, 2243-2261, https://doi.org/10.1029/2018GC007584, 2018.

Nelder, J. A., and Mead, R.: A simplex algorithm for function minimization, The Computer Journal, 7, 308-313, https://doi.org/10.1093/comjnl/7.4.308, 1965.

Pagani, M., Huber, M., Zhonghui, L., Bohaty, S. M., Henderiks, J., Sijp, W., Krishnan, S. and DeConto, R.M.: The Role of
Carbon Dioxide During the Onset of Antarctic Glaciation, Science, 334, 1261-1264, https://doi.org/10.1126/science.1203909,
2011.

Pohl, A., Donnadieu, Y., Le Hir, G., Ladant, J.-B., Dumas, C., Alvarez-Solas, J. and Vandenbroucke, T. R. A.: Glacial onset
predated Late Ordovician climate cooling, Paleoceanography, 31, https://doi.org/10.1002/2016PA002928, 2016.

Pollard, D.: A retrospective look at coupled ice sheet-climate modelling, Climatic Change, https://doi.org/10.1007/s10584-010-9830-9, 2010.

Seroussi, H., Nowicki, S., Payne, A. J., Goelzer, H., Lipscomb, W. H., Abe-Ouchi, A., Agosta, C., Albrecht, T., Asay-Davis,
X., Barthel, A., Calov, R., Cullather, R., Dumas, C., Galton-Fenzi, B. K., Gladstone, R., Golledge, N. R., Gregory, J. M.,
Greve, R., Hattermann, T., Hoffman, M. J., Humbert, A., Huybrechts, P., Jourdain, N. C., Kleiner, T., Larour, E., Leguy, G.
R., Lowry, D. P., Little, C. M., Morlighem, M., Pattyn, F., Pelle, T., Price, S. F., Quiquet, A., Reese, R., Schlegel, N.-J.,
Shepherd, A., Simon, E., Smith, R. S., Straneo, F., Sun, S., Trusel, L. D., Van Breedam, J., van de Wal, R. S. W., Winkelmann,
R., Zhao, C., Zhang, T., and Zwinger, T.: ISMIP6 Antarctica: a multi-model ensemble of the Antarctic ice sheet evolution over
the 21st century, The Cryosphere, 14, 3033–3070, https://doi.org/10.5194/tc-14-3033-2020, 2020.

Stap, L. B., van de Wal, R. S. W., de Boer, B. Bintaja, R. and Lourens, L. J.: The influence of ice sheets on temperature during
the past 38 million years inferred from a one-dimensional ice sheet-climate model, Climate of the Past, 13, 1243-1257,
https://doi.org/10.5194/cp-13-1243-2017, 2017.

Thompson, S. L. and Pollard, D.: Greenland and Antarctic Mass Balances for Present and Doubled Atmospheric CO2 from
the GENESIS Version-2 Global Climate Model, J. Clim., 10, 871-900, https://doi.org/10.1007/s00382-018-4205-4, 1997.

Tran, G. T., Oliver, K. I. C., Holden, P. B., Edwards, N. R., Sóbester, A. and Challenor, P.: Multi-level emulation of complex
climate model responses to boundary forcing data, Clim. Dyn., 52, 1505-1531, https://doi.org/10.1007/s00382-018-4205-4,
2019.

Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J.,
Farnsworth, A., Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V.,
Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models:
HadCM3@Bristol v1.0, Geosci. Model Dev., 10, 3715–3743, https://doi.org/10.5194/gmd-10-3715-2017, 2017.

Van Breedam, J., Goelzer, H., and Huybrechts, P.: Semi-equilibrated global sea-level change projections for the next 10 000
years, Earth Syst. Dynam., 11, 953–976, https://doi.org/10.5194/esd-11-953-2020, 2020.

Wilkinson, R. D: Bayesian Calibration of Expensive Multivariate Computer Experiments, in: Large-Scale Inverse Problems
and Quantification of Uncertainty, edited by: Biegler, L., Biros, G., Ghattas, O., Heinkenschloss, M., Keyes, D., Mallick, B.,
Marzouk, Y., Tenorio, L., van Bloemen Waanders, B. and Willcox, K., John Wiley & Sons, Ltd, Chichester, UK,
https://doi.org/10.1002/9780470685853.ch10, 2010.

Williams, K. D., Senior C. A., and Mitchell, J. F. B.: Transient Climate Change in the Hadley Centre Models: The Role of Physical Processes, Journal of Climate, 2659-2674, https://doi.org/10.1175/1520-0442(2001)014<2659:TCCITH>2.0.CO;2, 2001.

Wilson, D. S., Jamieson, S. R., Barrett, P. J., Leitchenkov, G., Gohl, K. and Larter, D.: Antarctic topography at the Eocene-Oligocene boundary, Paleogeography, Paleoclimatology, Palaeoecology, 335-336, 24-34, https://doi.org/10.1016/j.palaeo.2011.05.028, 2012.

Zhang, Y. G., Pagani, M., Liu, Z., Bohaty, S. M. and DeConto, R.: A 40-million-year history of atmospheric CO_2, Phil Trans R Soc A, 371, 20130096, http://doi.org/10.1098/rsta.2013.0096, 2013.

710

715

720

725

730

735

Table 1: The mean percentage of grid boxes predicted within one and two standard deviations for the four different emulators calibrated with a different ice sheet parameter. The values in bold are closest to the theoretical 1 σ of 68,3 % and 2 σ of 95,5 %,.

	EMULATOR_8		EMULATOR_12a		EMULATOR_12b		EMULATOR_20	
	1 σ	2 σ	1 σ	2 σ	1 σ	2 σ	1σ	2σ
Ice volume	78.8	97.4	**71.2**	**94.5**	78.4	97.0	77.3	96.9
Ice area	77.9	97.2	78.7	97.2	80.2	97.4	78.7	97.3
Ice volume + ice area	77.8	99.8	74.9	**96.3**	76.0	96.7	78.4	98.1

Table A1: Experiments with their name, orbital parameter values, CO_2 values, and ice level expressed in terms of ice volume (VOL) and ice area (AR) for EMULATOR_8 (EM_8), EMULATOR_12a (EM_12a), EMULATOR_12b (EM_12b) and EMULATOR_20 (EM_20).

Nr	name	e	$\tilde{\omega}$	ε	CO_2	EM_8 VOL/AR $[10^7 \text{ km}^3/\text{km}^2]$	EM_12a VOL/AR $[10^7 \text{ km}^3/\text{km}^2]$	EM_12b VOL/AR $[10^7 \text{ km}^3/\text{km}^2]$	EM_20 VOL/AR $[10^7 \text{ km}^3/\text{km}^2]$
1	xaemaa	0.0492	194.7	22.14	552.1	3.29/13.26	2.35/10.98	2.35/10.98	3.29/13.26
2	xaemab	0.0523	237.0	23.36	709.1	3.29/13.26	2.35/10.98	2.35/10.98	3.29/13.26
3	xaemac	0.0090	198.1	23.68	1143.9	3.29/13.26	2.35/10.98	2.35/10.98	3.29/13.26
4	xaemad	0.0610	276.3	22.07	804.3	3.29/13.26	2.35/10.98	2.35/10.98	3.29/13.26
5	xaemae	0.0514	302.4	23.12	946.1	3.29/13.26	2.35/10.98	2.35/10.98	3.29/13.26
6	xaemaf	0.0122	353.1	22.53	1140.8	3.29/13.26	2.35/10.98	2.35/10.98	3.22/13.17
7	xaemag	0.0425	19.2	23.22	679.0	3.29/13.26	2.35/10.98	2.35/10.98	3.22/13.17
8	xaemah	0.0264	168.7	22.47	562.4	3.29/13.26	2.35/10.98	2.35/10.98	3.22/13.17
9	xaemai	0.0273	245.8	24.17	727.9	3.29/13.26	1.71/8.04	1.71/8.04	3.22/13.17
10	xaemaj	0.0306	306.4	22.97	660.3	3.29/13.26	1.71/8.04	1.71/8.04	3.22/13.17
11	xaemak	0.0467	302.0	22.73	640.5	3.29/13.26	1.71/8.04	1.71/8.04	3.07/12.86
12	xaemal	0.0438	355.6	23.86	622.8	3.29/13.26	1.71/8.04	1.71/8.04	3.07/12.86
13	xaemam	0.0589	12.1	22.85	788.9	2.42/12.54	1.71/8.04	1.71/8.04	3.07/12.86
14	xaeman	0.0396	105.9	23.91	665.0	2.42/12.54	1.71/8.04	1.71/8.04	3.07/12.86
15	xaemao	0.0422	176.6	24.42	720.1	2.42/12.54	1.71/8.04	1.71/8.04	3.07/12.86
16	xaemap	0.0248	148.3	23.02	965.5	2.42/12.54	1.71/8.04	1.71/8.04	2.87/12.66
17	xaemaq	0.0078	167.7	22.34	579.9	2.42/12.54	1.71/8.04	1.71/8.04	2.87/12.66
18	xaemar	0.0145	183.1	23.39	918.5	2.42/12.54	1.44/7.59	1.44/7.59	2.87/12.66
19	xaemas	0.0348	212.5	22.91	589.2	2.42/12.54	1.44/7.59	1.44/7.59	2.87/12.66
20	xaemat	0.0309	276.1	23.14	690.0	2.42/12.54	1.44/7.59	1.44/7.59	2.87/12.66

