

A model intercomparison of radiocarbon-based marine reservoir ages during the last 55 kyr including abrupt changes in the Atlantic Meridional Overturning Circulation

Peter Köhler¹, Laurie Menviel², Frerk Pöppelmeier³, Timothy J. Heaton⁴, Edouard Bard⁵, and Luke C. Skinner⁶

¹Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung, P.O. Box 12 01 61, 27515 Bremerhaven, Germany

²Climate Change Research Centre, Australian Centre for Excellence in Antarctic Science, University of New South Wales, Sydney, NSW, Australia

³Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

⁴Department of Statistics, School of Mathematics, University of Leeds, Leeds, UK

⁵CEREGE, Aix-Marseille University, CNRS, IRD, INRAE, Collège de France, Technopole de l'Arbois BP 80, Aix en Provence Cedex 4, France

⁶Godwin Laboratory for Palaeoclimate Research, Earth Sciences Department, University of Cambridge, Downing Street, CB2 3EQ Cambridge, UK

Correspondence: Peter Köhler (peter.koehler@awi.de)

Abstract. Changes in the marine reservoir age (MRA) of the surface ocean are important information used for radiocarbon dating of marine sediment cores or archaeological artifacts. MRA changes are expressed relative to the atmosphere, and as such are dependent on the prevailing atmospheric radiocarbon calibration curve. The most recent estimate for evolving global average MRA for latitudes approximately $< 50^\circ$ is incorporated into the marine calibration curve Marine20, which was directly calculated from the atmospheric $\Delta^{14}\text{C}$ record, IntCal20, using the carbon cycle box model BICYCLE, taking into account observed changes in the carbon cycle. These simulations did not consider changes in the strength of the Atlantic meridional overturning circulation (AMOC) related to Dansgaard/Oeschger and Heinrich events. A recent study using the successor BICYCLE-SE suggested that abrupt AMOC changes would lead to changes in MRA of less than $100 \text{ }^{14}\text{C yr}$ in the non-polar surface ocean (about $< 50^\circ$). To better support previous model-based MRA and to further constrain the impact of AMOC changes on MRA, we here assess transient simulations of the last 55 kyr performed by two Earth System Models of Intermediate Complexity (EMICs), LOVECLIM and Bern3D, and compare them to the published BICYCLE-SE box model results and previous output from the [LSG-Large Scale Geostrophic \(LSG\)](#) ocean general circulation model. The setups within this MRA model intercomparison (MRA-MIP) are not identical, but all models are forced by atmospheric CO_2 and $\Delta^{14}\text{C}$ to have the surface ocean carbon cycle state as close as possible to reconstructions. Simulations with abrupt AMOC reductions during stadials display a rise in ~~in-MRA~~ in the surface northern Atlantic ($>50^\circ\text{N}$) and the deep Atlantic, for example ~~during Heinrich stadial 1 of reaching~~ [during Heinrich stadial 1](#) ~~300–1250 and 500–1300 $^{14}\text{C yr}$, respectively, roughly in agreement with their reconstructed rises by about 1200–1300 yr.~~ [during Heinrich stadial 1](#). We find that the changes in the mean non-polar surface MRA ($< 50^\circ$ latitude) during abrupt AMOC

changes in LOVECLIM are also in the order of ± 100 ^{14}C yr, while in Bern3D simulating changes are up to ± 200 ^{14}C yr. While the models tend to agree that a reduced AMOC leads to lower MRA of about 100–300 ^{14}C yr in the low-latitude surface ocean, under some conditions the opposite is found (e.g. simulations with LOVECLIM across Heinrich stadial 1). Spatially resolved results of the models show that changes in surface MRA during stadials depict the general pattern of a radiocarbon bipolar seesaw (older surface water in the high north, younger in the high south and in the Indo-Pacific), in agreement with previously published reconstructions, ~~but with model-specific details~~. However, some model-dependent differences remain in the non-polar Atlantic. Throughout the last 50 kyr, the change in the multi-model mean in non-polar MRA of the 2 EMICs when compared with Marine20 is less than 100 ^{14}C years and within the uncertainties of Marine20. Furthermore, changes in the MRA of the high latitude Southern Ocean ($> 50^\circ\text{S}$) are extremely model-dependent and ~~for most times~~, for much of the period between 18 and 43 kyr BP, the changes in the multi-model mean MRA are larger than the 95% confidence interval of the non-polar MRA depicted in Marine20, ~~making~~. These differences make the construction of a ~~similarly~~-numerical model-based calibration curve for ~~this region a challenging task~~ the high latitude Southern Ocean challenging.

30 1 Introduction

Radiocarbon (^{14}C), with a half-life of about 5700 years, is an ideal tool for determining the age of carbonaceous materials and tracing components of the global carbon cycle over the last $\sim 55,000$ years (Hajdas et al., 2021). However, the amount of ^{14}C in all reservoirs is not constant over time, but varies due to a changing carbon cycle and geomagnetic and solar effects on the ^{14}C production rates in the upper atmosphere (e.g. Heaton et al., 2021; Köhler et al., 2022). Therefore, radiocarbon dating needs to rely on calibration curves, that take these temporal changes into consideration.

Within the last iteration of these calibration curves the carbon cycle box model BICYCLE (Köhler et al., 2006) has been used to calculate $\Delta^{14}\text{C}$ in the ocean. For that effort the model has been forced by changes in atmospheric CO_2 (Figure 1a) as seen in ice cores (Köhler et al., 2017) and by changes in atmospheric $\Delta^{14}\text{C}$ (Figure 1b) as compiled within IntCal20 (Reimer et al., 2020). From $\Delta^{14}\text{C}$ in both the atmosphere and the surface ocean the marine reservoir age (MRA) of the non-polar surface ocean, also called the Marine20 calibration curve (Heaton et al., 2020), has been constructed (Figure 1c), ~~In general, the MRA is a measure for the level of the oceanic ^{14}C depletion with respect to the contemporaneous atmosphere, which is depending on time, geographic location, and when addressing not only the ocean surface, also on the water depth.~~ Here, BICYCLE was applied using a Monte-Carlo approach with 500 repetitions to account for uncertainties in data and the prescribed parametrization of ~~piston the air-sea gas exchange~~ velocity and the strength of the Atlantic meridional overturning circulation (AMOC), the two processes identified to be most important for simulated surface ocean $\Delta^{14}\text{C}$. So far, this approach ignored changes in the AMOC linked to the millennial-scale variability of Dansgaard/Oeschger and Heinrich events (Henry et al., 2016; Menviel et al., 2020), and thus climatic shifts observed in ice core records and marine sediment cores (e.g. Blunier and Brook, 2001; Davtian and Bard, 2023) (Figure 1d) were not included. However, considering abrupt AMOC changes in models can produce anomalies in radiocarbon age of the non-polar surface ocean of the order of ~ 100 ^{14}C years (Köhler et al., 2024a). Although this is within the 2σ uncertainty of Marine20, reconstructions of past MRA variability demonstrate the

occurrence of larger changes across the last deglaciation in association with millennial-scale climate anomalies (Skinner et al., 2019), ~~which~~. Taken together, this points to not yet fully considered errors in obtained simulations and larger uncertainties in marine radiocarbon calibrations around abrupt changes in AMOC strength.

To further explore the features and robustness of Marine20 — and its potential model-dependency — we here compare
55 changes in surface ocean MRA, mainly in the non-polar areas (50°S to 50°N), over the last 55 kyr as simulated in a box
model and two Earth system models of intermediate complexity (EMICs). We rely on initial results of BICYCLE-SE (Köhler
et al., 2024a) and add outputs from more complex EMICs, as they are fast enough to transiently simulate the last 55 kyr in a
reasonable amount of time. To this end we use the outputs from LOVECLIM (Menviel et al., 2014) and Bern3D (Pöppelmeier
et al., 2023b). We also add some results of the available outputs from the LSG Large Scale Geostrophic (LSG) ocean general
60 circulation model (OGCM) without abrupt AMOC changes, that have already been used within Marine20 (Butzin et al., 2020).
These LSG OGCM runs are transiently forced with variable atmospheric carbon records (CO_2 , $\Delta^{14}\text{C}$), but all else was kept
constant. While the focus is on the non-polar areas, here defined as about $< 50^{\circ}$ latitude roughly corresponding to the areal
extend of the surface ocean boxes in BICYCLE since this is what is covered by Marine20 so far, we also present and discuss
MRA changes in the surface polar oceans.

65 2 Methods

This is a “come-as-you-are” model intercomparison project on MRA (MRA-MIP) implying that model setups have not been
homogenised, but different groups, which have simulations of the last 55 kyr readily available (or easily produced), have been
asked to provide their calculated surface ocean MRA. Such a call offers the possibility of contributions from more groups, but,
also contains inter-model offsets that are based on the differences in the setup.

70 All models have been run with prescribed atmospheric carbon records, namely ice core CO_2 (Köhler et al., 2017) (Figure 1a)
and IntCal20-based $\Delta^{14}\text{C}$ (Reimer et al., 2020) (Figure 1b) in order to calculate MRA in the surface ocean as closely as possible
to the data. Additional forcings and exceptions to this rule are found in the subsections describing the individual model setups.

MRA (in units $^{14}\text{C yr}$) is calculated after

$$\text{MRA} = 8033 \cdot \ln \left(\frac{\frac{\Delta^{14}\text{C}_{\text{atmosphere}}}{1000} + 1}{\frac{\Delta^{14}\text{C}_{\text{sample}}}{1000} + 1} \right) \text{ } \underline{^{14}\text{C yr}} \quad (1)$$

75 as described in different format in Soulet et al. (2016). This equation is identical to what is used for Marine20, when the sample
is-corresponds to the non-polar surface ocean (Heaton et al., 2020) ~~and as~~. As summarized in Skinner and Bard (2022) it is
equivalent to the radiocarbon age difference between two contemporary marine and atmospheric signal carriers, e.g. as for B-
Atm radiocarbon age offsets (e.g. Skinner et al., 2023) for which benthic values are compared to the atmosphere. As elaborated
in detail elsewhere (e.g. Skinner and Bard, 2022), due to its dependency on atmospheric $\Delta^{14}\text{C}$, and therefore of changes in the
80 ^{14}C production rate and the carbon cycle, neither MRA nor B-Atm are perfect tracers of a water mass transient time or ‘ideal
age’.

The surface MRA was calculated from the surface boxes in the box model or roughly the top 50 m of the ocean in the EMICs and the OGCM. For comparisons, surface MRA from the model results have also been computed from roughly the top 100 m and 200 m. Results were reported from full vertical layers, which due to different model grids differed in detail from these depth horizons. An overview of the models and simulation scenarios used in this study and the model-specific depth of the surface layers from which data have been processed is compiled in Table 1. While we focus in the discussions of results and the plotted time series in the main text on changes in MRA with respect to pre-industrial (PI), since this is the target of the marine ^{14}C calibration curves, we show absolute values of MRA in time series in the SI, which might be more of an interest from a climate/carbon cycle perspective.

90 2.1 BICYCLE-SE

BICYCLE-SE is a carbon cycle box model consisting of 10 ocean boxes, a one box atmosphere and a 7 box terrestrial biosphere which also considers carbon exchange fluxes with the solid Earth by a sediment module, volcanic outgassing, weathering and coral reef growth. It is fully described in Köhler and Munhoven (2020) with the processes related to carbon isotopes being updated recently (Köhler and Mulitza, 2024). The simulation scenarios shown here have already been discussed in Köhler et al. (2024a). Scenario A0 is a run without abrupt changes in AMOC during Greenland stadials, while AMOC changes drastically in scenario A3. These AMOC changes are prescribed by the mean of two independent Iberian Margin sea surface temperature (SST) data sets (Davtian and Bard, 2023) which is then rescaled to obtain a minimal AMOC of 2 Sv during HS1, since this agrees best with ^{14}C reconstructions in the deep Atlantic Ocean (Köhler et al., 2024a). See Figure S1 for details on both the contained climate change and the simulated MRA with BICYCLE-SE.

100 2.2 LOVECLIM

LOVECLIM (Goosse et al., 2010) is an EMIC, which includes a quasi-geostrophic T21 atmospheric model, an ocean general circulation model ($3^\circ \times 3^\circ$, 20 vertical levels) coupled to a dynamic-thermodynamic sea-ice model, a vegetation model and a global carbon cycle model. The conditions at 54 kyr BP were obtained through a transient experiment starting at 140 kyr BP (Menviel et al., 2021). The model is forced with the transient evolution of orbital parameters (Berger, 1978), atmospheric greenhouse gas concentration (Köhler et al., 2017; Bereiter et al., 2015), and continental ice-sheet geometry and albedo. For the period 140 to 120 kyr BP, the evolution of the continental ice-sheet geometry and albedo are as described in the PMIP4 protocol of the penultimate deglaciation (Menviel et al., 2019). Between 120 kyr BP and 20 kyr BP, the continental ice-sheet geometry and albedo evolution are given by a 130 kyr BP off-line ice-sheet model simulation (Abe-Ouchi et al., 2007). Between 20 and 2 kyr BP, the model is forced by the continental ice-sheet geometry and albedo evolution from ICE4G (Peltier, 1994). At 54 kyr BP, the Bering Strait is closed, and is gradually opened between 11 and 10 kyr BP.

