Abrupt Climate and Weather Changes Across Time Scales

Gerrit Lohmann1,2, Martin Butzin1, Nina Eissner1, Xiaoxu Shi1, and Christian Stepanek1

1Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, 2Department of Environmental Physics, University of Bremen, Bremen, Germany

Abstract The past provides evidence of abrupt climate shifts and changes in the frequency of climate and weather extremes. We explore the nonlinear response to orbital forcing and then consider climate millennial variability down to daily weather events. Orbital changes are translated into regional responses in temperature, where the precessional response is related to nonlinearities and seasonal biases in the system. We question regularities found in climate events by analyzing the distribution of interevent waiting times. Periodicities of about 900 and 1,150 yr are found in ice cores besides the prominent 1,500 yr cycle. However, the variability remains indistinguishable from a random process, suggesting that centennial-to-millennial variability is stochastic in nature. New numerical techniques are developed allowing for a high resolution in the dynamically relevant regions like coasts, major upwelling regions, and high latitudes. Using this model, we find a strong sensitivity of the Atlantic meridional overturning circulation depending on where the deglacial meltwater is injected into. Meltwater into the Mississippi and near Labrador hardly affect the large-scale ocean circulation, whereas subpolar hosing mimicking icebergs yields a quasi shutdown. The same multiscale approach is applied to radiocarbon simulations enabling a dynamical interpretation of marine sediment cores. Finally, abrupt climate events also have counterparts in the recent climate records, revealing a close link between climate variability, the statistics of North Atlantic weather patterns, and extreme events.

Plain Language Summary Predicting the future spread of possible climates, the risk of climate extremes and the risk of rapid transitions is of high socioeconomic relevance. The past provides evidence of abrupt climate change and the frequency of extremes. This allows to separate anthropogenic signals from natural climate variability. Earth system models applied both to past and future scenarios will enhance our ability to detect regime shifts which are necessary to potentially predict climate extremes and transitions. We consider the response of the system to regular orbital forcing and then focus on shorter time scales down to weather. The appearance of precession is linked to nonlinear responses of the climate system to external orbital forcing. Furthermore, we find that centennial-to-millennial variability is stochastic in nature. We also discuss recent developments of climate models with superior resolution in typical retrieval regions of paleoclimate records, such as continental margins and coasts. Using this model, we find a strong sensitivity of the Atlantic meridional overturning circulation depending on where the deglacial meltwater is injected into. Meltwater into the Mississippi and near Labrador hardly affect the large-scale ocean circulation, whereas subpolar hosing related to icebergs yields a quasi shutdown. Our multiscale approach is applied to radiocarbon simulations enabling a dynamical interpretation of marine sediment cores. Finally, we explore a close link between climate variability, the statistics of weather patterns, and extreme events.

1. Introduction

Weather and climate vary on broad ranges of spatial and temporal scales. This is evident from observations and simulations of the present climate as well as from climate history as recorded in geological and glacio-logical archives (Bradley, 2014; Crowley & North, 1991; Mitchell, 1976; Saltzman, 2002). Climate and Earth system models are widely used to evaluate the impact of anthropogenic emissions on future climate. The validation of these models by simulating different climate scenarios is essential to understand the sensitivity of the climate system to external forcing. The models are clearly unrivaled in their ability to simulate a broad range of large-scale phenomena on seasonal to decadal time scales (Flato et al., 2013). However, the reliability of models to simulate climate variability on multidecadal and longer time scales requires additional
Figure 1. Schematic diagram of the spatiotemporal scales considered. DO = Dansgaard-Oeschger, H = Heinrich events; AMO = Atlantic Multidecadal Oscillation; PDO = Pacific Decadal Oscillation; ENSO = El Niño–Southern Oscillation. The annual and astronomical cycles are externally driven and have quasi-global impact.

Evaluation (Haywood et al., 2019). Climate records derived from paleoenvironmental parameters facilitate the testing of models out of the “comfort zone of present-day climate,” or, in other words, one application of paleoclimate is to validate state-of-the-art coupled climate models for past time slices and past climate transitions.

In this review, we will highlight key areas where major future efforts are needed to address fundamental questions related to abrupt changes of climate and weather patterns across a range of time scales (Figure 1). As climate can rapidly change, potential thresholds for the occurrence of climate extremes and transitions must be estimated. During the most recent glacial period, the Northern Hemisphere was subject to abrupt large-scale transitions between cold (stadial) and warmer (interstadial) conditions (Figure 2a). These shifts are well documented by isotope records from Greenland ice cores and are known as Dansgaard-Oeschger (DO) events (Dansgaard et al., 1993). They typically show an abrupt warming of up to 10–15 K within a few decades followed by a gradual cooling that stretches over several hundred to several thousand years. We evaluate the stochastic nature of this multicentennial-to-millennial climate variability and will present a Bayesian methodology to quantify the uncertainties. Such concepts are also useful to interpret decadal and multidecadal variability, linked to convection events in the Labrador Sea.

The pacing of regional to global climate variations on time scales of tens of thousands of years is related to changes in the seasonal and latitudinal sunlight distribution induced by variations in the Earth’s orbital parameters (e.g., Berger, 1978; Laskar et al., 2004; Milankovitch, 1941). A key element of the orbital theory is that summer insolation at high latitudes of the Northern Hemisphere determines glacial-interglacial transitions connected with the waxing and waning of large continental ice sheets (e.g., Imbrie & Imbrie, 1980). During the last two million years, these glacial-interglacial cycles provide the dominant signal in the climate record. The annual mean insolation forcing shows the obliquity cycle at a time scale of ~40 kyr (Figure 2b). While precessional forcing (~20 kyr variations; kyr = 1,000 yr) is clearly visible for seasonal forcing (Figure 2a), it is 0 for local annual mean insolation forcing (Figure 2b). Therefore, the precessional signal is an indicator for nonlinear responses to the forcing (such as stronger response in summer than winter) or for nonlocal responses, which raises the question of forcing and feedback mechanisms for paleoclimatic change (Brickman et al., 1999; Laepple & Lohmann, 2009; Short et al., 1991; Tziperman et al., 2006; Valdes & Glover, 1999). We are going to reveal the response to orbital forcing in a complex climate model. This model can be viewed as a benchmark test of the climate system, especially for the question of how orbital changes are translated into local temperature variations. This regular, external orbital forcing is strongly linked to changes in seasonality (Figure 1).
Paleoclimatic evidence suggests that some past climate shifts were associated with changes in North Atlantic Deep Water (NADW) formation at the end of the last ice age (Dansgaard et al., 1993; Lehman & Keigwin, 1992; Sarnthein et al., 1994). During the last deglaciation about 20–10 kyr before present, the climate has warmed due to insolation and rising carbon dioxide (CO₂) concentrations (Figures 2b and 2c). Considerable meltwater discharge from the large continental ice sheets over the Northern Hemisphere entered the North Atlantic. The most prominent meltwater event occurred at about 14.6 kyr BP (BP = before present) and was named Meltwater Pulse 1a (MWP-1a) (Fairbanks, 1989) with up to 0.5 Sv (1 Sv = 10⁶ m³ s⁻¹) (Clark et al., 2002) entering the North Atlantic with some possible contributions from the Arctic Ocean (Tarasov & Peltier, 2006) and Southern Ocean (Carlson & Clark, 2012; Weber et al., 2014). However, paleoclimate data indicate that NADW formation was not affected seriously (McManus et al., 2004), and the next major cooling phases occurred either earlier (during Heinrich-event 1 at about 17.5 kyr BP) or more than 1,000 yr later at the onset of the Younger Dryas between about 12.9 and 11.5 kyr BP (Clark et al., 2002). Analyses of coarse-grained ice-rafted debris and planktonic foraminifers revealed pronounced dropstone layers that have been deposited in the North Atlantic documenting ice rafting during cold events (Heinrich, 1988). At the end of the Younger Dryas, the North Atlantic realm again experienced rapid warming that ended up in Holocene climate conditions. In the discussion of probable causes, the melting and calving of continental
One of the greatest obstacles of resolving the leads and lags puzzle is the difficulty of developing an accurate time scale for long ice cores and marine sediments. The development of such a time scale would allow for testing of many climate forcing hypotheses, leading to significant advances in paleoclimate theory. In the late Quaternary, the backbone of such a time scale could be provided through radiocarbon ($^{14}$C) dating. Numerical models are key to quantify processes affecting $^{14}$C-based age models of climate records. This is particularly the case for marine climate records. While the ocean is the most important reservoir of heat and carbon regarding decadal to millennial-scale climate variability, a direct evaluation of $^{14}$C-dated marine records is hampered by various effects introducing dating uncertainties. Prior to the Holocene, these effects are poorly constrained and have to be inferred either from ad hoc assumptions or through modeling, which may involve ensemble approaches to explore its uncertainty range (e.g., Butzin et al., 2012a, 2017, 2020). To show an example, the blue curve in Figure 3 shows a conversion from $^{14}$C years to calendar years without the information guided by numerical modeling of the ocean-carbon system (e.g., Reimer et al., 2004). Using a $^{14}$C-equipped climate model where paleoceanographic information is assimilated (Butzin et al., 2012a), the black curve in Figure 3 displays a nonstationary relation between the simulated $^{14}$C years to calendar years relation. The difference between the resulting dates can be up to 1,000 yr and can vary in time. A further complication in radiocarbon dating is that marine records typically originate from continental margins, marginal seas, or tropical islands. Such sites may not be representative to mirror large- or global-scale processes of climatic change, demanding a valid extrapolation method to infer past global conditions, which makes a proper data-model comparison difficult (Lohmann et al., 2013a). High-resolution models are required to elucidate the causal chains in the climate system, notably during abrupt transitions of the last deglaciation, and provide a benchmark for future transitions under rapid CO₂ increase (cf. Figure 2c).

Another way to ascertain the extent of past changes is through the inspection of historical time series of direct temperature measurements or documentation of such environmental observations. The twentieth century showed a series of decadal-scale climate anomalies such as the “Great Salinity Anomaly” (GSA) observed in the subpolar North Atlantic in the late 1960s and the early 1970s (Dickson et al., 1988) or the Early Arctic Warming (EAW) in the 1920s and 1930s (Bengtsson et al., 2004; Schokalsky, 1936). The EAW might be seen as a regime shift in the North Atlantic ocean and atmosphere responsible for a regional warming (Bengtsson et al., 2004; Drinkwater, 2006; Tokinaga et al., 2017), whereas the GSA can be linked to a sea ice transport out of the Arctic through Fram Strait (Häkkinen & Geiger, 2000; Hilmer et al., 1998), which represents a major source of freshwater for the northern North Atlantic (Aagaard & Carmack, 1989; Schmith & Hansen, 2003). The GSA was accompanied by a subsequent decadal cooling in the 1970s over
the North Atlantic realm (Dima & Lohmann, 2007). These events could be considered as analogs to past Heinrich events which appeared during glacial times (Heinrich, 1988), whereas the EAW has a flavor of a DO-type warming over Greenland. The knowledge about potential drivers and the inherent dynamics is still poorly known. We show a potential link between synoptic atmospheric variability of days and decadal-to-multidecadal time scale variability, for example, the GSA, droughts and floods, extreme winters, and atmospheric blocking (Dickson et al., 1988). Increased sea ice export from the Arctic Ocean stabilizes the upper water column in the North Atlantic, diminishes the production of intermediate and deep water masses through ocean convection, and can influence the large-scale ocean circulation (Häkkinen, 1999). An important part of sea ice export from the Arctic is forced by specific atmospheric structures (Cavalieri, 2002; Häkkinen & Geiger, 2000; Hilmer et al., 1998; Ionita et al., 2016). These can be parts of coupled ocean-atmosphere modes of climate variability with large-scale projections.