21	xaemau	0.0490	322.6	24.26	770.8	2.42/12.54	1.44/7.59	1.44/7.59	2.42/12.54
22	xaemav	0.0442	38.7	23.94	686.0	2.42/12.54	1.44/7.59	1.44/7.59	2.42/12.54
23	xaemaw	0.0199	0.2	22.10	879.2	2.42/12.54	1.44/7.59	1.44/7.59	2.42/12.54
24	xaemax	0.0564	136.4	24.47	824.5	2.42/12.54	1.44/7.59	1.44/7.59	2.42/12.54
25	xaemay	0.0557	149.6	22.25	912.8	2.42/12.54	1.44/7.59	1.44/7.59	2.42/12.54
26	xaemaz	0.0499	230.0	22.28	1012.5	1.85/11.70	1.29/6.48	1.20/6.21	2.15/12.08
27	xaemba	0.0272	269.2	23.26	727.5	1.85/11.70	1.29/6.48	1.20/6.21	2.15/12.08
28	xaembb	0.0627	85.3	24.28	635.1	1.85/11.70	1.29/6.48	1.20/6.21	2.15/12.08
29	xaembc	0.0134	25.5	22.16	567.3	1.85/11.70	1.29/6.48	1.20/6.21	2.15/12.08
30	xaembd	0.0552	88.7	22.78	699.7	1.85/11.70	1.29/6.48	1.20/6.21	2.15/12.08
31	xaembe	0.0467	90.4	24.08	1083.2	1.85/11.70	1.29/6.48	1.20/6.21	1.85/11.70
32	xaembf	0.0431	128.6	23.33	1132.4	1.85/11.70	1.29/6.48	1.20/6.21	1.85/11.70
33	xaembg	0.0220	153.6	22.48	1039.7	1.85/11.70	1.29/6.48	1.20/6.21	1.85/11.70
34	xaembh	0.0161	247.9	22.18	1036.1	1.85/11.70	1.04/6.19	1.08/5.74	1.85/11.70
35	xaembi	0.0165	270.7	23.32	888.9	1.85/11.70	1.04/6.19	1.08/5.74	1.85/11.70
36	xaembj	0.0281	78.9	23.97	1070.3	1.85/11.70	1.04/6.19	1.08/5.74	1.59/10.21
37	xaembk	0.0106	48.8	23.30	740.6	1.85/11.70	1.04/6.19	1.08/5.74	1.59/10.21
38	xaembl	0.0242	80.5	22.32	1025.3	1.33/9.18	1.04/6.19	1.08/5.74	1.59/10.21
39	xaembm	0.0543	157.1	22.39	751.0	1.33/9.18	1.04/6.19	1.08/5.74	1.59/10.21
40	xaembn	0.0267	167.0	24.48	748.6	1.33/9.18	1.04/6.19	1.08/5.74	1.59/10.21
41	xaembo	0.0230	166.8	23.63	772.9	1.33/9.18	1.04/6.19	1.08/5.74	1.33/9.18
42	xaembp	0.0026	210.6	23.06	628.6	1.33/9.18	0.90/5.68	0.90/5.68	1.33/9.18
43	xaembq	0.0443	184.2	22.28	1103.3	1.33/9.18	0.90/5.68	0.90/5.68	1.33/9.18
44	xaembr	0.0396	198.5	23.61	841.9	1.33/9.18	0.90/5.68	0.90/5.68	1.33/9.18
45	xaembs	0.0523	254.0	23.23	673.6	1.33/9.18	0.90/5.68	0.90/5.68	1.33/9.18
46	xaembt	0.0610	271.5	24.04	646.2	1.33/9.18	0.90/5.68	0.90/5.68	1.13/8.12
47	xaembu	0.0428	282.6	23.72	574.0	1.33/9.18	0.90/5.68	0.90/5.68	1.13/8.12
48	xaembv	0.0445	0.3	22.01	1005.6	1.33/9.18	0.90/5.68	0.90/5.68	1.13/8.12
49	xaembw	0.0398	3.3	24.06	782.6	1.33/9.18	0.90/5.68	0.90/5.68	1.13/8.12
50	xaembx	0.0561	93.7	22.21	1112.0	0.77/7.05	0.90/5.68	0.90/5.68	1.13/8.12
51	xaemby	0.0310	304.5	22.89	594.5	0.77/7.05	0.72/5.17	0.72/5.17	0.96/7.45
52	xaembz	0.0479	342.0	23.48	1121.1	0.77/7.05	0.72/5.17	0.72/5.17	0.96/7.45
53	xaemca	0.0270	359.6	23.80	556.1	0.77/7.05	0.72/5.17	0.72/5.17	0.96/7.45
54	xaemcb	0.0163	12.8	22.94	736.9	0.77/7.05	0.72/5.17	0.72/5.17	0.96/7.45
55	xaemcc	0.0253	42.1	23.55	812.4	0.77/7.05	0.72/5.17	0.72/5.17	0.96/7.45
56	xaemcd	0.0346	39.8	24.21	884.2	0.77/7.05	0.72/5.17	0.72/5.17	0.77/7.05
57	xaemce	0.0579	120.1	23.88	982.1	0.77/7.05	0.72/5.17	0.72/5.17	0.77/7.05
58	xaemcf	0.0591	161.4	24.33	1097.2	0.77/7.05	0.72/5.17	0.72/5.17	0.77/7.05
59	xaemcg	0.0481	184.6	23.67	1050.7	0.77/7.05	0.68/4.29	0.68/4.29	0.77/7.05
60	xaemch	0.0476	184.1	24.12	956.1	0.77/7.05	0.68/4.29	0.68/4.29	0.77/7.05

61	xaemci	0.0237	210.0	24.32	1087.7	0.77/7.05	0.68/4.29	0.68/4.29	0.57/5.85
62	xaemcj	0.0524	259.7	23.55	854.8	0.77/7.05	0.68/4.29	0.68/4.29	0.57/5.85
63	xaemck	0.0107	294.2	22.76	813.3	0.50/5.14	0.68/4.29	0.68/4.29	0.57/5.85
64	xaemcl	0.0389	0.9	22.68	614.4	0.50/5.14	0.68/4.29	0.68/4.29	0.57/5.85
65	xaemcm	0.0289	19.3	22.52	801.0	0.50/5.14	0.68/4.29	0.68/4.29	0.57/5.85
66	xaemcn	0.0279	96.1	22.71	604.7	0.50/5.14	0.68/4.29	0.68/4.29	0.50/5.14
67	xaemco	0.0407	95.4	22.03	650.0	0.50/5.14	0.68/4.29	0.68/4.29	0.50/5.14
68	xaemcp	0.0226	277.6	22.57	666.3	0.50/5.14	0.37/2.85	0.58/4.26	0.50/5.14
69	xaemcq	0.0311	348.4	24.02	874.0	0.50/5.14	0.37/2.85	0.58/4.26	0.50/5.14
70	xaemcr	0.0341	325.5	22.62	625.0	0.50/5.14	0.37/2.85	0.58/4.26	0.50/5.14
71	xaemcs	0.0133	305.8	23.98	904.8	0.50/5.14	0.37/2.85	0.58/4.26	0.44/4.58
72	xaemct	0.0300	22.0	22.82	959.2	0.50/5.14	0.37/2.85	0.58/4.26	0.44/4.58
73	xaemcu	0.0113	38.9	22.66	758.2	0.50/5.14	0.37/2.85	0.58/4.26	0.44/4.58
74	xaemcv	0.0495	81.6	24.44	830.4	0.50/5.14	0.37/2.85	0.58/4.26	0.44/4.58
75	xaemcw	0.0304	175.5	23.18	587.3	0.50/5.14	0.37/2.85	0.58/4.26	0.44/4.58
76	xaemcx	0.0411	216.0	23.08	575.0	0.24/3.05	0.20/2.01	0.31/2.53	0.33/3.71
77	xaemcy	0.0412	191.5	22.45	925.4	0.24/3.05	0.20/2.01	0.31/2.53	0.33/3.71
78	xaemcz	0.0374	215.0	24.19	597.9	0.24/3.05	0.20/2.01	0.31/2.53	0.33/3.71
79	xaemda	0.0150	270.7	22.36	856.4	0.24/3.05	0.20/2.01	0.31/2.53	0.33/3.71
80	xaemdb	0.0343	355.2	23.51	1054.3	0.24/3.05	0.20/2.01	0.31/2.53	0.33/3.71
81	xaemdc	0.0303	357.7	23.45	896.9	0.24/3.05	0.20/2.01	0.31/2.53	0.24/3.05
82	xaemdd	0.0091	311.1	23.16	716.8	0.24/3.05	0.20/2.01	0.31/2.53	0.24/3.05
83	xaemde	0.0109	80.9	23.73	993.9	0.24/3.05	0.20/2.01	0.31/2.53	0.24/3.05
84	xaemdf	0.0075	243.9	22.63	611.0	0.24/3.05	0.08/1.00	0.12/2.53	0.24/3.05
85	xaemdg	0.0609	267.5	24.23	844.0	0.24/3.05	0.08/1.00	0.12/1.45	0.24/3.05
86	xaemdh	0.0372	274.3	22.99	652.2	0.24/3.05	0.08/1.00	0.12/1.45	0.18/2.34
87	xaemdi	0.0214	314.1	22.86	1061.7	0.24/3.05	0.08/1.00	0.12/1.45	0.18/2.34
88	xaemdj	0.0432	76.8	24.14	868.4	0.24/3.05	0.08/1.00	0.12/1.45	0.18/2.34
89	xaemdk	0.0321	92.0	23.83	706.6	0.02/0.66	0.08/1.00	0.12/1.45	0.18/2.34
90	xaemdl	0.0389	109.5	23.03	691.8	0.02/0.66	0.08/1.00	0.12/1.45	0.18/2.34
91	xaemdm	0.0167	175.2	23.77	832.0	0.02/0.66	0.08/1.00	0.12/1.45	0.10/1.45
92	xaemdn	0.0406	103.9	22.42	791.0	0.02/0.66	0.08/1.00	0.12/1.45	0.10/1.45
93	xaemdo	0.0420	229.6	23.59	942.0	0.02/0.66	0.01/0.18	0.01/0.18	0.10/1.45
94	xaemdp	0.0335	250.9	23.46	988.7	0.02/0.66	0.01/0.18	0.01/0.18	0.10/1.45
95	xaemdq	0.0290	256.6	24.38	971.6	0.02/0.66	0.01/0.18	0.01/0.18	0.10/1.45
96	xaemdr	0.0080	76.4	24.37	765.1	0.02/0.66	0.01/0.18	0.01/0.18	0.02/0.66
97	xaemds	0.0183	53.5	22.60	607.5	0.02/0.66	0.01/0.18	0.01/0.18	0.02/0.66
98	xaemdt	0.0300	83.4	22.11	929.8	0.02/0.66	0.01/0.18	0.01/0.18	0.02/0.66
99	xaemdu	0.0416	96.3	23.80	1021.3	0.02/0.66	0.01/0.18	0.01/0.18	0.02/0.66
100	xaemdv	0.0073	126.9	23.41	562.2	0.02/0.66	0.01/0.18	0.01/0.18	0.02/0.66

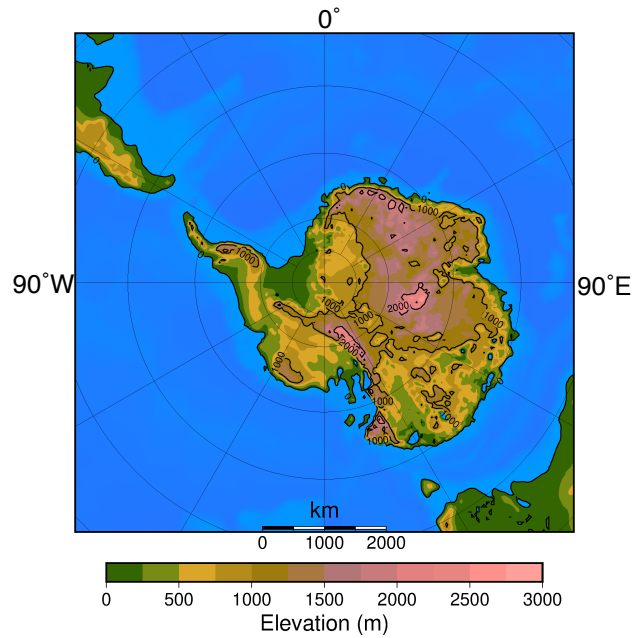



Figure 1: Antarctic bedrock topography after Wilson et al. (2012) as used in the simulations. Latitudes are given each 10˚ and longitudes each 30˚. Note the different paleogeographic position of the continents from today.

750

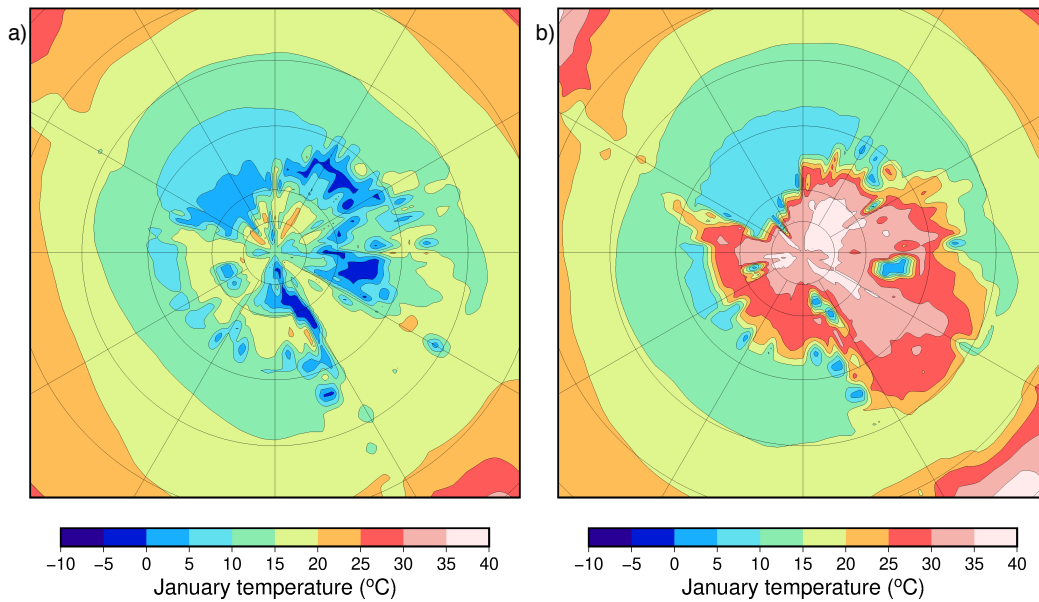



Figure 2: Simulated January mean surface air temperature (˚C) for (a) a cold orbital configuration and (b) a warm orbital configuration for a 3 x CO₂ scenario (840 ppmv).

