

Constraining ocean diffusivity from the 8.2 ka event

Alexander Lorenz · Hermann Held ·
Eva Bauer · Thomas Schneider von Deimling

Received: 17 September 2008 / Accepted: 29 April 2009 / Published online: 22 May 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Greenland ice-core data containing the 8.2 ka event are utilized by a model-data intercomparison within the Earth system model of intermediate complexity, CLIMBER-2.3 to investigate their potential for constraining the range of uncertain ocean diffusivity properties. Within a stochastic version of the model (Bauer et al. in *Paleoceanography* 19:PA3014, 2004) it has been possible to mimic the pronounced cooling of the 8.2 ka event with relatively good accuracy considering the timing of the event in comparison to other modelling exercises. When statistically inferring from the 8.2 ka event on diffusivity the technical difficulty arises to establish the related likelihood numerically per realisation of the uncertain model parameters: while mainstream uncertainty analyses can assume a quasi-Gaussian shape of likelihood, with weather fluctuating around a long term mean, the 8.2 ka event as a highly nonlinear effect precludes such an a priori assumption. As a result of this study the Bayesian Analysis leads to a sharp single-mode likelihood for ocean diffusivity parameters within CLIMBER-2.3. Depending on the prior distribution this likelihood leads to a reduction of uncertainty in ocean diffusivity parameters (e.g. for flat prior uncertainty in the vertical ocean diffusivity parameter is reduced by factor 2). These results highlight the potential of paleo data to constrain uncertain system properties and strongly suggest to make further steps with more complex models and richer data sets to harvest this potential.

Keywords Uncertainty · Ocean diffusivity · Bayesian analysis · 8.2 ka event · Paleo data

Abbreviations

AR(1)	One-dimensional autoregressive process
BP	Before present (before 1950)
CLIMBER	CLIMate-BiosphERE model
CS	Climate sensitivity
EMIC	Earth system model of intermediate complexity
GCM	Global Circulation Model
GICC05	Greenland Ice Core Chronology, 2005
GRIP	GRenland Ice core Project
iid	Identically independently distributed
IPCC	Intergovernmental panel on climate change
ka = kyr	Thousand years
LIS	Laurentic ice sheet
MOC	Meridional overturning circulation
NADW	North Atlantic deep water
NH	Northern hemisphere
pdf	Probability distribution function
ppm	Parts per million
SST	Sea surface temperature
std	Standard deviation

1 Introduction

Timing and magnitude of changes in atmospheric mean temperature in response to changes in greenhouse gas concentrations strongly depends on both, climate sensitivity and ocean heat uptake. The magnitude of climate sensitivity has been subject of intense research over the last decade (e.g. Forest et al. 2002; Hegerl et al. 2006; Knutti et al. 2002; Schneider von Deimling et al. 2006; Roe and Baker 2007; Allen and Frame 2007), and quite some effort has been spent on ocean heat uptake (Polzin et al. 1997;

A. Lorenz (✉) · H. Held · E. Bauer · T. S. von Deimling
Potsdam Institute for Climate Impact Research (PIK),
P.O. Box 60 12 03, 14412 Potsdam, Germany
e-mail: lorenz@pik-potsdam.de

Ledwell et al. 2000; Collins et al. 2006; Raper et al. 2002; Stouffer et al. 2006; Forest et al. 2008). But until now, climate sensitivity and secular ocean heat uptake are subject to large uncertainty. On the one hand, key processes in global circulation models need to be parameterised, giving room for only semi-determined parameter settings. On the other hand twentieth century's global warming signal makes it difficult to independently infer on climate sensitivity and secular ocean heat uptake. Both are strongly correlated as seen in twentieth century data.

In this situation it appears attractive to search for additional data sources that ideally were statistically independent from the anthropogenically induced warming signal. In (Schneider von Deimling et al. 2006), the relatively large signal-to-noise ratio of the glacial to interglacial climate transition comparison helped to reduce uncertainty in climate sensitivity, especially to rule out high sensitivity model versions as being inconsistent with reconstructed glacial cooling. The key idea was to utilize a climate model that represents both modern-day climate as well as an alternative climate state of the past without retuning the key uncertain parameters that would in turn affect the magnitude of climate sensitivity. In this article we consider whether the analogue approach could be implemented for constraining ocean heat uptake, which strongly depends on vertical mixing, in turn being represented in climate models as uncertain vertical ocean diffusivity parameter. We ask which past climatic event could have been strongly shaped by vertical mixing and would, therefore, possibly allow to infer on the related model parameters. We selected the so-called 8.2 ka event for our study as it represents a pronounced and well-dated transient climate signal that should be strongly influenced by vertical mixing in the Northern Atlantic Ocean.

The 8.2 ka event (or 8k event) refers to an outstanding cooling event in paleoclimate records at approximately 8,200 years before present [BP before 1950, that is 8240 before 2000 (b2k)] (Rohling and Pälike 2005; Alley and Ágústsdóttir 2005; Thomas et al. 2007). The event was first reported in the Greenland ice core records as an abrupt cooling of about $6 \pm 2^\circ\text{C}$ at summit, Greenland, which lasted roughly two centuries (Johnsen et al. 1992; Dansgaard 1993; Alley et al. 1997). Since then much has been published about the characteristics of this event, concerning the duration, the range, the driving mechanisms and the implications. Thomas et al. (2007) (where one can find a comprehensive overview over the discussion) describes the 8.2 ka event as a 160.5 years cold period (from about 8250 to 8090 BP), where decadal-mean oxygen isotopic values of a compound of four Greenland ice cores were below the early Holocene average (9.3–8.3 kyr BP). The minimum of $\delta^{18}\text{O}_{\text{ice}}$ is observed in the GRIP ice core at a calendar date of 8190 BP, dated on the GICC05

age scale (Rasmussen et al. 2006a). During the event $\delta^{18}\text{O}_{\text{ice}}$ drops about 1.5 per mille, which corresponds to a surface air temperature decrease of 3–6 K depending on the transformation method (e.g. Johnsen et al. 1995; Cuffey and Clow 1997; Johnsen et al. 2001). Besides reduced Greenland temperature the northern climate during the 8.2 ka event was characterized by a fresher and colder North Atlantic Ocean, drier and stronger winds over the northern Atlantic, drier monsoon regions and intensified North Atlantic trade winds, according to (Alley et al. 1997). A variety of additional paleoclimatic data from locations in the Northern Hemisphere (NH) show climate anomalies in the same time regime (overview of references from Bauer et al. 2004).

When utilizing paleo data from the 8.2 ka event for a model data intercomparison we have to choose an appropriate model representation of the event as well as an appropriate subset of the available paleo data. Therefore, we make several assumptions concerning the driving physical processes and the temporal and spatial extension of the event.

First of all we assume that the Earth System model of intermediate complexity CLIMBER-2.3 is in principle able to reproduce the 8.2 ka event. This assumption is based on the fact that simulated Greenland temperature (Bauer et al. 2004) following a realistic forcing scenario of the cold event agree reasonably with paleo data. The question of the cause of the 8.2 ka event has been addressed by different suggestions. The possible causes mainly discussed (e.g. in Kobashi et al. 2007) are changes in solar irradiation, as investigated by Muscheler et al. (2004) or Renssen et al. (2006), and freshwater fluxes, investigated by Wiersma and Renssen (2006), Wiersma et al. (2006), from the drainage of glacial lakes to the northern Atlantic. In the latter case the weakening of deep water formation in the northern Atlantic and, therefore, reduced northward heat transport by the Atlantic Thermohaline Circulation (THC) could have caused the cold event (Barber et al. 1999; Clark 2001; Rahmstorf 2002). This thesis is corroborated by the observation of the relatively long cold event duration of 160 years which point towards the involvement of oceanic processes. Numerous model simulations have been performed (e.g. Renssen et al. 2001; Renssen et al. 2002; Bauer et al. 2004; Wiersma and Renssen 2006; Wiersma et al. 2006; LeGrande et al. 2006) that have been able to reproduce an asymmetric cold event induced by freshwater pulses of different strength and duration. Evidence of the drainage of glacial lakes Agassiz and Ojibway in an outburst at about 8470 BP (^{14}C time) (Liccardi et al. 1999; Leverington et al. 2002; Teller et al. 2002) deliver a plausible scenario of a strong pulse-like freshwater forcing to the North Atlantic region for the time of interest. The causal link between the drainage of lake Agassiz and the

weakening of North Atlantic ocean circulation has recently been supported by proxy records taken in the Labrador basin (Kleiven et al. 2008). The uncertainties in both the timing of the cold event in ice core data and the timing of the drainage of Lake Agassiz leave space for a prompt response of the Greenland temperature cooling following a freshwater pulse as suggested from various climate model studies. Here we follow the experimental setting of Bauer et al. (2004), who, employing a climate model of intermediate complexity, were able to reproduce a cold event with a duration that exceeded the scale of the freshwater forcing considerably by using a pulse-like drainage of 2.6 Sv, released for 2 years.

Although we are aware of an oceanic role affecting the cold event by changes in the meridional overturning circulation leading to an hemispherical extension of climate changes around the 8.2 event we limit our model data intercomparison to the Greenland ice core data, due to the poor temporal and spatial resolution of data from outside of Greenland. Further we assume the event to be strongly influenced by vertical mixing, which is determined by ocean diffusivities. In reality the event might also be influenced by other processes, not represented in CLIMBER-2.3. This potential structural dependence of results on the setting needs to be countered by further, independent analysis with other models. Therefore, our results can only be taken as an upper bound of available information from the specific model-data setting.

Constrained by these assumptions, by utilizing Bayesian Analysis, we aim at an informative influence chain from the 8.2 ka event onto ocean diffusivity parameters in a dynamically consistent way, within the stylised world of CLIMBER-2.3. This is to be seen as an incremental progress in systematic analysis of causes and context of the 8.2 ka event, in particular in relation to ocean diffusivity, and its potential for future paleo data—GCM intercomparison projects.