Examples of abrupt transitions and extremes are known for the last century, and the multidecadal variability is coupled to the persistence of atmospheric blocking and synoptic variability (Pfahl et al., 2009; Rimbu & Lohmann, 2010, 2011; Rimbu et al., 2014; Rimbu, Lohmann, et al., 2016). Paleoclimate data allow to evaluate climate variability and extremes and their relation to the mean climate. A model-based approach will help to get a more consistent picture to obtain a common model-data interpretation of the underlying mechanisms and forcings of paleoweather events in the North Atlantic/European realm. High-resolution environmental archives bring the relatively short period of the instrumental record into a long-term context. Transitions can be detected with the highest possible resolution which can range from decadal up to seasonal or even synoptic time scales. We further propose future research directions for studying climate changes regarding sampling strategies, statistical tools, and the formulation of climate models, reviewing the challenges to understand the causes and contributors to climate shifts and extremes.

2. Temperature Signature Due to Orbital Variations

A fundamental feature of Earth’s climate during the last few million years is the regular fluctuation between different states. The driving effect is due to the geometry of the Earth’s orbit affecting the latitudinal and seasonal distribution of insolation (Berger, 1979; Laskar et al., 2004; Milankovitch, 1941). By computational analysis of the planetary system, these variations can be calculated to a high accuracy for the last millions of years (Laskar et al., 2004). Apart from shaping the seasonal cycle, the changes in the seasonal and latitudinal distribution of insolation are a primary driver for climate variability on multimillennial time scales (Huybers & Curry, 2006). This relationship has been hypothesized for a long time (Adhémar, 1842; Croll, 1875), and coherent variability of climate records and orbital parameters were found when geologists started to date climatic proxy records of ocean sediments (e.g., Broecker & van Donk, 1970; Hays et al., 1976). One classical concept for the response of the climate system on orbital forcing was proposed by Imbrie et al. (1992) in that insolation changes at high northern latitudes initiate the climate response. However, the question remains how the climate system reacts to the local insolation forcing.

To quantify the response of Earth’s climate to orbital forcing during the last 1 Myr (Myr =10⁶ yr), we produce a new numerical simulation of transient climate that is forced with the solution of orbital elements by Laskar et al. (2004). We employ the Community Earth System Models (COSMOS) framework to provide a numerical representation of the coupled atmosphere-ocean-land system, which is described in Appendix A. In order to compute monthly mean climate over the last 1 Myr, we prescribe a time series of eccentricity, obliquity, and precession of the Earth’s orbit. These values are sampled at a time resolution of 100 yr, which results in the acceleration of orbital forcing by a factor of 100. Such an acceleration method is necessary to derive the numerical solution of 1 Myr of climate in a coupled atmosphere-ocean general circulation model within a reasonable amount of time.

Our analysis is based on simulated changes in near surface air temperature (SAT), which reflect a direct response to orbitally modulated spatial and seasonal distribution of insolation at the top of the atmosphere. We study the climate’s response to forcing on orbital time scales. Hence, our focus is in particular on those modes of climate variability that are linked to the impact of precession, obliquity, and eccentricity of the Earth’s orbit around the Sun. We perform a spectral analysis of simulated transient near-surface temperature over the last 1 Myr and compute the integrated spectral power density in three different periodicity bands: precession band (18–26 kyr), obliquity band (38–56 kyr), and eccentricity band (90 to 110 kyr). The 420 kyr
orbital periodicity is not analyzed here because we concentrate on the last 1 Myr. A fast Fourier transform is applied on detrended time series at grid resolution of the atmosphere model across the entire globe. We provide three different graphical representations of the power spectral density for the last 1 Myr. We illustrate the latitudinal dependency of the zonally averaged power spectrum (Figure 4). This illustration highlights the substantial dependence of the climate’s response to latitude and periodicity of orbital forcing. The frequencies of precessional response are detected at ~23 and ~19 kyr, together with a first harmonic with half of the years lined to the semi-precession cycle with much lower amplitude. The precessional response is strongest in the subtropics but also visible at polar latitudes in the Northern Hemisphere and at high latitudes in the Southern Hemisphere. The precessional response is related to nonlinearities and/or seasonal biases in the climate system. Furthermore, we see power in the obliquity band (41, 29, and 54 kyr), mostly at high latitudes with a small maximum in the subtropics.

In order to see the frequency bands explicitly, geographical maps highlight regions that are particularly sensitive to a specific periodicity of orbital forcing (Figure 5). We provide the power spectrum density over periodicity and latitude, which is done for the entire surface of the planet as well as for land and ocean separately. Figure 5a shows that the precessional response dominates in subtropical Africa, India, South America, and some regions around Greenland. Obliquity (Figure 5b) shows a similar pattern, but with more emphasis on the region around Greenland and suppressed response across South America as compared to precession. The high latitude response is not that different over land and ocean. The 100 kyr band (Figure 5c) emphasizes again the subtropical region in northern Africa and some regions in the tropics and midlatitudes, but with a much smaller amplitude than precession and obliquity. The changes in eccentricity of the Earth’s orbit have a small effect on the global mean insolation but modulate the strength of the seasonality and therefore have an effect through the precession. The 100 kyr frequency band in part simply originates from an amplitude cycle of precession cycles, a linkage that may control the similarities shown by the local power spectrum of annual mean SAT (Figures 5a and 5c).

The frequency analysis is also applied to other climate phenomena like the ocean circulation (not shown). The wind-driven ocean circulation is dominated by obliquity (more a linear response), whereas the Atlantic meridional overturning and heat transport show pronounced precessional and obliquity peaks due to nonlinearities and seasonal biases toward winter. Our model integrations and analysis confirm that convection and deepwater formation serve to rectify the zero annual-mean precessional forcing, resulting in precessional energy in the ocean. The Antarctic Bottom Water also shows eccentricity peaks indicating rectification mechanisms.

The local approach complements the global Milankovitch/Imbrie hypothesis (climate is driven by northern summer insolation) in explaining observed climate variability and potentially offers new insights in interpreting paleoclimate records, especially beyond the 100 kyr responses. It shows the limitations of global calibration curves based on Northern Hemisphere insolation changes at high latitudes. The pronounced high-latitude response both in precession and obliquity emphasizes the important feedbacks and nonlinearities including the ocean circulation (deep water formation is happening in winter) and land (ice-albedo feedback). Actually, the ice-albedo feedback is underrepresented in our long-term simulation, as ice sheets are static. The climate-vegetation feedback has so far not been analyzed this in the 1 Myr simulation. Interesting is the variability in the subtropics. The insolation forcing favors a north-south seasonal march of the ITCZ (Intertropical Convergence Zone) and the inland penetration of Afro-Asian monsoon precipitation during boreal summer. Surface temperature is reduced in regions where precipitation is enhanced due to the combination of increased cloud cover and increased surface evaporation (Braconnot et al., 2019; Herold & Lohmann, 2009). Braconnot et al. (2019) demonstrate that the projection of temperature onto the insolation curve is larger in the Northern than in the Southern Hemisphere, especially between 10°N and 40°N where about 80% of the temperature signal is a direct response to the local insolation forcing. A completely different approach is followed by the template model (Laeppele & Lohmann, 2009) where insolation-driven temperature variability on orbital time scales relies on the modern relationship between insolation and temperature.
Figure 5. Power spectrum density for different frequency bands of annual mean SAT. (left) Zonal mean with solid lines for land and ocean, dashed line for land, and dashed-dotted lines for the ocean. (right) Local power spectrum of annual mean SAT. Note the different scales. The precessional band is dominated by a peak over North Africa (a), obliquity for a region around Greenland and the Barents Sea (b), and eccentricity with a much lower amplitude over North Africa and the Sea of Okhotsk (c).

throughout the year. Over extratropical continental areas, a linear response is detected, whereas over the ocean, the subtropics, and sea ice-sensitive regions, precessional bands at 19 and 23 kyr are visible related to nonlinearities such as a seasonal mixing regime. Such an empirical approach includes all fast feedback processes from the seasonal time scale. In a similar way, the climate sensitivity can be constrained from the seasonal cycle in temperature (Knutti et al., 2005). Qualitatively, COSMOS and the template model show similar patterns. Of course, the full climate model has “longer-term” feedbacks like the ice/snow-albedo
feedback on millennial time scales and nonlocal effects giving higher responses at most latitudes. However, there are also clear differences. The full climate model shows a strong response in the obliquity band in northern high latitudes, which is absent in the template model without long-term feedbacks. Our results are different to the results of Short et al. (1991), who used a linear two-dimensional energy balance model to study the spatially resolved temperature response of the last 800 kyr. However, as their model was linear, precession and eccentricity only appeared when using the maximum temperature as diagnostics. Our study shows that the climate archives recording annual mean temperature will also reflect these frequencies caused by nonlinearities in the climate model already.

3. Millennial Time Scale Variability

DO and Heinrich events are pronounced climatic changes over the last 120,000 yr. Although many of their properties were derived from climate reconstructions, the associated physical mechanisms are not yet fully understood. We are interested in the exact position in time where a change in temperature occurred on the given time scale. According to different DO criteria, we determined several sequences of DO onsets and are interested in any potential periodic recurrence. A standard technique would be a spectral analysis of the detrended time series and the identification of significant peaks in the frequency spectrum (e.g., Grootes & Stuiver, 1997). Here, we focus on the pacing of discrete DO onsets which is the reason why we apply a different measure of periodicity based on Huybers and Wunsch (2005) and Ditlevsen et al. (2007). It allows us to evaluate the phase stability of the discrete event onsets although we have a small sample size (Huybers & Wunsch, 2005). This of course also applies for quasiperiodic systems with an unpredictable component where events repeat themselves irregularly or are skipped. In Appendix B, we list the timing of the DO events for the last glacial (Table B1). There, we describe the threshold approach where we consider change points when a certain threshold criterion is fulfilled. To quantify such uncertainties, we furthermore provide a probabilistic methodology and use a Bayesian approach to identify the DO events in time. The algorithm in use for this study was provided by Ruggieri (2013) and has the advantage of quantifying uncertainties in number and location of warming events (Appendix B). Figure 6 shows the $\delta^{18}O$ signal of the NGRIP ice core from Greenland covering the last glacial-interglacial period covering the past 120 kyr and change point statistics.

