Dynamics of abrupt glacial climate change

Johannes Lohmann

September 30, 2018

University of Copenhagen
Faculty of Science
Niels Bohr Institute
Center for Ice and Climate

PhD Thesis

Dynamics of abrupt glacial climate change

Johannes Lohmann

Supervisor Peter Ditlevsen

This thesis has been submitted to the PhD School of The Faculty of Science,
University of Copenhagen
September 30, 2018

Johannes Lohmann
Dynamics of abrupt glacial climate change
PhD Thesis, September 30, 2018
Reviewers: Aslak Grinsted , Christian Franzke, and Thomas Laepple
Supervisor: Peter Ditlevsen

University of Copenhagen
Center for Ice and Climate
Niels Bohr Institute
Faculty of Science
Juliane Maries Vej 30
2100 Copenhagen

Abstract

The pre-industrial interglacial climate has been characterized by warm and stable temperatures, without major abrupt changes. This has been different during the last glacial period, which was much colder, but has been frequently interrupted by abrupt climate changes, the so-called Dansgaard-Oeschger (DO) events. These events are most pronounced in Greenland ice core records where the temperature proxy suggests jumps in between a cold (stadial) and a warm (interstadial) state with temperature increases of up to 16 K within few decades. Their cause is unknown and most realistic climate models do not produce the observed behavior. Uncovering how these models need to be changed in order to reproduce past climate changes will greatly increase faith in their projections of present day anthropogenic climate change. Nevertheless, a small number of realistic models, and a larger number of conceptual ones, have been proposed that show variability resembling DO events. They represent different dynamical mechanisms, ranging from transitions in between multiple stable climate states, due to stochastic perturbations or changes of a control parameter, to periodic oscillations. Most comparisons of models to proxy data are qualitative, and do not address the complex temporal pattern of DO events.

This thesis presents a characterization of DO events by robust statistical analysis of the ice core record, which can serve as a basis for comparisons of model output and data. Different simple statistical and dynamical models are tested against the data, using statistical hypothesis tests and Bayesian model comparison, with the aim to distinguish different dynamical mechanisms underlying DO events. The temporal variability in the DO warming and cooling event series is too pronounced to originate from two independent random processes that are stationary and memoryless. Instead, there is evidence for external modulation of DO events, with distinct factors influencing stadials and interstadials. However, external modulation, such as changes in solar radiation, is not sufficient to explain the entire variability. Additionally, the stadial and interstadial durations have distinct statistics, indicating different mechanisms that cause the respective transitions. Bayesian model comparison on the basis of summary statistics shows that different stochastic dynamical systems can equally well explain the data, because the dynamics are dominated by high intensity noise in order to reproduce the large temporal variability of DO events. The corresponding dynamical regimes include noise-induced transitions in a bi-stable potential, excitability and relaxation oscillations. This thesis provides new insights into the dynamical mechanisms underlying DO events, and constitutes a statistical basis for future quantitative comparisons of proxy records and realistic climate models that exhibit DO-type variability.

Resumé på Dansk

Den varme før-industrielle mellemistidsperiode var relativt stabil, uden store pludselige klimaforandringer. Det var anderledes under den sidste istid, som udover at være meget koldere også blev afbrudt af pludselige og voldsomme klimaforandringer, de såkaldte Dansgaard-Oeschger (DO) begivenheder. Disse klimabegivenheder ses tydeligst i iskernemålinger fra Grønlands iskappe. Her viser isotopvariationer, som er proxy for temperaturen, spring mellem en kold glacial (stadial) tilstand til en varmere (interstadial) tilstand med en opvarmning på op til 16 grader indenfor få årtier. Årsagen til DO begivenhederne er ukendt, og de fleste realistiske klimamodeller har hidtil ikke været i stand til at reproducere dette variable istidsklima. Hvis vi kan finde ud af hvad der skal til for at få klimamodellerne til at beskrive den observerede klimaopførsel vil i høj grad styrke vores tiltro til at modellerne er i stand til realistisk at forudsige de fremtidige klimaforandringer som følge af den menneskeskabte drivhusopvarmning. Dog er der foreslået få realistiske og den del konceptuelle klimamodeller, der faktisk er i stand til at simulere klimavariationer, der minder om de paleoklimatiske observationer. Disse modeller foreslår forskellige dynamiske mekanismer som årsag til DO begivenhederne: Det kunne være at der er flere stabile klimaer, hvor spring imellem disse er forårsaget af tilfældige forstyrrelser eller at stabiliteten af de enkelte klimatilstande ændrer sig, ved en langsom ændring af en ydre faktor, en såkaldt kontrolparameter. Alternativt, kunne klimaet ændre sig med periodiske oscillationer. De fleste sammenligninger mellem proxy observationerne og modelklimaet er kvalitativ og adresserer ikke den komplekse tidslige struktur i forekomsten af DO begivenheder.