Between 140 and 54 kyr BP, the atmospheric $\Delta^{14}\text{C}$ is set constant at 0‰. The model is then re-equilibrated with atmospheric $\Delta^{14}\text{C}$ at 54 kyr BP for 5000 years. From 54 kyr BP the model is forced by the transient evolution of atmospheric $\Delta^{14}\text{C}$ using the IntCal20 data (Reimer et al., 2020). To allow the ocean to fully equilibrate the oceanic $\Delta^{14}\text{C}$ can be used in this context here from about 50 kyr BP onwards.

115 To simulate the millennial-scale variability of the last glacial period and the impact of deglacial ice-sheet disintegration (scenario fwf), meltwater ~~of up to 0.3 Sv~~ is added into the North Atlantic (50–60°N, 10–60°W) with the timing these events being based on Iberian Margin SST from Martrat et al. (2007). In detail, 0.3 Sv was added to achieve a collapsed AMOC during Heinrich stadials (HS) 1-4, while only 0.2 Sv was added during HS5. A triangular pulse of up to 0.3 Sv was also added between 13 and 12 ka to simulate a reduced AMOC during the Younger Dryas. For comparison, a simulation without such
120 meltwater fluxes and related millennial-scale changes is performed (scenario nofwf). See Figure S2 for details on both the contained climate change and the simulated MRA with LOVECLIM.

2.3 Bern3D

The Bern3D model is a coarse resolution ($68 \times 46 \times 40$ irregular spaced grid in longitude, latitude and depth) Earth system model (or EMIC) that couples a frictional-geostrophic ocean to 2D energy-moisture balance atmosphere. In contrast to the detailed
125 description of the model in Pöppelmeier et al. (2023b), we here employ it without the dynamical ice-sheet component. Instead, ice-sheet evolution over the last 55 kyr is prescribed. For this, the Last Glacial Maximum (LGM) (~~ICE-6G (Peltier et al., 2015)~~) (ICE-6G, Peltier et al., 2015) and modern ice-sheet extents are linearly interpolated based on the benthic $\delta^{18}\text{O}$ record of Lisiecki and Raymo (2005) for the last 55 kyr. The biogeochemical cycle including the implementation of radiocarbon is described in Parekh et al. (2008) and Tschumi et al. (2008).

130 In addition, to changes in orbital configuration (Berger, 1978), greenhouse gases (CO_2 , CH_4 , and N_2O (Köhler et al., 2017)), and ice-sheets, aerosol radiative forcing due to the changing atmospheric dust load is prescribed as well for the transient simulations. The temporal evolution of the aerosol radiative forcing follow the EPICA Dome C dust record (Lambert et al., 2012), which has been normalised to values between 0 during the late Holocene and -2 W m^{-2} , which corresponds to an average radiative forcing of about -1 W m^{-2} at the LGM (see Pöppelmeier et al., 2023b). In scenario PallSTD, additional
135 freshwater fluxes are applied into a box in the North Atlantic (45–70°N, 66°W–14°E) during all stadials, with a maximum freshwater flux of 0.4 Sv during ~~Heinrich stadials (HS)~~ HS and 0.2 Sv during non-Heinrich stadials. The timing of stadials is, similarly as in BICYCLE-SE, based on the Iberian Margin SST as published in Davtian and Bard (2023).

The model was first spun up to PI conditions, followed by a 10 kyr spinup to the conditions of 65 kyr BP. The model was then run transiently from 65 kyr BP to present day, with the time from 65 kyr BP to 55 kyr BP, serving as an additional transient
140 spinup period (with atmospheric $\Delta^{14}\text{C}$ fixed at its IntCal20 value for 55 kyr BP) to reach a realistic radiocarbon content in the deep ocean. The additional scenario Pnofwf is in all but the missing freshwater fluxes during stadials identical to PallSTD. ~~The Note that the~~ simulations shown here are due to the different spinup (and missing flux corrections) different to the run have different spin-ups—and contain no flux corrections from the North Atlantic to the North Pacific (45–70°N)—and so provide different results to the runs published in Pöppelmeier et al. (2023a) ~~, which has been compared to which considered~~
145 proxies during Termination ~~I-1~~. See Figure S3 for details on both the contained climate change and the simulated MRA with Bern3D.

2.4 LSG

Output from the ~~Hamburg Large Scale Geostrophic OGCM (LSG-OGCM)~~ LSG OGCM has already been used within IntCal20 (Butzin et al., 2020; Heaton et al., 2020; Reimer et al., 2020). LSG has been used with a horizontal resolution of 3.5° and a vertical resolution of 22 unevenly spaced levels. Further setup details are contained in Butzin et al. (2005, 2020). No additional simulations (or outputs) have been performed. The model has been run with temporally constant ocean circulation ~~in three different climate setups for three different atmospheric forcings. The three different climate setups which differed in three climate setups.~~ They contain a scenario that mimics present-day climate background conditions approximating the Holocene and interstadials. One glacial scenario aims at representing the LGM, features a shallower and by about 30% weaker AMOC compared to interglacial. A second glacial climate scenario mimics cold stadials with further AMOC weakening by about 60%. As described in detail in Butzin et al. (2005) these different climate setups have been generated by forcing the surface ocean with output from an atmospheric model, that itself was driven by different SST reconstructions. Each of these three climate setups has been performed for three different atmospheric $\Delta^{14}\text{C}$ forcings ending in nine simulations. The three atmospheric $\Delta^{14}\text{C}$ forcings contain the posterior mean estimate for the Hulu-based ^{14}C atmospheric reconstruction and two time series that cover its 95% confidence interval (CI) (mean $\pm 2\sigma$). Here, the median MRA from these nine simulations is shown, which is also what has been used within IntCal20.

2.5 Marine ^{14}C data

For an initial evaluation of model-based surface MRA we compare simulations with data-based MRA from the Global Ocean Data Analysis Project (GLODAP) (Key et al., 2004), which is a synthesis of ocean sampling expeditions carried out in the 1990s within the World Ocean Circulation Experiment (WOCE), the Joint Global Ocean Flux Study (JGOFS), and the Ocean Atmosphere Carbon Exchange Study (OACES). Here, we calculate

$$\text{MRA} = 8033 \cdot \ln \left(\frac{\frac{\Delta^{14}\text{C}_{\text{atmosphere}}}{1000} + 1}{\frac{\text{pre-bomb DI}^{14}\text{C}_{\text{abiotic}}}{\text{PI DIC}_{\text{abiotic}}}} \right) \text{ } \underline{^{14}\text{C yr}} \quad (2)$$

in units $^{14}\text{C yr}$ based on the approach used for the ocean model intercomparison project (Orr et al., 2017). Abiotic PI dissolved inorganic carbon (DIC) has been derived from the atmospheric CO_2 concentration of 284 ppm and pre-bomb abiotic dissolved inorganic ^{14}C (DI ^{14}C) from $\Delta^{14}\text{C}_{\text{atm}} = -24\text{‰}$. According to Graven et al. (2017), the latter represents the mean value of the decade before the onset of the bomb-peak in mid-1955, which has been used as pre-bomb baseline to reconstruct marine ^{14}C from potential alkalinity (Rubin and Key, 2002). This so-called “natural” GLODAP-based MRA will always be a compromise since ^{14}C is only corrected for bomb- ^{14}C , but not for fossil fuels, the so-called ^{14}C -Suess effect (Suess, 1955; Stuiver and Quay, 1981).

To evaluate this natural GLODAP-based MRA we calculate local anomalies in MRA, the so-called ΔR by subtracting the Marine20-based global MRA for 1950 CE ($407 \text{ }^{14}\text{C yr}$), and compare the resulting map of ΔR with the entire content of the ΔR data base found at <http://calib.org/marine/> (Reimer and Reimer, 2001). We take the 2000 entries of this data base which were available with stated uncertainties on 2nd October 2025. The data were collected mostly between the years 1729 and 1959

CE with 3 entries from the 17th century, one from 1512 CE and one from before CE. ΔR is calculated from the difference of
180 the reverse-calibrated collection year using Marine20 and the measured ^{14}C age and is assumed to stay constant in time. We
average the data in a spatial resolution of 2° in both latitude and longitude ending with 609 values. Here, we use (as in the
application of the online scripts of the data base) weighted means for averaging with the mean of ΔR ($\overline{\Delta R}$) given by

$$\overline{\Delta R} = \frac{\sum_i \frac{\Delta R_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}} \quad (3)$$

where σ_i is the uncertainty in ΔR_i and the reported uncertainty σ of $\overline{\Delta R}$ is the maximum of the standard deviation of $\overline{\Delta R}$
185 and the weighted uncertainty in $\overline{\Delta R}$ (see <http://calib.org/marine/AverageDeltaR.html> or Bevington (1969) for details). We plot
not only $\overline{\Delta R}$, but also a minimum and a maximum version with $\overline{\Delta R} \pm 1\sigma$ (Figure S4). Independent of which version of the
data-based ΔR we take we find in general a good agreement between them and natural GLODAP (differences of typically up
to 100 ^{14}C years) with the exception of single data points and the entire the west coast of North America, where values in the
data base are more than 100 ^{14}C years older than in GLODAP, potentially caused by coastal effects.

190 In order to compare model outputs with observations, we make use of compiled deglacial marine radiocarbon data from
Skinner et al. (2023) and regional time-series splines [based on planktic foraminifera](#) from Skinner et al. (2019). The latter
include regional time series [from 25 sites](#) for the surface northeast Atlantic ($>52^\circ\text{N}$ and east of 24°W) and the Iberian Margin
($\sim 38^\circ\text{N}$, 10°W). We use the deep Atlantic MRA (i.e. B-Atm, deeper than 2 km water depth) based on [34 sites that contributed](#)
[to](#) the baseline selection of Skinner et al. (2023) in the realisation of Köhler et al. (2024a). We do not compare model results to
195 Stern and Lisiecki (2013), since they used age models based on IntCal09 and the resulting 42 kyr-long time series of surface
MRA cover $0\text{--}65^\circ\text{N}$ in the Atlantic, a value which cannot be extracted from the box model simulations.

Additionally, point-wise surface MRA for the time slices of the LGM (19–21.8 kyr BP) and the HS1 (15–17.5 kyr BP) from
the baseline selection of the data base of Skinner et al. (2023) are used for further model evaluation. Here, multiple MRA
entries for the same sites are averaged using weighted means (similar to Eq. 3) reducing 67 entries for the LGM to 19 values
200 (89 entries for the HS1 to 22 values), from which 13 exist for the same location in both periods making the calculation of
HS1–LGM differences in MRA possible. Here, one entry for [the](#) LGM was rejected as outlier since its MRA was > 1000
 ^{14}C yr offset from four alternative MRAs for the same site. Most surface MRA data are from the northeast Atlantic (between
Iberian Margin and Iceland) which are complemented with data from single sites in the Southern Ocean, the tropical East- and
West Pacific and the tropical Atlantic.

205 3 Results and ~~Discussions~~[Discussion](#)

3.1 Pre-industrial MRA compared to GLODAP

Estimates of changes in surface MRA are necessary for ^{14}C dating of marine carbon-containing material. For radiocarbon
dating of marine organisms, such as foraminifera, it is important to take into account the organism's seasonal/depth habitat,
e.g. for the estimation of appropriate local variations in reservoir age, that is ΔR (Reimer and Reimer, 2001), or for comparison

210 with model-based MRA estimates. Planktonic foraminifer habitats are species-specific, and often related to the mixed layer depth, but sometimes extend to greater depths (Kimoto, 2015). The mixed layer depth has a seasonal cycle, but for modern day latitudes $< 50^\circ$, it is typically less than 200 m (de Boyer-Montégut et al., 2004), which in our two EMICs is indeed the case for the annual mean mixed layer depths during the PI (Figure S5). Note, that due to a lack of more recent results from LOVECLIM, results for 2 kyr BP are chosen to stand in for the PI reference. The mean surface MRA calculated for the top 50 m
215 within IntCal20 from LSG output (Butzin et al., 2020) might be relevant to foraminifer data from the lower end of the [mixed layer](#) depth range, while the 100 m thick non-polar surface ocean boxes of the BICYCLE model used within Marine20 were probably in the middle of the relevant [mixed layer](#) depth range. From the GLODAP data we calculate a mean surface MRA for latitudes $< 50^\circ$ of 291, 318 and 356 ^{14}C yr, when considering data from the top 50, 100, and 200 m water depth, respectively, illustrating the centennial-scale of the uncertainty related to the assumed habitat depth range. Furthermore, calculated surface
220 MRA for different depth ranges from the transient simulations provide insights into how this depth-dependency might vary with time (Figures S2–S3). Here, the depth ranges over which surface MRA have been calculated differ in detail (Table 1), depending on model grid, since we only consider full vertical layers. Differences between MRA based on roughly the top 50 m or 100 m are similar, and in agreement with GLODAP, while those based on roughly the top 200 m are about 50 and 100 ^{14}C yr higher than MRA of top 50 m in Bern3D and LOVECLIM, respectively.