2 Methods

2.1 The bayesian algorithm

A general overview of application of Bayesian Analysis within climate science is given in Appendix 1. For our special case of CLIMBER-2.3 the application of Bayesian inference reads as follows: comparing numerous model realizations of the 8.2 ka event produced by one and the same climate system model, only differing in the values of a number of model parameters (ocean diffusivities and experiment related parameters) which have a high influence on the model performance at the cold event, to the paleo records some parameter combinations might result in

an appropriate representation of the cold event while others can be ruled out as the model output is inconsistent with the paleo data. As the time dependence of the simulated temperature response depends on stochastic freshwater forcing, that means the resulting cold event differs in duration for each single realization of noisy freshwater forcing, a combination of model parameters cannot simply said to be ruled out but every parameter value is assigned a certain likelihood of reproducing the correct cold event duration seen in the data. Repeating this procedure for a whole ensemble of prior-weighted parameter values one ends up with a distribution function on the space of parameters that represents the probability of a certain parameter value given the information of the 8.2 ka event.

More formally spoken the output of the model of intermediate complexity CLIMBER-2.3 is compared to Greenland ice-core data displaying the 8.2 ka event to reduce uncertainty of model parameters α (i.e. a vector). The model parameters α are chosen to contain the horizontal and vertical ocean diffusivity (a_{hoc} , a_v), which are supposed to have strong influence on the model performance at reproducing the 8.2 ka event. The comparison is complicated by the fact that the 8.2 ka event in CLIMBER-2.3 does not only depend on α but also on a particular realization η of noisy freshwater forcing. So several transformations are applied after which the model output of “CLIMBER-2.3n” (the noisy version of CLIMBER-2.3) can be compared to observations, which themselves are aggregated to CLIMBER box scale. Bayes’ formula then reads:

$$P_{\text{post}}(\alpha) = \frac{P_{\text{prior}}(\alpha)P(y|\alpha)}{\int d\alpha' P_{\text{prior}}(\alpha')P(y|\alpha')}. \quad (1)$$

y denotes the observational spatiotemporal data in terms of CLIMBER-2.3 scale aggregated fields.

Applying this method to the 8.2 ka event involves several challenges. (1) The likelihood $P(y|\alpha)$ for given α is not known a priori for CLIMBER-2.3n and therefore has to be estimated by running an ensemble of test runs of the 8.2 ka event (i.e. realizations η of noisy freshwater forcing). The complexity of this estimation in terms of necessary numbers of ensemble members rises by a factor of order $10^n - 1$ if n is the dimension of y . To reduce the complexity of comparison the information contained in y is reduced by nonlinear projection of both the data and the model output on the (scalar) duration of the cold event measured by a least square fit of a trapezoid function to data and model output (see Fig. 1). That duration encapsulates a major fraction of the information contained in the original time series y . In Bayes formula (1) the likelihood is replaced according to $P(y|\alpha) \rightarrow P(T|\alpha)$, whereby T denotes the extracted duration of the cold (8k) event, obtained from trapezoidal fitting. (2) The observations are only on proxies

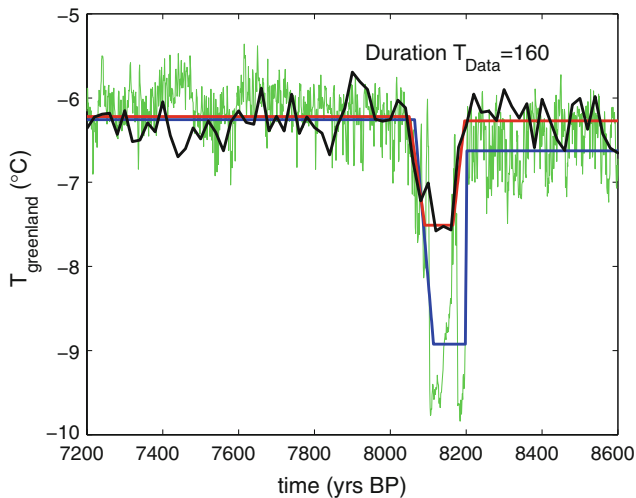


Fig. 1 The nonlinear trapezoid fitting procedure to estimate the duration of the cold event in Greenland data and model output. *Black curve*: the Kriging Mean of the Greenland ice core data (in $\delta^{18}\text{O}_{\text{ice}}$ offset added to be displayed on temperature scale together with CLIMBER-2.3 output), *green curve*: example CLIMBER-2.3 Greenland temperature of the cold event with added weather noise, *red curve*: trapezoid fit to Greenland data, *blue curve*: trapezoid fit to CLIMBER-2.3 output

of temperature instead of temperature itself and (3) the observations are obscured by local weather noise not represented by η . Item (2) is simultaneously addressed with item (1) by projecting onto T as the duration of the event is not affected by the proxy-temperature transfer functions. Item (3) is taken into account by inclusion of additive weather noise: $y \rightarrow y' = y + \zeta$, $\zeta_i \text{ iid}^1 \sim N(0, \sigma)$ and $T = T(y')$. Thereby the amplitude of this Greenland weather noise σ_{weather} can be derived from the Kriging (Wackernagel 1995) of the Greenland ice core data (for a detailed description of this noise model see Appendix 2).

2.1.1 Implementation of bayesian updating

The likelihood function $P(T|\alpha)$ is reconstructed for every α only in a vicinity of the specific T found in the Greenland data by a histogram of bin-size ΔT . Per i , indicating the sampling of $\alpha_1, \dots, \alpha_i, \dots, \alpha_I$ in α -space the posterior probability reads:

$$P_{\text{post}}(\alpha_i) \approx \frac{P_{\text{prior}}(\alpha_i) f_i}{\sum_i P_{\text{prior}}(\alpha_i) f_i} \quad (2)$$

where

$$f_i = \frac{1}{N_i} \sum_{jk} \text{Ind}(T_{ijk} \in [T - \Delta T/2, T + \Delta T/2]) \quad (3)$$

¹ Identically independently distributed.

with j, k denoting the index of η and ζ sampled in a factorial design² (as application of weather noise to the CLIMBER output is computational cheap), N_i the number of η and ζ realizations for each α_i . ‘Ind’ is the indicator function that is 1 if the Boolean argument is true and 0 otherwise, $P(\alpha_i)$ denotes the finite probability for the realization α_i (the α domain is coarsely resolved), while p represents respective densities in analytic form.

As one necessary test on convergence of the procedure and optimal choice of ΔT , a bootstrapping (see Efron and Tibshirani 1993) approach is implemented³. This approach leads to optimal bin sizes $\Delta T \approx 5$ years. But the resolution of Greenland ice core data of only 20 years prohibits smaller bin sizes. Therefore, we adjusted the bin size ΔT to this data induced minimum value of 20 years.

2.2 Model and data

2.2.1 The Greenland ice core data

The European Greenland ice core Project (GRIP) (GRIP-Project-Members 1993), the parallel US Greenland ice sheet project 2 (GISP2) (Mayewski et al. 1994) and the Dye 3 (Dansgaard 1985) and North GRIP (NGRIP) ice cores (NGRIP-Project-Members 2004) all represent the 8.2 ka event. Thomas et al. (2007) used different isotope data to determine the duration and the structure of the 8.2 ka event. In this study only the $\delta^{18}\text{O}_{\text{ice}}$ data were taken into account, synchronized to the GICC05 age scale with a resolution of 20 years as presented by Rasmussen et al. (2006b). For the application of the Bayesian Analysis the data are aggregated to CLIMBER-2.3 box scale⁴ by “Kriging” their mean (Wackernagel 1995). That method generates the mean value of the CLIMBER-2.3 box under consideration by weighting the data sets according to their covariance matrix. Roughly, it allocates the more weight to

² Implementing a factorial design means that in each parameter dimension a sample is chosen according to an appropriate distribution and the experimental units are then chosen as all possible combination of values from the single dimensions; therefore a factorial design might also be called a fully crossed design.

³ A bootstrapping of the the following ensemble is performed: the α_i whereby for any i , α_i is found $\text{Ind}(T_{ijk} \in [T - \Delta T/2, T + \Delta T/2])$ times in the ensemble. From that bootstrapping, the variance of the final output (posterior of diffusivities or posterior in derived observables like ocean heat uptake or climate sensitivity) z is plotted over ΔT , and the minimum of that curve is identified. In order to minimise twofold use of statistical information, sampling of CLIMBER-2n is repeated after ΔT has been fixed.

⁴ The CLIMBER-2.3 ocean submodel uses 20 uneven vertical layers and three longitudinal ocean boxes (Atlantic, Indic, Pacific) and has a latitudinal resolution of 2.5°. For atmosphere and land modules the latitudinal resolution is the same (10°). Atmosphere and land modules consist of seven equal longitudinal sectors of 51°.

a data source the more it is statistically independent from the other sources.

Naturally the next step would be the transformation of this Greenland-wide $\delta^{18}\text{O}_{\text{ice}}$ time series into a temperature record as the CLIMBER-2.3 output is in units of temperature. As the transformation functions from $\delta^{18}\text{O}_{\text{ice}}$ to temperature are highly uncertain, the absolute value of a so transformed record would be of no use for comparison to model output. Therefore, a different approach is used: Both the model output and the aggregated data are transformed to the duration of the event by a nonlinear fitting procedure of a trapezoid function, which takes the asymmetric evolution of the cold event into account. Thereby it is assumed that the duration T is roughly invariant under uncertainties of the $\delta^{18}\text{O}_{\text{ice}} \rightarrow T$ transformation. For the aggregated data this fitting is shown in Fig. 1. The non-linear fitting procedure was chosen for several reasons: First the assumption of an equilibrium state before and long after the cold event, only slightly disturbed by freshwater noise, naturally leads to a linear fitting of these periods. Second, the trapezoid fitting serves to classify of the “event-no event” border more than to rightly represent the whole time series. As for some realisations of freshwater noise the end of the cold event is disturbed by several fluctuations between the cold and warm states, the trapezoid fitting is more robust than pure smoothing methods like nonparametric fitting or running mean. The fitting was performed by using a local minimization algorithm within MATLAB in combination with an iterated Monte Carlo Sampling of starting points to address the problem of local optima. Although a global optimum can not be guaranteed, the results proved robust within the temporal resolution of the ice-core data. The resulting duration of the 8.2 ka event in the Greenland ice core data of 160 years is in good agreement with the findings of Thomas et al. (2007).