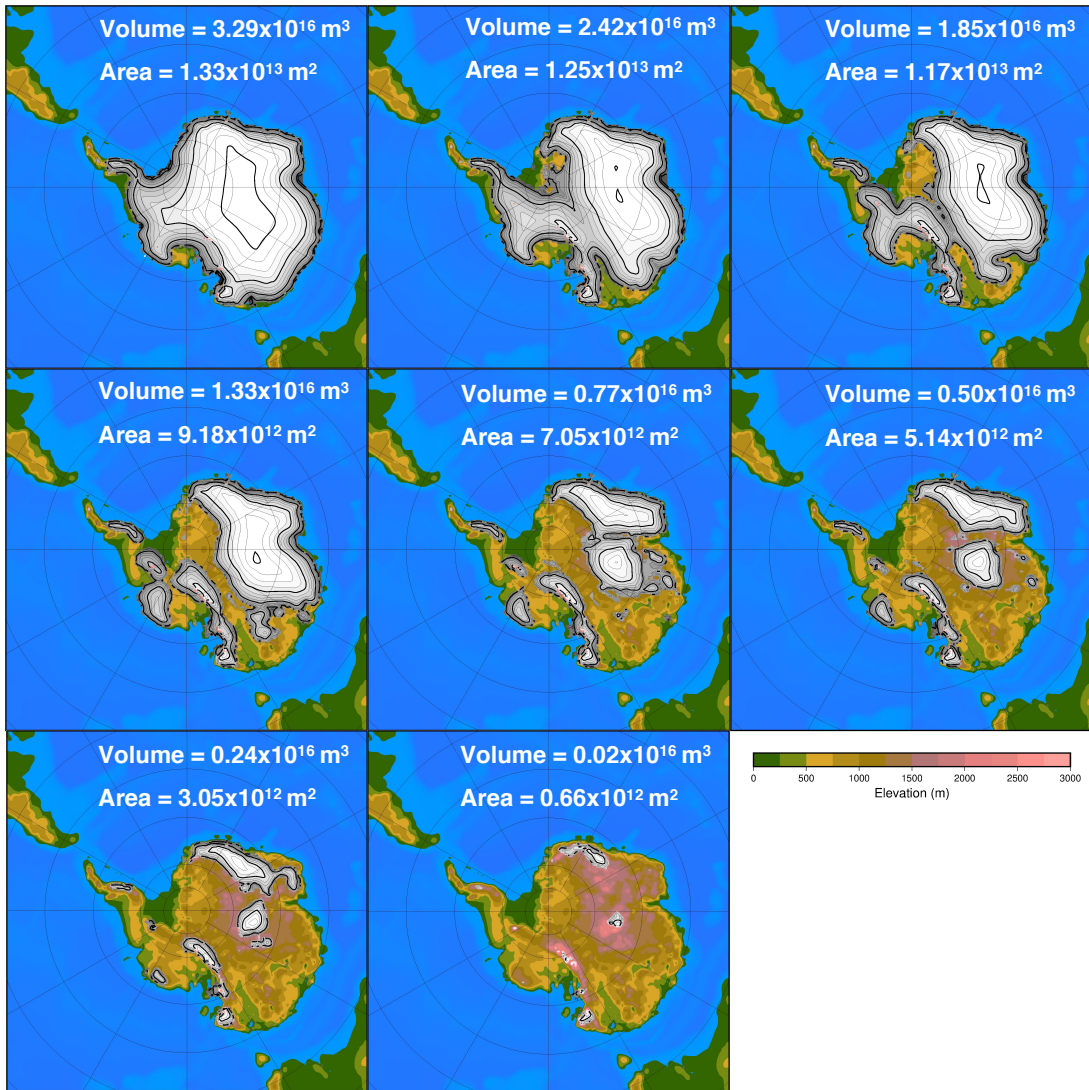



Figure 3: Eight different ice sheet geometries with their respective ice sheet volume and ice sheet area as input to EMULATOR_8. Ice sheet contour lines are given each 250 m and thick contour lines each 1000 m.

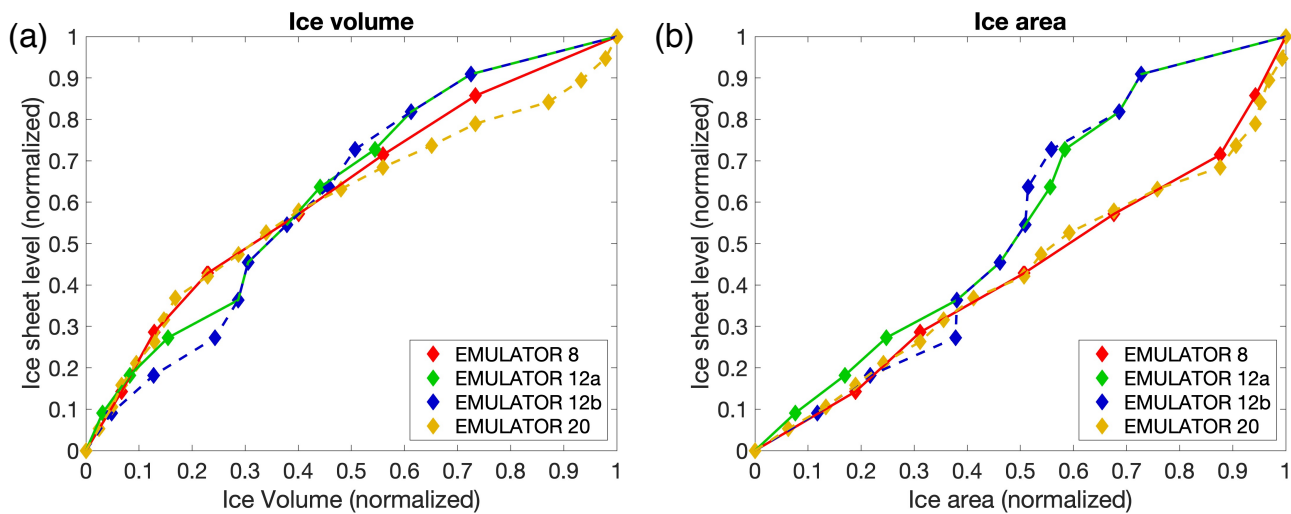



Figure 4: Spread of the ice sheet parameter defined by ice volume and ice area for the four different emulators. Note that EMULATOR_8 has only 8 predefined ice sheet geometries, while EMULATOR_12a and EMULATOR_12b have 12 predefined ice sheet geometries and EMULATOR_20 has 20.

760

765

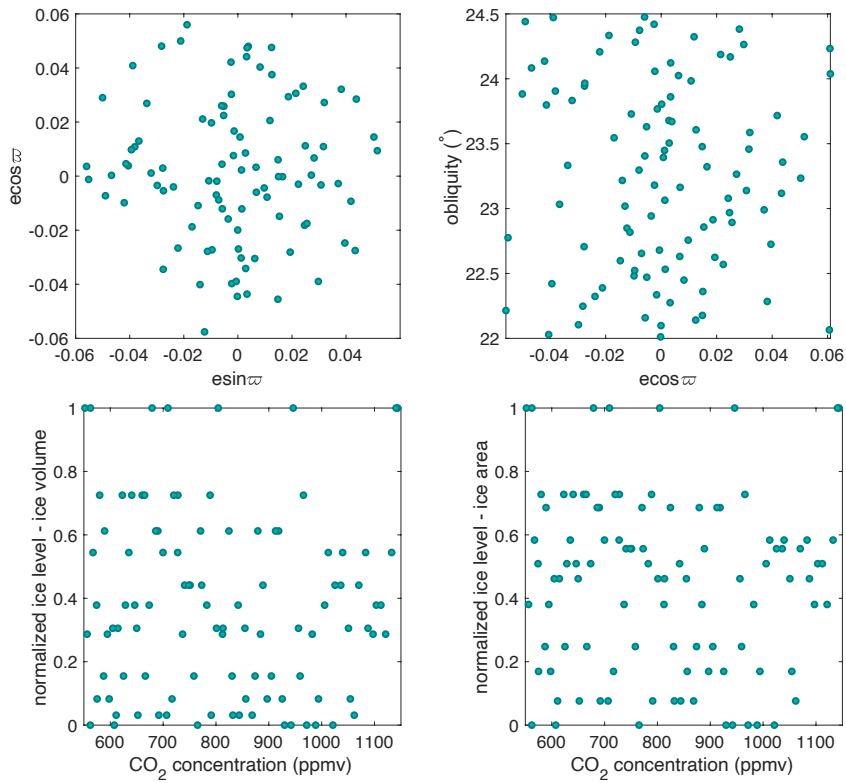
770

775

Figure 5: Latin hypercube model design of EMULATOR_12a showing the values for the orbital forcing, carbon dioxide forcing and ice sheet parameter forcing (defined by ice volume). Note that each dot represents one experiment from a total of 100 experiments.

780

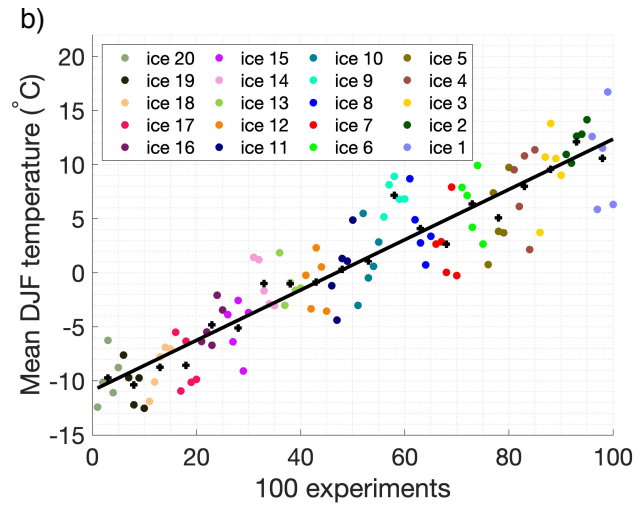
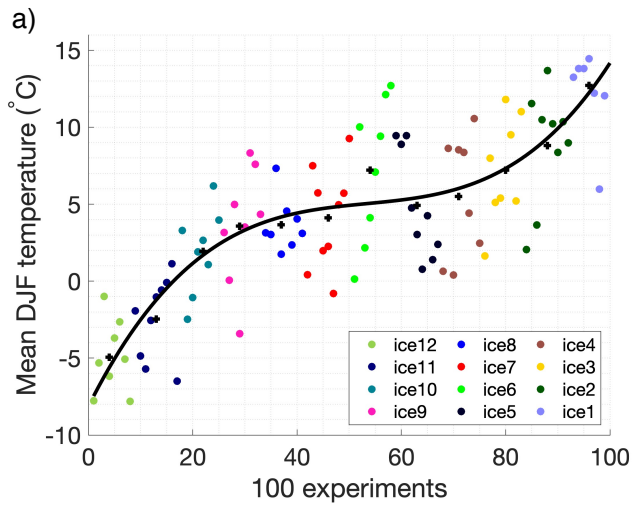



Figure 6: Mean austral summer temperature (December, January, February) for (a) EMULATOR_12a and (b) EMULATOR_20. Each dot represents the output from one GCM run. The output is grouped for each input ice sheet geometry. The mean austral summer for each input ice sheet geometry is given by the black cross and the best fit is given by the black line.