225 Maps of surface MRA for 2 kyr BP (our PI reference) compare well with MRA based on natural ^{14}C in GLODAP (Figure 2). [The area-weighted root mean square errors of the residuals of the model-based differences to natural \$^{14}\text{C}\$ in GLODAP are 202, 133, 86 and 265 \$^{14}\text{C}\$ yr for BICYCLE-SE, LOVECLIM, Bern3D and LSG, respectively, which are all substantially smaller than the 399 \$^{14}\text{C}\$ yr of the area-weighted root mean square of the natural \$^{14}\text{C}\$ in GLODAP evidencing the explanatory power of all the four models.](#) Note, that atmospheric $\Delta^{14}\text{C}$ at 2 kyr BP was with -16‰ only slightly different from the pre-bomb value
230 of -25‰ in 1950 CE (0 kyr BP). However, be aware that this comparison has certain weaknesses due to the compromises in the GLODAP data (not free of the ^{14}C -Suess effect, see Methods section for details) and our choice of using simulation results at 2 kyr BP. For latitudes $< 50^\circ$ the simulated surface MRAs are ~~around 500~~ [461, 401, 344 and 443 \$^{14}\text{C}\$ yr](#) ~~and for BICYCLE-SE, LOVECLIM, Bern3D and LSG, respectively and~~ the differences from GLODAP are for all models typically less than 300 ^{14}C yr with BICYCLE-SE and LSG showing predominantly larger values, while the two EMICs (especially Bern3D) seem to be
235 closer to GLODAP. [The surface MRA in LOVECLIM in the North Atlantic seemed be an exception here showing a difference of more than 200 \$^{14}\text{C}\$ yr, a larger offset to GLODAP than the other models.](#) LSG contains remarkably large positive offsets from GLODAP of more than 600 ^{14}C yr in high latitudes probably related to the simulated sea ice. BICYCLE-SE shows a negative offset from GLODAP in the Southern Ocean of around 500 ^{14}C yr. ~~Although it is tempting to suggest that this~~ [This](#) offset might be caused by ~~the BICYCLE-SE's~~ [1000 m deep surface ocean boxes in the polar regions,](#) ~~However, since~~ a similar
240 offset to GLODAP is missing in the [northern high latitudinal boxes, which are](#) also 1000 m deep ~~northern high latitudinal boxes, pointing to an alternative explanation, some counteracting processes might be at work here.~~

3.2 MRA variations across the last glacial period

As a metric of the simulated climate variations in the various models we here show changes in global mean sea surface temperature (GMSST). The simulated GMSST dropped during the LGM, depending on model, between -1.5 K and -2.5 K with respect to PI (Figure 3a). While the -2.5 K cooling (BICYCLE-SE) is in the range covered by data-based studies (e.g. Tierney et al., 2020; Annan et al., 2022) the upper end of the range of -1.5 to -2.0 K (Bern3D, LOVECLIM) is probably under-estimating climate change during full glacial conditions.

The timing of AMOC reductions in the model simulations was prescribed to coincide with the timing of Greenland stadials expressed in the ice-core record of $\delta^{18}\text{O}$ (e.g. NGRIP Members, 2004), and was also captured in the SST reconstruction at the Iberian Margin (Martrat et al., 2007; Davtian and Bard, 2023). Here, the precise timing in BICYCLE-SE and Bern3D is based on Davtian and Bard (2023), while in LOVECLIM data from Martrat et al. (2007) have been used, giving more or less ~~the~~ similar results. In Bern3D the applied freshwater fluxes were adjusted according to its known freshwater sensitivity (Pöppelmeier et al., 2023b) to achieve a virtually collapsed AMOC during Heinrich stadials and a strongly reduced AMOC during non-Heinrich stadials. ~~Similarly, in LOVECLIM, while in LOVECLIM AMOC collapsed due to the applied freshwater fluxes were adjusted to achieve collapsed AMOC during Heinrich stadials 1–4 (only a slightly reduced AMOC during forcing only in HS1–4, but has been slightly reduced in HS5), but freshwater was applied only during some non-Heinrich stadials, e.g. and~~ the Younger Dryas.

In BICYCLE-SE sensitivity tests and a comparison of simulated deep ocean MRA with reconstructions has been used to choose the applied AMOC changes (Köhler et al., 2024a). The resulting changes in simulated AMOC strength across models (Figure 3b) varied widely, e.g. 16 – 26 Sv for PI, 10 – 20 Sv for the LGM and 2 – 5 Sv for HS1 and the Younger Dryas, summarizing results for runs with abrupt AMOC shutdown. Note, that the AMOC strength was in each model calculated differently (see caption to Figure 3 for details). For earlier HS, most models show smaller AMOC reductions than for HS1 and LOVECLIM also contains a strong but short reduction in AMOC during the 8.2 kyr event not covered by the other models. Notably, the AMOC at the LGM in Bern3D, which has been evaluated in a multi-proxy approach to about 11 Sv (Pöppelmeier et al., 2023a) is now ~ 15 Sv, since flux corrections applied in the other study have been neglected here.

The simulated non-polar MRA ($< 50^\circ$) contains remarkable model-specific differences. From the absolute values we can identify the model-specific offsets with Bern3D simulating the lowest PI MRA, while BICYCLE-SE simulates the highest PI MRA (Figure 3d). On the other hand, at the LGM (20 kyr BP), the smallest MRA is simulated by LSG, followed by Bern3D, LOVECLIM, and BICYCLE-SE.

Across the last glacial period, Bern3D simulates the largest changes in MRA, with a 1000 ^{14}C yr anomaly with respect to PI during the Laschamps geomagnetic excursion around 42 kyr BP (Simon et al., 2020), while LOVECLIM and LSG display the smallest MRA anomalies with respect to PI (Figure 3e). The differences to Marine20 — when plotted with respect to PI — show ~~no consistent pattern, but~~ to a large degree model-specific responses which nearly all fall in their sizes within the 95% CI of Marine20 (Figure 3f). One more general pattern is the decrease of this difference in results from BICYCLE-SE and Bern3D during Greenland stadials not similarly seen in LOVECLIM.

Only when we calculate differences from simulations with and without abrupt AMOC shutdown during stadials for the same model [do](#) we find the tendency of smaller non-polar surface MRA during stadials ~~with the notable exception of (Figure 3g).~~ [However, a notable exception here is HS1 in LOVECLIM showing which shows](#) a more complex (rather opposite) dynamic ~~(Figure 3g).~~ Bard (1988) showed that Δ MRA remains small (~ 100 ^{14}C yr) when changing the eddy diffusivity in the box-diffusion model (Oeschger et al., 1975) that has been widely used to convert atmospheric $\Delta^{14}\text{C}$ changes in terms of oceanic changes (Stuiver et al. (1986) and subsequent IntCal calibration iteration until 2013). The main point outlined by Bard (1988) is that the global average surface MRA and mean deep ocean ^{14}C age vary in opposite directions when eddy diffusivity changes, mimicking global overturning variations (all else being equal). Interestingly, scaling the eddy diffusivity to our AMOC proxy curve (SST record by Davtian and Bard, 2023) leads to a similar Δ MRA pattern in the ~~outerop-diffusion~~ [box-box-diffusion](#) model as in the other models when they are averaged in the ~~low-latitude regime low and mid-latitudes~~ (Figure 3g). [The linear scaling of the eddy diffusivity assumes a modern value of 4000 \$\text{m}^2/\text{yr}\$ and a minimum value of 500 \$\text{m}^2/\text{yr}\$ at ca. 40 kyr BP during the coldest interval of HS4 \(\$\Delta\$ SST of \$-10\$ K, Fig. 1d\). Instantaneous steady state MRA values are then calculated with the analytical equation 4 derived by Bard \(1988\). In the box-diffusion model, the \$^{14}\text{C}\$ production is mainly balanced by radioactive decay in the deep ocean reservoir. Reducing or stopping the exchange with the mixed layer implies](#) [that \$^{14}\text{C}\$ remains confined to the atmosphere and the mixed layer. These two boxes, which contain comparable total carbon inventories, would tend to homogenize, thus reducing the MRA.](#)

When plotting changes in MRA reconstructions for the LGM (Skinner et al., 2023) together with model outputs we find that most models underestimate for most locations the changes in surface MRA (Figure 4). The best agreement, especially in the high latitude, is obtained for LSG. This comparison is rather limited, due to the availability of only 19 data points. It has as additional caveat that simulations are focused on 20 kyr BP, while the data cover the wider time window of 19–21.8 kyr BP. Changes in models and data only roughly agree in a few areas (Iberian Margin, Caribbean) but elsewhere differ widely, with models underestimating the MRA derived from proxy records. Note that the about 10 data point in the northeast Atlantic show partly very different changes, ranging from a reduction in MRA to a rise by more than 1000 ^{14}C yr. This local diversity of the reconstructions challenges our model-data comparison since small-scale localised effects seen in the data might not be contained in the coarsely resolved models. Notably, there is also significant disagreement between models in particular regions. One reason for the data-model offsets might be the habitat depth of the planktonic foraminifera, even though the EMICs give surface MRA that differ by less than 50 ^{14}C yr when based on roughly the top 200 m instead of the top 50 m used so far (Figure S6). Interestingly, annual mean mixed layer depth at 20 kyr BP in the two EMICs differ in the non-polar regime only very little from the PI control (Figure S5). Another reason may be that the models are all missing key processes that might influence vertical diffusivity in the upper ocean or air-sea gas exchange at the surface, particularly at high latitudes, or that they tend to simulate a different radiocarbon distribution in the interior ocean than prevailed in the past. [Going here into more detail, e.g. by an extended deep ocean \$^{14}\text{C}\$ model-data comparison is beyond the scope of the present study, but a potential promising avenue for future studies.](#)

3.3 Details of the impacts of AMOC weakening on MRA

310 To further elucidate the model-specific responses for HS1, we compare simulations with reconstructions for that time interval (Figure 5). Here, we focus on the HS1–LGM difference. Since the simulations with the LSG OGCM were not containing abrupt AMOC changes during Greenland stadials, the output from this model is not included here. For this comparison, 13 data points exist in the data base of Skinner et al. (2023). Both EMICs show a general MRA decrease at mid and low latitudes notably in the Indo-Pacific. This almost global MRA reduction is compatible with that obtained with the BICYCLE-SE model. In addition to this common pattern, the models exhibit ~~dynamical behaviour~~ different dynamical behaviour (e.g. a different ocean overturning or a different air-sea gas exchange of ^{14}C) that is responsible for an increase of MRA in the northern North Atlantic, with varying intensity and spatial extent: confined to the northern North Atlantic (+Arctic) in BICYCLE-SE, intermediate in Bern3D, and widespread to most of the Atlantic basin in LOVECLIM. Thus, the different response in the Atlantic in HS1 readily explains LOVECLIMs different dynamics when comparing non-polar averages of runs with and without abrupt AMOC shutdown (Figure 3g). There are areas where models disagree with available data. In particular, the Iberian margin data, e.g. shown previously in Skinner et al. (2014, 2019, 2021), exhibit a clear MRA increase during HS1, as already acknowledged in Heaton et al. (2020), which is covered in LOVECLIM, but not in Bern3D. Both EMICs show the existence of a radiocarbon bipolar seesaw pattern, with widespread decrease in the Southern Ocean, not restricted to the Atlantic sector. This radiocarbon bipolar seesaw pattern with MRA increases in the northern Atlantic ~~, which is connected with reduced MLD there (Figure S5, right column),~~ and MRA reductions in the Southern Ocean is very reminiscent of the thermal bipolar seesaw (Stocker and Johnsen, 2003) available from paleo-SST data (e.g. Barker et al., 2009; Davtian and Bard, 2023) and from numerous model experiments (e.g. Zhang et al., 2014; Pedro et al., 2022). Indeed, this seesaw pattern in MRA variability is directly consistent with what has been previously observed in direct MRA reconstructions (Skinner et al., 2014, 2019, 2021, 2023; Skinner and Bard, 2022), where it has been hypothesized to relate to sea-ice variability in each hemisphere. Interestingly, the mixed layer depth is reduced in both the North Atlantic and the Southern Ocean between HS1 and LGM, especially in Bern3D (Figure S5, right column). This in-phase change of the mixed layer depth with the MRA in the north together with their anti-phase change in the south suggests that other processes are more important for explaining surface MRA in connection with AMOC weakening.

~~For some areas where the models disagree~~In some of the regions where the two EMICs disagree, reconstruction-based evidence ~~exist, especially from~~ also exists—in particular for the Iberian Margin. However, ~~for most of in~~ the equatorial Atlantic, where ~~both models show the opposite change in MRA~~ the models show opposite changes in MRA, such direct measurements are not available. This prevents a data-based evaluation of the ~~simulations is due to a lack of data still missing two diverging simulations.~~

An alternative to calculating the MRA anomaly during HS1 is based on the different surface MRA at 16 kyr BP from simulations with and without abrupt AMOC shutdowns (Figure S7, right column). However, such a model-based anomaly cannot directly be compared with reconstructions. Here, results from all contributing models (BICYCLE-SE, LOVECLIM,

Bern3D) show more positive Δ MRA than in the calculations based on the HS1–LGM difference, but contain similar patterns. Increases in surface MRA in the non-polar Atlantic basin are now also contained in BICYCLE-SE.