2.2.2 CLIMBER-2.3

In this study the climate model of intermediate complexity CLIMBER-2 version 3 is employed. CLIMBER-2.3 (CLIMATE-BiospERE model) is a 2.5-dimensional, low resolution climate system model designed for simulation of large-scale processes on time scales from seasonal to millennia and longer (Petoukhov et al. 2000). It consists of modules describing atmosphere, ocean, sea ice, land surface processes, and terrestrial vegetation cover. The atmosphere module is a dynamical-statistical 2.5-dimensional atmosphere model as the vertical structure of the atmosphere and the synoptic-scale activity are parameterised. The ocean component is composed of zonal mean ocean basins as used by (Schmittner and Weaver 2001). The submodels are coupled interactively without flux adjustments through

fluxes of heat and water and momentum is transferred from the atmosphere to the ocean.

CLIMBER-2 has been evaluated against data in various ways. The simulated climate characteristics of the atmosphere and the ocean for the preindustrial climate state agree well with observational data (Petoukhov et al. 2000). Several sensitivity studies have been performed (Ganopolski et al. 2001) to compare the model response to changes in solar insolation, carbon dioxide, freshwater flux and land cover with results of GCMs. The model response, e.g. to a CO_2 concentration increase, closely agrees with results of GCMs. A third possible method of model testing is the comparison of model output to paleoclimatic data. Driven by natural and anthropogenic forcings, the temperature variations of the last millenium were reproduced (Bauer et al. 2003). Aspects of glacial (21 kyrs BP) and mid-Holocene (6 kyrs BP) climate seen in paleo-data have successfully been reproduced (Ganopolski et al. 1998). Even abrupt climate changes can be reproduced (Ganopolski and Rahmstorf 2001). Nevertheless it has to be mentioned that the reproductions of aspects of paleo climate are not fully robust within the possible parameter ranges and are valid in face of large uncertainties about paleo climatic data. Therefore, large efforts in increasing both the quality of paleo data and parameterisation of models have to be undertaken. Bauer et al. (2004) used CLIMBER-2 with different (solar, freshwater) forcing mechanisms including noisy freshwater fluxes as a substitute for natural variability to reproduce a cold event in a climate state corresponding to early Holocene conditions around 9 kyr BP. By applying a freshwater forcing into the northern Atlantic basin consisting of a freshwater pulse, additive noise and different baseline fluxes which are constrained by proxy data and modelling studies, Bauer et al. (2004) could reproduce the amplitude and the centennial duration of the cold event. They found a dependency of the cold event duration on the realization of noisy freshwater forcing and suggested that the cold event duration can be considerably lengthened by natural freshwater noise forcing after preconditioning by a freshwater pulse and optional baseline fluxes. The essential finding is the existence of a metastable state of the overturning circulation inbetween the ON mode with present day characteristics of the circulation and the OFF mode without MOC. The INT state has nearly the same characteristics as the transient cooling signal from the 8.2 ka event, but is stable against small distortions in freshwater forcing within a hysteresis experiment.

The low computational costs of CLIMBER-2.3 allows the creation of huge ensemble climate scenarios necessary for the ensemble operationalisation of a Bayesian approach. The CLIMBER-2.3 model was used by Schneider von Deimling et al. (2006) in this way to constrain eleven

internal parameters which are most influential on climate sensitivity. The uncertainty reduction effect was propagated to climate sensitivity, to a range similar to the IPCC estimate (1.5–4.5°C) and thereby ruled out much higher estimates from other simulations. Here, we combine the methods of Bauer et al. (2004) and Schneider von Deimling et al. (2006) to systematically compare the model output containing the 8.2 ka event to the Greenland ice core data.

3 Model simulations

3.1 Experimental setup

Following Bauer et al. (2004) the transient climate simulations for the 8.2 ka event are started from a near equilibrium state adapted to the boundary conditions for 9 ka BP. These are the orbital parameters affecting solar irradiance (eccentricity, obliquity, and precession) (Berger 1978), the atmospheric CO₂ concentration of 261 ppm and a remnant Laurentide ice sheet on the North American continent (Marshall and Clarke 1999). The resulting 9 kyr climate state, reached by a 3 kyr equilibrium run per parameter setting, is characterized by nearly the same global and hemispherical temperatures in the annual mean as in the preindustrial state with 280 ppm but the seasonal temperature cycle is stronger than in the preindustrial state. For a detailed comparison of the 9 kyr state to the preindustrial state in CLIMBER-2.3 see (Bauer et al. 2004). In the simulation runs, the cold event is then forced at 8,200 years BP by a freshwater pulse released to the northern Atlantic Ocean,⁵ representing a pulse-like drainage of melt water from the Lake Agassiz through Hudson Bay as suggested by Teller et al. (2002) and Leverington et al. (2002). This pulse has a volume of $1.6 \times 10^{14} \text{ m}^3$ and was released very quickly (<1 year; Teller, 2007 personal communication). For numerical stability of derivatives within CLIMBER-2.3 the pulse duration was taken to be 2 years; that corresponds to a freshwater flux of 2.6 Sv. Sensitivity experiments by Bauer et al. (2004) have shown, that the cold event duration within CLIMBER-2.3 is only weakly affected by changes in volume of the freshwater pulse. Changes in the duration of the pulse from 1 year up to 30 years can not reproduce cold events of appropriate duration without inclusion of background fluxes or freshwater noise.

To lengthen the cold event duration to a sensible range (see Bauer et al. 2004) and to account for short term variability in the runoff, a noise model for natural

freshwater fluctuations and a baseline flux are added to the surface freshwater fluxes computed by the model. The noise is generated by a white noise model with adjustable standard deviation (σ) and a different seed for the noise generator is chosen for each realization of a simulation with a certain setting of parameters. Bauer et al. (2004) showed that by including this noise, the model's temperature response strongly depends on a certain realization of the noise. Thus this noisy version of the 8.2 ka event in CLIMBER-2.3 calls for an ensemble approach to estimate the influence of the different parameters.

The additional baseline flux represents enhanced runoff from the two possible runoff routes: Hudson Bay and St. Lawrence strait. In CLIMBER-2.3 these routes are represented by introducing additional fluxes in the Atlantic grid cells between 50°–70°N (Hudson) and 40°–50°N (St. Lawrence). There exist different estimates for the strength and the duration of these additional fluxes (Teller et al. 2002; Clark 2001). For practical reasons, that means reduction of dimensions, in this study only one additional baseline flux in the grid cells between 50°–70°N (Hudson) is introduced. As Bauer et al. (2004) showed, such an additional baseline can prolong the duration of the cold event considerably. The baseline flux can alter in duration and strength and the noise may vary in amplitude (std). So the experimental setup of the 8.2 ka simulation introduces at least three additional uncertain parameters to deal with (four if the uncertain early Holocene background freshwater forcing is also taken into account). The freshwater forcing components introduced in the experimental 8.2 ka setup are displayed schematically in Fig. 2 and all relevant experimental parameters are listed in Table 1.

3.2 Sampling strategy

Within CLIMBER-2.3, 11 uncertain parameters strongly influence key climate state properties. In principle, our Bayesian analysis would have to address that 11 D parameter space. However, as in this conceptual study we address ocean properties (in particular the 8.2 ka event) only, for the sake of transparency we confine the analysis to the 2 D parameter space of ocean diffusivities. In the following we describe how we numerically address the three ingredients of the Bayesian formula: prior, likelihood, and integrated probability (i.e. the denominator) of observing the climate state that nature displays.

We construct the prior in two steps. (1) First the space of physically reasonable values for the diffusivities is chosen as the most conservative constraint. These ranges of values are given as expert knowledge by the constructors of CLIMBER (see Schneider von Deimling et al. 2006). The horizontal diffusivity at near surface depths $k_H = 200\text{--}5,000$ {standard value = 2,000} m^2/s is directly addressed

⁵ Within CLIMBER this part of the northern Atlantic is represented by the ocean grid cells from 50–70°N.

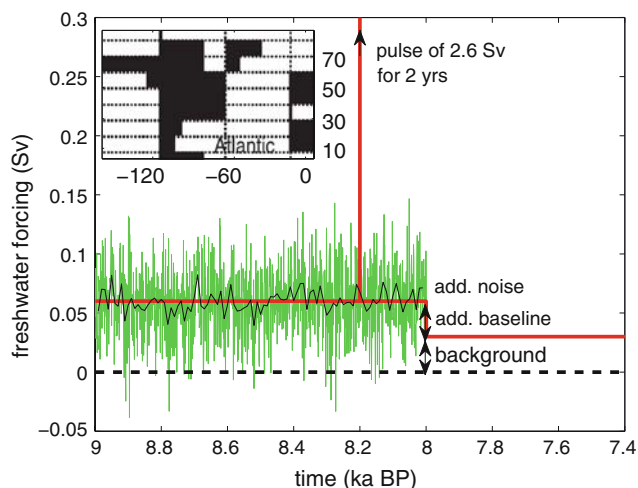


Fig. 2 Components of northern Atlantic freshwater forcing within the 8.2 ka experiment setup (all in Sv): an unknown (but relatively constant) background freshwater forcing of 0–0.1 Sv is complemented by an additional baseline stemming from advanced freshwater runoff before and during the drainage of Lake Agassiz that ended approx. 200 years after the pulse like drainage, that consists of a flow of 2.6 Sv for 2 years at 8,200 years BP. The freshwater forcing is blurred by noise (green). The resulting 5 years running mean is shown in black

by the CLIMBER-2.3 variable a_{hoc} . The vertical diffusivity is taken to follow a vertical profile after Bryan Lewis with CLIMBER-2.3 variable $a_{\text{Kv}} = 0.5\text{--}1.5 \times 10^{-4} \{ \text{standard value} = 0.8 \times 10^{-4} \} \text{m}^2/\text{s}$ addressing the diffusivity at the turning point of the profile. We call this space the Physically plausible Domain. (2) As a second step of including prior knowledge the insights of Schneider von Deimling et al. (2006) are used. They applied constraints on the present day performance of the model to reduce uncertainty of 11 model parameters (including the ocean diffusivities). As an auxiliary step we would like to obtain a qualitative impression on the shape of the 2D domain of diffusivity parameters that comply with those present-day climate constraints, being a subset of the Physically plausible Domain. Accordingly an ensemble of 1,000 members is created according to a Monte Carlo scheme with values for a_{hoc} and a_{Kv} sampled on a logarithmic scale according to a beta distribution⁶ (indicated by β below) within the bounds of the Physically plausible Domain. In order to test whether a parameter combinations is in accordance with Present Day constraints, for any such ensemble member an equilibrium run of 3,000 years is performed under the boundary conditions of present day climate. Seven of the resulting climate characteristics are

⁶ The distribution chosen as factorial in the a_i and is of the form $P(x) \approx x^{(1-\alpha_i)} * (1-x)^{(1-\beta_i)}$, hereby x being an affine transform of $\log a_i$ such that $x \in [0, 1]$. The α_i and β_i are chosen such that the distribution is maximal at the standard values of a_i , leading to a nearly flat distribution in the Present Day Domain.

tested with respect to a set of requirements defined in Schneider von Deimling et al. (2006) to represent tolerable present day climate states. They contain intervals for the annual mean values that encompass corresponding empirical estimates.⁷ In Fig. 3 the parameter settings that pass all seven constraints are indicated by green dots. The resulting domain is called Present Day Domain. We now assume as prior probability density: $P(\alpha) = \beta(\alpha) \times \text{ind}(\alpha \in \{\text{Present Day Domain}\})$ (hereby “ind” is 1 if and only if α lies in the Present Day Domain, and 0 otherwise). This study now investigates a possible reduction of the boundaries of the diffusivities with respect not only to the Physically plausible- but also to the markedly stronger confined Present Day Domain.