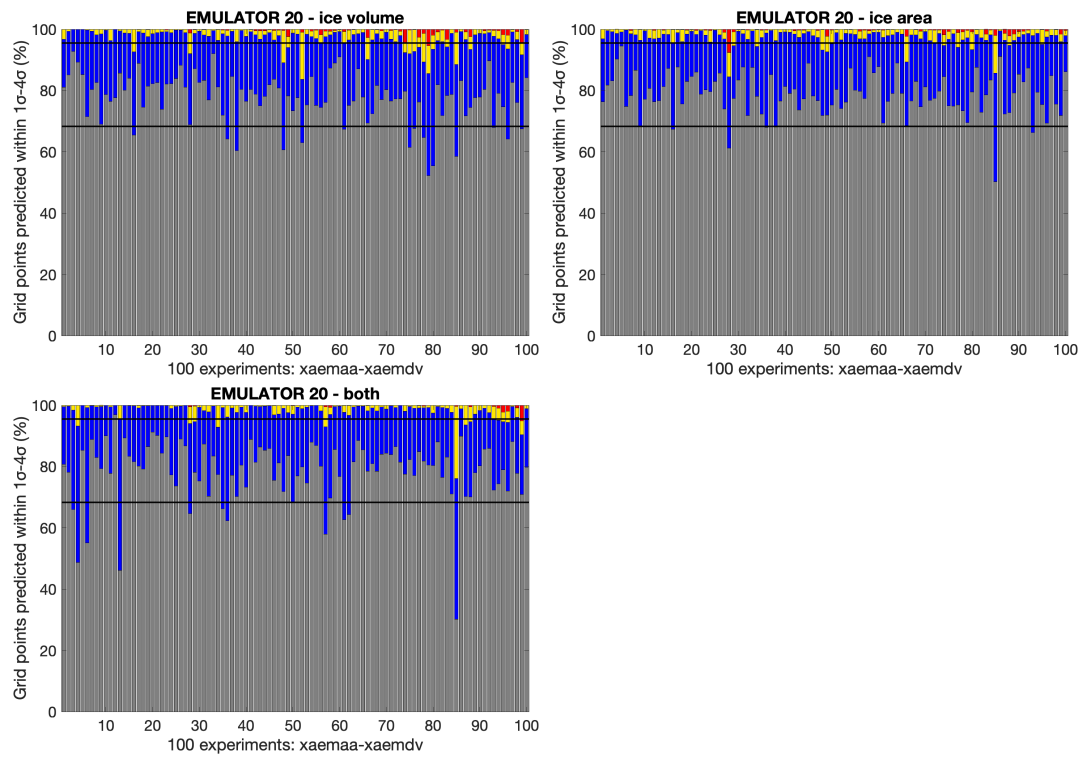



Figure 7: Calibration of EMULATOR_20 for the ice sheet parameter defined by the ice volume (a), by the ice area (b) and by both ice volume and ice area (c). The bars indicate the percentage of grid points where the emulator predicts the January temperature above the Antarctic continent within 1, 2, 3, or 4 standard deviations for each of the 100 experiments.

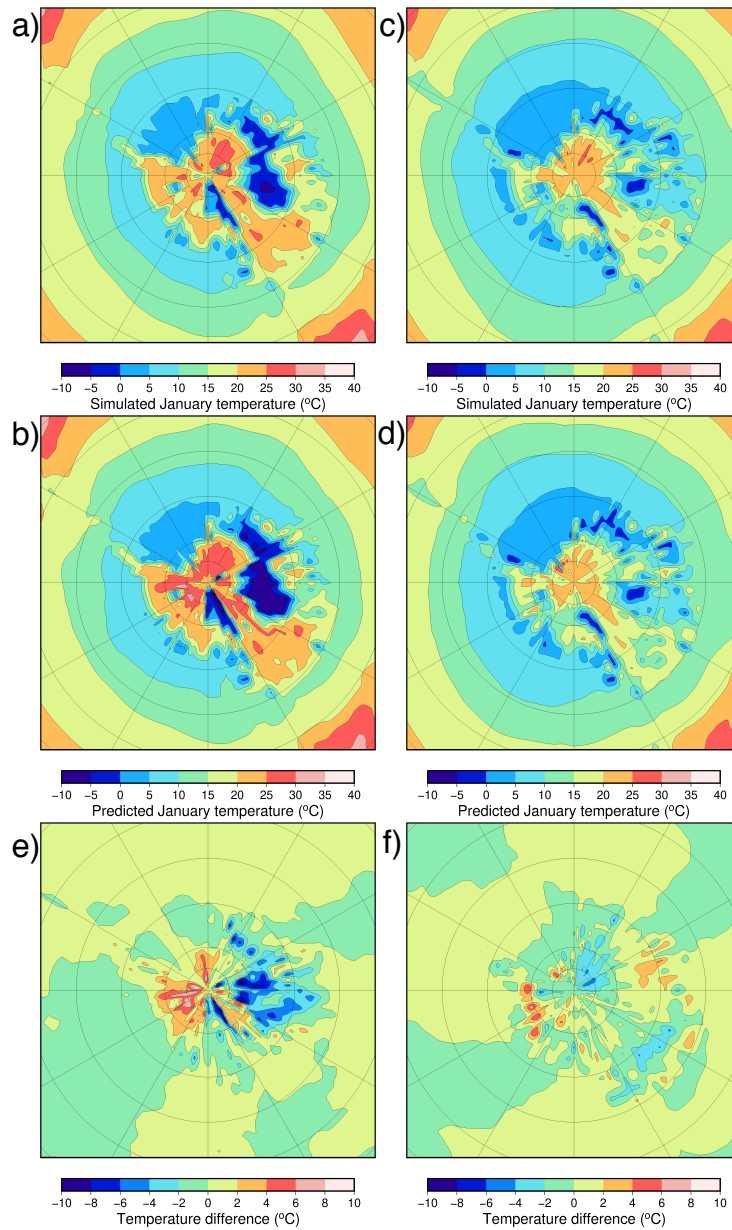



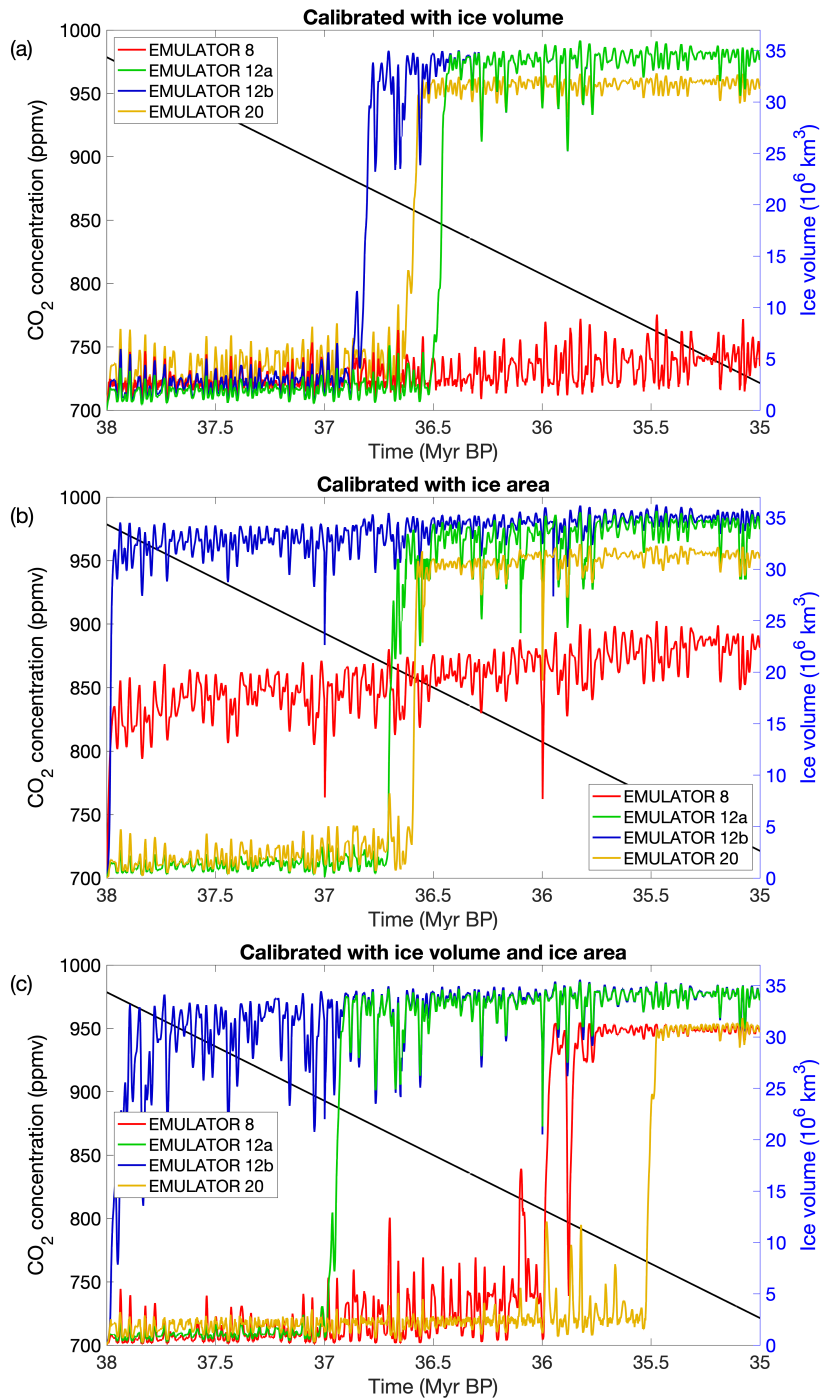
Figure 8: (a) Simulated and (b) predicted (with a leave-one-out experiment) January temperatures for the emulator showing a very poor- performing experiment (*xaemdk*) using EMULATOR_12b. (c) Simulated and (d) predicted January temperatures for a well-performing experiment (*xaemdv*) using EMULATOR_12b. Difference between the simulated and predicted temperature fields for (e) experiment *xaemdk* and (f) experiment *xaemdv*.

805

Figure 9: Ice sheet evolution during a 3 million year period forced by the orbital parameters from Laskar et al. (2004) and linearly declining CO₂ concentrations from ~980 ppm to ~720 ppm. Ice sheet evolution for the four different emulators calibrated based on (a) ice volume, (b) ice area and (c) both ice volume and ice area.