As a final step, we compare simulated and reconstructed time series of surface northern Atlantic ($>50^\circ$ N) and deep Atlantic MRA — all with respect to PI (Figure 6). The maximum lengths of reconstructed timeseries are limited to 25–30 kyr, but they provide useful insights for Termination I, including HS1. As expected, simulations with abrupt AMOC changes contain higher Δ MRA during HS1 than those without these abrupt changes. However, only LOVECLIM reached changes of $1300 \text{ }^{14}\text{C yr}$ in the surface northern Atlantic as found in the reconstructions, and only Bern3D reaches the HS1 peak of $1200 \text{ }^{14}\text{C yr}$ in the deep Atlantic, although some thousand years later. The model-data offset at the surface northern Atlantic might be contributed to aliasing of spatial patterns, whereby the compiled data are largely restricted to east of 24° W in the northern Atlantic $>42^\circ$ N whereas model results include the whole northern Atlantic and Arctic Ocean ($>50^\circ$ N). However, if this was the case, MRA variability in the western North Atlantic (and Arctic) would need to have been greatly reduced compared to the eastern part of the basin, which is opposite to what the model simulations tend to suggest (Figure 5). Note that in BICYCLE-SE, Δ MRA in both surface northern Atlantic and deep Atlantic are very similar probably linked to the box model geometry and prescribed fluxes which contain a direct connection and vigorous exchanges between both boxes. All models tend to agree on having peaks in both surface northern Atlantic and deep Atlantic Δ MRA connected with millennial-scale climate change. However, amplitudes are model-specific with Bern3D simulating the largest amplitudes which even appear, although in smaller size, in the simulation without abrupt AMOC changes. This indicates (as already notified in Köhler et al., 2024a) that they are partly related to the variability contained in either atmospheric $\Delta^{14}\text{C}$ or CO_2 and not only to abrupt AMOC changes, though the impact of such atmospheric effects is likely limited to less than $\sim 300 \text{ }^{14}\text{C years}$ equivalent for the global deep ocean and far less for the surface ocean (Skinner et al., 2023). Possible explanations might be related to changes in the terrestrial carbon cycle (e.g. Menking et al., 2022; Wu et al., 2022).

3.4 Towards global MRA

Finally, we calculate changes in surface MRA from a multi-model mean (MMM) based on low latitude ($<50^\circ$) results of scenarios with abrupt AMOC changes in the two EMICs, which are then compared with Marine20 (Figure 7a,b). While the full range of simulation results is nearly always within the 95% CI of Marine20, it is on its lower edge since the LGM. Furthermore, the MMM is consistently about $100 \text{ }^{14}\text{C yr}$ smaller than Marine20. We tested if this offset might be based on the chosen depth of the analysed surface MRA. However, we only found a slightly better agreement between MMM and Marine20 when using results from the mean of the top 200 m instead of the top 50 m as done in the default case (Figure 7a). This offset between MMM and Marine20 again illustrates that the pre-bomb non-polar MRA, which was about $407 \text{ }^{14}\text{C yr}$ for the year 1950 CE when using BICYCLE within Marine20 (Heaton et al., 2020) is model-specific.

In a very last step we additionally calculate changes in the surface MRA of the high latitudes ($>50^\circ$) and of the global mean in order to understand how much they differ from the MRA of the non-polar areas (Figure S8). The MRA in both high northern and high southern latitudes in LSG contain a jump by more than $1000 \text{ }^{14}\text{C yr}$ at 10.7 kyr BP; however, this jump is related to an arbitrary change to an ‘interglacial’ overturning scenario from the reduced ‘glacial’ overturning scenario in the LSG model.

As the much larger MRA at the LGM in high latitudes in LSG agreed better with reconstructions than the smaller values of the EMICs (Figure 4), this raises some questions: if these sparse LGM-based reconstructions are broadly representative of glacial conditions, one might ask if the EMICs are missing or inadequately resolving key processes in the high latitudes during glacial times. Alternatively, if the EMIC-based glacial polar MRA turn out to be more accurate than the LSG-based results any polar age calibration based on the latter (Heaton et al., 2023) are then potentially biased towards too old values. For the two EMICs, the global mean MRA and the non-polar MRA are very similar (Figure S8) suggesting that the non-polar MRA might indeed be of global applicability. However, when looking to the details we find that the range and the MMM of the changes in the northern polar MRA (which is similar, but not identical to the MRA of the northern North Atlantic discussed earlier) is — apart from certain HSs with several centuries older MRA — indeed comparable with the 95% CI of Marine20 (Figure 7c). The changes in the southern polar MRA are much higher than in Marine20, both in range and in MMM (Figure 7d), especially in Bern3D although still by a few centuries younger than in LSG (Figure S8). Thus, a global usage of the non-polar MRA would lead especially in the Southern Ocean (but also at the Iberian Margin; Skinner et al., 2019) to underestimations of MRA by at least one century leading to an overestimate of calendar age in probes whose radiocarbon age calibration is based on them. This Southern Ocean age offset in the models is probably related to the higher glacial sea ice coverage (Figure S9) and the larger glacial MLD-mixed layer depth (Figure S5). Here, larger glacial summer sea ice in the Southern Ocean and larger glacial MLD-mixed layer depth in Bern3D compared to LOVECLIM potentially explain at least part of the model-specific Δ MRA in the Southern Ocean.

4 Conclusions

Comparing for the first time results from transient simulations of two EMICs, 1 OGCM and 1 box model over the radiocarbon time window we find that abrupt AMOC changes might introduce anomalies of up to ~~100-200~~100-200 ^{14}C yr to the changes in the mean non-polar surface MRA. Our results show, that especially the surface Δ MRA in the non-polar Atlantic is model-specific and potentially different from the Indo-Pacific. ~~Although existing MRA data for evaluating model behaviour are relatively sparse for Heinrich stadials (here, especially~~ Directly observed information on MRAs over Heinrich stadials is relatively sparse which limits independent evaluation of model behaviour. The direct data that do exist during HS1 ~~) (e.g. Skinner et al., 2019, 2023), they suggest that the~~ (e.g. Skinner et al., 2019, 2023) suggest that no one model provides simulations with uniformly greater model-to-data accuracy. The ~~simulated Δ MRA during HS1 in LOVECLIM appear to be closer to the reconstructions than in~~ in LOVECLIM may be more accurate when compared to data-based reconstructions in the surface northern Atlantic; however ~~Bern3D .The model-intercomparison shows where different models have strengths and weaknesses, but leaves still open, which might be the best approach in the next iteration of IntCal, e. g. calculating multi-model means or relying on individual models.~~ may be more accurate in other locations. The very different changes in glacial polar MRA (especially in the Southern Ocean) in the two EMICs and LSG indicates that our ability to simulate radiocarbon in the polar regions remains limited and a robust conclusion on their MRA changes remains dependent on a relatively sparse observational database. The model-intercomparison shows where different models have strengths and weaknesses, but does

410 not yet determine, what the best approach in the next iteration of IntCal might be, e.g. calculating multi-model means or
relying on individual models. While there are indeed important differences between the models that need to be understood, the
overall level of agreement is fairly remarkable. This lends substantial support to the marine calibration endeavour and should
strengthen community faith in the reliability of marine radiocarbon calibration curves (at least at low-latitudes in the open
oceans).

415 For marine radiocarbon calibrations temporal changes in global average surface MRA (e.g. as provided by models such
as in this study) may be combined with regional ΔR estimates (for details see Heaton et al., 2020). However, this method
requires that the regional ΔR s remain constant over time, whereas a growing body of observations suggests that this is not
the case over the last deglaciation and during HS1 in particular (e.g. Siani et al., 2013; de la Fuente et al., 2015; Skinner
et al., 2014, 2019, 2023). Our ~~mini-MRA-MIP~~ MRA-MIP supports this observation (as yet based on relatively sparse data),
and suggests that during Greenland stadials MRA in the Atlantic varies differently than in the Indo-Pacific. While this points
420 to the potential usefulness of regional calibration curves (Skinner et al., 2019; Mârza et al., 2024), an alternative and more
immediately practical approach might be to continue to use one global calibration curve (as done so far) and apply larger and
appropriately structured age uncertainties for different regions and time periods, such as the northern Atlantic during Greenland
stadials.

Data availability. All simulation results are available from PANGAEA: BICYCLE(-SE), LSG and marine data splines: Butzin et al. (2019);
425 Heaton et al. (2020); Köhler et al. (2024b). New results from Bern3D and LOVECLIM (this study): Köhler et al. (2026).

Author contributions. PK designed the study, performed the BICYCLE-SE simulations, made all figures and led the writing of the draft.
LM performed LOVECLIM simulations. FP performed Bern3D simulations. EB performed box diffusion model simulations and contributed
knowledge on SST data, marine ^{14}C data and MRA. LCS contributed knowledge on marine ^{14}C data and MRA. TH contributed insights
how the findings might be used within the next iteration of IntCal. All authors discussed the results.

430 *Competing interests.* LM and LCS are members of the editorial board of *Climate of the Past*. Apart from that, the authors declare no
competing interests.

Acknowledgements. We thank Martin Butzin for digged out details on natural ^{14}C within GLODAP and for ongoing discussions on ^{14}C and
on LSG output and comments on the draft. We thank Florian Adolphi for comments on an earlier version of the draft. LM acknowledges
funding from the Australian Research Council (ARC) grant SR200100008. EB is funded by the ANR project MARCARA. FP is funded
435 by the European Union's Horizon Europe research and innovation program under Grant Agreement 101137601 (ClimTip) and 101184070

(Past-To-Future). LCS is acknowledges funding and support from the Royal Society and NERC grant NE/V011464/1. [We thank Patrick Rafter and an anonymous reviewer for their comments.](#)

References

- Abe-Ouchi, A., Segawa, T., and Saito, F.: Climatic Conditions for modelling the Northern Hemisphere ice sheets throughout the ice age cycle, *Climate of the Past*, 3, 423–438, <https://doi.org/10.5194/cp-3-423-2007>, 2007.
- Annan, J. D., Hargreaves, J. C., and Mauritzen, T.: A new global surface temperature reconstruction for the Last Glacial Maximum, *Climate of the Past*, 18, 1883–1896, <https://doi.org/10.5194/cp-18-1883-2022>, 2022.
- Bard, E.: Correction of accelerator mass spectrometry ^{14}C ages measured in planktonic foraminifera: Paleoceanographic implications, *Paleoceanography*, 3, 635–645, <https://doi.org/10.1029/PA003i006p00635>, 1988.
- Barker, S., Diz, P., Vantravers, M. J., Pike, J., Knorr, G., Hall, I. R., and Broecker, W. S.: Interhemispheric Atlantic seesaw response during the last deglaciation, *Nature*, 457, 1007–1102, <https://doi.org/10.1038/nature07770>, 2009.
- Bereiter, B., Eggleston, S., Schmitt, J., Nehrbass-Ahles, C., Stocker, T. F., Fischer, H., Kipfstuhl, S., and Chappellaz, J.: Revision of the EPICA Dome C CO_2 record from 800 to 600 kyr before present, *Geophysical Research Letters*, 42, 542–549, <https://doi.org/10.1002/2014GL061957>, 2015.
- Berger, A. L.: Long-term variations of daily insolation and Quaternary climatic changes, *Journal of Atmospheric Sciences*, 35, 2362–2367, [https://doi.org/10.1175/1520-0469\(1978\)035<2362:LTVODI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2), 1978.
- Bevington, P. R.: *Data Reduction and Error Analysis for the Physical Sciences*, McGraw Hill, 1969.
- Blunier, T. and Brook, E. J.: Timing of millennial-scale climate change in Antarctica and Greenland during the last glacial period, *Science*, 291, 109–112, <https://doi.org/10.1126/science.291.5501.109>, 2001.
- Butzin, M., Prange, M., and Lohmann, G.: Radiocarbon simulations for the glacial ocean: the effects of wind stress, Southern Ocean sea ice and Heinrich events, *Earth and Planetary Science Letters*, 235, 45–61, <https://doi.org/10.1016/j.epsl.2005.03.003>, 2005.
- Butzin, M., Heaton, T. J., Köhler, P., and Lohmann, G.: Marine radiocarbon reservoir ages simulated for IntCal20, link to model results in NetCDF format, PANGAEA, [data set], <https://doi.org/10.1594/PANGAEA.902301>, 2019.
- Butzin, M., Heaton, T. J., Köhler, P., and Lohmann, G.: A short note on marine reservoir age simulations used in IntCal20, *Radiocarbon*, 62, 865–871, <https://doi.org/10.1017/RDC.2020.9>, 2020.
- Cooper, A., Turney, C. S. M., Palmer, J., Hogg, A., McGlone, M., Wilmshurst, J., Lorrey, A. M., Heaton, T. J., Russell, J. M., McCracken, K., Anet, J. G., Rozanov, E., Friedel, M., Suter, I., Peter, T., Muscheler, R., Adolphi, F., Dosseto, A., Faith, J. T., Fenwick, P., Fogwill, C. J., Hughen, K., Lipson, M., Liu, J., Nowaczyk, N., Rainsley, E., Bronk Ramsey, C., Sebastianelli, P., Souilmi, Y., Stevenson, J., Thomas, Z., Tobler, R., and Zech, R.: A global environmental crisis 42,000 years ago, *Science*, 371, 811–818, <https://doi.org/10.1126/science.abb8677>, 2021.
- Davtian, N. and Bard, E.: A new view on abrupt climate changes and the bipolar seesaw based on paleotemperatures from Iberian Margin sediments, *Proceedings of the National Academy of Sciences*, 120, e2209558 120, <https://doi.org/10.1073/pnas.2209558120>, 2023.
- de Boyer-Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology, *Journal of Geophysical Research: Oceans*, 109, <https://doi.org/10.1029/2004JC002378>, 2004.
- de la Fuente, M., Skinner, L., Calvo, E., Pelejero, C., and Cacho, I.: Increased reservoir ages and poorly ventilated deep waters inferred in the glacial Eastern Equatorial Pacific, *Nature Communications*, 6, 7420, <https://doi.org/10.1038/ncomms8420>, 2015.
- Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, A., Selten, F., Barriat, P.-Y., Campin, J.-M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I., Loutre, M.-F., Morales Maqueda, M. A., Opsteegh, T., Mathieu, P.-P., Munhoven, G., Pettersson, E. J., Renssen, H., Roche, D. M., Schaeffer, M., Tartinville, B., Timmermann, A., and Weber, S. L.: De-