Denne afhandling præsenterer en karakteristik af DO begivenhederne ved hjælp af en robust statistisk analyse af iskernemålingerne. Dette kan bruges som basis for sammenligning af modelklimaer med paleoklimatisk data. Forskellige enkle statistiske og dynamiske modeller bliver testet ved sammenligning med data. Dette gøres ved brug af statistisk hypotese testning og Baysiansk model sammenligning, med det formål at kunne skelne mellem forskellige foreslåede dynamiske mekanismer styrende for DO begivenheder. Den tidslige variation i perioderne mellem de pludselige opvarmninger og afkølinger er af sådan en beskaffenhed at de ikke kan beskrives som to uafhængige stationære tilfældige processer uden hukommelse. I stedet peger analysen på at der er en ekstern modulation af DO begivenheder, som er styret af forskellige faktorer i stadiale og interstadiale perioder. Dog er eksterne modulationer, såsom variationerne i solindstråling, ikke tilstrækkeligt til at forklare hele variationen. Oveni modulationerne observerer vi at varigheden af stadialperioder og interstadialperioder har forskellig statistik, hvilket kunne betyde at afbrydelsen af henholdsvis kolde og varme perioder kunne skyldes forskellige

mekanismer. Baysiansk modelsammenligning på baggrund af statistiske karakteristika (summary statistisk) viser at forskellige foreslåede stokastisk dynamiske modeller lige godt forklarer data. Dette skyldes at dynamikken i modellerne er domineret af den tilfældige støj, hvis de skal være i stand til at reproducere den store tidslige variation af de observerede DO begivenheder. De korresponderende dynamiske parameter regimer i modellerne inkluderer støj-drevne overgange i et bi-stabilt potentiale og forskellige former for oscillationer. Afhandlingen giver ny indsigt i den dynamik som DO begivenhederne skyldes. Det giver en statistisk basis for fremtidige kvantitative sammenligninger mellem proxy målingerne og avancerede klimamodeller, som reproducerer DO-lignende variationer.

Acknowledgements

The first thanks goes to my supervisor Peter Ditlevsen for being a great mentor. You were always ready for a scientific discussion during which I learned a lot, never failed to ask good questions, and faithfully carried out the administrative duties of a supervisor. Next up is my office mate Karl Nyman, who shall be credited for countless discussions on dynamical systems and paleoclimate, which was a town in Russia to both of us at the start of the PhD. These discussions helped form and make clear a lot of the concepts that ended up in this thesis. Another thanks goes out to all of my PhD colleagues from the CRITICS network for the company during our numerous workshops, for important debates on the question of what constitutes a critical transition, and, more urgently, 3am discussions about filtrations. Furthermore, I want to thank Sune Rasmussen for very valuable comments on some of my work, and Anders Svensson for helping me out with the ice core data on many occasions. Another very warm thanks goes to the whole Center for Ice and Climate for its nice atmosphere and inspiring characters, and for the amazing opportunity to experience the field work at the EGRIP camp on the Greenland ice sheet. Finally, the biggest thanks goes to my family and friends for the vital support and encouragement during the years of this thesis and beyond.