	EMULATOR_8	EMULATOR_12a	EMULATOR_12b	EMULATOR_20

a) Warm orbit / $CO_2 \sim 970$ ppmv

January temperature (°C): −10 −5 0 5 10 15 20 25 30 35 40

b) Cold orbit / $CO_2 \sim 970$ ppmv

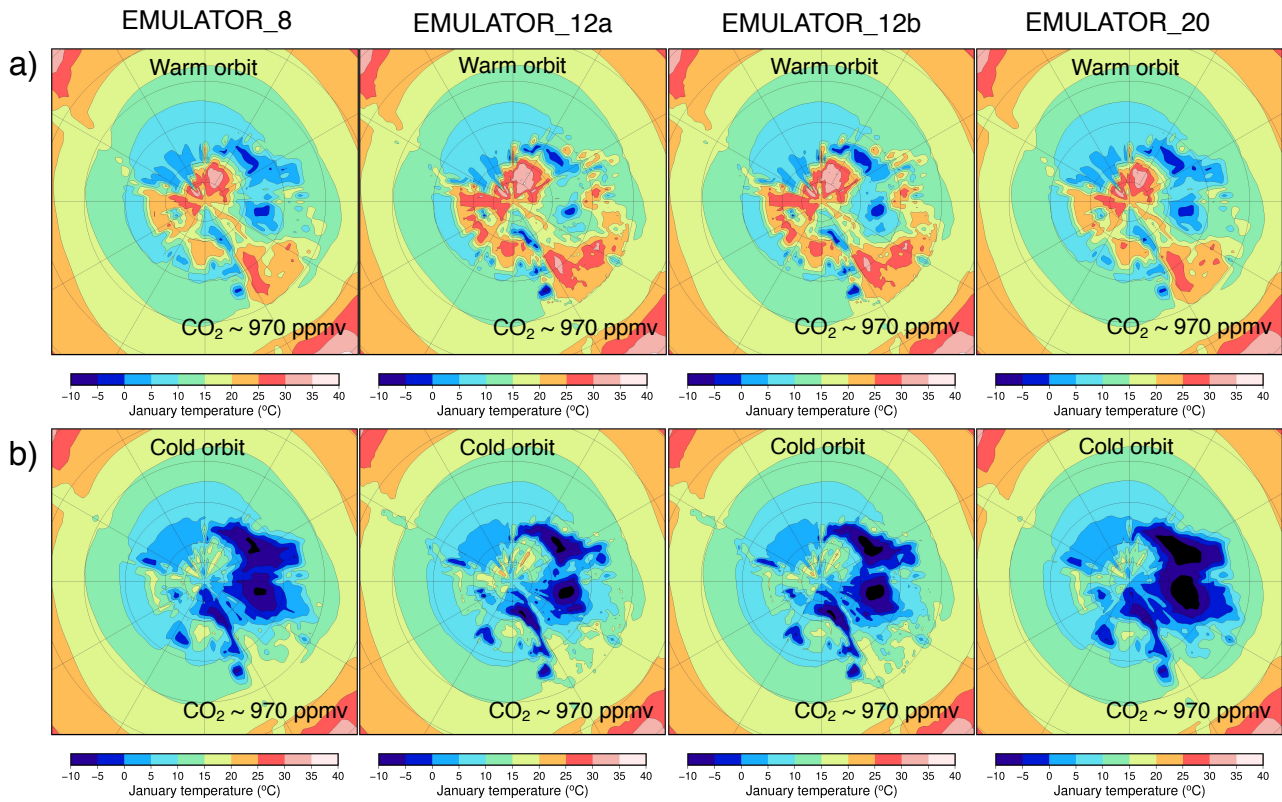
January temperature (°C): −10 −5 0 5 10 15 20 25 30 35 40

810

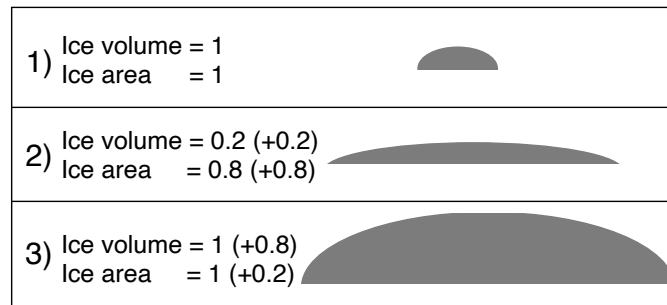
Figure 10: Simulated January temperature (˚C) with the different emulators starting from an ice-free continent at 38 Ma after (a) 48,000 years (the first high insolation maximum) and after (b) 60,000 years (the first high insolation minimum) for the ice sheet parameter defined as ice volume.

815

Figure 11: Three schematic ice sheet geometries with corresponding ice sheet volume and ice sheet area (normalized units).

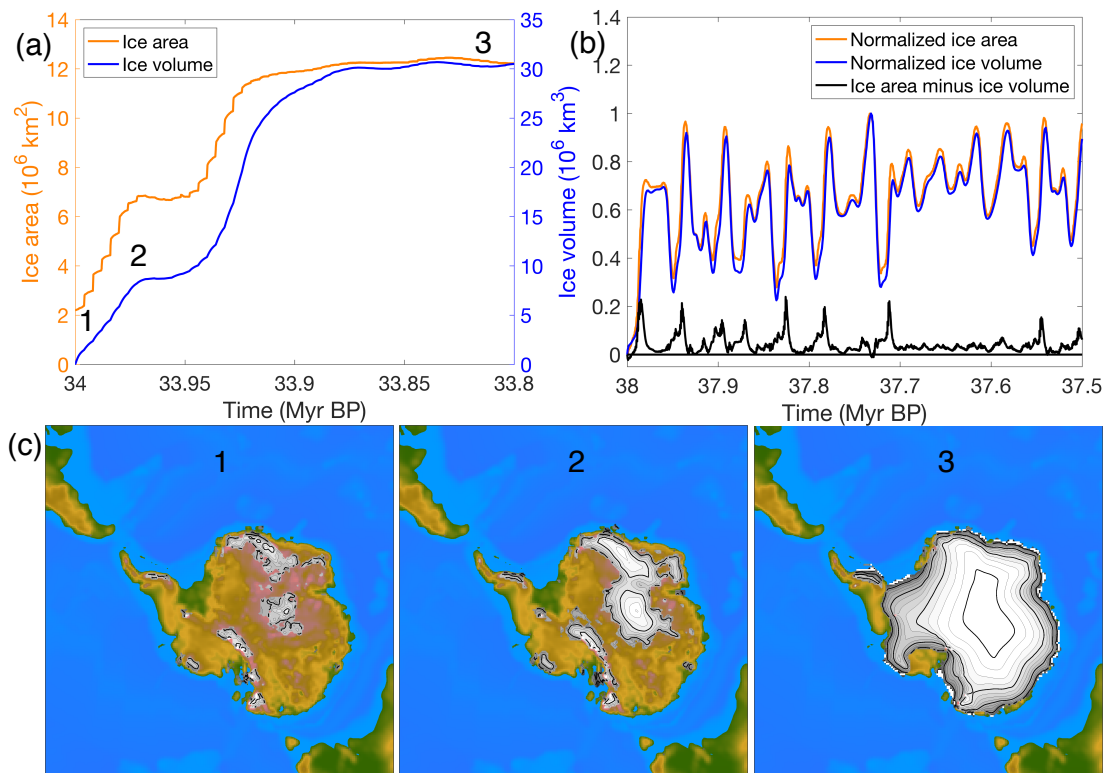



Figure 12: (a) Ice sheet area and ice sheet volume for an ice sheet that grows rapidly to a continental scale ice sheet. **(b)** Normalized ice area , ice volume and the difference between both for an ice sheet that grows and melts in response to the orbital forcing. **(c)** Ice sheet geometry during three snapshots for the run when the ice sheet grows to a continental scale. The numbers in (a) and (c) indicate the time at which the snapshots are taken.

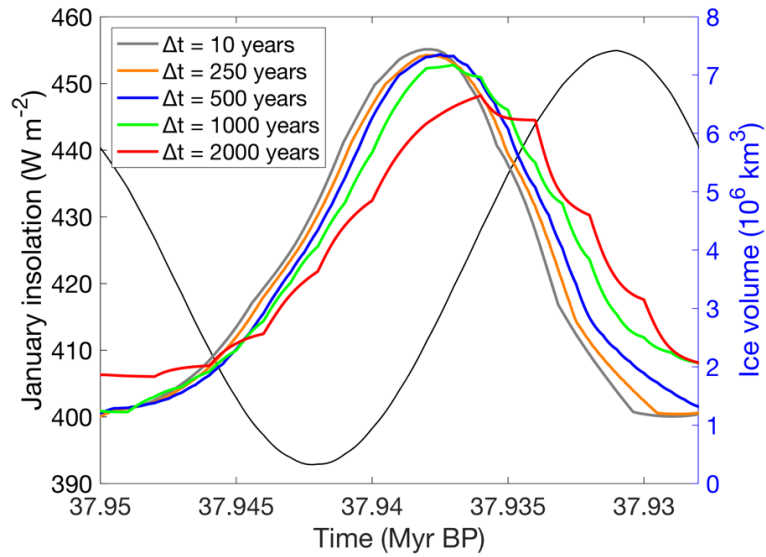



Figure 13: Illustration of the ice sheet volume evolution for different coupling timesteps during a precession cycle (~23,000 years) at the beginning of the simulations. The mean January insolation is given by the thin black line.

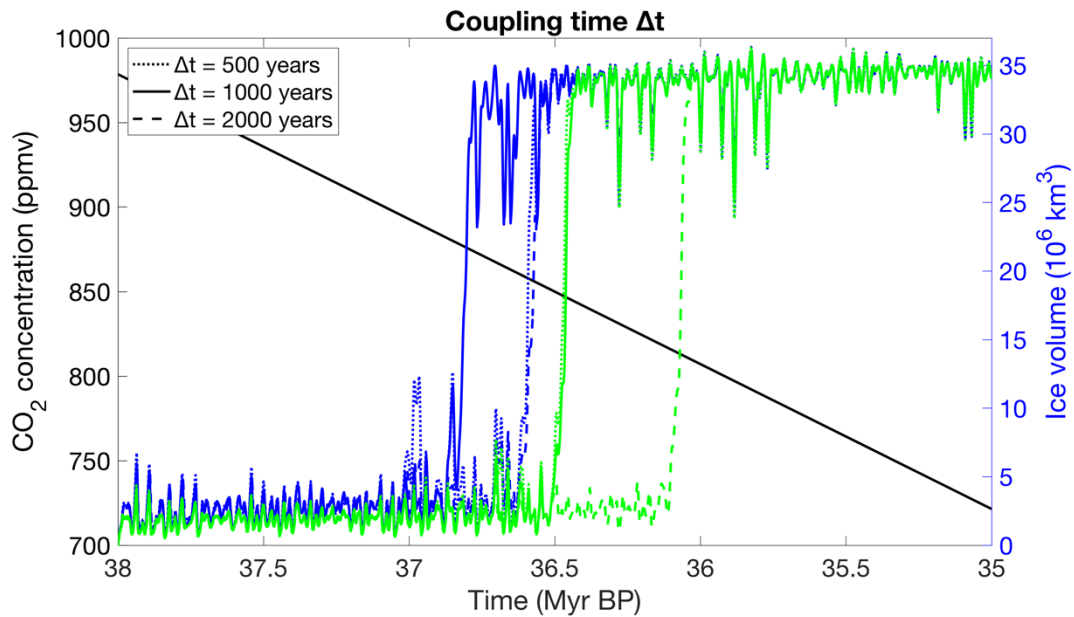



Figure 14: Sensitivity of the ice sheet evolution to the coupling time using the emulator calibrated with ice volume for EMULATOR_12a (green) and EMULATOR_12b (blue).

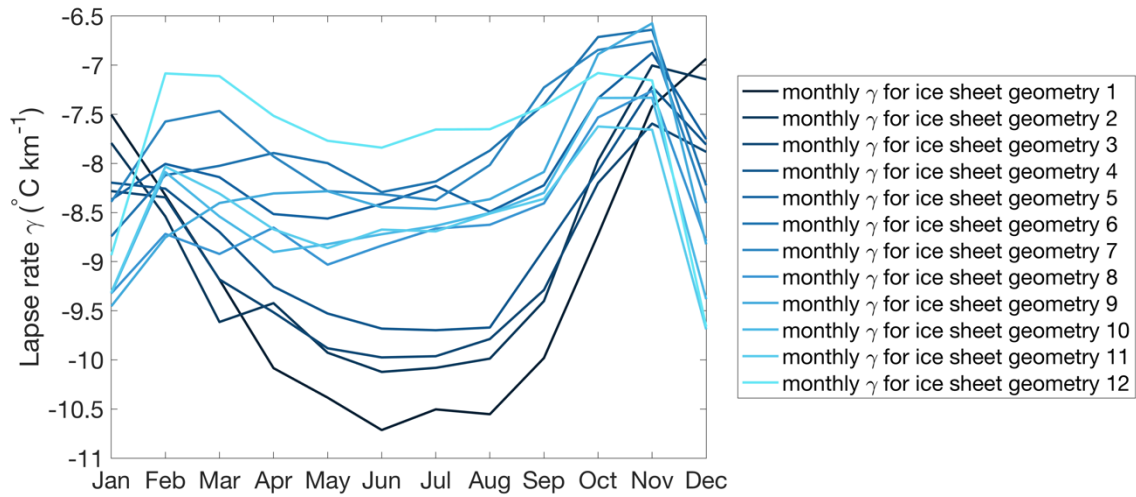



Figure 15: Average monthly near-surface lapse rate in HadSM3 over the Antarctic continent for each of the 12 ice sheet geometries. Ice sheet geometry 1 is the largest ice sheet and ice sheet geometry 12 is the smallest ice sheet geometry.