- 475 description of the Earth system model of intermediate complexity LOVECLIM version 1.2, *Geoscientific Model Development*, 3, 603–633, <https://doi.org/10.5194/gmd-3-603-2010>, 2010.
- Graven, H., Allison, C. E., Etheridge, D. M., Hammer, S., Keeling, R. F., Levin, I., Meijer, H. A. J., Rubino, M., Tans, P. P., Trudinger, C. M., Vaughn, B. H., and White, J. W. C.: Compiled records of carbon isotopes in atmospheric CO₂ for historical simulations in CMIP6, *Geoscientific Model Development*, 10, 4405–4417, <https://doi.org/10.5194/gmd-10-4405-2017>, 2017.
- 480 Hajdas, I., Ascough, P., Garnett, M. H., Fallon, S. J., Pearson, C. L., Quarta, G., Spalding, K. L., Yamaguchi, H., and Yoneda, M.: Radiocarbon dating, *Nature Reviews Methods Primers*, 1, 62, <https://doi.org/10.1038/s43586-021-00058-7>, 2021.
- Heaton, T. J., Köhler, P., Butzin, M., Bard, E., Reimer, R. W., Austin, W. E., Ramsey, C. B., Grootes, P. M., Hughen, K. A., Kromer, B., Reimer, P. J., Adkins, J. F., Burke, A., Cook, M. S., Olsen, J., and Skinner, L. C.: Marine20 — the marine radiocarbon age calibration curve (0–55,000 cal BP), simulated data for IntCal20, PANGAEA, [data set], <https://doi.org/10.1594/PANGAEA.914500>, 2020.
- 485 Heaton, T. J., Köhler, P., Butzin, M., Bard, E., Reimer, R. W., Austin, W. E. N., Ramsey, C. B., Grootes, P. M., Hughen, K. A., Kromer, B., Reimer, P. J., Adkins, J., Burke, A., Cook, M. S., Olsen, J., and Skinner, L. C.: Marine20 — the marine radiocarbon age calibration curve (0–55,000 cal BP), *Radiocarbon*, 62, 779–820, <https://doi.org/10.1017/RDC.2020.68>, 2020.
- Heaton, T. J., Bard, E., Ramsey, C. B., Butzin, M., Köhler, P., Muscheler, R., Reimer, P. J., and Wacker, L.: Radiocarbon: a key tracer for studying the Earth’s dynamo, climate system and carbon cycle and Sun, *Science*, 374, eabd7096, <https://doi.org/10.1126/science.abd7096>,
- 490 2021.
- Heaton, T. J., Butzin, M., Bard, E., Bronk Ramsey, C., Hughen, K. A., Köhler, P., and Reimer, P. J.: Marine Radiocarbon Calibration in Polar Regions: A Simple Approximate Approach using Marine20, *Radiocarbon*, 65, 848–875, <https://doi.org/10.1017/RDC.2023.42>, 2023.
- Henry, L. G., McManus, J. F., Curry, W. B., Roberts, N. L., Piotrowski, A. M., and Keigwin, L. D.: North Atlantic ocean circulation and abrupt climate change during the last glaciation, *Science*, 353, 470–474, <https://doi.org/10.1126/science.aaf5529>, 2016.
- 495 Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., Millero, F. J., Mordy, C., and Peng, T.-H.: A global ocean carbon climatology: results from global data analysis project (GLODAP), *Global Biogeochemical Cycles*, 18, GB4031, <https://doi.org/10.1029/2004GB002247>, 2004.
- Kimoto, K.: Planktic Foraminifera, pp. 129–178, Springer Japan, Tokyo, https://doi.org/10.1007/978-4-431-55130-0_7, 2015.
- Köhler, P. and Mulitza, S.: No detectable influence of the carbonate ion effect on changes in stable carbon isotope ratios ($\delta^{13}\text{C}$) of shallow dwelling planktic foraminifera over the past 160 kyr, *Climate of the Past*, 20, 991–1015, <https://doi.org/10.5194/cp-20-991-2024>, 2024.
- 500 Köhler, P. and Munhoven, G.: Late Pleistocene carbon cycle revisited by considering solid Earth processes, *Paleoceanography and Paleoclimatology*, 35, e2020PA004 020, <https://doi.org/10.1029/2020PA004020>, 2020.
- Köhler, P., Muscheler, R., and Fischer, H.: A model-based interpretation of low frequency changes in the carbon cycle during the last 120,000 years and its implications for the reconstruction of atmospheric $\Delta^{14}\text{C}$, *Geochemistry, Geophysics, Geosystems*, 7, Q11N06, <https://doi.org/10.1029/2005GC001228>, 2006.
- 505 Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H.: A 156 kyr smoothed history of the atmospheric greenhouse gases CO₂, CH₄, and N₂O and their radiative forcing, *Earth System Science Data*, 9, 363–387, <https://doi.org/10.5194/essd-9-363-2017>, 2017.
- Köhler, P., Adolphi, F., Butzin, M., and Muscheler, R.: Toward reconciling radiocarbon production rates with carbon cycle changes of the last 55,000 years, *Paleoceanography and Paleoclimatology*, 37, e2021PA004 314, <https://doi.org/10.1029/2021PA004314>, 2022.
- 510 Köhler, P., Skinner, L. C., and Adolphi, F.: Radiocarbon cycle revisited by considering the bipolar seesaw and benthic ^{14}C data, *Earth and Planetary Science Letters*, 640, 118 801, <https://doi.org/10.1016/j.epsl.2024.118801>, 2024a.

- Köhler, P., Skinner, L. C., and Adolphi, F.: Splines of reconstructed deep ocean ^{14}C ages over that last 25 kyr and simulation results of the carbon and radiocarbon cycle across the last 55 kyr, PANGAEA, [data set], <https://doi.org/10.1594/PANGAEA.967149>, 2024b.
- 515 Köhler, P., Menviel, L., Pöppelmeier, F., Heaton, T. J., Bard, E., and Skinner, L. C.: Simulation results of 2 EMICs for marine reservoir age determination using ^{14}C during the last 55 kyr, PANGAEA, [data set], <https://doi.org/10.1594/PANGAEA.987388>, 2026.
- Lambert, F., Bigler, M., Steffensen, J. P., Hutterli, M., and Fischer, H.: Centennial mineral dust variability in high-resolution ice core data from Dome C, Antarctica, *Climate of the Past*, 8, 609–623, <https://doi.org/10.5194/cp-8-609-2012>, 2012.
- Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records, *Paleoceanography*, 20, PA1003, <https://doi.org/10.1029/2004PA001071>, 2005.
- 520 Martrat, B., Grimalt, J. O., Shackleton, N. J., de Abreu, L., Hutterli, M. A., and Stocker, T. F.: Four climate cycles of recurring deep and surface water destabilizations on the Iberian Margin, *Science*, 317, 502–507, <https://doi.org/10.1126/science.1139994>, 2007.
- Mârza, A.-C., Menviel, L., and Skinner, L. C.: Towards the construction of regional marine radiocarbon calibration curves: an unsupervised machine learning approach, *Geochronology*, 6, 503–519, <https://doi.org/10.5194/gchron-6-503-2024>, 2024.
- Menking, J. A., Shackleton, S. A., Bauska, T. K., Buffen, A. M., Brook, E. J., Barker, S., Severinghaus, J. P., Dyonisius, M. N., and Petrenko, V. V.: Multiple carbon cycle mechanisms associated with the glaciation of Marine Isotope Stage 4, *Nature Communications*, 13, 5443, <https://doi.org/10.1038/s41467-022-33166-3>, 2022.
- 525 Menviel, L., Timmermann, A., Friedrich, T., and England, M. H.: Hindcasting the continuum of Dansgaard/Oeschger variability: mechanisms, patterns and timing, *Climate of the Past*, 10, 63–77, <https://doi.org/10.5194/cp-10-63-2014>, 2014.
- Menviel, L., Capron, E., Govin, A., Dutton, A., Tarasov, L., Abe-Ouchi, A., Drysdale, R. N., Gibbard, P. L., Gregoire, L., He, F., Ivanovic, R. F., Kageyama, M., Kawamura, K., Landais, A., Otto-Bliesner, B. L., Oyabu, I., Tzedakis, P. C., Wolff, E., and Zhang, X.: The penultimate deglaciation: protocol for Paleoclimate Modelling Intercomparison Project (PMIP) phase 4 transient numerical simulations between 140 and 127 ka, version 1.0, *Geoscientific Model Development*, 12, 3649–3685, <https://doi.org/10.5194/gmd-12-3649-2019>, 2019.
- 530 Menviel, L., Govin, A., Avenas, A., Meissner, K. J., Grant, K. M., and Tzedakis, P. C.: Drivers of the evolution and amplitude of African Humid Periods, *Communications Earth & Environment*, 2, 237, <https://doi.org/10.1038/s43247-021-00309-1>, 2021.
- 535 Menviel, L. C., Skinner, L. C., Tarasov, L., and Tzedakis, P. C.: An ice–climate oscillatory framework for Dansgaard–Oeschger cycles, *Nature Reviews Earth & Environment*, 1, 677–693, <https://doi.org/10.1038/s43017-020-00106-y>, 2020.
- NGRIP Members: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, *Nature*, 431, 147–151, <https://doi.org/10.1038/nature02805>, 2004.
- Oeschger, H., Siegenthaler, U., Schotterer, U., and Gugelmann, A.: A box diffusion model to study the carbon dioxide exchange in nature, *Tellus*, 27, 168–192, <https://doi.org/10.1111/j.2153-3490.1975.tb01671.x>, 1975.
- 540 Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., Doney, S. C., Dunne, J. P., Dutay, J.-C., Graven, H., Griffies, S. M., John, J. G., Joos, F., Levin, I., Lindsay, K., Matear, R. J., McKinley, G. A., Mouchet, A., Oschlies, A., Romanou, A., Schlitzer, R., Tagliabue, A., Tanhua, T., and Yool, A.: Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP), *Geoscientific Model Development*, 10, 2169–2199, <https://doi.org/10.5194/gmd-10-2169-2017>, 2017.
- 545 Parekh, P., Joos, F., and Müller, S. A.: A modeling assessment of the interplay between aeolian iron fluxes and iron-binding ligands in controlling carbon dioxide fluctuations during Antarctic warm events, *Paleoceanography*, 23, PA4202, <https://doi.org/10.1029/2007PA001531>, 2008.