Now the case is further complicated as the likelihood of interest does not only depend on the 11 parameters of the standard version of CLIMBER-2.3, here reduced to two parameters, but also on three further parameters in our stochastically extended version of CLIMBER-2.3: the noise amplitude, the duration and strength of the baseline flux. Within our incremental approach of analysis, we would like to strictly stick to an only two-dimensional framing of the problem. Hence we keep those three additional parameter (that we denote as $\gamma_1, \gamma_2, \gamma_3$) fixed as we do for the other 9 ($9 = 11 - 2$) standard CLIMBER-2.3 parameters. We decide to choose γ_i such that they maximise the likelihood function for the standard values of α (i.e. γ as “maximum likelihood value”). Now we need to establish the likelihood function L that is not analytically given for CLIMBER-2.3. In an auxiliary precursory step, L is utilised to fix γ . In principle, for any parameter combination (α, γ) , a histogram of duration T of the 8.2 ka event would have to be obtained. However, in the vicinity of the standard value for α , we numerically establish the following approximation: $L(\alpha, \gamma) \approx L(\alpha) \times N(\gamma_1) \times N(\gamma_2) \times N(\gamma_3)$ (N denoting a Gaussian). From that approximation we deduce γ as Gaussian means and display their numerical values in Table 1. Independently the amplitude of freshwater noise σ is bounded from below by data from Walsh and Portis (1999) who delivered estimates for the standard deviation of fluctuations in northern Atlantic freshwater budget from evaporation and precipitation. Rescaled to the North Atlantic region in CLIMBER-2.3 this corresponds to a minimum standard deviation of $\sigma = 0.02$ Sv (as lower bound being consistent with what we obtained by our maximum likelihood estimate). The histogram of cold

⁷ i.e. Surface Air Temperature 13.1–14.1°C; area of sea ice in the Northern Hemisphere 6–14 mil km² and in the Southern Hemisphere 6–18 mil km²; total precipitation rate 2.45–3.05 mm/day; maximum Atlantic northward heat transport 0.5–1.5 PW; maximum of North Atlantic meridional overturning streamfunction 15–25 Sv; volume averaged ocean temperature 3–5°C; for references see Sect. 7.2 in Schneider von Deimling et al. (2006).

Table 1 Experiment parameters of the 8.2 ka event simulation: shown are (a) the relatively certain and (b) the uncertain nuisance freshwater forcing parameters that are taken as their (a) known or (b) maximum likelihood values; (c) finally the uncertain internal parameters are listed with their initial quantiles

(a) Uncertain forcing parameter (maximum likelihood)			
Amplitude of freshwater noise	σ	0.05 Sv	
Strength of freshwater baseline	FW	0.03 Sv	
Duration of freshwater baseline after 9000 BP	D	1,000 years	
Early holocene background forcing relative to present day		0 Sv	
(b) Certain forcing parameter			
Duration of freshwater pulse		2 years	
Amplitude of freshwater pulse		2.6 Sv	
Timing of freshwater pulse		8200 BP	
(c) Uncertain internal parameter			
Horizontal ocean diffusivity	a_{hoc}	200–5,000 m ² /s	
Vertical ocean diffusivity	a_{Kv}	0.5–1.5 × 10 ⁻⁴ m ² /s	

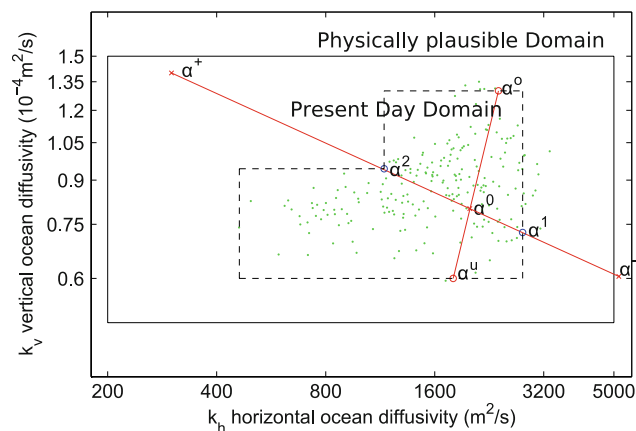


Fig. 3 The two-dimensional parameter space of horizontal ocean diffusivity (k_h) and vertical ocean diffusivity at the turning point of the Bryan Lewis profile (k_v) with different constrained domains: Physically plausible Domain represents the ranges of parameters for which the model is feasible, that is the largest physically feasible domain. The green dots represent that part of an equilibrium run ensemble under present day conditions which passes all of the seven present day constraints imposed by Schneider von Deimling et al. (2006); the resulting domain in diffusivity space is called Present Day domain. The parameter values marked by α^i represent a choice of loglinear combinations α of diffusivity parameters along a dimension $[\alpha^+ - \alpha^-]$ that is most influential on the Atlantic overturning circulation, as pointed out by Held and Kleinen (2004) (and along a dimension orthogonal to $[\alpha^+ - \alpha^-]$)

event durations (see Fig. 5) shows a dependence on the noise amplitude as a higher noise amplitude smoothes the histogram leading to higher likelihood for the right duration of the cold event with a maximum likelihood value at $\sigma = 0.05$ Sv. The maximum likelihood values of duration and strength of the additional freshwater baseline were

found as $D = 1,000$ years and $FW = 0.03$ Sv. As both parameters have a potentially strong influence on the cold event duration in further investigations this maximum likelihood choice has to be replaced by a more systematic approach.

Having fixed γ , we proceed to numerically establish $L(\alpha)$. We now utilise a problem-adjusted version of importance sampling Robert and Casella 1999 (i.e. denser sampling where we can expect L to be larger) along the following line of reasoning: about 20 samples are taken within the 2D Present Day subdomain of the ocean diffusivities ($a_{\text{hoc}}, a_{\text{Kv}}$), primarily along the line (in log space) between the parameters α^- , α^0 and α^+ . This follows the construction of a parameter α out of the diffusivities which is most influential on the Atlantic overturning stream function and therefore most likely also on the 8.2 ka event as done by Held and Kleinen (2004). To cover two dimensions, samples are also taken along the direction about orthogonal to α . The sampling then is iterated to resolve more closely the small domain in which the likelihood is nonzero. This strategy led to an overall sampling of 24 samples within the two dimensional diffusivity space.

For these points the likelihood of correctly reproducing the 8.2 ka event has been established by running the 8.2 ka scenario about 300 times. This represents a total computational cost of 24 (samples) × 300 (runs per sample) × 3 (CPU hours per run) ≈ 22,000 CPU hours. Under usage of the standard approach for estimating the denominator of the Bayesian formula (Robert and Casella 1999) and the assumptions of quasi linear prior and the resulting gaussian likelihood in logarithmic diffusivity space the likelihood can simply be normalized by the sample size to derive the posterior distribution.

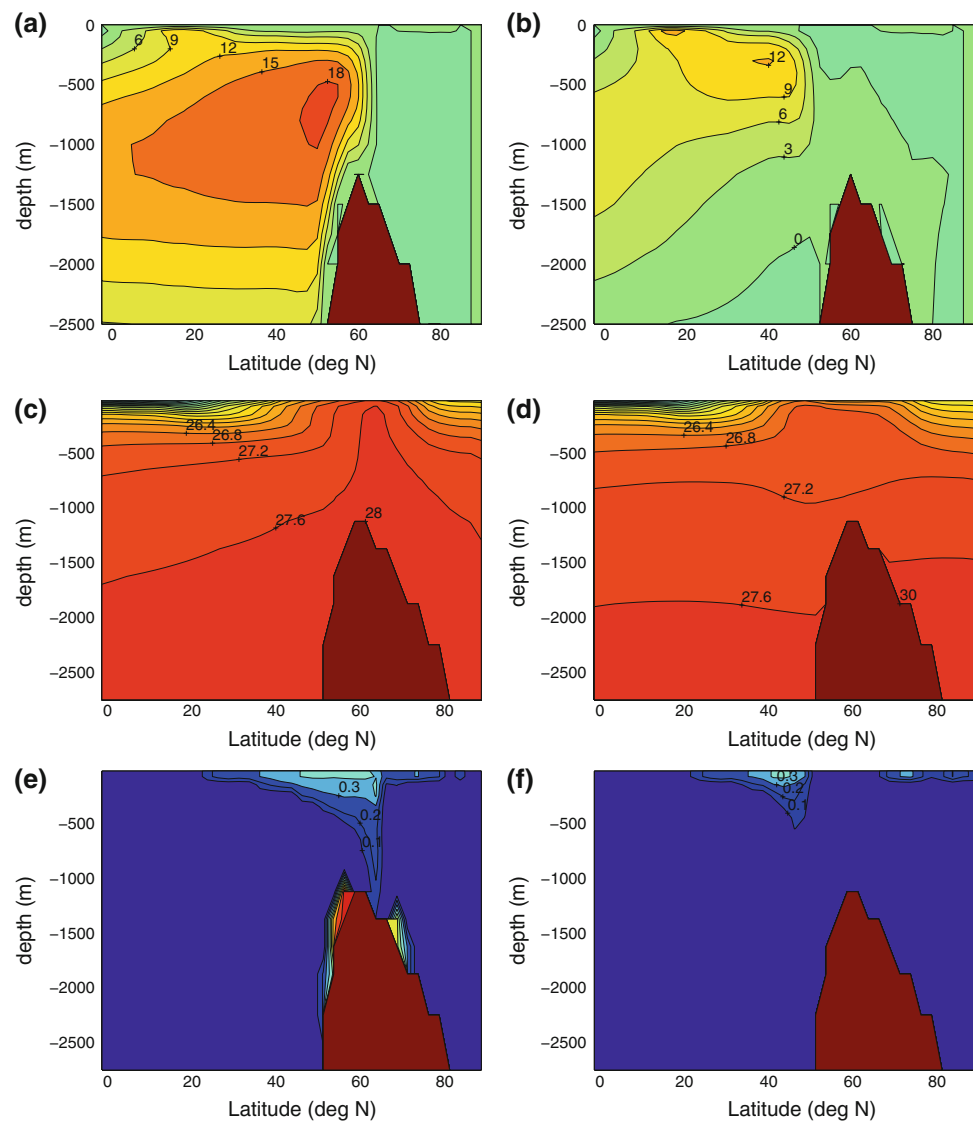
4 Results

4.1 Interpretation of the 8.2 ka event in CLIMBER-2.3

The simulation of the 8.2 ka event is performed according to the experimental setup described above. The field output of meridional overturning circulation (MOC), potential density and Frequency of occurrence of convection events both, before and during the cold event are shown in Fig. 4.