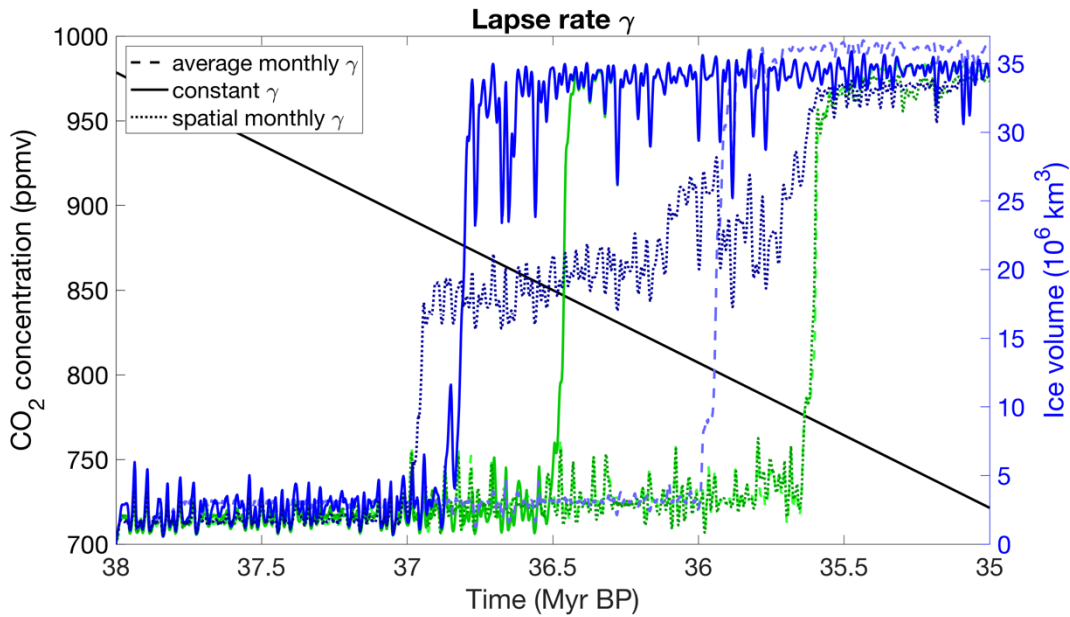
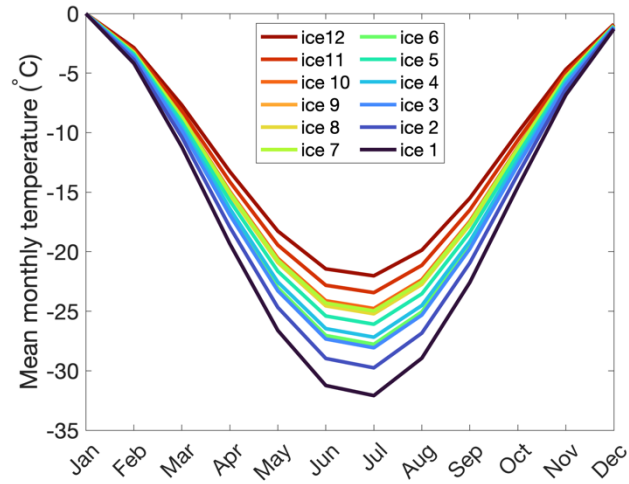



Figure 16: Sensitivity of the ice sheet evolution to the application of the lapse rate correction using the emulator calibrated with ice volume for EMULATOR_12a (green) and EMULATOR 12b (blue).

845 **Figure 17: (a) Mean monthly temperature for each of the 12 input ice sheet geometries over the Antarctic continent. (b) Annual cycle for each of the 12 input ice sheet geometries with respect to January temperatures.**

850

855

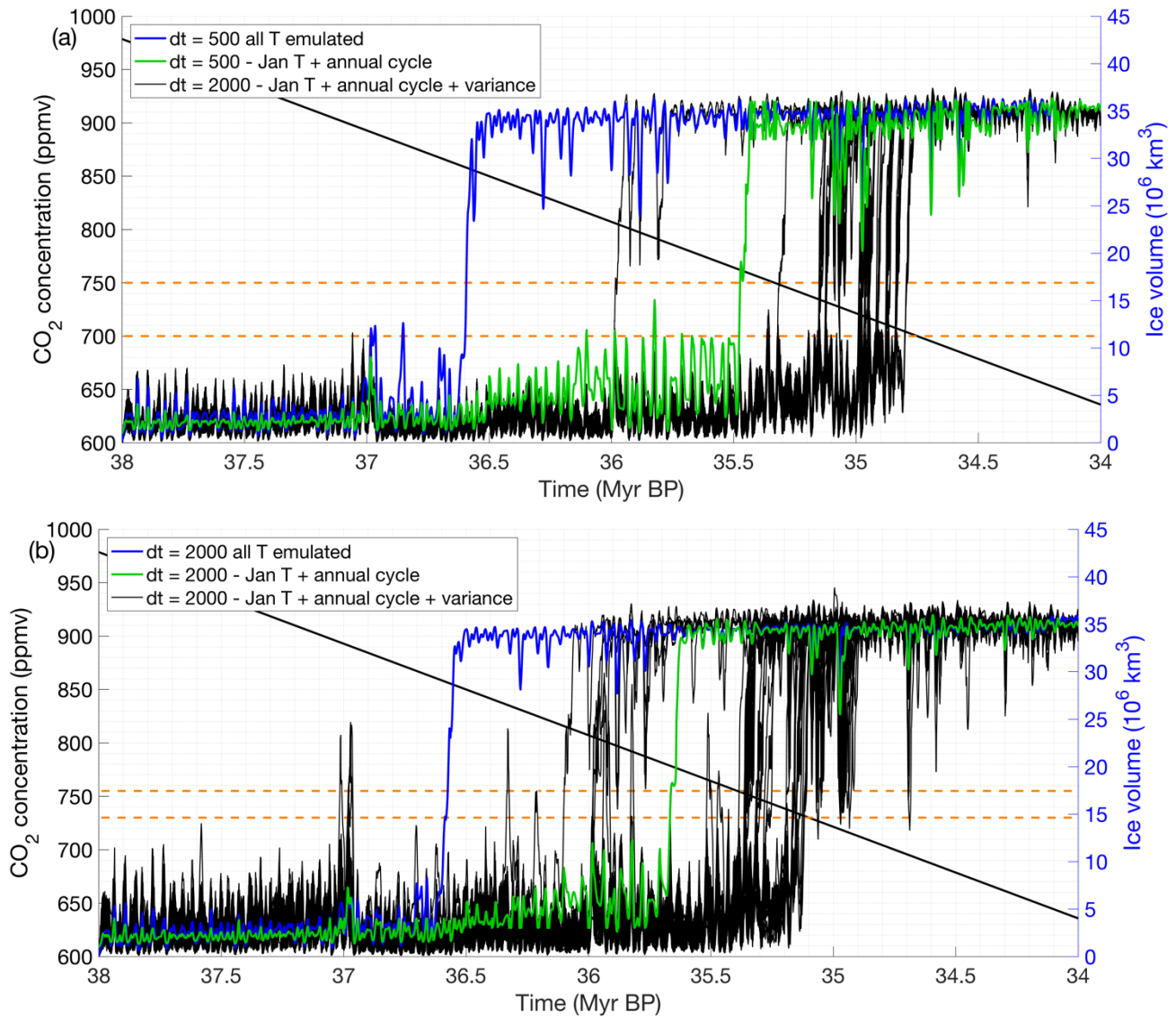



Figure 18: Monte Carlo simulations showing the ice sheet evolution for a coupling step of (a) 500 years and (b) 2000 years. The simulation ran with the emulated January temperatures and the annual cycle is shown by the green line and the original run is shown by the blue line. Horizontal orange dashed lines indicate the interval at which most experiments reach the glaciation threshold.

860

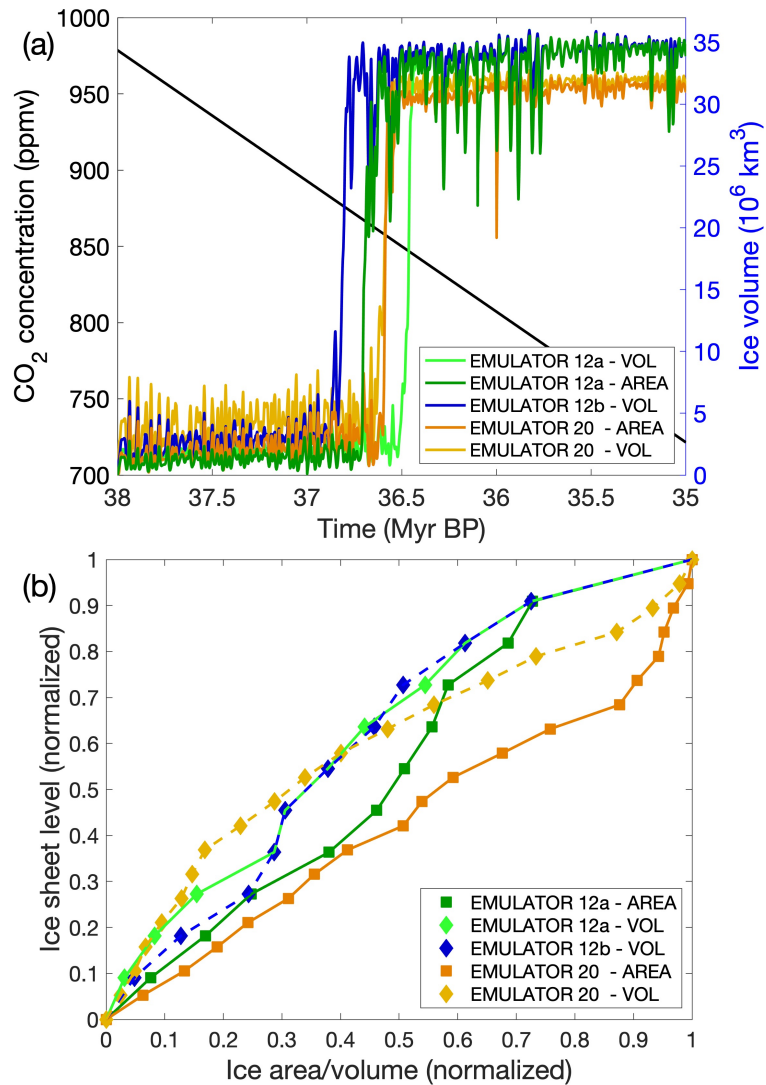



Figure 19: (a) Ice sheet evolution for EMULATOR_~~100a~~ 12a tuned on ice area, EMULATOR_~~100a~~ 12a tuned on ice volume and EMULATOR_ -12~~00~~b tuned on ice volume. (b) Comparison of the input ice sheet geometry spacing when looking at ice area and ice volume for EMULATOR_~~100a~~ 12a and EMULATOR_ ~~100b~~12b.