Repeatedly, a periodicity of about 1,500 yr is reported in the literature, starting with the spectral analysis of the GISP2 and GRIP by Grootes and Stuiver (1997) with a peak at 1,470 yr (Yiou et al., 1997). Also with a focus on the recurrence interval of partially skipped warming events in the Greenland record an interevent waiting time of 1,470 yr is announced (Alley et al., 2001; Rahmstorf, 2003; Schulz, 2002). Ditlevsen et al. (2007) include also the Bolling/Allerød period and Event 9 which was reported by Rahmstorf (2003) after applying a low pass filter and defining a 2‰ anomaly in a 200 yr time interval as a DO criterion.

Figure 6. Bayesian change point analysis for the $\delta^{18}O$ signal of the NGRIP ice core from Greenland based on GICC05 time scale (Andersen et al., 2007). Note that time evolves from right to left in accordance to geological convention. Red spikes indicate the probabilities for a change point at each time point proposed by the Bayesian Change Point Algorithm (Ruggieri, 2013). Warming and cooling change points are detected. Blue triangles mark the warming events with a probability higher then 0.5 (dotted red line), numbered at the bottom of the graph and used for further analysis.
Table 1
Different Proposed DO Sequences (“Seq.”) in Question for the Last Glacial Period and the 42–10 kyr Interval for the NGRIP Record

<table>
<thead>
<tr>
<th>Seq.</th>
<th># Ev.</th>
<th>Estimate</th>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
<th>waiting time/R-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past 120 kyr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ras14</td>
<td>26</td>
<td>4,147</td>
<td>649</td>
<td>596</td>
<td>1,282</td>
<td>1,799</td>
<td>519</td>
<td>519 yr R 0.487 0.463 0.427 0.413 0.397</td>
</tr>
<tr>
<td>bark11</td>
<td>22</td>
<td>4,526</td>
<td>871</td>
<td>573</td>
<td>1,305</td>
<td>967</td>
<td>763</td>
<td>763 yr R 0.622 0.477 0.453 0.452 0.446</td>
</tr>
<tr>
<td>thresh</td>
<td>24</td>
<td>4,369</td>
<td>827</td>
<td>956</td>
<td>1,020</td>
<td>546</td>
<td>510</td>
<td>510 yr R 0.563 0.522 0.482 0.474 0.47</td>
</tr>
<tr>
<td>bayes</td>
<td>21</td>
<td>5,248</td>
<td>517</td>
<td>4,446</td>
<td>958</td>
<td>637</td>
<td>1,380</td>
<td>1,380 yr R 0.525 0.499 0.459 0.443 0.438</td>
</tr>
<tr>
<td>Period of 10–42 kyr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ras14</td>
<td>11</td>
<td>2,976</td>
<td>948</td>
<td>1,488</td>
<td>594</td>
<td>1,574</td>
<td>649</td>
<td>649 yr R 0.66 0.659 0.621 0.596 0.559</td>
</tr>
<tr>
<td>dit07</td>
<td>12</td>
<td>2,705</td>
<td>1,487</td>
<td>1,579</td>
<td>946</td>
<td>1,137</td>
<td>593</td>
<td>593 yr R 0.688 0.611 0.54 0.532 0.51</td>
</tr>
<tr>
<td>bark11</td>
<td>10</td>
<td>3,311</td>
<td>937</td>
<td>873</td>
<td>711</td>
<td>1,364</td>
<td>574</td>
<td>574 yr R 0.745 0.614 0.573 0.573 0.572</td>
</tr>
<tr>
<td>thresh</td>
<td>11</td>
<td>2,682</td>
<td>1,475</td>
<td>546</td>
<td>866</td>
<td>1,134</td>
<td>594</td>
<td>594 yr R 0.598 0.574 0.558 0.54 0.536</td>
</tr>
<tr>
<td>bayes</td>
<td>9</td>
<td>3,908</td>
<td>548</td>
<td>1,482</td>
<td>1,572</td>
<td>652</td>
<td>1,167</td>
<td>1,167 yr R 0.718 0.692 0.689 0.629 0.616</td>
</tr>
</tbody>
</table>

Note: “#Ev.” = number of events considered for the analysis; “Estimate” = the individual estimates of interevent waiting times in the exponential distribution which is used for Monte Carlo simulations. P1 to P5 = the five highest R scored periods for each sequence in decreasing order. Highlighted in red/blue/green, we find the periods in a 10% range around the period of interest 14,700/900/1,150. Gray values indicate out of confidence interval R values.

The warming events for the different DO sequences are listed in Table B1. For the Rayleigh test itself only the R_{max} = R(P1) is considered.

Rahmstorf (2003) supported the idea of regular paced discrete events by a period of 1,470 yr, by showing that 13 of his defined events fall at least in an interval of 20% of multiples of this period. Here we present a simple stochastic process given random input for the occurrence of DO events (Ditlevsen et al., 2007). Consequently, we are testing the interevent waiting times in obeying the exponential distribution which is by definition of a Poisson process (Kroese et al., 2011) and can formally be stated as our null hypothesis.

We will show that the analysis is statistically robust with regard to the exact methodology for event identification as we base our analysis on the dating reported in previous studies (Barker et al., 2011; Ditlevsen et al., 2007; Rasmussen et al., 2014) (cf. Table B1). Using Monte Carlo simulations we generate 1,000 randomly sampled time sequences of events with a waiting time distribution. We then calculate the test statistics of each Monte Carlo simulation which estimates the distribution under our null hypothesis and then tests whether the score of the data (different DO time sequences) lies within a 90% confidence interval, corresponding to a conservative testing. The Rayleigh R measure is chosen because it captures information about the regularity which is in question here (Mardia & Jupp, 2009). For unperturbed periodic event occurrence, it results in a maximum value of 1. In contrast to Ditlevsen et al. (2005), we will also report the periods for which we obtain the five most dominant peaks (Table 1). We evaluate the test for the whole period (Table 1, past 120 kyr) and the 42–10 kyr BP interval (Table 1, period of 10–42 kyr) of the NGRIP record. For the shorter interval, all R values for major periods are greater than 0.5. The prominent 1,470 yr period results in a R_{max} = R(P1) for the two threshold sequences and is rated as the second-highest period for the other DO sets. For Barker et al. (2011) and Rasmussen et al. (2014) data sets, the dominant peak in the Rayleigh R test is for 940 yr. Generally, Barker et al. (2011) (bark11) and and Bayes (bayes) result in higher R values because fewer events are considered. The Bayes sequence attains its R_{max} for a period of 548 yr, with respect to the high interevent waiting time of 3,908 yr (“Estimate” in Table 1, period of 10–42 kyr) which would mean that
the recurrence is highly irregular and several periodic beatings are skipped. Regardless of the period analyzed, the measure for periodicity for all sequences falls within a high likelihood region of a stochastic driven process such that we cannot reject our null hypothesis of statistical random occurrence of DO events. What is interesting here is that the periods 1,470, 900, and 1,150 yr (colored red, blue, and green in Table 1) also obtain very high Rayleigh $R$ values, which fall well in a 90% confidence interval for the probability density from exponentially distributed waiting times. This is because the Rayleigh $R$ values for the dominant periods for NGRIP (42–10 kyr) do not deviate significantly from the true $R$ for random event time sequences with exponentially distributed waiting times.

One additional circumstance that needs to be contemplated is the small sample size. Ditlevsen et al. (2007) only considered a maximum number of 12 events. During the most recent glacial cycle we consider only 18–29 events which might be one reason why we find those striking periodicities (high $R$ values). If DO events were really drawn from a random process, then we would expect a decreasing Rayleigh $R$ (decreasing measure for periodicity) with increasing sample size. That is why we are also interested in periodic components on a longer time scale. Although the abrupt warmings are best dated and researched on a shorter interval (50–10 kyr) they are encountered throughout the whole last glacial and beyond. We see that the Rayleigh $R$ values decrease which is consistent with our null hypothesis. Interestingly, the periods $\sim$1,500 and $\sim$1,150 yr do not dominate the range of periods which result in high $R$ values, more outstanding is the period of $\sim$900 yr (Table 1, past 120 kyr). To overcome the problem of small sample size, one could also perform the analysis for the predicted DO event occurrence of the synthetic Greenland record of Barker et al. (2011). The difficulty is then that the dating becomes more uncertain. Low $R$ values with $R_{\text{max}} < 0.4$ indicate a weak periodicity for all periods when using Barker et al. (2011). High-resolution and accurate dating seems to be necessary to get a more coherent view of the full system. Similar statistical data analysis is necessary to build hypotheses related to events on decadal and centennial time scales. In the following, we will examine some of the major challenges in Earth System Models to be applied for long-term climate and weather changes.

4. Methodological Challenges of Earth System Models

The continuous evolution of climate models over recent decades has been enabled by a considerable increase in computational capacity, with supercomputer speeds increasing by more than a factor of $10^6$ since the 1970s (IPCC, 2007). This computational progress has permitted a corresponding increase in model complexity (by including more and more Earth system components and relevant processes), in the achievable length of the simulations, and in spatial resolution (Claussen et al., 2002). To significantly improve the simulation of climate by general circulation models, systematic errors in representations of relevant processes must first be identified, and then reduced. This endeavor demands that the parameterizations of unresolved processes, in particular, should be tested over a wide range of time scales (weather, climate, and paleoclimate).