The left column represents the state of the northern Atlantic ocean before the freshwater pulse is applied. The well known North Atlantic conveyor belt is well represented in the meridional stream function. The relatively warm and saline water is transported north by the near surface North Atlantic current. The potential density $\rho = f(T, S)$ (Fig. 4c, d) that depends on temperature and salinity shows a vertical instability as the isolines proceed vertically, thus downward convection takes place. The

Fig. 4 Characterization of the cold event in various state variables in the northern Atlantic. From *top to bottom*: Atlantic meridional stream function in Sv, potential density in kg m^{-3} above 10^3 kg m^{-3} and frequency of occurrence of convection events without (*left*) and with (*right*) cold event in transient 8.2 ka event simulation. Isolines are in steps of 3 Sv, 0.4 kg m^{-3} and 0.1, respectively



water sinks down at the Iceland–Greenland ridge and flows southward as North Atlantic Deep Water. Normally, that means in the standard Holocene setting denoted as ON mode of the North Atlantic Overturning Circulation, the maximum of this circulation is located slightly north of the Iceland–Greenland ridge at a depth between 500 and 1,000 m.

At 8.2 ka BP an enormous amount of freshwater is released into the surface layer of the northern Atlantic. Thus the density of the surface layer is lower than the deeper ocean, the water column becomes stable (Fig. 4d) and the deep convection stops immediately. As the northward transport of warm saline water does not stop, the overturning is not turned off completely but shifted southward; the surface water now sinks at a latitude of 40–50°N (see Fig. 4b). During the event the convection is shifted south and consists of a purely wind driven part at the surface and a slowed overturning that reaches only

500 m downwards. The OFF mode only shows the wind driven surface current without any convection events (not shown).

The vertical diffusivity influences the rate of occurrence and the vertical range of mixing events. Therefore, a higher vertical diffusivity smoothes the gradient in potential density and reduces the instability that drives the overturning circulation. Thus the MOC is weaker for higher diffusivities and recovers more slowly from the 8.2 ka cold event. As the overturning is weakened at 8.2 ka BP, the northward heat transport is also reduced and thus the temperature in the northern hemisphere decreases whereas the southern hemisphere becomes warmer due to the so called seesaw effect (Crowley 1992). As the overturning does not stop completely, a northward transport of warm water below the surface continues and warm water accumulates north of the original overturning area. Caused by relatively small perturbations (from synoptic scale

freshwater fluxes) this warm water can then restart the circulation very quickly. This process may explain the fast recovery of the deep overturning at the end of the cold event. The cooling is seen strongest in the North Atlantic region (about 5°C in Greenland temperature. The cooling is accompanied by lesser precipitation. In this example the cold period lasts about 250 years.

4.2 Histogram of cold event durations

Running the 8.2 ka scenario several hundred times 150–350 for each combination of diffusivities and considering only the duration of the cold event (computed by a non-linear trapezoid fit) one ends up with a histogram of cold event duration (see Fig. 5). The histogram reveals a system of at least two different modes of duration: A short mode around 80 years and a longer mode centred about 30 years after the termination of the additional baseline flux used in the experiment. This points to coexisting physical effects as origin for the modes, represented by different factors in the experimental setup. Sensitivity analysis on the experiment related parameters (strength and duration of freshwater baseline flux, amplitude of noise) and comparison of different model output (density- and salinity field output, stream function) point to the following explanation: The short mode represents the mean lifetime of the shortening of overturning circulation, which is not altered considerably by different values of freshwater strength unless the baseline gets strong enough to completely shut down the circulation. The second mode is clearly triggered by the termination of the baseline flux. It seems clear that a

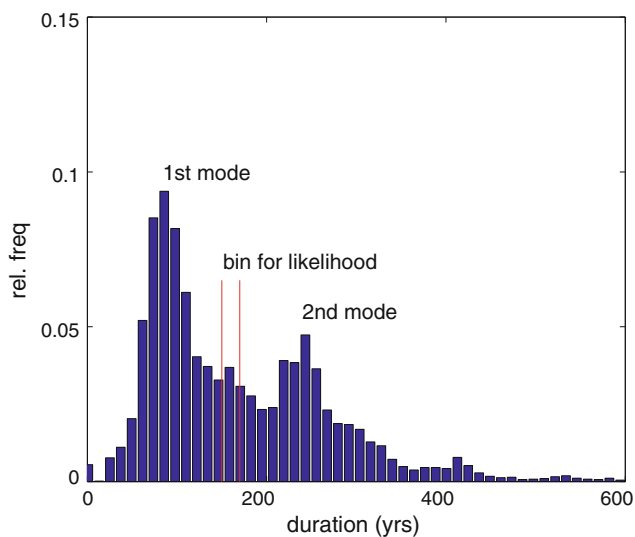


Fig. 5 Example of a histogram of durations T of the cold event for different realizations of noise η for one parameter setting: $a_{\text{hoc}} = 2,000 \text{ m}^2/\text{s}$, $a_v = 0.8 \times 10^{-4} \text{ m}^2/\text{s}$, $\sigma_{\text{noise}} = 0.06 \text{ Sv}$, $D_{\text{baseline}} = 1,000 \text{ years}$

continuing inflow of freshwater hinders the overturning from recovering as it smooths the gradient in density.

The first mode centred around 80 years represents a centennial time scale. The analysis of typical scales of advective processes influencing the upper Atlantic provides the appropriate centennial time scale. The characteristic time scale for the decay of regional distortions only reaches from annual to decadal scales and the diffusive scale of the Atlantic reaches 1,000 years. As the advective processes that seem to be responsible for the duration of the cold event (i.e. they have a centennial time scale) at least have a hemispheric spatial scale this points to an at least hemispherical impact of the 8.2 ka event in CLIMBER-2.3.

Besides the resulting mean of cold event duration in mode one is too short by a factor 2 compared to the Greenland ice core data. As this mode was not to be altered by adjusting the experiment related parameters it follows that the single-pulse scenario in CLIMBER-2.3 is not able to produce sensible high likelihood for a duration of 160 years without additional baseline flux. This leads to different possible conclusions: Under the assumption that nature has realized a highly probable state during the 8.2 ka event this means that either the model setting is in general unable to realistically represent the 8.2 ka event or if not so the one-pulse scenario can only lead to a high likelihood by introducing additional baseline fluxes (as done in this experiment). As an alternative one would have to consider a multi-pulse scenario. Although the forcing used here leads to an interesting exited mode of MOC with a timescale indicating at least hemispherical range of the event the question of the physical processes, structural design and environmental conditions behind this time scale remain open to further studies.

4.3 Uncertainty reduction in ocean parameters

From this nonlinear fits of cold event duration the corresponding likelihood was computed according to Eq. 2 (see Appendix) for all α_i . The resulting data of empirical likelihood are well represented by a 2D Gaussian Least Square fit. Figure 6 shows the (red) area in diffusivity space with likelihood above 1/20 of maximum likelihood (for a Gaussian distribution this corresponds to the 95% quantile of the distribution). The point of maximum likelihood and the error bars in the diffusivities can directly be extracted: The maximum likelihood is found at $\alpha = 2,265, 0.75 \times 10^{-4} \text{ (m}^2/\text{s)}$. The 95% quantiles lc arise as $a_{\text{hoc}} = 1,100\text{--}3,300 \text{ m}^2/\text{s}$ and $a_v = 0.58\text{--}0.88 \times 10^{-4} \text{ m}^2/\text{s}$. The values of a_v at the turning point of the Bryan Lewis profile (at depth of 2,500 m) correspond to an interval of $1\text{--}4 \times 10^{-5} \text{ m}^2/\text{s}$ in the upper ocean layer (against $1\text{--}8 \times 10^{-5} \text{ m}^2/\text{s}$ prior to the experiment). This error bar enables different interpretations: First the error bar lc can be interpreted as ratio

of likelihood without need of any prior distribution, thereby losing a probabilistic measure but gaining objectivity. Second, the Gaussian shape of the likelihood (and thereby also of the posterior distribution) and the assumption of a quasi-uniform prior distribution⁸ allows to interpret the error bar as posterior quantile. The fitted 2D Gaussian function encloses a probability of 95% within the part above the 1/20-level. Taking the Present Day Domain to enclose $\leq 95\%$ probability this represents a reduction of uncertainty in a_{hoc} and a_v of about a factor 2 (on logarithmic scale) or larger against the ranges of the Present Day Domain (assuming locally approximately flat prior).

5 Discussion

Comparing our resulting ranges in the vertical ocean diffusivity parameter with constraints contemplated by Forest et al. (2008) we find that the spread both before (0.5–1 cm²/s) and after (0.5–0.7 cm²/s) including the 8.2 ka information is quite small and lies within the confidence region of their posterior distribution (0.2–2 cm²/s) for global mean parameter for diffusivity of mixing anomalies. This far smaller spread may be explained by the fact that our ensemble created to include pre-industrial equilibrium climate as constraint was produced by only varying the diffusivity parameters whereas Forest et al. (2008) simultaneously vary the diffusivity parameter and both equilibrium and effective climate sensitivity. And as another point of concern the principle comparability between the parameters in the two models may be questioned as they arise from different assumptions and the ocean models differ in processes they resolve. In comparison to the GCMs of the current IPCC assessment report our range of diffusivity parameters lies within the region of extremely small values. But Forest et al. (2008) found that these small values are highly probable given the twentieth century temperature data.