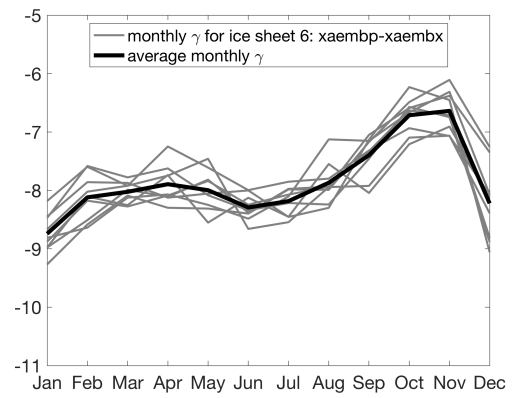
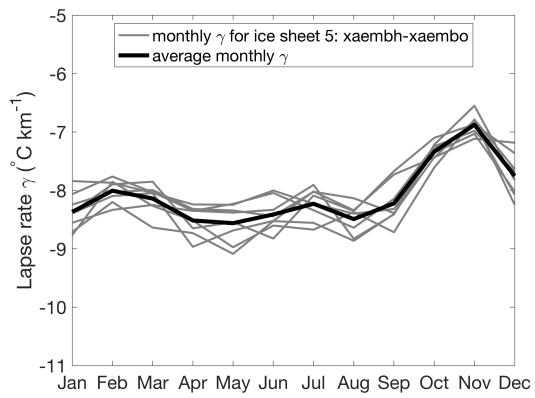
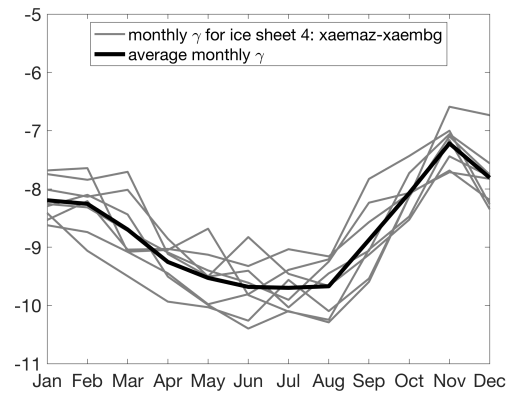
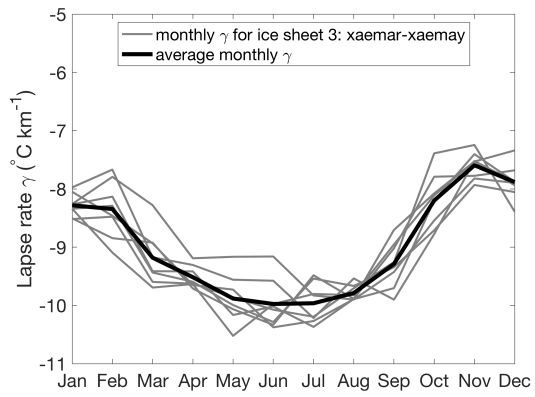
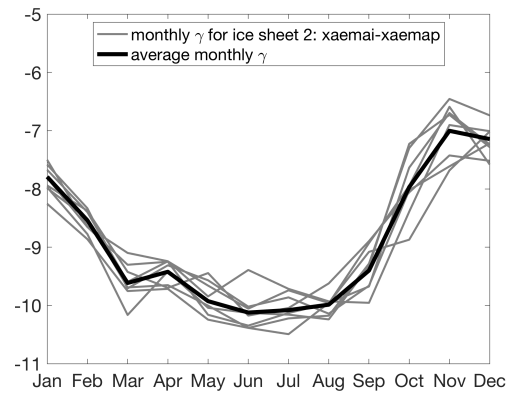
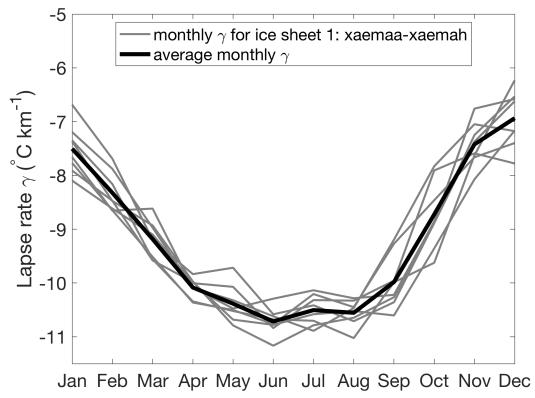
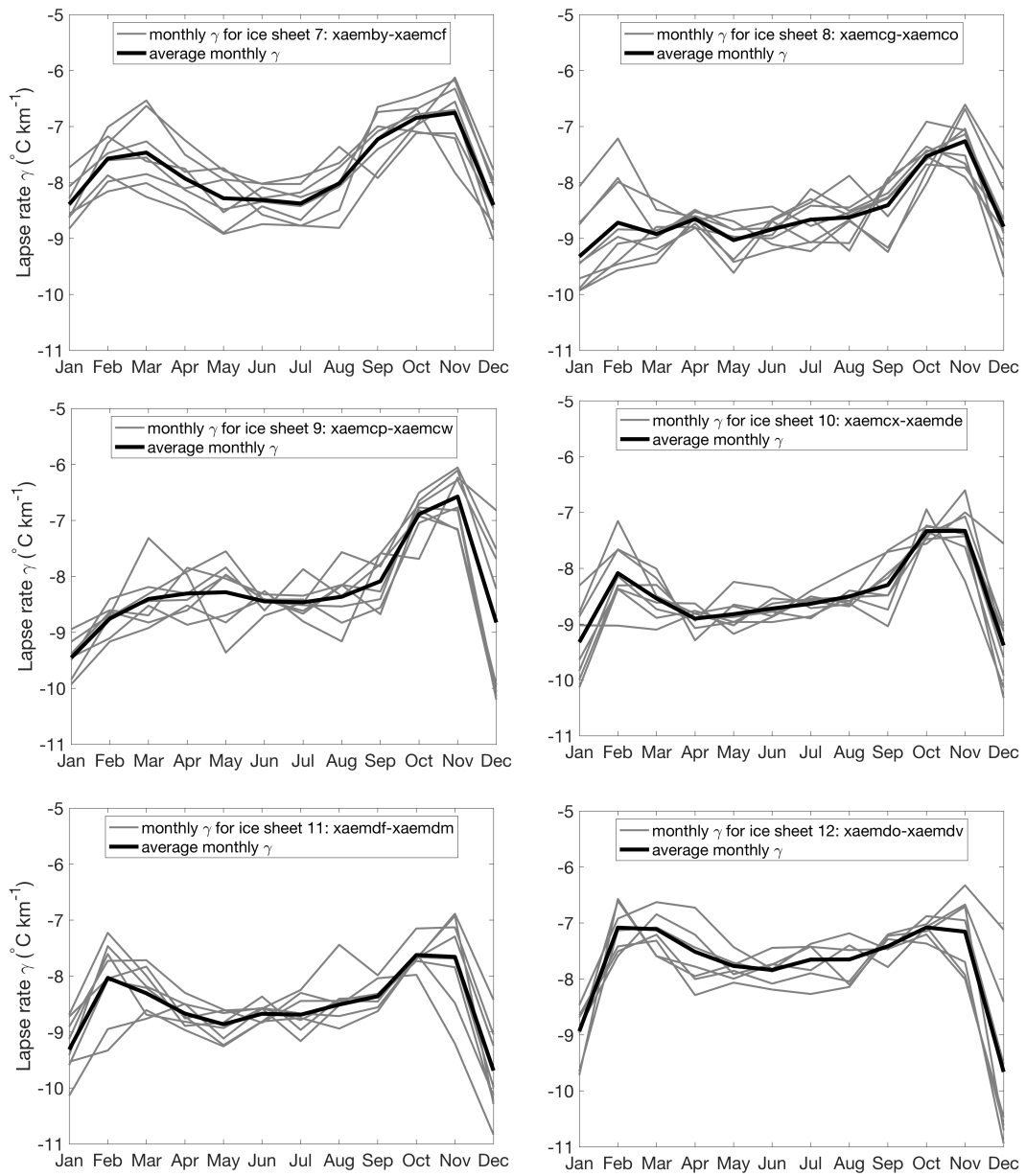



Figure A1: Monthly average lapse rate for the twelve different ice sheet geometries calculated from the 100 GCM runs for EMULATOR_~~100b~~12b.