4.1. Internal Variability: Challenges in Models

Different flavors of DO-type oscillation are reported for models of different complexity (Kim et al., 2012; Li & Born, 2019; Peltier & Vettoretti, 2014; Wang & Mysak, 2006; Winton, 1993; Yin & Sarachik, 1995). The underlying causes responsible for millennial-scale variability are still debated. Ice core data suggests a strong coupling of the largest millennial-scale warm events in Antarctica with the longest DO events in Greenland through the Atlantic meridional overturning circulation (AMOC) (Barbante et al., 2006; Dima et al., 2018). A quasi-global impact of the events has been reported in the literature (Voelker, 2002). Marine data suggest an antiphase correlation between AMOC and Atlantic subsurface temperatures under glacial climate conditions, suggesting that the vertical temperature structure and associated changes in AMOC are a key element governing DO cycles (Kim et al., 2012; Sadatzki et al., 2019). As intermediate ocean layers are gaining buoyancy as a result of downward mixing of the warmer overlying water, a point is reached at which water at the surface becomes denser than those in the middepth and deep ventilation is initiated (Kim et al., 2012). Sea ice expansion preceded the buildup of a deep oceanic heat reservoir (Sadatzki et al., 2019). During a transition to a strong AMOC, former sites of maximum convection are still there, and the Labrador Sea and the North Atlantic open ocean appear as additional sites of deep convection. During the strong mode, the North Atlantic basin gets colder by convection and fresher until the heat import from lower latitudes cannot overcome the freshening anymore. Thus, the strong convection is inhibited and the weak mode of the oscillations is reactivated. In a recent comprehensive review (Li & Born, 2019), the authors examine evidence that DO
events are an unforced or spontaneous oscillation of the coupled atmosphere–ice-ocean system comprising the subpolar gyre as a key region producing DO-like events. Interestingly, similar behavior is reported under preindustrial conditions in a long control integration (Sidorenko et al., 2015) (Figure 7a). In a preindustrial control setup these events could be judged as undesirable. Yet, ultimately this result shows value in suggesting that on the one hand current models include the physics that are necessary to produce abrupt climate transitions, but on the other hand—as Li and Born (2019) emphasized—exhibit incorrect sensitivity to the boundary conditions. Through our time series analysis, we have now another constraint to the models, namely, that the oscillations should be irregular with random waiting times. On shorter time scales Danek et al. (2019) find that the resolution in the ocean circulation model matters for the Labrador Sea convection (Figure 7b). The high-resolution run shows general agreement to observations (Good et al., 2013). A basic feature is a sporadic deepening of the mixed layer depth, related to turbulent processes on the mesoscale and submesoscale within the Labrador Sea interior. As one driver, we identify the atmospheric circulation with strong winds over the Labrador Sea area (Figure 7c). More research is required to study the physical processes affecting the mixing at these sensitive sites, including submesoscale turbulence (Callies et al., 2015). It seems that irregularities on millennial as well decadal time scale are an important feature of the convection and thus of the large-scale overturning circulation, again bridging time and spatial scales (Figure 1).

In further model developments (Sidorenko et al., 2018), the large oscillations in the Labrador Sea mixing were reduced compared to the old experiments. However, it might be that the centennial-to-millennial oscillations are required to explain climate variability as expressed, for example, by the Little Ice Age and the Medieval Warm Period during the last 1,000 yr (Bradley, 2014; Broecker, 2000). Recent model development is mainly done with a particular focus on the present and future climate. It could be that a detuning of the models is necessary in order to exhibit irregular oscillations on centennial-to-millennial time scales. A systematical analysis of the mechanisms leading to centennial-to-millennial variability remains open.

Numerical experiments (Danek et al., 2019) suggest that at least in the Labrador Sea and other sensitive areas high resolution can play an important role in realistically simulating the variability in the mixed layer depth affecting AMOC.

Besides the vertical dipole, there seems a bipolar signature of the DO events like for the termination (Broecker, 1998; Steig & Alley, 2002). During the first stage of the DOs, Antarctica steadily warmed, but little change occurred in Greenland. Thereafter, at the time when Greenland’s climate underwent an abrupt warming, the warming in Antarctica stopped and reversed. During the final phase, a sudden decrease in the northward heat transport cools the north (Dima et al., 2018). CO2 variations are an active part in the system on millennial time scales. Changes are dominated by slow changes in the deep ocean inventory of carbon, correlated with Antarctic temperature and Southern Ocean stratification, while AMOC changes are more responsible for fluctuations on millennial time scales (Schmittner & Galbraith, 2008). Moreover, carbon dissolved in the global ocean may also be coupled to geosphere-climate interactions (Hasenclever et al., 2017; Wallmann et al., 2018).

4.2. Deglacial Meltwater and Climate

Sea level reconstructions indicate a massive input of meltwater during the early stage of the last deglaciation. During the Bølling period (14.7–14.1 kyr before present) up to ~0.5 Sv (1 Sv =10⁶ m³ s⁻¹) of meltwater entered the ocean (Fairbanks et al., 1992). Most of the meltwater likely originated from the decaying Laurentide ice sheet; the inferred freshwater pulse out of the Mississippi River entered the Atlantic via the Gulf of Mexico (Fairbanks et al., 1992). Various modeling studies indicate that such massive meltwater input should lead to the cessation of NADW, a reduction in large-scale AMOC, and a cooling (e.g., Kageyama et al., 2013; Manabe & Stouffer, 1997; Rahmstorf, 1996; Stocker et al., 1992). A significant drop in AMOC is, however, not supported by paleoclimate data (McManus et al., 2004; Sartheim et al., 1994). In a quite different approach for the last deglaciation, Knorr and Lohmann (2007) illustrated that the onset of a strong AMOC is inevitable and that only the timing of AMOC resumption is triggered by melting water, indicating that the freshwater flux plays only a secondary role for the termination of ice sheets. On the contrary, a more recent study by Liu et al. (2009) emphasizes a huge influence of freshwater perturbation to the AMOC history where a negative freshwater input was applied to get a warm Bolling period.
Sensitivity experiments to high-latitude freshwater perturbations started with climate model scenarios of the polar halocline catastrophe by Bryan (1986). As stability and response of the AMOC to identical freshwater perturbation scenarios vary greatly between different models (e.g., Kageyama et al., 2013; Manabe & Stouffer, 1997; Rahmstorf, 1996), we concentrate here on the dynamics of the deglacial meltwater in climate. Again, by applying our multiresolution approach in the Finite Element Sea ice Ocean Model FESOM (Sidorenko et al., 2015; Wang et al., 2014), we will be able to zoom into regions of interest (coasts, high latitudes, and equator) while keeping the resolution sufficiently low in other areas (Figure 8). The model includes furthermore a land surface scheme with interactive vegetation dynamics, ensuring that climate and vegetation are consistent with each other. The model is described in the Appendix C.

Before exploring the effects of different regions of freshwater forcing, recall that deglacial freshwater fluxes are by iceberg calving and surface runoff from the ice sheets. Such meltwater fluxes were simulated by the ice sheet model of Zweck and Huybrechts (2005), shown in Figure 9. For the period 14.7–14.1 kyr BP, most meltwater entered as liquid water, whereas during the Heinrich-1 event (ca. 17.5–14.8 kyr BP) the freshwater entered as solid water in terms of calving. The calved icebergs bring the freshwater very effectively into the subpolar region (Heinrich, 1988), and it is a common approach to mimic this freshwater input as a broad anomaly into the subpolar North Atlantic Ocean (e.g., Rahmstorf et al., 2005).
Figure 8. Resolution of the ocean model component in our Earth system model AWI-ESM. Example of an unstructured mesh for our FESOM model with horizontal resolution between 12 and 135 km (CORE-2 mesh Sein et al., 2018; Sidorenko et al., 2011).

We start now with water hosing experiments based on a glacial background state which has been obtained after 840 yr of model integration. After the spin-up, we insert a 150 yr meltwater flux of 0.2 Sv perturbation at different locations (Mississippi, Mackenzie, or local coastal runoff close to Labrador). Significant differences are found in the oceanic responses to these different locations of freshwater discharge (Figure 10a). The AMOC index refers to the maximum of the AMOC in the North Atlantic Ocean. The Mississippi (orange) and the Mackenzie (blue) runoff have little influence, in contrast to the subpolar forcing (red) where the freshwater flux is equally distributed in the 50–70°N-band in the North Atlantic Ocean. As an additional experiment, the hosing is inserted at the coast close to the Labrador Sea between 50°N and 70°N (green). The different response in the AMOC can be attributed to the negative sea surface salinity anomalies lowering

Fresh water flux (Sv)

Figure 9. Components of the total freshwater flux (in Sv =10⁶ m³ s⁻¹) originating from the Northern Hemisphere ice sheets at the end of the last ice age based on Zweck and Huybrechts (2005).
the density (Figure 10b). The low salinities prevent deep water formation and can lead to a partial collapse of the AMOC, especially in the subpolar setup (red).

The weakened overturning circulation further suppresses heat exchange between the North Atlantic surface and subsurface water masses, leading to a cooling in the upper ocean and a pronounced warming in the subsurface layer (Figure 11). The signature bears similarities with the antiphase correlations between AMOC and Atlantic subsurface temperatures for the DO events. Again, the intermediate ocean is densified by the downward mixing. The resumption of the deep ventilation (Figure 10a) is the strongest, the more effective the subsurface warming is (Figure 11). In the case of the subpolar hosing, the overshoot is around the model year 1100. Note that the smallest response and the least efficient hosing is in the Gulf of Mexico, a potential region for a major MWP-1a discharge (Aharon, 2003, 2006). The results are in line with previous uncoupled ocean model experiments (Condron & Winsor, 2011, 2012) showing that the narrow coastal boundary currents play an important role efficiently preventing the freshwater entering into the deep-water formation areas.

Figure 12 shows the near-SAT anomaly for the different glacial and deglacial scenarios. The temperature response at the Mackenzie, Mississippi, and coastal scenarios are relative insensitive compared to the subpolar hosing. Interestingly, the subpolar shows a pronounced overshoot after the hosing, releasing heat stored at the subsurface (Figure 11). Figure 13 shows the associated streamfunction of AMOC, indicating the strong changes for the subpolar scenario both during the hosing and in the overshoot phases. The AMOC in other scenarios is rather insensitive to the freshwater forcing. As an alternative explanation for the insensitivity of the Bølling freshening, it has been suggested that the ocean circulation is too sensitive to surface freshening when underestimating the overflow contribution of the deep water seeping from the Nordic Seas (Lohmann, 1998), and that a most likely northern source of deep water formation can stabilize the system during the Bølling (Lohmann & Schulz, 2000). Another possibility follows the approach of Roche et al. (2007), where the glacial meltwater pulse could have sneaked unnoticed into the deep ocean during the last glacial due to the effect of hyperpycnal flow. The exact dynamics of the AMOC perturbation is not elaborated here and is subject of a subsequent analysis. It can be shown that connections between latitudes are due to boundary-trapped Kelvin waves propagating along western boundaries. For interannual or larger periods, the midlatitude Kelvin waves are replaced by long Rossby waves on the eastern boundary and viscous boundary waves on the western boundary. The localization of the viscous boundary waves is on the so-called Munk boundary layer scale which is much larger than the Rossby radius (Johnson & Marshall, 2002, 2004; Kawase, 1987).

Source and causal mechanism of MWP-1a are still being explored. Based on corals drilled offshore from Tahiti, Deschamps et al. (2012) show that MWP-1a started no earlier than 14,650 yr ago and ended before 14,310 yr ago, making it coeval with the Bølling, but most likely including a significant meltwater contribution from the Southern Hemisphere (see also Weber et al., 2014, for a discussion). Gregoire et al. (2012) found in their ice-sheet modeling simulation that the separation of the Laurentide and Cordilleran ice sheets in North America produces a meltwater pulse corresponding to MWP-1a. Stanford et al. (2006) derive that MWP-1a did not coincide with the sharp Bølling warming, but instead with the abrupt cooling of the Older Dryas with only a relatively minor 200 year weakening of NADW flow, which is very much in line with our findings.