- Pedro, J., Andersson, C., Vettoretti, G., Voelker, A., Waelbroeck, C., Dokken, T., Jensen, M., Rasmussen, S., Sessford, E., Jochum, M., and Nisancioglu, K.: Dansgaard-Oeschger and Heinrich event temperature anomalies in the North Atlantic set by sea ice, frontal position and thermocline structure, *Quaternary Science Reviews*, 289, 107–599, <https://doi.org/10.1016/j.quascirev.2022.107599>, 2022.
- 550 Peltier, W. R.: Ice Age Paleotopography, *Science*, 265, <https://doi.org/10.1126/science.265.5169.195>, 1994.
- Peltier, W. R., Argus, D. F., and Drummond, R.: Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model, *Journal of Geophysical Research: Solid Earth*, 120, 450–487, <https://doi.org/10.1002/2014JB011176>, 2015.
- Pöppelmeier, F., Jeltsch-Thömmes, A., Lippold, J., Joos, F., and Stocker, T. F.: Multi-proxy constraints on Atlantic circulation dynamics since the last ice age, *Nature Geoscience*, 16, 349–356, <https://doi.org/10.1038/s41561-023-01140-3>, 2023a.
- 555 Pöppelmeier, F., Joos, F., and Stocker, T. F.: The Coupled Ice Sheet–Earth System Model Bern3D v3.0, *Journal of Climate*, 36, 7563–7582, <https://doi.org/10.1175/JCLI-D-23-0104.1>, 2023b.
- Rasmussen, S. O., Bigler, M., Blockley, S. P., Blunier, T., Buchardt, S. L., Clausen, H. B., Cvijanovic, I., Dahl-Jensen, D., Johnsen, S. J., Fischer, H., Gkinis, V., Guillevic, M., Hoek, W. Z., Lowe, J. J., Pedro, J. B., Popp, T., Seierstad, I. K., Steffensen, J. P., Svensson, A. M., Vallelonga, P., Vinther, B. M., Walker, M. J., Wheatley, J. J., and Winstrup, M.: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy, *Quaternary Science Reviews*, 106, 14–28, <https://doi.org/10.1016/j.quascirev.2014.09.007>, 2014.
- 560 Reimer, P. J. and Reimer, R. W.: A Marine Reservoir Correction Database and On-Line Interface, *Radiocarbon*, 43, 461–463, <https://doi.org/10.1017/S0033822200038339>, 2001.
- 565 Reimer, P. J., Austin, W. E. N., Bard, E., Bayliss, A., Blackwell, P. G., Bronk Ramsey, C., Butzin, M., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Hajdas, I., Heaton, T. J., Hogg, A. G., Hughen, K. A., Kromer, B., Manning, S. W., Muscheler, R., Palmer, J. G., Pearson, C., van der Plicht, H., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Turney, C. S. M., Wacker, L., Adolphi, F., Büntgen, U., Capano, M., Fahrni, S., Fogtmann-Schulz, A., Friedrich, R., Köhler, P., Kudsk, S., Miyake, F., Olsen, J., Reinig, F., Sakamoto, M., Sookdeo, A., and Talamo, S.: The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP), *Radiocarbon*, 62, 725–757, <https://doi.org/10.1017/RDC.2020.41>, 2020.
- 570 Rubin, S. I. and Key, R. M.: Separating natural and bomb-produced radiocarbon in the ocean: The potential alkalinity method, *Global Biogeochemical Cycles*, 16, 52–1–52–19, <https://doi.org/10.1029/2001GB001432>, 2002.
- Siani, G., Michel, E., De Pol-Holz, R., DeVries, T., Lamy, F., Carel, M., Isguder, G., Dewilde, F., and Laurantou, A.: Carbon isotope records reveal precise timing of enhanced Southern Ocean upwelling during the last deglaciation, *Nature Communications*, 4, 2758, <https://doi.org/10.1038/ncomms3758>, 2013.
- 575 Simon, Q., Thouveny, N., Bourlès, D. L., Valet, J.-P., and Bassinot, F.: Cosmogenic ^{10}Be production records reveal dynamics of geomagnetic dipole moment (GDM) over the Laschamp excursion (20–60 ka), *Earth and Planetary Science Letters*, 550, 116–547, <https://doi.org/10.1016/j.epsl.2020.116547>, 2020.
- Skinner, L. C. and Bard, E.: Radiocarbon as a Dating Tool and Tracer in Paleoceanography, *Reviews of Geophysics*, 60, e2020RG000720, <https://doi.org/10.1029/2020RG000720>, 2022.
- 580 Skinner, L. C., Waelbroeck, C., Scrivner, A. E., and Fallon, S. J.: Radiocarbon evidence for alternating northern and southern sources of ventilation of the deep Atlantic carbon pool during the last deglaciation, *Proceedings of the National Academy of Sciences*, 111, 5480–5484, <https://doi.org/10.1073/pnas.1400668111>, 2014.

- 585 Skinner, L. C., Muschitiello, F., and Scrivner, A. E.: Marine Reservoir Age Variability Over the Last Deglaciation: Implications for Marine Carbon Cycling and Prospects for Regional Radiocarbon Calibrations, *Paleoceanography and Paleoclimatology*, 34, 1807–1815, <https://doi.org/10.1029/2019PA003667>, 2019.
- Skinner, L. C., Freeman, E., Hodell, D., Waelbroeck, C., Vazquez Riveiros, N., and Scrivner, A. E.: Atlantic Ocean Ventilation Changes Across the Last Deglaciation and Their Carbon Cycle Implications, *Paleoceanography and Paleoclimatology*, 36, e2020PA004074, <https://doi.org/10.1029/2020PA004074>, 2021.
- 590 Skinner, L. C., Primeau, F., Letsch-Thömmes, A., Joos, F., Köhler, P., and Bard, E.: Rejuvenating the ocean: mean ocean radiocarbon, CO₂ release, and radiocarbon budget closure across the last deglaciation, *Climate of the Past*, 19, 2177–2202, <https://doi.org/10.5194/cp-19-2177-2023>, 2023.
- Soulet, G., Skinner, L. C., Beaupré, S. R., and Galy, V.: A Note on Reporting of Reservoir ¹⁴C Disequilibria and Age Offsets, *Radiocarbon*, 58, 205–211, <https://doi.org/10.1017/RDC.2015.22>, 2016.
- 595 Stern, J. V. and Lisiecki, L. E.: North Atlantic circulation and reservoir age changes over the past 41,000 years, *Geophysical Research Letters*, 40, 3693–3697, <https://doi.org/10.1002/grl.50679>, 2013.
- Stocker, T. F. and Johnsen, S. J.: A minimum thermodynamic model for the bipolar seesaw, *Paleoceanography*, 18, 1087, <https://doi.org/10.1029/2003PA000920>, 2003.
- Stuiver, M. and Quay, P. D.: Atmospheric ¹⁴C changes resulting from fossil fuel CO₂ release and cosmic ray flux variability, *Earth and Planetary Science Letters*, 53, 349–362, [http://dx.doi.org/10.1016/0012-821X\(81\)90040-6](http://dx.doi.org/10.1016/0012-821X(81)90040-6), 1981.
- 600 Stuiver, M., Pearson, G. W., and Braziunas, T.: Radiocarbon age calibration of marine samples back to 9000 cal yr BP, *Radiocarbon*, 28, 980–1021, <https://doi.org/10.1017/S0033822200060264>, 1986.
- Suess, H. E.: Radiocarbon Concentration in Modern Wood, *Science*, 122, 415–417, <https://doi.org/10.1126/science.122.3166.415-a>, 1955.
- Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., and Poulsen, C. J.: Glacial cooling and climate sensitivity revisited, *Nature*, 605, 584, 569–573, <https://doi.org/10.1038/s41586-020-2617-x>, 2020.
- Tschumi, T., Joos, F., and Parekh, P.: How important are Southern Hemisphere wind changes for low glacial carbon dioxide? A model study, *Paleoceanography*, 23, PA4208, <https://doi.org/10.1029/2008PA001592>, 2008.
- Wu, J., Mollenhauer, G., Stein, R., Köhler, P., Hefter, J., Fahl, K., Grotheer, H., Wei, B., and Nam, S.-I.: Deglacial release of petrogenic and permafrost carbon from the Canadian Arctic impacting the carbon cycle, *Nature Communications*, 13, 7172, <https://doi.org/10.1038/s41467-022-34725-4>, 2022.
- 610 Zhang, X., Lohmann, G., Knorr, G., and Purcell, C.: Abrupt glacial climate shifts controlled by ice sheet changes, *Nature*, 512, 290–294, <https://doi.org/10.1038/nature13592>, 2014.

Table 1. Overview on used models and simulation scenarios. The control (CTRL) scenarios have no abrupt AMOC changes, but AMOC changes rapidly in “abrupt” scenarios. The “top layers” indicate for which top surface water depth the MRA have been calculated. The target was 50, 100 and 200 m, but results differ due to different model grids.

Model	Type	Scenario		Used Top Layers and Comment	Citation
		CTRL	abrupt		
BICYCLE-SE	box model	A0	A3	100 m (non-polar); 1000 m (polar)	Köhler et al. (2024a)
LOVECLIM	EMIC	nofwf	fwf	54 m, 104 m, 252 m	this study
Bern3D	EMIC	Pnofwf	PallSTD	64 m, 99 m, 216 m	this study
LSG	OGCM	med	-	50 m, median of nine simulations without abrupt AMOC changes	Butzin et al. (2020)

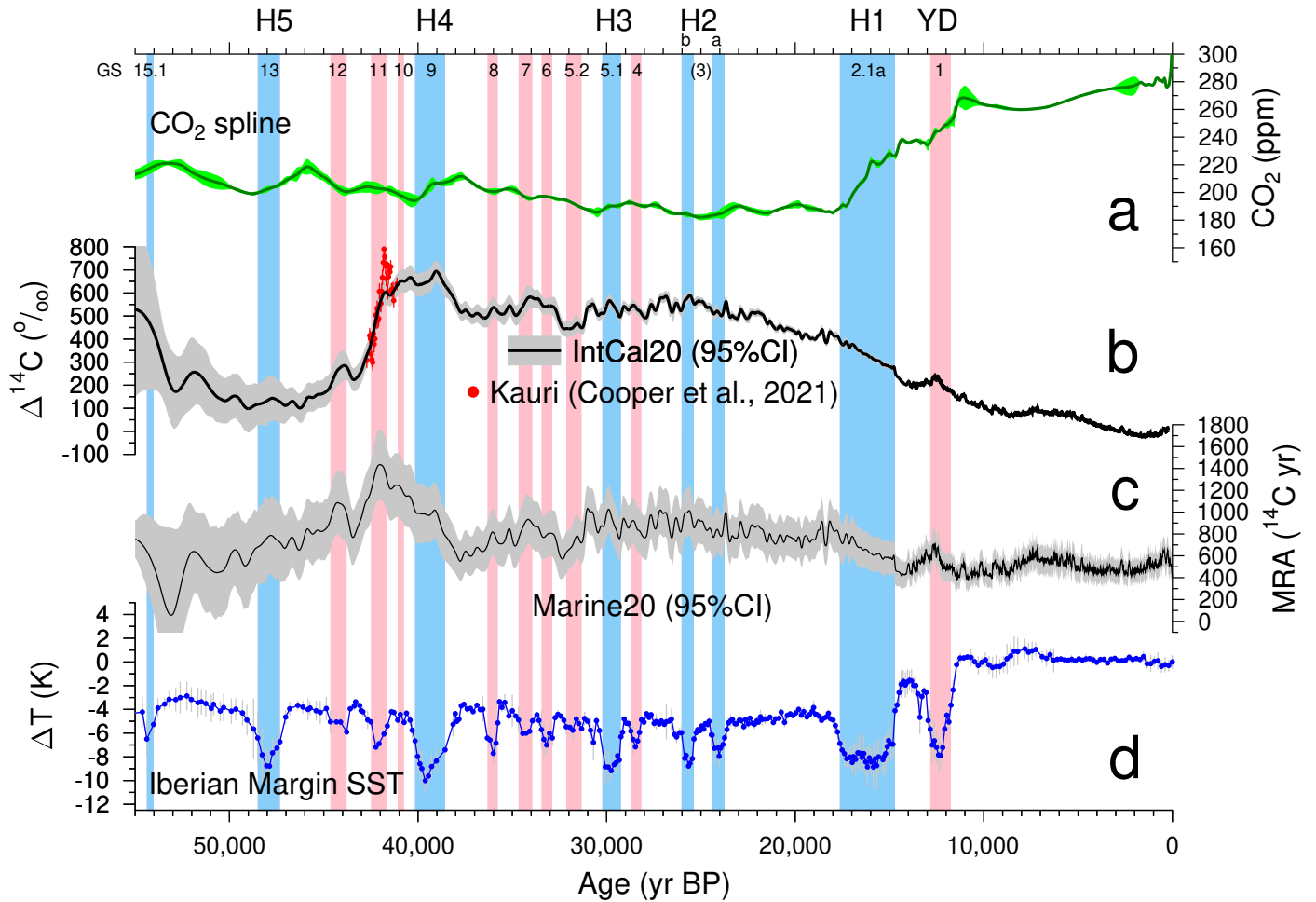


Figure 1. Relevant time series across the last 55 kyr. (a) Atmospheric CO₂ spline based on multi-records (Köhler et al., 2017) as used by the models. (b) Atmospheric $\Delta^{14}\text{C}$ from IntCal20 (Reimer et al., 2020) and (red points) an added extension from new kauri-based data around 42 cal kBP (Cooper et al., 2021). (c) Non-polar marine reservoir age (MRA) Marine20 (Heaton et al., 2020). For CO₂, IntCal20 and Marine20 the mean values and the 95% CI are plotted. Vertical bands mark Heinrich events (blue) or non-Heinrich stadials (pink) defined by (d) Iberian Margin SST (mean of UK37' and RI-OH' SST records) (Davtian and Bard, 2023). Heinrich events and the Younger Dryas (YD) and Greenland stadials (GS, Rasmussen et al. (2014)) with their numbers are labelled on the top ignoring GS 2.1b and 2.1c which fall into the LGM.