The success of the learning from the 8.2 ka event is limited by different imperfections that lead to an over/under-estimation of the learning effect and the remaining uncertainty: (1) The number of parameters to learn on had to be constrained leading to an underestimation of remaining uncertainty as additional learning parameters would add their own uncertainty. In a first iteration the method was demonstrated by choosing only 2D learning on the ocean diffusivities as key parameters associated with ocean circulation changes, here with abrupt ocean

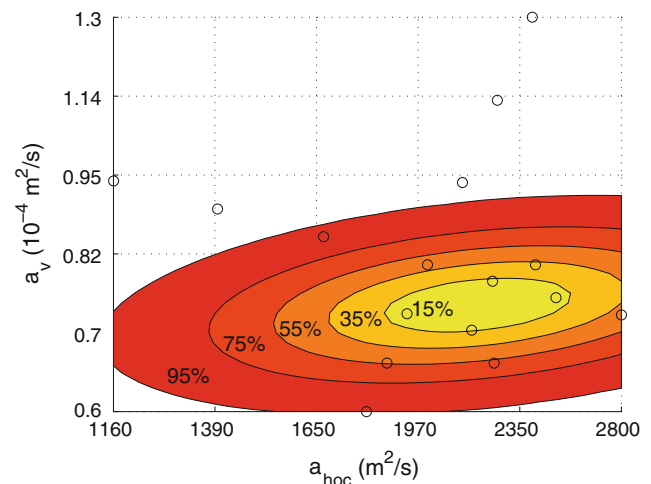


Fig. 6 Quantiles of the Gaussian fit to the empirical likelihood [which is identical in shape to the posterior distribution $P_{\text{post}}(\alpha)$ for flat prior] after including the knowledge stemming from the 8.2 ka event in comparison to the right part of the Present Day domain. The black points represent sampling points for which the likelihood was established by a 300 member ensemble of different noise realizations. The coloured areas represent the quantiles of the posterior distribution. The real value of diffusivities lies within the yellow domain with 15% probability (given the experimental setup and the prior knowledge). Thus the outer red domain represents the 95% quantile of diffusivity values after the 8.2 ka experiment

circulation changes, with all other parameters taken as known constants. (2) The basis of comparison between model and data was chosen as one-dimensional output, namely the duration of the cold event as seen in the Greenland ice core data. This approach potentially overestimates the remaining uncertainty as not all available information about the event is used. (3) The strength of the approach of directly estimating the likelihood from ensemble runs is that no specific functional form has to be assumed a priori; but leading to higher computational effort.

To overcome imperfection (1) a next step would be the inclusion of at least all of the experiment related parameters (duration and strength of additional baseline flux, amplitude of freshwater noise) and other parameters potentially influencing the result, like the depth of the mixed ocean layer or sea ice extend. Here the limitation of the model data comparison can be seen clearly: We expanded the likelihood to a third dimension by including the strength of the baseline flux as an additional parameter at the cost of 20 additional shots in the now three-dimensional parameter space. As a result the learning effect on the horizontal diffusivity vanishes as it strongly depends on the baseline flux. This could have been expected as CLIMBER-2.3 only provides a 2D ocean with an averaged longitudinal dimension. However the learning effect on the vertical diffusivity is only slightly weakened. Also a

⁸ Assuming quasi-uniform prior distribution the posterior probability distribution function $P_{\text{post}}(\alpha)$ can be gained by simply normalizing the likelihood by the number of samples α_i .

change towards a multipulse drainage of Lake-Agassiz is possible. This would surely change the histograms of cold event durations by adding other modes and therefore would also change the resulting posterior pdf. Actually one could suppose that the learning effect from the 8.2 ka event would be diminished by assuming a multipulse freshwater scenario with uncertain timing of the pulses as any duration of cooling is achievable even without adding noise or an additional baseline flux just by the appropriate series of freshwater pulses. Therefore, the route towards a multi pulse scenario is another possible path to be taken in future research and hopefully the uncertainty about the freshwater forcing in general will be reduced by further hydrological and glaciological investigations.

It is obvious that when adding more uncertain parameters the uncertainty in each single parameter rises. The information contained in the data of the 8.2 ka event can only be allocated amongst the parameter under consideration. An extended comparison of time series of data fields would preserve more information of the 8.2 ka event. For instance the inclusion of additional data like reconstructions of monsoon precipitation patterns, sea surface temperature or sea ice extend would increase the basis of information available [thereby addressing (2) and (1) simultaneously].

Hereby the key question is whether the uncertainty in the freshwater forcing required to reproduce the 8.2 ka event can be reduced by new modelling exercises (e.g. ice-sheets, lakes, etc.) or by new data. This would highly increase the potential of reducing important model parameter uncertainty within more complex climate models by making the harvest of data on the cold event more effective.

The uncertainty reduction in ocean diffusivities can be linked to other important parameters like the overall freshwater input from North America into the Atlantic and the distance of the North Atlantic Thermohaline Circulation (THC) to a shut down. The linkage has to be established as functional dependence of the quantity in question on the diffusivity space. A potential linkage between ocean diffusivity, freshwater input and distance to THC shutdown $\Delta\mu$ will open different valuable possibilities for constraints on uncertainties: Following Sect. 3, namely that the overall freshwater forcing (in terms of the 8.2 ka event the known background flux) is constant the linkage could be used to further reduce the uncertainty of the diffusivities. Alternatively one could use the link to transfer the uncertainty reduction effect on the diffusivities to a reduction effect on the $\Delta\mu$ and thus constraining the proximity to a THC breakdown. Those links provide the possibility to freely choose the parameter that is most suitable to be measured and the one to perform a Bayesian analysis about indirectly. Of course such links are only valid if one trusts the

model to rightly represent all processes involved in this causal chain.

Finally, to solve (3), the usage of more complex models, advanced sampling schemes and improved insights in the processes involved in the 8.2 ka event would probably allow to estimate the likelihood less costly (in terms of computational cost). Whatever the state of development, for each implementation of Bayesian Analysis a balance between complexity in model-data comparison and complexity in the uncertainty space must be found; limited by the available computational power and the information content in the data. When interpreting the results one has also to keep in mind that the employed model and data themselves set limits on how close the result can get to reality. The results are a priori only valid inside the stylised world of CLIMBER-2.3.

6 Summary

Employing CLIMBER-2.3 a scheme how to extract information through Bayesian analysis from paleo-data containing the 8.2 ka event was implemented to constrain model parameters representing ocean diffusivities. Ensemble simulations of the 8.2 ka cold event in CLIMBER-2.3 revealed a time scale of cooling that points towards an at least hemispherical spread of the event. The inability of CLIMBER-2.3 to reproduce the right duration of cooling within a one-pulse scenario emphasizes the importance of including additional continental runoff around the 8.2 ka event. Within affordable costs of computation the likelihood of the diffusivity parameters was estimated from ensemble runs of the noisy version of CLIMBER-2.3. The method led to considerable reductions of uncertainty in the vertical ocean diffusivity (factor 2 vs. prior knowledge). Sensitivity tests on forcing parameters revealed weaknesses in the method and hampered the uncertainty reduction effect on horizontal diffusivity.

The limited availability of computational power rises constraints on the dimension of model-data comparison and the dimension of parameter space to be investigated by Bayesian Analysis. Due to this imperfections all results presented here prove valid in the stylised CLIMBER-2.3 world only. The dependence of the results on the specific modelling framework of CLIMBER-2.3 needs to be assessed by a modelling comparison exercise. Besides this structural uncertainty, the main assumption within CLIMBER-2.3, namely that the models performance at the 8.2 ka event is fully determined by ocean diffusivity parameters, needs to be validated by including other uncertain model parameters, especially the atmospheric parameters and their influence on the sea-ice extent. Therefore this study can only be seen as a preliminary and conceptual

investigation of the feasibility and value of the Bayesian assimilation scheme integrating the 8.2 ka event. A more sophisticated treatment of the subject by using more complex models (paleo-GCMs) and data (e.g. SST reconstructions for equatorial Atlantic, reconstructions of precipitation in monsoon regions) will help in better evaluating the potential of the 8.2 ka event for constraining important model parameters.

Acknowledgments We wish to thank Andrey Ganopolski for valuable advice and inspiring discussions. We are grateful to Stefan Rahmstorf and Gerald Haug for helpful and constructive comments on an earlier version of this paper. We thank two anonymous referees for their valuable comments that helped us to refocus the paper and clarify some major issues in an earlier version.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendices

Bayesian analysis

Over the last decades, climate science has employed two main statistical concepts to infer from data on model (or, generally, system) properties: (1) the frequentist and (2) the Bayesian approach (for an overview on the various pros and cons of both approaches, see e.g. Berger 1985). For either approach we would need to establish the likelihood function $L(\alpha|y)$ that delivers the relative plausibility of parameter value α in view of the evidence (observational manifestation) y . Hereby L can be chosen as $L(\alpha|y) = P(y|\alpha)$, i.e. the probability for the observation y under the parameter assumption α .

Then from the frequentist approach we can derive intervals of confidence (or, in a multivariate setting, areas or volumes of confidence) on α , given the observation y , L and a desired level of confidence. The Bayesian approach requires (in addition to L , y) the specification of a prior probability distribution on α , $P(\alpha)$. In general, $P(\alpha)$ represents subjective prior knowledge on α ; hereby “subjective” meant in the sense of “relating to the mind of a (well-informed) subject” Rougier 2007 (rather than indicating distorted, non-scientific information). Prior knowledge P and additional data represented by L are merged to the posterior distribution (representing knowledge after the observation is made) by:

$$P_{\text{post}}(\alpha) = \frac{P_{\text{prior}}(\alpha)P(y|\alpha)}{\int d\alpha' P_{\text{prior}}(\alpha')P(y|\alpha')}. \quad (4)$$

Bayes (1783), Berger (1985), O’Hagan and Forster (2004) (“Bayes’ formula”). Compared to the frequentist approach, the Bayesian approach comes with the following

advantages: (1) the Bayesian posterior carries the maximum information, given the observation (Berger 1985), (2) the output (the posterior) is a probability distribution rather than an interval, hence it (a) can be interpreted as new prior if additional, statistically independent observations become available, allowing for repeated application of Bayes’ formula, (b) can be transformed into other quantities of interest without loss of information, (3) the statistical procedure (i.e. Eq. 4) is unique while in the frequentist case, the statistic becomes (if at all) unique only after additional assumptions are made, (4) if L is non-analytic (in particular if α is related to data by a nonlinear mapping) then it may happen that a well-established frequentist statistic may come of a lot of loss of information or a case-optimised frequentist statistic is by definition not well-established and hence lacks a common interpretation in the community—while Bayes’ formula holds in any case.