Summarizing, we argue that the way how meltwater enters into the ocean (liquid or solid) provides a first order control for the sensitivity of the ocean circulation and thus the deglacial climate. For future studies, a fully coupled climate–ice sheet model capturing the drainage chronology of the Laurentide Ice Sheet is being developed to shed more light on the role of deglacial meltwater on climate. A high resolution at the coasts seems to be an important feature for a proper understanding of the coastally confined freshwater pathways. Indeed, using a coarse-resolution version of the Condron and Windsor (2011) model, it was demonstrated that the spreading of freshwater across the subpolar North Atlantic results from the inability of numerical models of this resolution to accurately resolve narrow coastal flows, producing instead a diffuse circulation that advects freshwater away from the continental boundaries. In order to capture the climatic impact of freshwater released in the past and possible future, the ocean needs to be modeled at a resolution sufficient to resolve the dynamics of narrow, coastal buoyant flows. Another difficulty in resolving the history of abrupt climate change during the last glacial period is beyond the sensitivity of climate models to freshwater and to the location of the freshwater source. It is directly related to the dynamic behavior of the ice sheets.
Figure 10. (a) AMOC indices of North Atlantic hosing for different hosing areas. Units are Sv. Black line represents the unperturbed LGM experiment. Hosing is for the period 840–990. (b) Annual mean sea surface salinity anomaly between LGM and the perturbation experiment LGM with 0.2 Sv for the model years 900–950. Dotted areas are significant on a 95% confidence level (t test).

Figure 11. Annual mean Atlantic Ocean temperature anomaly (K) for the different hosing scenarios. (a) Subpolar minus LGM, (b) coastal minus LGM, (c) Mississippi minus LGM, and (d) Mackenzie hosing minus LGM. Shown are model averages for the years 900–950. Dotted areas are significant on a 95% confidence level (t test).
To this end, proper modeling of the ice-substrate interface is an important boundary condition for ice sheet modeling. The substrate affects the ice sheet by allowing sliding through sediment deformation and accommodating the storage and drainage of subglacial water (e.g., Gowan et al., 2019). Temperate ice sheets, such as the Laurentide and Eurasian ice sheets behaved differently depending on whether or not there were thick, continuous unconsolidated sediments underneath the ice (Clark & Walder, 1994), providing an uncertainty in the ice sheet history of deglacial climate.
Figure 13. AMOC (Sv) for (a) LGM, (b) Mackenzie hosing, (c) Mississippi hosing, (d) coastal hosing, (e) subpolar hosing, and (f) subpolar overshoot. Shown are model averages for the years 940–990, and for the overshoot the years 1075–1125.
Figure 14. Radiocarbon age in the late Holocene. Filled areas are simulation results using a $^{14}$C-equipped version of the global multiresolution ocean model (Appendix C1). (a) At the sea surface where dots indicate historical values collected at the $^{14}$CHRONO Marine Reservoir Database (see http://calib.org/marine/ and further references there). (b) Meridional distribution of simulated $^{14}$C ages in the Central Atlantic along 30°W.

4.3. Radiocarbon Dynamics for Chronology

Marine climate records are frequently dated by means of stratigraphic approaches such as orbital tuning or the synchronization with other cross-dated geological records. A more physical approach is radiocarbon ($^{14}$C) dating which takes advantage of the half-life of radioactive $^{14}$C (5,700 yr Audi et al., 2003; Bé et al., 2013). Radiocarbon is produced through cosmogenic radiation in the upper atmosphere from where it enters the global carbon cycle as $^{14}$CO$_2$. Cosmogenic (Hain et al., 2014) $^{14}$C production has been varying (e.g., Adolphi et al., 2018; Channel et al., 2018), and in addition, the partitioning of $^{14}$C between the various carbon reservoirs has been subject to past climate-carbon cycle interactions (e.g., Hesse et al., 2011; Köhler et al., 2006).

The $^{14}$C uptake by the oceans is slow (~10 yr) and superimposed by mixing with water from the $^{14}$C-depleted deep sea (Broecker & Peng, 1974). Therefore, marine $^{14}$C concentrations are systematically depleted with respect to $^{14}$C in the overlying contemporaneous atmosphere. The concentration difference translates into an apparent surface water $^{14}$C age (or “marine reservoir age”) ranging from about 300 yr in the subtropics to up to 1,000 yr in high latitudes during the late Holocene (e.g., Key et al., 2004). Marine reservoir ages can have even a much broader range of 100–2,500 yr at sites far offshore (e.g., Cook & Keigwin, 2015; Sarnthein et al., 2015; Skinner et al., 2010).
A further complication is that marine $^{14}$C and $^{14}$C-dated climate records typically originate from continental margins, marginal seas, or tropical islands (see Figure 14a). Such sites may not be representative to mirror large- or global-scale processes of climatic change, demanding a valid extrapolation method to infer past global conditions. The spatial gaps can be filled employing ocean-climate circulation models. However, the current generation of $^{14}$C-equipped models suffers from insufficiently low resolution near the ocean margins (e.g., Butzin et al., 2017), and conventional high-resolution ocean modeling approaches are computationally too expensive when it comes to long-term simulations covering the entire time scale accessible to $^{14}$C dating (approximately the past 55 kyr). This problem can be solved by means of global multiresolution models using unstructured meshes (Appendix C). Figure 14 shows a first simulation result for the late Holocene obtained with a $^{14}$C-equipped version of the global multiresolution model FESOM2 (Danilov et al., 2017). With this approach we will be able to zoom into regions of interest while keeping the resolution sufficiently low in other areas (Figure 8). By means of conventional ocean models, it is difficult to resolve important interbasin exchange fluxes through narrow key seaways like the Bering Strait, Strait of Gibraltar, Indonesian Throughflow, and Canadian Arctic Archipelago. Unstructured meshes have the potential to represent such throughflows with higher accuracy.

A major challenge in paleoclimatology is to disentangle the relationship between climatic leads and lags during the past 50 kyr (e.g., Shakun et al., 2012). In the case of radiocarbon dating this implies to quantify the uncertainties associated with climatic variability and cosmogenic $^{14}$C production changes, and that the time axis might be changed (Figure 3). Here, we consider an ensemble of simulated marine surface water $^{14}$C concentrations (Figure 15), using a more efficient model version (Butzin et al., 2017). The atmospheric $^{14}$C forcing is based upon the Hulu Cave speleothem record including uncertainties (Butzin et al., 2020; Cheng et al., 2018; Southon et al., 2012). We combine three different ocean climates and $^{14}$C forcing scenarios to create an ensemble of nine experiments. Each experiment is spun up over 20 kyr with fixed atmospheric values of $^{14}$C and CO$_2$ at 54 kyr before present, and is then run transiently forced by time-variable $^{14}$C (Figure 15a) and CO$_2$. Marine reservoir ages for the upper 50 m are calculated afterward (Appendix C). Radiocarbon concentrations in terms of the fractionation-corrected $^{14}$C/$^{12}$C ratio of marine surface water normalized to the corresponding ratio in the atmosphere ($^{14}$C) are shown in Figure 15b. Figure 15c shows the $^{14}$C age averaged between 50°N and 50°S, and Figure 15d a snapshot for the Last Glacial Maximum at 21 kyr BP. Ensemble approaches provide a new direction in evaluating dating uncertainties on regional scale and a dynamical interpretation of paleoclimate data (Alves et al., 2018), and thus substantially enhance our understanding of the Earth’s climate system.

5. Variability of Extreme Climate and Weather Over Europe

In the atmosphere, preferred modes of climate variability at different time scales exist, which often span large geographical areas. Some regions may be cooler or perhaps drier than average, while at the same time, thousands of kilometers away, warmer and wetter conditions prevail. These simultaneous variations in climate, often of opposite sign, over distant parts of the globe, are commonly referred to as teleconnections in the literature (e.g., Barnston & Livezey, 1987; Trenberth et al., 1998; Wallace & Gutzler, 1981). Examples are the Atlantic Multidecadal Oscillation (AMO), Pacific Decadal Oscillation (PDO), El Niño–Southern Oscillation (ENSO) (cf. Figure 1). The teleconnections are an important tool for the interpretation of proxy data, as they link a time series at one point to large-scale patterns. Below, we will discuss the mechanisms behind decadal to multidecadal variability and extreme climate during the recent past and their possible dynamics (e.g., Otto, 2017; Pfahl et al., 2009; Rimbu & Lohmann, 2010, 2011; Rimbu et al., 2014; Rimbu, Lohmann, et al., 2016; Sippel et al., 2020). Our analysis is based on advanced statistical techniques applied to observational data, high-resolution proxy records, and climate simulations. Extreme climate events are typically associated with specific weather patterns and coherent atmospheric structures, for example, cyclones and blockings, but also to large-scale and long-term variations. We show again a linkage of the dynamics across time scales.

5.1. Decadal Time Scales: The GSA

Beyond freshwater variations during the last deglaciation, that had an impact on circulation patterns in the North Atlantic Ocean as outlined above, changes in North Atlantic freshwater have also been observed for recent climate. Major freshening episodes have been described in recent decades (Curry & Mauritzen, 2005). The first freshening event has happened in the 1970s, the so-called the GSA (Dickson et al., 1988). Interestingly, the GSA which was accompanied with a decadal cooling over the North Atlantic realm, might have
some analogs to past Heinrich events which appeared during glacial times (Heinrich, 1988). The GSA in the subpolar North Atlantic in the 1970s, is also linked to droughts and floods, extreme winters, and atmospheric blocking. Aagaard and Carmack (1989) reported that the origin of this event must have been an anomalously large Arctic freshwater discharge through Fram Strait, later confirmed by Häkkinen (1993) using a coupled ocean-sea ice model. The 1970s GSA was followed by similar events (but weaker in amplitude) in the 1980s (Belkin et al., 1998) and the 1990s (Belkin, 2004). The events have also been linked to atmospheric variability (Dima & Lohmann, 2007; Haak et al., 2003; Karcher et al., 2005; Mauritzen et al., 2012).