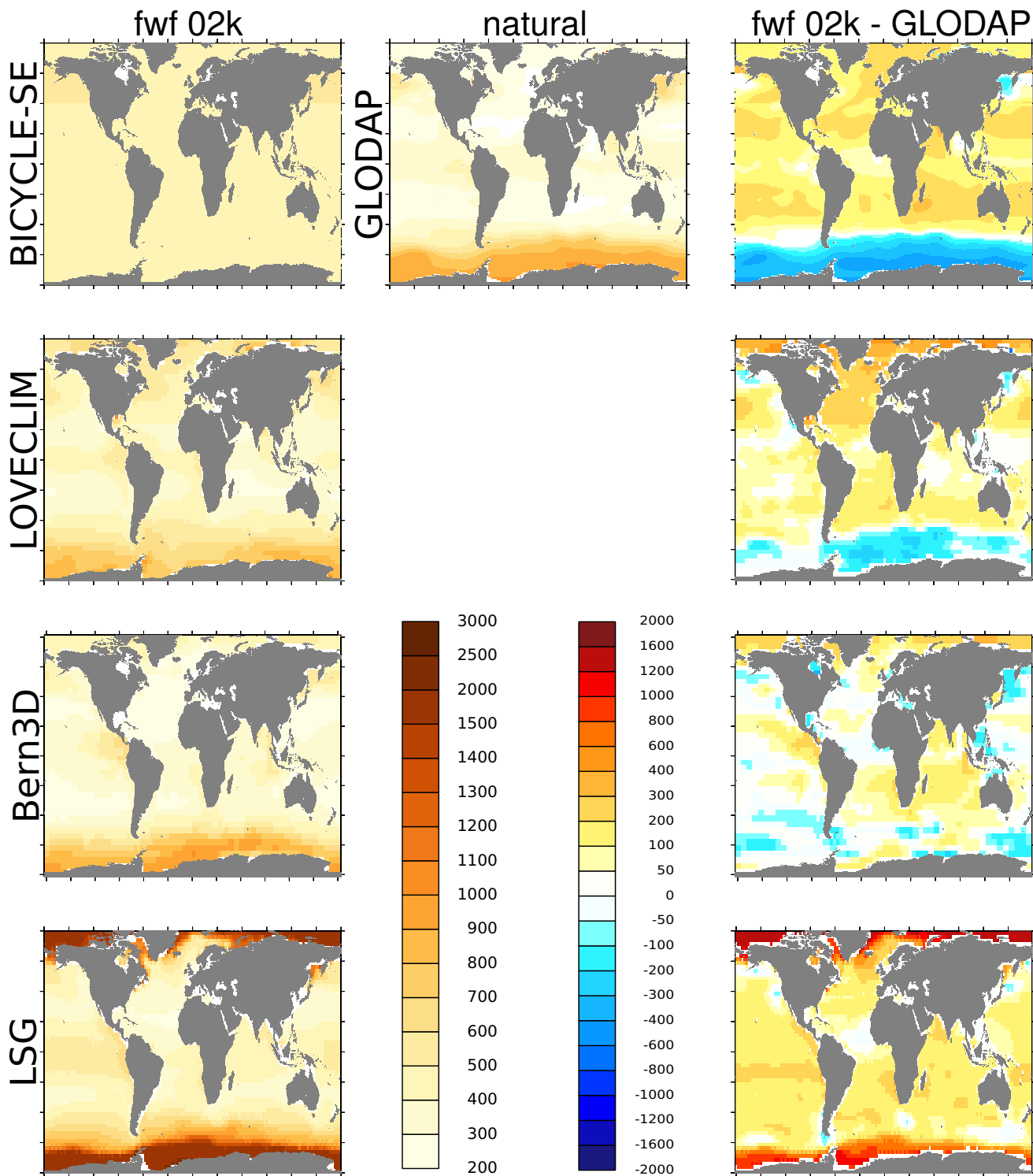


Figure 2. Surface MRA (^{14}C yr) from all models (left) at 2 kyr BP for runs with forced abrupt AMOC changes (scenarios: A3@BICYCLE-SE, fwf@LOVECLIM; PallSTD@Bern3D) and med@LSG; (right) differences in surface MRA from natural GLODAP (top middle). BICYCLE-SE: surface boxes; all else: mean values of roughly the top 50 m. Use left color-code (brownish) for absolute values and right color-code (blue-to-red) for differences.

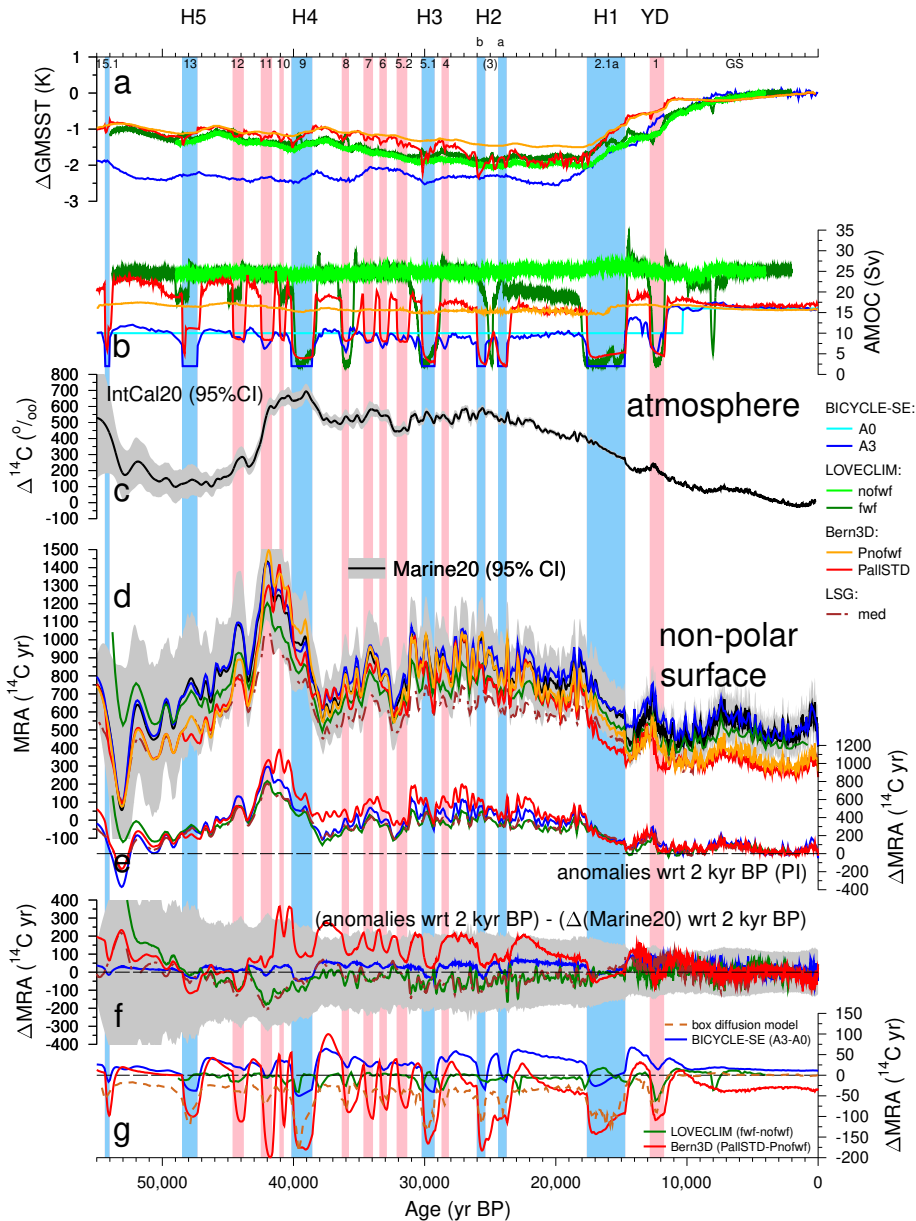


Figure 3. Combining model outputs with focus on non-polar surface ocean (see SI for time series sorted by model). Changes in (a) global mean sea surface temperature (GMSST) with respect to (wrt) to the most recent points of the time series and (b) AMOC strength in the investigated [model-runs](#) models. [Note that the way they are calculated is model-dependent. It is the prescribed strength of the overturning cell in the Atlantic in BICYCLE-SE and the maximum of the meridional overturning streamfunction in the North Atlantic in LOVECLIM and Bern3D, but for the latter only at depths greater than 400 m to exclude the wind driven \(sub\)surface part.](#) (c) Atmospheric $\Delta^{14}\text{C}$ from IntCal20 (Reimer et al., 2020) for comparison. (d) MRA of the non-polar surface ocean (surface box or roughly the top 50 m) ranging from about 50°S to 50°N for the different simulations and Marine20 (Heaton et al., 2020). Differences in simulated non-polar surface MRA (e) wrt 2 kyr BP (PI) and (f) to Marine20 wrt 2 kyr BP. Grey background is the 95% CI of Marine20. (g) Differences in non-polar surface MRA between simulations with and without abrupt AMOC changes for BICYCLE-SE, LOVECLIM, Bern3D. Additionally, ΔMRA from a box diffusion model (Bard, 1988) in which eddy diffusivity is linearly related to our AMOC proxy (Iberian Margin SST of Davtian and Bard, 2023). Annual output from LOVECLIM for GMSST and AMOC is plotted as 50-yr running mean, output of Bern3D comes in timesteps of 50 yr. Vertical bands mark Heinrich events (blue) or non-Heinrich stadials (pink), see caption to Figure 1 for details. Results from simulations without abrupt AMOC changes have been omitted in (d–f) for clarity, but are found in the SI (Figures S1–S3).

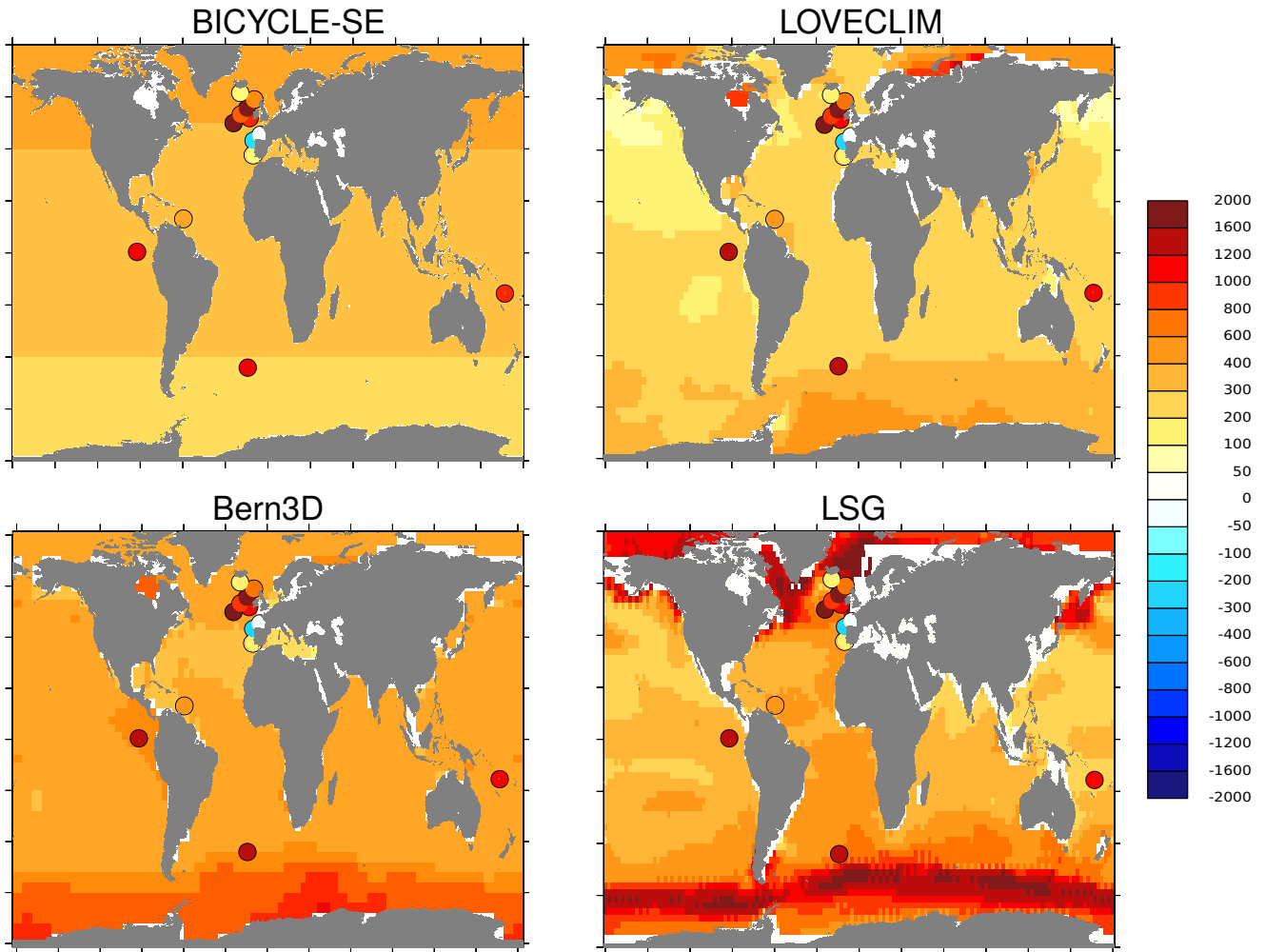


Figure 4. Difference in surface MRA (^{14}C yr) in all models (BICYCLE-SE: surface box; all else: roughly the top 50 m) between 20 kyr BP and 02 kyr BP (20 – 02 kyr BP) for runs with abrupt AMOC changes (scenarios: A3@BICYCLE-SE, fwf@LOVECLIM; PallSTD@Bern3D) and med@LSG. The 19 points in each panel are the difference for the LGM (19–21.8 kyr BP) based on reconstructions (Skinner et al., 2023) from natural GLODAP.