As drawback of the Bayesian scheme it is often articulated that it needs subjective knowledge and that subjective knowledge may generically be poorly represented by a prior distribution Walley, 1991. Quite the contrary, the frequentist approach does not require input of a prior (generically subjective) distribution.

The climate community has utilised both approaches. Over the last decades, when attributing global warming to anthropogenic CO₂ emissions, mainly the frequentist approach was followed (for an overview on concept, see e.g. Allen and Tett 1999). Global mean temperature rise was interpreted as a linear superposition of forcings from competing agents, including CO₂. Under the assumption of Gaussian model—data discrepancies, ellipses of confidence on the transfer coefficient vector were derived, thereby representing error bars of linear regression analysis.

Quite the contrary, considering climate model output as a function of uncertain climate model parameters, one generically deals with a nonlinear, rather complex relation, hence several “pro-Bayesian items” of above list hold. Forest et al. (2002) established a “perturbed physics ensemble” (PPE) by sampling multivariate parameter space and comparing time series of several output quantities to observation data. They used a frequentist approach to infer on climate sensitivity, but in addition, obtained more informative results from an approximate Bayesian approach. Thereafter a dozen of further studies (for an overview, see Collins and Knight 2007) Bayesian analysed PPEs of their models. Hereby, in the practical application of Bayes’ formula, for any of the three items of the right hand side of the equation, conceptual and technical challenges emerge.

On the choice of the prior distribution, rather than eliciting an expert’s prior, most authors pragmatically

followed the ideology of “objective Bayesianism” (Rosenkrantz 1977), i.e. that in case of low prior knowledge, consistent results are obtained if a uniform density distribution or at least a distribution with a rather flat maximum is used (“non-informative prior”). However, Bertrand (1889) argues that this approach may lead to paradoxical situations, in our view only to be resolved within a generalised Bayesian concept (Walley 1991). Roughly speaking, weak prior knowledge may be more adequately modelled by classes of priors rather than a single prior. Within climate science, Kriegler and Held (2005), Tomassini et al. (2007) utilised such generalised procedures resulting in larger spans of posterior uncertainty. For the remainder of this article, however, we stick to the main stream of applied Bayesianism in climate science that utilises quasi “non-informative” priors in cases where explicit expert interviews on prior knowledge are lacking. One can interpret posteriors obtained from “low information” priors as low-uncertainty bounds (Held and von Deimling 2006) on results that would have been obtained from the more sophisticated, generalised, yet computationally much more demanding procedures described and demonstrated in (Kriegler and Held 2005; Tomassini et al. 2007; Walley 1991).

Most Bayesian studies in climate science derive the likelihood L from Gaussian weather fluctuations around a slowly increasing temperature mean, estimated from GCM runs, on the order of a thousand model years long (see, e.g. Forest et al. 2002). The Gaussian assumption allows to analytically express L in terms of the difference of model and observation temperature signal (the “first moment”). Hereby the covariance structure (i.e. the second moment) is silently assumed as independent of model parameters, and model parameter values would influence the first moments only. A more consistent approach is followed by Wigley et al. (2005) for the special case of an analytically tractable climate model. In the present study, we have no reasons to assume a Gaussian likelihood apriori, as the 8k event is highly nonlinear a mechanism. Hence we choose to numerically approximate L . As we will outline below, we identify y with the duration T of a cold event, subject to stochastic realisations in CLIMBER-2.3. For any value of α we need to generate a histogram as a proxy for $P(T|\alpha) \propto L(\alpha)$. We would like to stress that making up the likelihood from a nonlinear model involves an additional layer of numerical complexity as well as conceptual self-consistency, not addressed in most Bayesian climate model analyses so far. Finally, Bayes’ formula requires to solve the multivariate integral $\int d\alpha' P_{\text{prior}}(\alpha') P(y|\alpha')$. The standard approach is to interpret the PPE as a sample over the multivariate parameter space such that it can be used as a numerical approximation of that integral (Robert and

Casella 1999), an approach we also follow here. Alternatively, one may explicitly involve an “emulator” as an approximation of the complex climate model, allowing for a more guided selection of test runs of the complex model (Rougier and Sexton 2007).

Applying weather noise to CLIMBER-2.3 output

To ensure the comparability of CLIMBER-2.3 output to Greenland paleo-data, the model output has to be adapted to the data by applying local weather noise⁹, not represented in the CLIMBER-2.3 output. In this study an additive weather noise model ζ is suggested. The CLIMBER-2.3 output is modified in the way: $y' \rightarrow y := y' + \zeta$, where ζ is drawn *iid* from a normal distribution $N(0, \sigma)$. The main assumption of this transformation is that noise-driven CLIMBER-2.3 has all the necessary variability beyond weather, reddened¹⁰ by ocean dynamics. The noise amplitude σ is transferred from the observable weather noise in Greenland which is in terms of $\delta^{18}\text{O}_{\text{ice}}$ derived from the covariance of the different data stations in Greenland to σ in units of temperature T via $\sigma^2 = x^2 \times \sigma_{\text{weather}}^2$ constrained by the following assumption:

$$\frac{\sigma_{\text{weather}}^2}{(\text{var}(\delta^{18}\text{O}_{\text{ice}}) - \sigma_{\text{weather}}^2)} = \frac{x^2 \times \sigma_{\text{weather}}^2}{\text{var}(y')}, \quad (5)$$

with y' the stationary CLIMBER-2.3 output, $\delta^{18}\text{O}_{\text{ice}}$ the stationary (without long term trends) Greenland data, x the transfer factor from $\delta^{18}\text{O}_{\text{ice}}$ to temperature (under the assumption of a simple linear transfer function).

That means the ratio of weather noise to the variance of the timeseries without weather noise should be the same in the model output and the data. The difference in the Greenland data on short time scales (after long term trends are removed) should only originate from local weather noise. So the amplitude of this Greenland weather noise σ_{weather} can be derived as variance from the Kriging (see Wackernagel 1995) of the Greenland ice core data. Hereby it is important not only to transfer σ in unit but also to rescale the amplitude properly according to the different time resolutions Δt in data and model output (Kleinen et al. 2003):

$$\frac{\sigma_{\Delta t}}{\sqrt{\Delta t}} = \frac{\sigma_{\Delta t'}}{\sqrt{\Delta t'}}. \quad (6)$$

Therefore using σ_{weather} and Eq. 6 the transfer parameter x is tuned to fulfill (5). The resulting time series σ (in units

⁹ Hereby denoting the fraction of variability that does not correlate with ocean variability.

¹⁰ While for white noise the amplitude is generated by a normal distribution for each single timestep, red noise contains a memory of amplitude of distortion from earlier time steps.

of temperature change per year) can either be directly applied to the CLIMBER-2.3 output for each timestep, or one takes only the mean of σ as a global noise amplitude. In this study we used the latter approach by applying a global amplitude of weather noise (in units of temperature change per year).