875

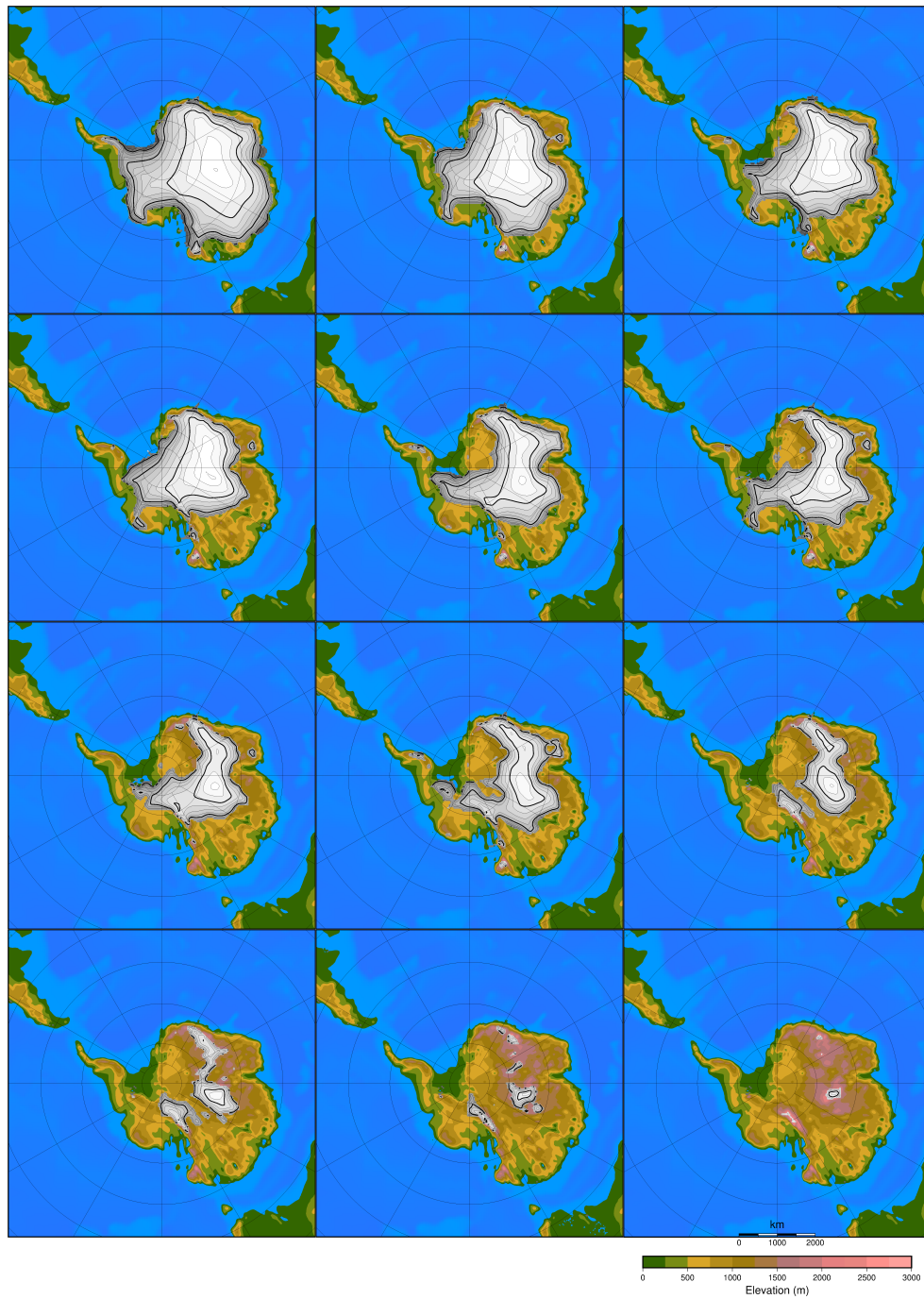
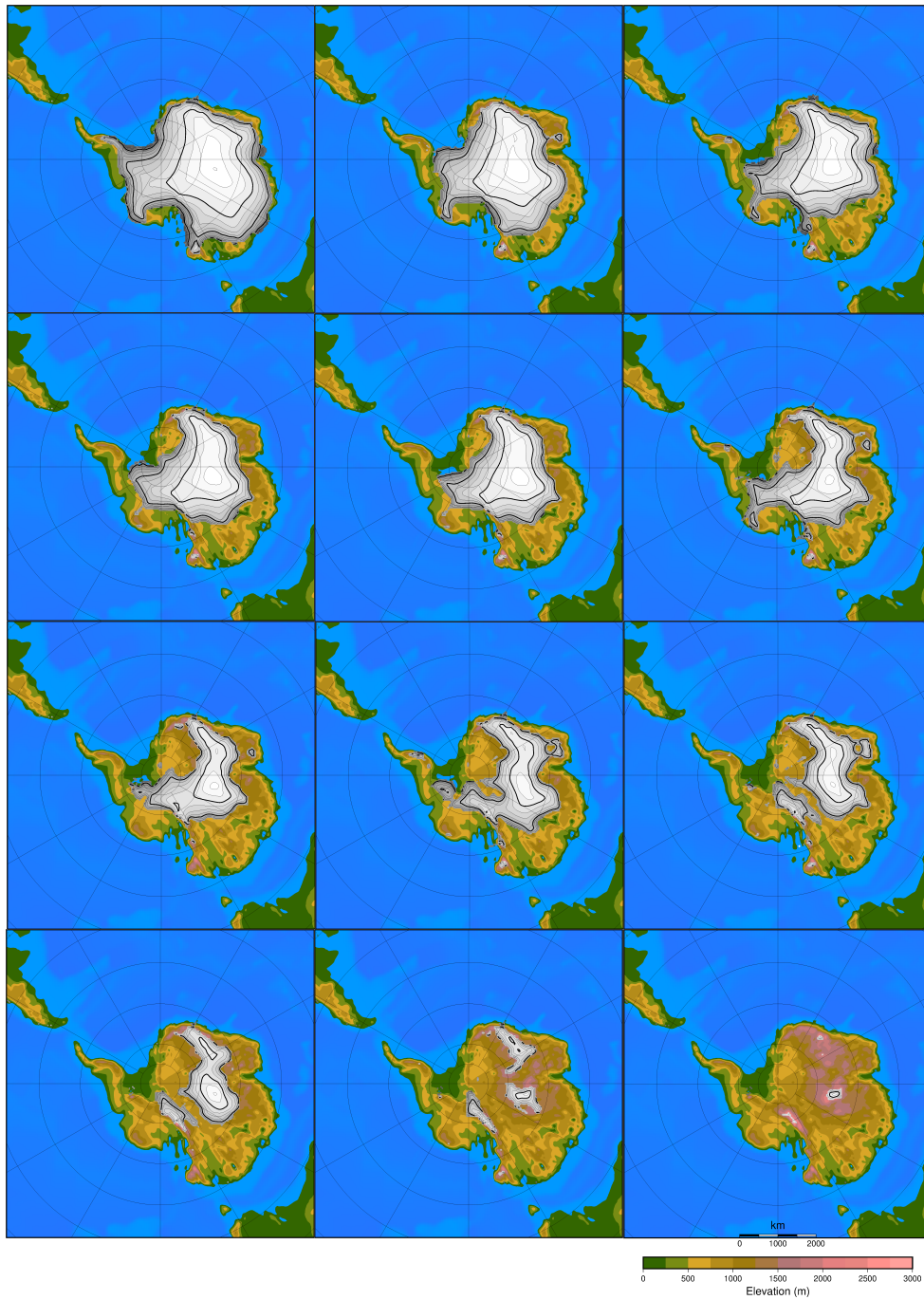



Figure A2: 12 ice sheet geometries as input to EMULATOR_12a.

880 **Figure A3: 12 ice sheet geometries as input to EMULATOR_12b.**

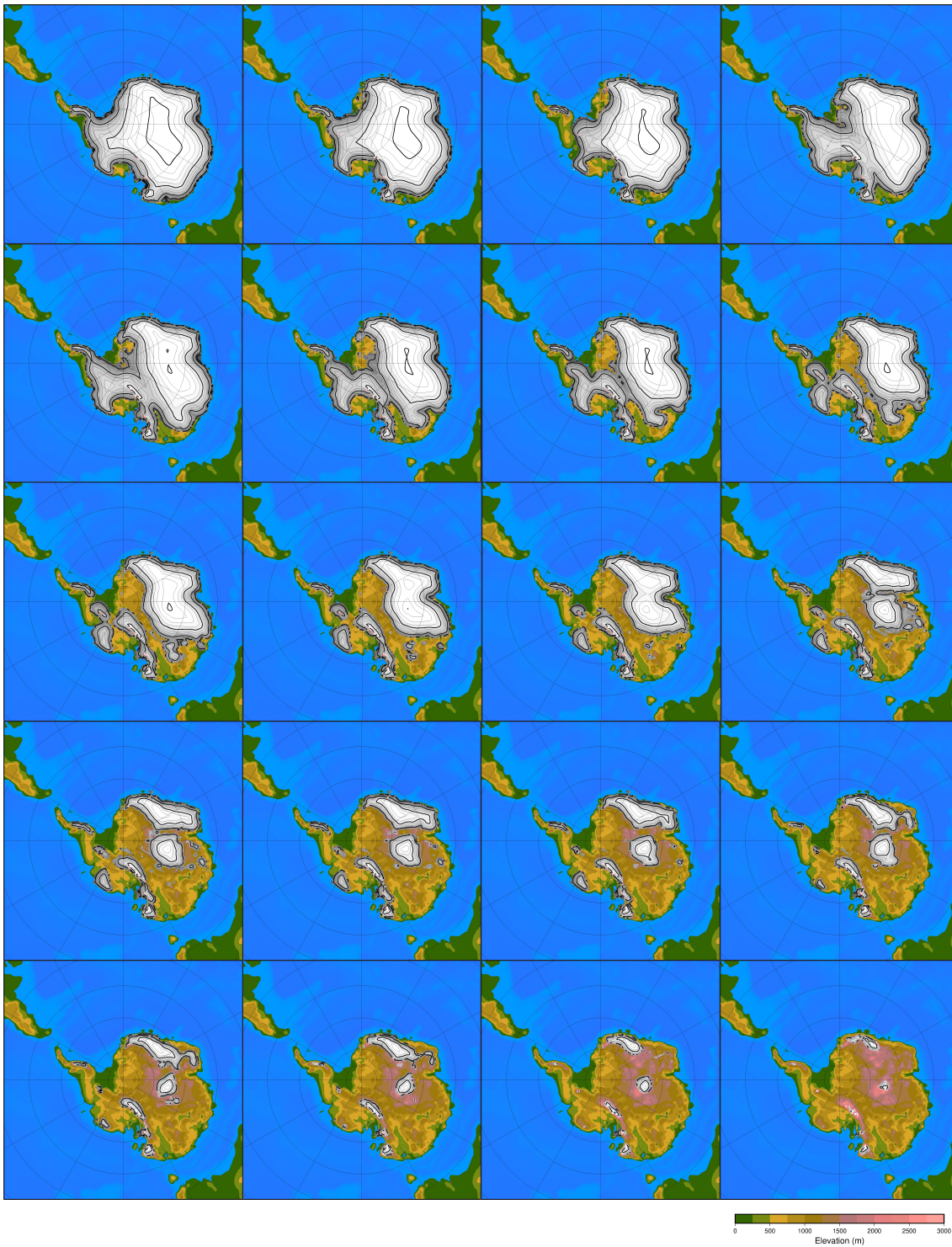



Figure A4: 20 ice sheet geometries as input to EMULATOR_20.

885

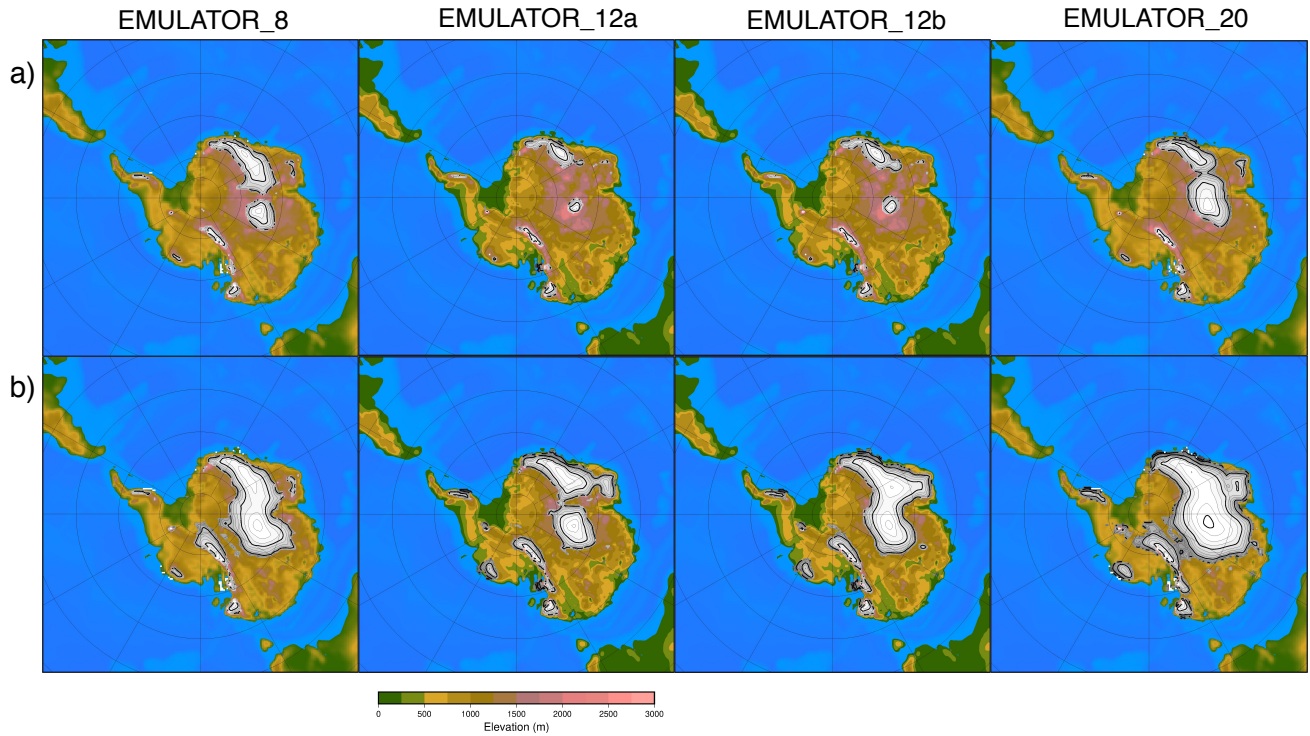



Figure A5: Snapshots of the ice sheet geometry after (a) 47,000 years and (b) 60,000 years (corresponding to the
890   **temperature fields in Fig. 10) for EMULATOR_8, EMULATOR_12a, EMULATOR_12b and EMULATOR_20**
calibrated with ice volume.

895