Figure 16 indicates a potential link of synoptic variability to decadal-to-multidecadal time scale variability (Ionita et al., 2016). During times of enhanced blocking between Greenland and Svalbard (e.g., in 1962–1966), sea ice first accumulated in the Arctic, later leading to an enhanced sea ice release into the North Atlantic, and with a delay reducing Labrador Sea salinity. In the following decade, the freshening causes a weakening of the large-scale AMOC. The AMOC goes to a weaker state several years after the GSA in the late 1960s, consistent with earlier finding of a fast response of the meridional overturning to North Atlantic forcing (Eden & Willebrand, 2001). From this perspective, the 1960s GSA induces an AMOC shift to a new weaker state caused by the significant fresh water forcing related to Fram Strait Sea Ice Export (Figure 16). For more recent times, a close link between decadal atmospheric blocking is proposed (Häkkinen et al., 2011), but knowledge about the inherent atmosphere-ocean-sea ice dynamics and their potential drivers is still incomplete. When comparing the three major freshwater events in the Northern North Atlantic...
Figure 16. (a) Winter Fram Strait sea ice export time series (normalized): model simulation (solid line) and observations. (b) Atmospheric blocking frequency composite maps with respect to the simulated Fram Strait sea ice export: Maps for above and below 75% standard deviation indicating different regimes of blocking in the North Atlantic realm. (c) The 3 yr running mean winter sea surface salinity in the central Labrador Sea with three major freshwater events, based on Ionita et al. (2016).

(Figure 16c), the Fram Strait Sea Ice Export and atmospheric blocking are not the only causes for large-scale ocean circulation changes on these time scales.

Decadal-to-multidecadal variability is coupled both to synoptic time scales, such as atmospheric blocking and long-term background conditions. As an example, the jet stream in the Euro-Atlantic region and associated blocking activity is linked to Atlantic multidecadal ocean variability and AMOC, implying the need for a proper representation of synoptic scale variability in coupled climate models. Model simulations are advancing toward increased resolution, the addition of feedback mechanisms between different components of the earth/climate system, and, in particular, the added ability of considering the dynamic simulation of isotopes. Future research will use a hierarchy of global climate simulations and all available types of observational, (paleo-)environmental and reanalysis data to evaluate variability patterns and extreme events on synoptic to millennial time scales. This integrated approach provides a long-term context beyond our limited view from the observational period. The combined analysis of data and models is essential for interpreting paleoclimate data and related weather and climate patterns, which is the topic of the next subsection.

5.2. Changes in the Intensity and Frequency of Extremes: Atmospheric Rivers
Changes in the intensity and frequency of extremes have drawn much attention during recent decades (Coumou et al., 2014; Horton et al., 2016; IPCC, 2012), mainly due to their large impacts on natural environment, economy and human health (Ciais et al., 2005; Kovats & Kristie, 2006). Due to the inherent nature of extreme events, the variability of high-temperature extremes differs from that of mean temperature (Raible et al., 2007; Schär, 2015). Therefore, we need to understand the causes of climate extremes in the twentieth century and in a more distant past. Paleoclimate data and modeling allow an assessment of climate variability and extremes and their relation to the average climate.

One particular example of extreme events is linked to atmospheric rivers (AR) which are narrow regions responsible for the majority of the poleward water vapor transport across the midlatitudes. They are characterized by high water vapor content and strong low level winds of extratropical cyclones (e.g., Namias,
Newell et al. (1992) termed these long (∼2,000 km) and narrow (300–500 km wide) bands of enhanced water vapor flux tropospheric rivers (Figure 18b). They used the term rivers because they transport water at volumetric flow rates similar to those of the world’s largest rivers, through narrow corridors (Zhu & Newell, 1998). Although the meridional water vapor transport within ARs is critical for water resources, ARs can also cause disastrous floods especially when encountering mountainous terrain (Pasquier et al., 2018). Here, we describe future direction and research needs of the synoptic interpretation of paleoclimate data.

Rimbu, Czymzik, et al. (2016) show that lake sediment records from Ammersee (southern Germany) are a good proxy not only for local summer extreme precipitation. The significant peaks of flood frequency (Figure 18a) characterize also extreme local precipitation variability as well as extreme high temperatures over northeastern Europe. River Ammer floods are associated with enhanced moisture transport from the Atlantic Ocean and the Mediterranean toward the Ammer region, a pronounced trough over western Europe as well as enhanced potential vorticity at upper levels. Interannual to multidecadal increases in flood frequency are associated with a wave train pattern extending from the North Atlantic to western Asia, with a prominent negative center over western Europe. This link between low and midlatitudes is explained by the atmospheric circulation pattern associated with AR (Figure 18b). It has been shown that a large portion of the most intense precipitation events (and of course of their associated floods) in western Europe are objectively associated with the occurrence of ARs, both in Great Britain (Lavers et al., 2012) and in the Iberian Peninsula (Ramos et al., 2015). As ice cores, corals, as well as lake sediments go back in time from centuries to several millennia with high, that is, annual or even seasonal resolution, we can use them to reconstruct past variability of related climate extremes. This is useful not only for knowledge of past extreme climate variability, but also to anticipate future natural variability of climate extremes.

6. Conclusions and Outlook

The Earth’s climate is characterized by many modes of variability in the atmosphere, ocean, cryosphere, and biosphere. The generation of low-frequency variability in the climate system is crucial to allow a separation of anthropogenic signals from natural variability, thereby increasing the ability to recognize and improve the attribution of climate and weather events to anthropogenic climate change. In the paper, we first consider the response of the system to regular multimillennia orbital forcing and then consider shorter time scales down to daily weather events (Figure 1).

Milankovitch (1941) initially suggested that the critical factor for ice sheet formation and decay is total summer insolation at about 65°N, because for an ice sheet to grow some additional ice must survive each successive summer. In contrast, the Southern Hemisphere is limited in its response because the expansion of ice sheets is curtailed by the Southern Ocean around Antarctica. Thus, the conventional view of glaciation is that low summer insolation in the temperate Northern Hemisphere leads to perennial continental snow triggering Northern Hemisphere glaciation. Here, we have evaluated how the climate system creates nonlinear responses to local insolation forcing. Our model results show a pronounced polar amplification of the orbital forcing. In the frequency domain, temperature (and precipitation) variability at midlatitudes are dominated by precession, while high latitudes are dominated by obliquity. The precessional response is related to nonlinearities and/or seasonal biases in the climate system. Additional to the insolation forcing, greenhouse gases provide a net forcing driving glacial-interglacial variability. The order of magnitude of glacial-interglacial changes is similar in magnitude to our current anthropogenic induced radiative forcing of ∼2 W m⁻², which will increase in the future. Surface temperatures are currently not in equilibrium due to the thermal inertia of the ocean, and the Earth’s present energy imbalance of 0.5–1 W m⁻² (von Schuckmann et al., 2016) provides an indication of how much additional global warming is still in the pipeline if climate forcings remain unchanged (Hansen et al., 2016). We can learn about potential thresholds in the system from the past (e.g., Kwasniok & Lohmann, 2009; Robinson et al., 2012; Sutter et al., 2016). As an example, the Mid-Pleistocene Transition at about 1 Myr marks a transition from a 40 kyr variability to a dominant 100 kyr variability in the climate system. Lowering of global CO₂ below some critical threshold, conceivable triggers a transition from a 40 to a 100 kyr resonance (e.g., Farmer et al., 2019; Raymo et al., 1998; Tziperman et al., 2006). Further insight may come from the analysis of the oldest ice in Antarctica, providing information of CO₂ and Southern Hemisphere temperatures (Fischer et al., 2013), as well as through modeling of ice sheets that are truly coupled into Earth system models.
**Figure 17.** Schematic representation of atmospheric rivers (atmospheric conveyor belt) associated with extreme precipitation events, for example, in the United Kingdom in 2009.

**Figure 18.** (a) Time series of the observed mean daily river Ammer runoff during the period 1926–2006. The dashed line is for daily discharge of 125 m$^3$ s$^{-1}$. (b) Composite map of vertically integrated water vapor transport for flood days (vector) and its magnitude (color). Units: kg m$^{-2}$ s$^{-1}$ for runoff higher than 125 m$^3$ s$^{-1}$. The geographical Ammer runoff location is inserted as a gray dot. Based on Rimbu, Czyszczik, et al. (2016).
We have questioned regularities found in DO events occurrence and statistically analyzed the distribution of interevent waiting times. Our assumption was that DO events represent repeated oscillations between interstadial and stadial conditions with a threshold crossing process involved (Alley et al., 2003; Ditlevsen et al., 2007; Steffensen et al., 2008). That assumption is not only based on the general shape of DO events as rapid transitions between binary states but also concluded from model simulations (Ganopolski & Rahmstorf, 2001) that examined the stability of the glacial climate with the usage of ocean-atmosphere model of intermediate complexity. This has been recently confirmed in a more comprehensive model (Zhang et al., 2017). It was reported that DO-like events could be caused by a threshold-driven process, as a result of changes in the Atlantic freshwater balance combined with the ice sheet height, consistent with the data analysis of Schulz et al. (1999). Therefore, we have considered a threshold criterion for the definition of DO events and used a Bayesian methodology for accurate change point detection that allows us to quantify the uncertainties about the total number of change points and their locations (Partin et al., 2015; Ruggieri, 2013). For a given time interval each event occurs stochastically independent, meaning that the probability of one abrupt warming event does not affect the probability of any other warming event in the same interval. We find that the observed DO recurrence intervals in the NGRIP ice core proxy data are consistent with exponentially distributed waiting times not only in the interval of 42–11 kyr BP (as proposed by Ditlevsen et al., 2007) but also during the whole last glacial period in the last 120 kyr. This finding is independent from the dating methodology. Additionally, novel periodicities of ~900 and ~1,150 yr in the NGRIP record are reported besides the prominent 1,500 yr cycle. Although a high periodicity reflected in a high Rayleigh $R$ can be found in the data, it remains indistinguishable from a simple stationary random Poisson process. These are quite remarkable findings as ~1,500 and ~900 yr periods are visible throughout the Holocene (Bond, 1997; Rimbu et al., 2004; Xu et al., 2015).

Another feature for millennial climate variability comes from the spatiotemporal structure of DO events and modeling results. DO-like variability is seen in a handful of model simulations, including even some preindustrial simulations. As a mechanism, the subsurface is warmed by the downward mixing of the warmer overlying water during an AMOC weak state, until the surface becomes more dense than at middepth and deep ventilation is initiated. A similar mechanism is observed in the overshoot dynamics of AMOC.