References

- Allen MR, Frame DJ (2007) Call off the quest. *Science* 318:582–583
- Alley R, Mayewski P, Sowers T, Stuiver M, Taylor K, Clark P (1997) Holocene climate instability: a prominent widespread event 8,200 year ago. *Geology* 25:483–486
- Alley RB, Ágústadóttir AM (2005) The 8.2k event: cause and consequences of a major holocene abrupt climate change. *Quat Sci Rev* 24:1123–1149
- Barber DC, Dyke A, Hillaire-Marcel C, Jennings AE, Andrews JT, Kerwin MW, Bilodeau G, McNeely R, Southon J, Morehead MD, Gagnon JM (1999) Forcing of the cold event of 8,200 years ago by catastrophic drainage of laurentide lakes. *Nature* 400:344–348
- Bauer E, Claussen M, Brovkin V, Hühnerbein A (2003) Assessing climate forcings of the earth system for the past millenium. *Geophys Res Lett* 30:1276
- Bauer E, Ganopolski A, Montoya M (2004) Simulation of the cold climate event 8,200 years ago by meltwater outburst from lake agassiz. *Paleoceanography* 19:PA3014. doi:10.1029/2004PA001030
- Bayes T (1783) An essay towards solving a problem in the doctrine of chances. *Philos Trans R Soc* 53:370–418
- Berger AL (1978) Long-term variations of daily insolation and quaternary climate changes. *J Atmos Sci* 35:2362–2367
- Berger JO (1985) *Statistical decision theory and Bayesian analysis*. Springer, New York
- Bertrand J (1889) *Calcul des probabilités*. Gauthier-Villars, Paris
- Clark PU (2001) Freshwater forcing of abrupt climate change during the last glaciation. *Science* 293:283–287
- Collins M, Knight S (2007) Ensembles and probabilities: a new era in the prediction of climate change. *Phil Trans R Soc A* 1857:1955–2191
- Collins M, Booth BBB, Harris GR, Murphy JM, Sexton DMH, Webb MJ (2006) Towards quantifying uncertainty in transient climate change. *Clim Dyn* 27:127–147. doi:10.1007/s00382-006-0121-0
- Crowley TJ (1992) North atlantic deep water cools the southern hemisphere. *Paleoceanography* 7:489–497
- Cuffey KM, Clow GD (1997) Temperature, accumulation, and ice sheet elevation in central greenland through the last deglacial transition. *J Geophys Res* 102:26,383–26,389
- Dansgaard W (1985) Greenland ice core studies. *Paleogeogr Paleoclimaol Paleocol* 50:185–187
- Dansgaard W (1993) Evidence for general instability of past climate from a 250 kyr ice record. *Nature* 364:218–220
- Efron B, Tibshirani R (1993) *An introduction to the bootstrap*. Chapman and Hall, New York
- Forest CE, Stone PH, Sokolov AP, Allen MR, Webster MD (2002) Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science* 295:113–117
- Forest CE, Stone PH, Sokolov AP (2008) Constraining climate model parameters from observed 20th century changes. *Tellus A* 60:911–920
- Ganopolski A, Rahmstorf S (2001) Rapid changes of glacial climate simulated in a coupled climate model. *Nature* 409:153–158
- Ganopolski A, Rahmstorf S, Petoukhov V, Claussen M (1998) Simulation of modern and glacial climate with a coupled model of intermediate complexity. *Nature* 391:351–356
- Ganopolski A, Petoukhov VK, Rahmstorf S, Brovkin V, Claussen M, Eliseev A, Kubatzki C (2001) Climber-2: a climate system model of intermediate complexity. Part ii: sensitivity experiments. *Clim Dyn* 17:735–751
- GRIP-Project-Members (1993) Climate instability during the last interglacial period recorded in the grip ice core. *Nature* 364:203–207
- Hegerl GC, Crowley TJ, Hyde WT, Frame DJ (2006) Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature* 440:1029–1032. doi:10.1038/nature04679
- Held H, von Deimling TS (2006) Transformation of possibility functions in a climate model of intermediate complexity. *Adv Soft Comput* 6:337–345
- Held H, Kleinen T (2004) Detection of climate system bifurcations by degenerate fingerprinting. *Geophys Res Lett* 31:L23,207. doi:10.1029/2004GL020972
- Johnsen SJ, Clausen HB, Dansgaard W, Fuhrer K, Gundestrup N, Hammer CU, Iversen P, Jouzel J, Stauffer B, Steffensen JP (1992) Irregular glacial interstadials recorded in a new greenland ice core. *Nature* 359:311–313
- Johnsen SJ, Dahl-Jensen D, Dansgaard W, Gundestrup N (1995) Greenland paleo temperatures derived from grip borehole temperature and ice core isotope profiles. *Tellus Ser B* 47:624–629
- Johnsen SJ, Dahl-Jensena D, Dansgaard W, Gundestrup N, Steffensen JP, Clausen HB, Miller H, Masson-Delmotte V, Sveinbjrnsdóttir E, White J (2001) Oxygen isotope and paleo temperature records from six greenland icecore stations: Camp century, dye-3, grip, gisp2, renland and northgrip. *J Quat Sci* 16:299–307
- Kleinen T, Held H, Petschel-Held G (2003) The potential role of spectral properties in detecting thresholds in the earth system: application to the thermohaline circulation. *Ocean Dyn* 53:53–63
- Kleiven HKF, Kissel C, Laj C, Ninnemann US, Richter TO, Cortijo E (2008) Reduced north atlantic deep water coeval with the glacial lake agassiz freshwater outburst. *Science* 319:60–64. doi:10.1126/science.1148924
- Knutti R, Stocker TF, Joos F, Plattner GK (2002) Constraints on radiative forcing and future climate change from observations and climate model ensembles. *Nature* 416:719–723
- Kobashi T, Severinghaus JP, Brook EJ, Barnola JM, Grachev AM (2007) Precise timing and characterization of abrupt climate change 8200 years ago from air trapped in polar ice. *Quat Sci Rev* 26:1212–1222. doi:10.1029/2006GL027242
- Kriegler E, Held H (2005) Utilizing random sets for the estimation of future climate change. *Int J Approx Reason* 39:185–209
- Ledwell J, Montgomery E, Polzin K, Laurent LS, Schmitt R, Toole J (2000) Evidence for enhanced mixing over rough topography in the abyssal ocean. *Nature* 403:179–182
- LeGrande AN, Schmidt GA, Shindell DT, Field CV, Miller RL, Koch DM, Faluvegi G, Hoffmann G (2006) Consistent simulations of multiple proxy responses to an abrupt climate change event. *Proc Natl Acad Sci USA* 103:837–842
- Leverington DW, Mann JD, Teller J (2002) Changes in the bathymetry and volume of glacial lake agassiz between 9200 and 7700 ¹⁴c year bp. *Quat Res* 57:244–252
- Liccardi JM, Teller JT, Clark PU (1999) Freshwater routing by the laurentide ice sheet during the last deglaciation. In: Clark UP, Webb RS, Keigwin LD (eds) *Mechanisms of global climate change at millennial time scales*. American Geophysical Union, Washington DC. *Geophysical Monograph* 112:177–201
- Marshall SJ, Clarke GKC (1999) Modelling north american freshwater runoff through the last glacial cycle. *Quat Res* 52:300–315

- Mayewski PA, Meeker LD, Whitlow S, Twickler MS, Morisson MC, Bloomfield P, Bond GC, Alley RB, Gow AJ, Grootes PM, Meese DA, Ram M, Taylor KC, Wumkes W (1994) Changes in atmospheric circulation and ocean ice cover over the north atlantic during the last 41,000 years. *Science* 263:1747–1751
- Muscheler R, Beer J, Vonmoos M (2004) Causes and timing of the 8200 year bp event inferred from the comparison of the grip be-10 and the tree ring delta c-14 record. *Quat Sci Rev* 23:2101–2111
- NGRIP-Project-Members (2004) High-resolution record of northern hemisphere climate extending into the last interglacial period. *Nature* 432:147–151
- O'Hagan A, Forster J (2004) Bayesian inference. Volume 2b of Kendall's advanced theory of statistics, 2nd edn. Edward Arnold, London
- Petoukhov V, Ganopolski A, Brovkin V, Claussen M, Eliseev A, Kubatzki C, Rahmstorf S (2000) Climber-2: a climate system model of intermediate complexity. Part i: model description and performance for present climate. *Clim Dyn* 16:1–17
- Polzin K, Toole J, Ledwell J, Schmitt R (1997) Spatial variability of turbulent mixing in the abyssal ocean. *Science* 276:93–96
- Rahmstorf S (2002) Ocean circulation and climate during the past 12,000 years. *Nature* 419:207–214
- Raper SCB, Gregpny JM, Stouffer RJ (2002) The role of climate sensitivity and ocean heat uptake on AOGCM transient temperature response. *J Clim* 15:124–130
- Rasmussen S, Andersen K, Svensson A, Steffensen J, Vinther B, Clausen H, Siggaard-Andersen ML, Johnsen S, Larsen L, Bigler M, Röthlisberger R, Fischer H, Goto-Azuma K, Hansson M, Ruth U (2006a) A new greenland ice core chronology for the last glacial termination. *J Geophys Res* 111:D06, 102
- Rasmussen SO, Andersen KK, Svensson AM, Steffensen JP, Vinther BM, Clausen HB, Siggaard-Andersen ML, Johnsen SJ, Larsen LB, Bigler M, Röthlisberger R, Fischer H, Goto-Azuma K, Hansson ME, Ruth U (2006b) A new greenland ice core chronology for the last glacial termination. *J Geophys Res* 111:D06,102
- Renssen H, Goose H, Fichetfet T, Campin J (2001) The 8.2 kyr bp event simulated by a global atmosphere–sea–ice–ocean model. *Geophys Res Lett* 28:1567–1570
- Renssen H, Goose H, Fichetfet T (2002) Modeling the effect of freshwater pulses on early holocene climate: the influence of high-frequency climate variability. *Paleoceanography* 17:1020
- Renssen H, Goose H, Muscheler R (2006) Coupled climate model simulation of holocene cooling events: oceanic feedback amplifies solar forcing. *Clim Past* 2:79–90
- Robert CP, Casella G (1999) Monte Carlo statistical methods. Springer, New York
- Roe GH, Baker MB (2007) Why is climate sensitivity so unpredictable? *Science* 318:629–632
- Rohling EJ, Pälike H (2005) Centennial-scale climate cooling with a sudden cold event around 8,200 years ago. *Nature* 434:975–979
- Rosenkrantz RD (1977) Inference, method and decision: towards a Bayesian pPhilosophy of science. Springer, New York
- Rougier J (2007) Probabilistic inference for future climate using an ensemble of climate model evaluations. *Clim Chang* 81:247–264
- Rougier J, Sexton DMH (2007) Inference in ensemble experiments. *Phil Trans R Soc A* 365:2133–2144
- Schmittner A, Weaver AJ (2001) Dependence of multiple climate states on ocean mixing parameters. *Geophys Res Lett* 28:1027–1030
- Schneider von Deimling T, Held H, Ganopolski A, Rahmstorf S (2006) Climate sensitivity estimated from ensemble simulations of glacial climate. *Clim Dyn* 27:149–163 doi:[10.1007/s00382-006-0126-8](https://doi.org/10.1007/s00382-006-0126-8)
- Stouffer RJ, Russell J, Spelman MJ (2006) Importance of oceanic heat uptake in transient climate change. *Geophys Res Lett* 33:L17,704. doi:[10.1029/2006GL027242](https://doi.org/10.1029/2006GL027242)
- Teller JT, Leverington DW, Mann JD (2002) Freshwater outbursts to the oceans from glacial lake agassiz and their role on climate change during the last deglaciation. *Quat Sci Rev* 21:879–887
- Thomas ER, Wolff EW, Mulvaney R, Steffensen JP, Johnsen SJ, Arrowsmith C, White JW, Vaughn B, Popp T (2007) The 8.2k event from greenland ice cores. *Quat Sci Rev* 26:70–81
- Tomassini L, Reichert P, Knutti R, Stocker TF, Borsuk ME (2007) Robust bayesian uncertainty analysis of climate system properties using markov chain monte carlo methods. *J Clim* 20:1239–1245
- Wackernagel H (1995) Multivariate Geostatistics, 1st edn. Springer, Berlin
- Walley P (1991) Statistical reasoning with imprecise probabilities. Chapman and Hall, London
- Walsh JE, Portis DH (1999) Variations of precipitation and evaporation over the north atlantic ocean, 1958–1997. *J Geophys Res* 104:16,613–16,631
- Wiersma AP, Renssen H (2006) Model-data comparison for the 8.2ka bp event: confirmation of a forcing mechanism by catastrophic drainage of laurentide lakes. *Quat Sci Rev* 25:63–88
- Wiersma AP, Renssen H, Goose H, Fichetfet T (2006) Evaluation of different freshwater forcing scenarios for the 8.2 ka bp event in a coupled climate model. *Clim Dyn* 27:831–849. doi:[10.1007/s00382-006-0166-0](https://doi.org/10.1007/s00382-006-0166-0)
- Wigley TML, Ammann CM, Santer BD, Raper SCB (2005) Effect of climate sensitivity on the response to volcanic forcing. *J Geophys Res* 110:D09107