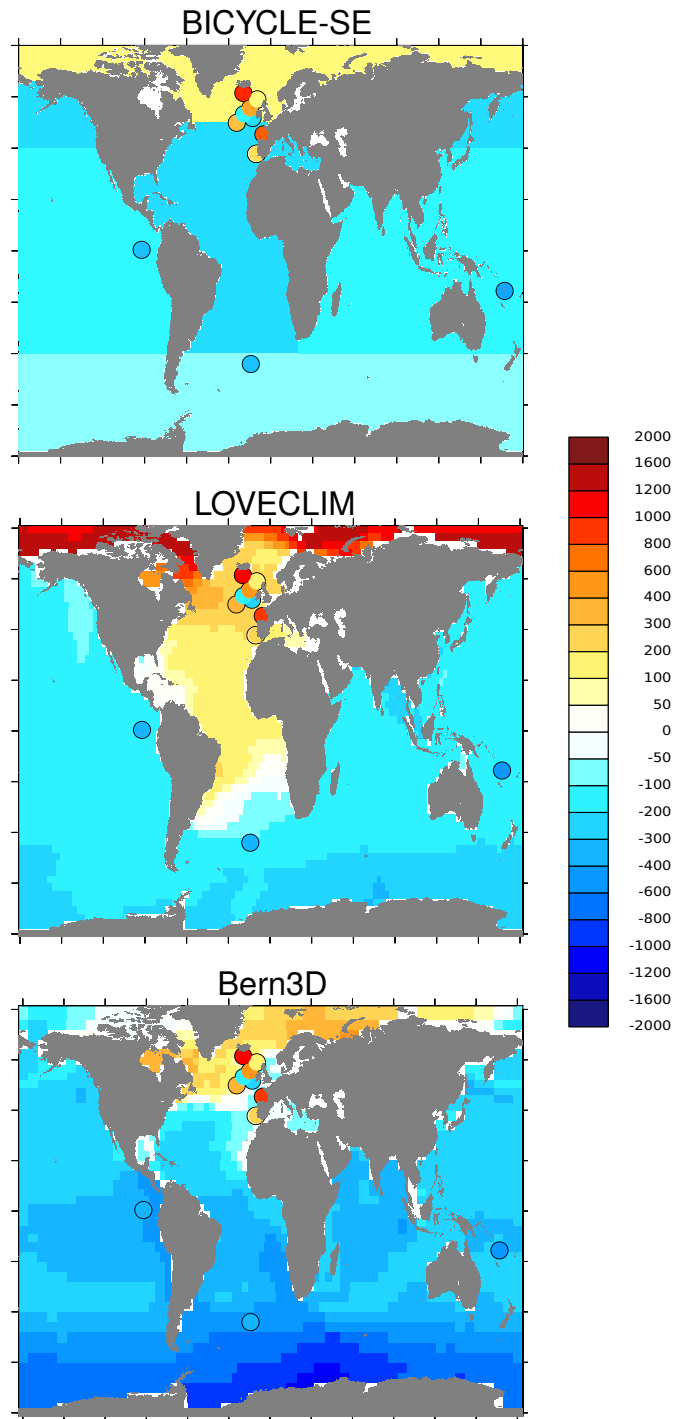


Figure 5. Differences in surface MRA (^{14}C yr) from all models but LSG (BICYCLE-SE: surface box; all else: roughly the top 50 m) between 16 kyr BP and 20 kyr BP (16 – 20 kyr BP) for runs with abrupt AMOC changes (scenarios: A3@BICYCLE-SE, fwf@LOVECLIM; PallSTD@Bern3D). The 13 points in each panel are differences in MRA for HS1 (15–17.5 kyr BP) – LGM (19–21.8 kyr BP) based on reconstructions (Skinner et al., 2023).

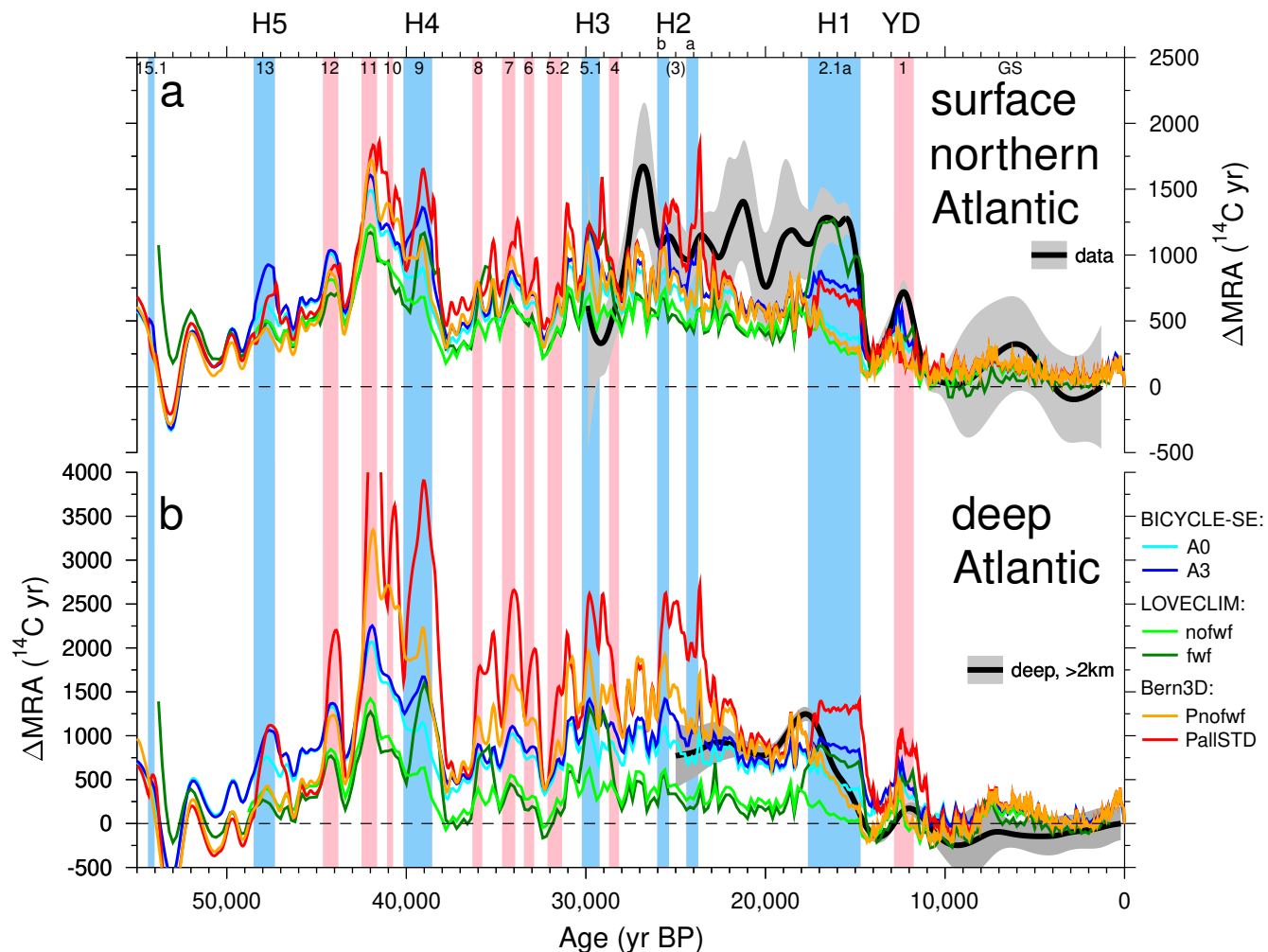


Figure 6. MRA changes with respect to the most recent points of the time series in simulated (a) surface northern ($>50^\circ$ N) and deep (b) Atlantic. Combining all model outputs with focus on the Atlantic. The surface North Atlantic MRA (surface box or roughly the top 50 m) covers all north of 50°N including the Arctic Ocean. The deep Atlantic contains water below 2 km water depth. Data source: surface North Atlantic MRA: Skinner et al. (2019); deep Atlantic: Skinner et al. (2023) as compiled in Köhler et al. (2024a). Vertical bands mark Heinrich events (blue) or non-Heinrich stadials (pink), see caption to Figure 1 for details.

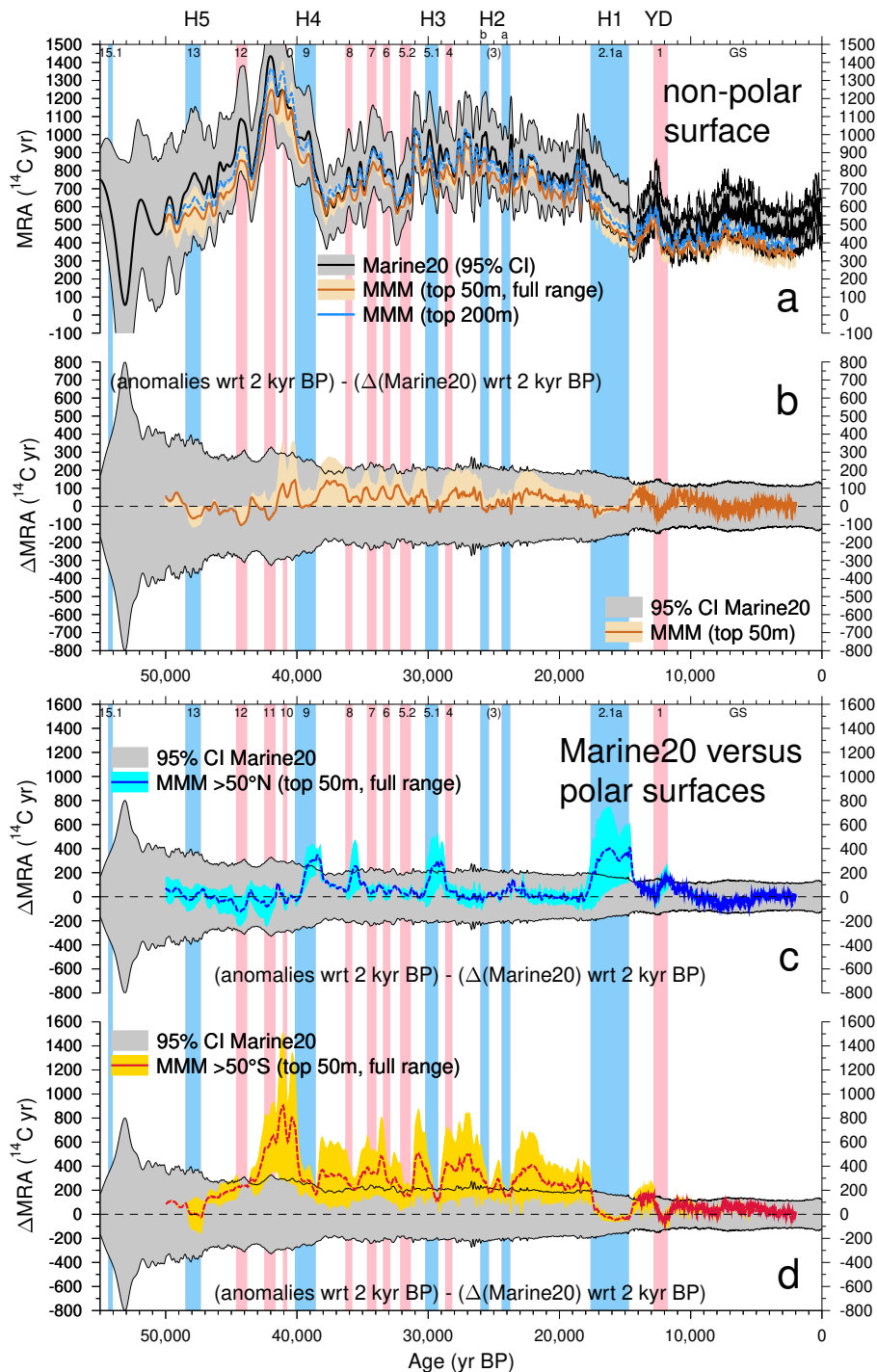


Figure 7. (a) Multi-model mean (MMM) of non-polar surface MRA from the 2 EMICs with simulations containing abrupt AMOC changes (fwf@LOVECLIM; PallSTD@Bern3D) in comparison to Marine20. Difference of MMM with respect to (wrt) 2 kyr BP (PI) of (b) non-polar surface MRA and of (c) northern and (d) southern polar surface ocean to changes in MRA of the mean of Marine20 also wrt 2 kyr BP. Light coloured areas show 95% CI for Marine20 and the full range of model results for the MMM. MMM is based on MRA results of roughly the top 50 m of the water column. Subfigure (a) contains also a version of the MMM based on roughly the top ~200 m. The MMM is only shown between 50 kyr BP and 2 kyr BP, since before 50 kyr BP spin-up effects take places, and at 2 kyr BP the LOVECLIM simulation ends. Vertical bands mark Heinrich events (blue) or non-Heinrich stadials (pink), see caption to Figure 1 for details.