-Decaying Northern Hemisphere ice sheets during the last deglaciation affected the high latitude hydrological balance in the North Atlantic and hence the ocean circulation after the Last Glacial Maximum. Surprisingly, geological data suggest that meltwater fluxes of about 14–20 m sea level equivalent flushed into the North Atlantic without significantly influencing the AMOC. Global sea level reconstructions indicate abrupt changes within several hundred years and rising the sea level during the Bølling/Allerød warm interval (Clark et al., 1996, 2002). Using a climate model with high resolution near the coast, we have investigated the response of the ocean circulation to deglacial freshwater discharges. In our experiments, we find a strong sensitivity of the ocean circulation depending where the deglacial meltwater is injected. Meltwater injections via the Mississippi and near-Labrador coast hardly affect the AMOC. The reduced sensitivity of the overturning circulation against freshwater perturbations following the Mississippi route provides a consistent representation of the deglacial climate evolution. However, a subpolar freshening of the North Atlantic, mimicking a transport of water by icebergs, leads to a quasi-shutdown of AMOC. Therefore, during Heinrich events the freshwater injection is more effective. Future model developments shall therefore concentrate on modeling of ice calving (Alley et al., 2008; Levermann et al., 2012) and iceberg transport in the ocean (Stern et al., 2019; Rackow, Wesche, et al., 2018). Most ocean climate models do not represent ice shelf calving in a physically realistic way (and even not properly parameterized), although calving of icebergs is a major component of the mass balance for present Antarctic ice shelves. It is remarkable that the vertical heat distribution matters for the AMOC dynamics. The strongly weakened AMOC (as in the case of subpolar hosing) leads to upper ocean cooling and pronounced warming in the subsurface layer (Rühlemann et al., 2004). In this particular scenario, we see an AMOC overshoot, also seen during the Bølling/Allerød. At millennial scales, paleoceanographic evidences have characterized abrupt changes in the other regions as well, for example, the subpolar North Pacific Ocean (Gong et al., 2019; Lembke-Jene et al., 2018; Lohmann et al., 2019; Maier et al., 2018; Rae et al., 2014). Such rapid shifts affect processes like surface stratification, formation and dynamics of sea ice, feedbacks with atmospheric patterns like the Aleutian Low, as well as deep and intermediate water formation across time scales. Studies of abrupt events are of great interest because they could help to indicate whether rapid reorganizations of the ocean circulation might occur in the future, and how they might affect climate in general (Thirumalai, 2018).
One of the greatest obstacles to resolve the leads and lags-puzzle is the difficulty of developing an accurate time scale for long ice cores and marine sediments. The development of such a time scale would allow testing of many climate forcing hypotheses, leading to significant advances in paleoclimate theory. In the late Quaternary, the backbone of such a time scale could be provided through radiocarbon dating (e.g., Stein et al., 2017). To improve the quality of the interpretation of the available radiocarbon records, high spatial model resolution is required around the relevant locations, which makes traditional ocean-climate models with their uniform meshes impractical. Here, traditional high-resolution modeling approaches, involving uniform meshes, would result in prohibitive computational costs. To overcome this drawback, one can apply a coupled climate-radiocarbon circulation model using innovative and scalable numerical schemes on supercomputing facilities (e.g., Koldunov et al., 2019). This issue deserves further efforts combining simulations and reconstructions, all the more, as uncertainties in reconstructions are large for times prior to about 30 kyr before present, where the marine reservoir variability is poorly constrained through reconstructions. Data assimilation approaches could provide a physically consistent climate-carbon isotope history. The data assimilation methodology combines the data (observations, reconstructions) and the underlying dynamical principles governing the system to provide a state estimate of the system which is better than what could be obtained using just the data or the model alone. Applying an ensemble Kalman filter approach, one could use the advances in a parallel data assimilation framework with relatively moderate increase in computation time (Nerger & Hiller, 2013).

As changes at high latitudes are particularly sensitive to forcing, the detection and understanding of polar-to-subpolar changes and their influence on global climate is of central importance. Model simulations can aid the interpretation of the drivers and causes of observed variations in paleoclimate quantities. Analyzing proxy-reconstructed paleoclimate records and models in tandem allows for the evaluation of forcing and feedback mechanisms under climate change. Conversely, model simulations can aid in the interpretation of the causes of observed variations in paleoclimate data. One important application of paleoclimate data is to validate state-of-the-art coupled climate models for past climate transitions, which allows to gain insight into possible future climate states that may be notably different from present day conditions. Past time periods provide the means for evaluating the performance of general circulation models for interglacials (e.g., Lohmann et al., 2013a; Pfeiffer & Lohmann, 2016; Shi & Lohmann, 2016; Stein et al., 2017). Key findings are that the models do not capture the magnitude of the sea surface temperature anomalies derived from marine proxy records for interglacials and that they underestimate the variance (Laeppe & Huybers, 2014a, 2014b). The connection between climate sensitivity and variance is consistent with a possible connection of the linear response to perturbations and the fluctuations (Lohmann, 2018). Indeed, Nijssse et al. (2019) emphasized in a similar way that models that are more sensitive to greenhouse gas emissions (i.e., higher equilibrium climate sensitivity) also have higher temperature variability on time scales of several years to several decades (Lenton, 2011). Therefore, high-sensitivity climates are more prone to rapid warming but also to hiatus periods.

Observed and simulated statistics of extreme events shall be compared between different model configurations and observations to evaluate if, and under what conditions climate models are reliable for simulating future changes in climate extremes. Through a combination of paleoclimate records with the analysis of climate-model simulations, we can contribute to a better understanding of the long-term evolution of climate variability modes and their linkage to synoptic events. Such strategy can overcome our current knowledge gaps, with special emphasis on the role of ocean circulation, climate modes of variability, and regional extreme climatic events. Numerical climate models operate on increasingly finer grid sizes as the performance of parallelized super computers increases. Whether a model can represent a geophysical process depends on the model formulation and discretization. Since the spatial scale of oceanic eddies is O(1–100) km (first baroclinic Rossby radius of deformation), the ocean model grid resolution needs to be on the same scale to represent these ubiquitous small-scale features (Chelton et al., 2011). Alternatively, their effects must be parameterized as in current state-of-the-art general circulation models, where the typical horizontal resolutions of the ocean component is about 1°. Practically, given the present high-performance computer capacities, efficient and parallelized model codes, it is now possible to conduct isotope-enabled (e.g., Cauquoin et al., 2019; Werner et al., 2018) simulations for 50–100 model years per day even with a multiscale ansatz (Shi et al., 2020).

Recent developments have considerably improved the computational efficiency and scalability of unstructured-mesh approaches on high-performance computing systems (Danilov et al., 2017). The surface ocean current
in such high-resolution simulation (Figure 19) has a completely different structure including eddies than the structure in coarse-resolution model. Conceptual and statistical approaches are used to explore climate variability from the persistence of atmospheric blocking to nonlinear responses to external forcing to multimillennial variability. Sea ice decline has a potential strong influence on the midlatitude climate. Results from recent modeling experiments indicate that the influence of Arctic sea ice decline is not sufficiently strong to explain anomalously cold winters in Europe, but that other regions such as Russia and North America may be more strongly affected than Europe (Semmler et al., 2016). The uncertainty surrounding modeled Arctic-midlatitude linkages has recently led to the Polar Amplification Model Intercomparison Project (Smith et al., 2019) which could be extended to past intervals which would be a step beyond current activities in the paleoclimate modeling community (Kageyama et al., 2017; Otto-Bliesner et al., 2017).

The simulated interglacial periods provide the means for evaluating the performance of general circulation models in representing sea surface temperature (SST) anomalies and trends (e.g., Lohmann et al., 2013a; Pfeiffer & Lohmann, 2016; Stein et al., 2017). However, the models do not capture the magnitude of the derived SSTs from marine proxy records in all climate simulations of the Holocene (Lohmann et al., 2013a) and Last Interglacial (Pfeiffer & Lohmann, 2016). It is suggested that a part of such discrepancies can be caused either by too simplistic interpretation of the proxy data (including dating uncertainties and seasonal biases) and/or by underestimated regional responses in climate models. By using long-term multimillennial climate model runs and paleoclimate data, a discrepancy with a too low-low-frequency variance is detected also with respect to variability (Laeppe & Huybers, 2014a, 2014b). The model-data differences on long time scales suggest that relevant feedback mechanisms and internal variability may not be well represented in current climate models (Lohmann, 2018).

One of the critical feedback mechanisms for long-term variability is the interaction of Arctic sea ice and the global ocean circulation, which can affect climate variability on short time scales and might lead to relatively rapid climate change (Häkkinen et al., 2011; Ionita et al., 2016). Improving our understanding of the sea ice feedback, Müller and Stein (2014) reconstructed the sea ice conditions in the Fram Strait at the end of the last glaciation at centennial to millennial time scales. For more recent times, sea ice can accumulate in the Arctic due to atmospheric blocking over Greenland (Ionita et al., 2016), leading finally to an enhanced sea ice release into the North Atlantic reducing Labrador Sea salinity affecting AMOC. It is likely that new model approaches are necessary to realistically obtain persistent blocking in conjunction with large-scale changes in the ocean circulation and sea ice (Kinter III et al., 2013). We suggest a close link between climate variability modes and the statistics of weather patterns in the North Atlantic realm. Such features provide a challenge of numerical models and the interpretation of long-term climate records. We provide an example of AR and extreme events in Europe as evaluated from paleoclimate data (Rimbu, Czymzik, et al., 2016). AR account for nearly all of the extreme flooding events with potentially huge economic losses (e.g., Dominguez et al., 2018). The potential effects of extreme events, such as heat waves or droughts, do not only depend on their abundance, but also on “how these extremes occur,” including the sequence of events. These events are unexplored for both past and future changes, and they require sophisticated analysis methods (Sedlmeier et al., 2016). Weather and climate extremes are an inherent part of climate and variability on synoptic time scales can have a large effect on decadal-to-multidecadal variability and vice versa.

Looking at the development of science in the twentieth century one can distinguish two trends, which Weisskopf (1963) called “intensive” and “extensive” research. Intensive research goes for the fundamental laws, extensive research goes for the explanation of phenomena in terms of known fundamental laws. There is always much less intensive research going on than extensive. Once new fundamental laws are discovered, a large and ever increasing activity begins in order to apply the discoveries to hitherto unexplained phenomena. Climate science is mostly extensive and most fundamental laws were discovered decades ago (Landau & Lifshitz, 1987) (although there are Navier-Stokes existence and smoothness problems, see
The Clay Mathematics Institute, http://www.claymath.org/millennium-problems/). However, many modeling and methodological issues of obtaining new data are related to current intensive research or at least intensive research of yesterday. Examples are recent breakthroughs in computational fluid mechanics (e.g., Danilov et al., 2017) and in high-precision radiocarbon dating (e.g., Wacker et al., 2010). Kuhn (1962) argues in terms of the evolution of scientific theories that different phases do not emerge from the straightforward accumulation of facts, but rather from a set of changing intellectual circumstances and possibilities. Climate science already went through a preparadigm phase in the sense of Kuhn (1962), in which there is no consensus on any particular approach. Somehow, climate and general Earth System Science are partly in a second phase already, in which the open questions are solved within the context of the established methods documented in review articles, books and reports (e.g., Berger, 1988; Holton & Hakim, 2012; McGuffie & Henderson-Sellers, 2014; Marshall & Plumb, 2007; Peixoto & Oort, 1992). However, note that it has been argued that ocean carbon models might be still in the illusion and preparadigm phase (Le Quere, 2011). From our discussions above we conclude that more scientific efforts yield to more complex views, which can be even contradictory. This indicates that at least some underlying assumptions might need to be revisited, reflecting Phase 3 of Kuhn (1962) theory of scientific development. In modeling, there is a clear “dilemma of growth” in that the models are required for understanding the dynamics of a system as complex as the Earth System. On the other hand, the growing complexity of the models themselves seems to jeopardize understanding (Gramelsberger et al., 2020; Held, 2005). Therefore, different levels of complexity are required to understand the Earth system on long time scales (e.g., Claussen et al., 2002; Ganopolski & Rahmstorf, 2001; Lohmann & Ditlevsen, 2019; Timmermann & Lohmann, 2000; Tziperman et al., 2006). Future research may be enhanced along three directions data, statistics, and models, leading to an increase in the current knowledge about the climate evolution. From an epistemological viewpoint, it is crucial that researchers deepen or acquire the ability to integrate all three directions into their arsenal of methods. Since it is hard to disagree with the simple statement “more data are better,” the task here is rather to identify those dimensions in the data space where invested resources may yield to a maximum of new information. In this way, data assimilation techniques could help for an estimate of the state of the system and also its uncertainty (Burgers et al., 1998; Kalman, 1960; Nerger & Hiller, 2013). Finally, a detailed analysis of paleoclimatology shall test the assumptions about the climate recorder systems regarding leads, lags, and other filter properties (Laepple et al., 2011; Lohmann et al., 2013b).

Appendix A: COSMOS Model

The atmosphere is represented in our model by means of the ECHAM5 atmosphere general circulation model (Roeckner et al., 2003). ECHAM5 is based on a spectral dynamical core and resolves in our model setup 19 vertical hybrid sigma-pressure levels. The series of spectral harmonics is curtailed via triangular truncation at wave number 31 (approximately 3.75 × 3.75). Beyond climate system components that are conventionally included in atmosphere general circulation models, ECHAM5 resolves, via its land surface component JSBACH (Raddatz et al., 2007), various land surface processes and a representation of vegetation dynamics (Broykin et al., 2009). Ocean circulation and sea ice dynamics are computed by the MPIOM component JSBACH (Raddatz et al., 2007) and previous (Pfeiffer & Lohmann, 2016) interglacials as well as of the Last Glacial Maximum (Zhang et al., 2013). The coupled model system ECHAM5/MPIOM is described by Jungclaus et al. (2006). A concise description of the application of the COSMOS for paleoclimate studies is given by Stepanek and Lohmann (2012). The COSMOS version we are using has proven to be a suitable tool for the study of the Earth’s past climate, from orbital to tectonic time scales. Examples for applications within the time period of the last 1 Myr include orbitally forced climate simulations of the current (Lohmann et al., 2013a; Wei et al., 2012; Wei & Lohmann, 2012) and previous (Pfeiffer & Lohmann, 2016) interglacials as well as of the Last Glacial Maximum (Zhang et al., 2013). While COSMOS is numerically more efficient than current state-of-the-art climate models, its accelerated integration still takes more than 10 real-time calendar months. We note that when exposing the COSMOS to a 100 year sampling of orbital parameters we derive a deep ocean that is less equilibrated to the applied orbital forcing than it actually has been the case during the past. Yet, as stated by Lorenz and Lohmann (2004) who describe and apply this method, even for 100 times accelerated orbital forcing, the necessary conditions for simulating a sensible climate response are fulfilled: First, the applied accelerated changes in seasonal insolation are small compared to the seasonal cycle. Second, the modes of internal variability in the
atmosphere and in the surface ocean are faster than variations in the orbital forcing. The standard model code of the COSMOS version COSMOS-landveg r2413 (2009) is available upon request from the Max Planck Institute for Meteorology in Hamburg (https://www.mpimet.mpg.de). The temperature data of COSMOS related to Figures 4 and 5 are available through PANGAEA (Stepanek and Lohmann, 2020).

**Appendix B: Defining DO Events With the Bayesian Change Point Analysis**

For the identification of DO event dating of the last glacial is based on

(1) a time sequence based on a threshold criterion defined by Ditlevsen et al. (2005), followed by
(2) a Bayesian approach which allows us to verify choices of events with uncertainties.

---

**Table B1**

 права: 2021 AGU. All rights reserved.

Timing of the DO Events for the Last Glacial Found in the NGRIP Ice Core of Greenland (GICC05 Time Scale)

<table>
<thead>
<tr>
<th>Event</th>
<th>DO time sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ras14</td>
</tr>
<tr>
<td>0</td>
<td>11,703</td>
</tr>
<tr>
<td>A</td>
<td>13,130</td>
</tr>
<tr>
<td>1</td>
<td>14,692</td>
</tr>
<tr>
<td>B</td>
<td>16,520</td>
</tr>
<tr>
<td>C</td>
<td>20,740</td>
</tr>
<tr>
<td>2</td>
<td>23,340</td>
</tr>
<tr>
<td>3</td>
<td>27,780</td>
</tr>
<tr>
<td>4</td>
<td>28,900</td>
</tr>
<tr>
<td>5</td>
<td>32,500</td>
</tr>
<tr>
<td>6</td>
<td>33,740</td>
</tr>
<tr>
<td>7</td>
<td>35,480</td>
</tr>
<tr>
<td>8</td>
<td>38,220</td>
</tr>
<tr>
<td>9</td>
<td>40,160</td>
</tr>
<tr>
<td>10</td>
<td>41,460</td>
</tr>
<tr>
<td>11</td>
<td>43,340</td>
</tr>
<tr>
<td>12</td>
<td>46,860</td>
</tr>
<tr>
<td>13</td>
<td>49,280</td>
</tr>
<tr>
<td>14</td>
<td>54,220</td>
</tr>
<tr>
<td>15</td>
<td>55,800</td>
</tr>
<tr>
<td>16</td>
<td>58,280</td>
</tr>
<tr>
<td>17</td>
<td>59,440</td>
</tr>
<tr>
<td>18</td>
<td>64,100</td>
</tr>
<tr>
<td>19</td>
<td>72,340</td>
</tr>
<tr>
<td>20</td>
<td>76,440</td>
</tr>
<tr>
<td>21</td>
<td>85,060</td>
</tr>
<tr>
<td>22</td>
<td>90,040</td>
</tr>
<tr>
<td>23</td>
<td>104,520</td>
</tr>
<tr>
<td>25</td>
<td>115,370</td>
</tr>
<tr>
<td>Total number</td>
<td>26</td>
</tr>
</tbody>
</table>

*Note.* The timings are taken from Rasmussen et al. (2014), Ditlevsen et al. (2007) and Barker et al. (2011). Additionally, the timing obtained from thresholding and Bayesian estimation are provided. For the total number of events we do not consider the start of the Holocene (event 0).
We will consider exact dating sequences used in previous studies. As DO events are similar in shape and characterized by the abrupt warming we will reduce the single events to single time points of warming onset. Criterion (1) goes along with the assumption that DO events represent repeated oscillations between interstadial and stadial conditions with a threshold crossing process involved (Alley et al., 2003; Ditlevsen et al., 2007; Steffensen et al., 2008). We consider a threshold criterion for the definition of DO events. In accordance with Ditlevsen et al. (2005) we define a time point an event at the first upcrossing of a higher threshold (initiation of an interstadial state), if it is consequently followed by an up-crossing of a lower threshold (relaxation process to stadial mode). A trend for the scaled NGRIP ice core is estimated by single spectral analysis (SSA) and subtracted, the higher threshold is set to 1, and the lower threshold is chosen to be \(-0.9\) anomaly from the estimated trend. The asymmetry is justified by the difference in state duration, stadial modes persist longer (Ditlevsen et al., 2005).

The Bayesian change point analysis provides us with uncertainty estimates of the number of change points and allows us to input a priori knowledge (Ruggieri, 2013). It splits the problem of possible placements of change points into sub problems of smaller data chunks and uses a recursive dynamic programming for reduction in computational costs. For this study we only consider those time points as DO events that result a certainty of location greater than 0.5. For the NGRIP record this results in the choice of 22 events (Figure 6 and Table B1).

### Appendix C: The Multiscale Ocean Circulation-Radiocarbon Model

The AWI Earth System Model (AWI-ESM) is an extension of the AWI-CM (Sidorenko et al., 2015, 2019) for earth system modeling (https://fesom.de/models/awi-esm/). The atmospheric module is represented by the general circulation model ECHAM6 (Stevens et al., 2013) (here in T63 horizontal resolution, \(-180\) km) including a land surface model (JSBACH) which is based on a tiling of the land surface and includes dynamic vegetation with 12 plant functional types and two types of bare surface (Raddatz et al., 2007). The ocean and sea ice component is the Finite Element Sea Ice-Ocean Model (FESOM) (Danilov et al., 2004; Rackow, Goessling, et al., 2018; Rackow et al., 2019; Sidorenko et al., 2011; Timmermann et al., 2009) which is discretized on a triangular grid with a continuous conforming representation of model variables. The mesh nodes are vertically aligned to avoid difficulties in resolving the hydrostatic balance. The model uses variable resolution, which can reach \(20\) km in the Arctic and along coastlines. A no-slip boundary condition along the coast is implemented in the model (Figure 8). Surface stress and buoyancy fluxes are derived from the ice-ocean coupling. FESOM model has been validated in Timmermann et al. (2009) and Scholz et al. (2013), and the coupled AWI climate model in more recent studies (Rackow et al., 2019; Sidorenko et al., 2015). The source code of the climate model is available from the AWI based svn repository (https://swrepo1.awi.de/projects/awi-cm/).

The radiocarbon-climate model is based on FESOM2 (Danilov et al., 2017). This model has been used as the ocean-sea ice component in the AWI climate model (Sidorenko et al., 2019). The current version of FESOM2 is available from the public GitHub repository at https://github.com/FESOM/fesom2 under the GNU General Public License (GPLv2.0). Radiocarbon is treated as \(^{14}\)C following Toggweiler et al. (1989), with an is available from the public GitHub repository at https://github.com/FESOM/fesom2 under the GNU General Public License (GPLv2.0). Radiocarbon is treated as \(^{14}\)C following Toggweiler et al. (1989), with an

\[ t_{1/2} = \frac{1}{\lambda} \ln 2 \]

where \(t_{1/2}\) is the half-life of \(^{14}\)C, and \(^{14}\)C/\(^{12}\)C is the age fractionation corrected \({}^{14}\)C/\(^{12}\)C ratios in atmosphere and ocean. High radiocarbon concentrations in water translate into low radiocarbon ages and vice versa. Note that “conventional” radiocarbon age values are based on the “Libby half-life” of \(5,568\) yr (Stuiver & Polach, 1977), while the most recent estimate of the true half-life is \(5,700\) yr (Adieu et al., 2003; Bé et al., 2013).

### References


References


Stepanek, C., & Lohmann, G. (2020). Orbitally forced simulated surface air temperature of the last 1,000,000 years. PANGAEA. https://doi.pangaea.de/10.1594/PANGAEA.913450