List of manuscripts

1. Johannes Lohmann and Peter Ditlevsen, '*Objective extraction and analysis of statistical features of Dansgaard-Oeschger events*', prepared for submission

2. Johannes Lohmann and Peter Ditlevsen, '*Random and externally controlled occurrence of Dansgaard-Oeschger events*', Climate of the Past **14**, 609-617, 2018, https://doi.org/10.5194/cp-14-609-2018

3. Johannes Lohmann and Peter Ditlevsen, '*A consistent model selection analysis of abrupt glacial climate changes*', under review in 'Climate Dynamics'

Contents

1 Introduction **1**

1.1 Evidence of abrupt glacial climate change 2
1.2 Hypotheses to explain abrupt glacial climate change 6
1.3 Dynamics of Dansgaard-Oeschger events 7
1.4 Scientific questions and methodology 10

2 Objective extraction and analysis of statistical features of Dansgaard-Oeschger events **13**

2.1 Introduction 13
2.2 Methods and Materials 15
2.2.1 Data 15
2.2.2 Fitting routine 16
2.2.3 DO event features and data analysis 19
2.3 Results 21
2.3.1 Piecewise-linear fit 21
2.3.2 DO event features 23
2.3.3 Uncertainty of fit parameters and features 24
2.3.4 Statistical analysis of DO event features 28
2.3.4.1 Interstadial periods 28
2.3.4.2 Stadial periods 34
2.3.4.3 Abrupt warming periods 36
2.4 Discussion and Conclusion 39
2.A Iterative algorithm to fit piecewise-linear model 43
2.B Convergence of iterative fitting routine 44
2.C Uncertainty estimation of fitting parameters 45

3 Random and externally controlled occurrence of Dansgaard-Oeschger events **49**

3.1 Introduction 49
3.2 Methods and Models 52
3.3 Results 56
3.4 Discussion and Conclusions 59

4 A consistent model selection analysis of abrupt glacial climate changes **65**

4.1 Introduction . . . 65
4.2 Models . . . 67
4.3 Materials and methods . . . 70
4.3.1 Data . . . 70
4.3.2 Summary statistics . . . 70
4.3.3 Inference and model comparison . . . 73
4.4 Results . . . 75
4.4.1 Synthetic data study . . . 75
4.4.1.1 Parameter inference . . . 76
4.4.1.2 Model comparison . . . 79
4.4.2 NGRIP data study . . . 80
4.4.2.1 Parameter inference . . . 80
4.4.2.2 Model comparison . . . 83
4.4.2.3 Time reversal asymmetry . . . 84
4.5 Discussion and Conclusion . . . 86
4.A ABC-PMC algorithm . . . 89

5 Conclusion **91**

6 Outlook **97**

A Supplemental material to 'Objective extraction and analysis of statistical features of Dansgaard-Oeschger events' **99**

B Supplemental material to 'A consistent model selection analysis of abrupt glacial climate changes' **103**

Bibliography **107**

1 Introduction

The current geological period, referred to as the Quaternary, began roughly 2.6 Million years ago and is characterized by a cyclic growth and disappearance of large continental ice sheets, known as glacial and interglacial periods. These cycles imply huge global-scale changes in climate, vegetation and fauna, and are well-documented in all climate archives that span this period, such as ocean sediment core proxies for global land ice extent [RL05]. It is generally believed that this climate variability is due to changes in the distribution of incoming solar radiation (*insolation*) caused by cyclic variations of the earth's orbit, known as Milankovich cycles [Mil30]. This insolation forcing has to be amplified by feedback mechanisms in order to explain the large magnitude of the climate cycles. The details of this phenomenon, such as the change in periodicity from approximately 40 kyr (kyr = one thousand years) to 100 kyr that occurred after the so-called middle Pleistocene transition about 1 Million years ago, are still a matter of scientific debate. Nevertheless, the insolation forcing as the driver is widely accepted. The situation is different for the next most pronounced mode of climate variability other than glacial cycles, which happens on time scales just below the variations in insolation. These so-called Dansgaard-Oeschger (DO) events are abrupt, large-scale climate changes that have been recorded in the last glacial period (approx. 12 - 120 kyr BP, BP = before present), and generally lack a consensus regarding their cause. First discovered in Greenland ice cores, they have been shown to have a large amplitude, which is about half of the mean glacial-interglacial temperature difference in Greenland. They are, furthermore, very abrupt, unfolding within a few decades. Since their discovery in ice cores, they have been indentified in various other terrestrial and marine climate archives all over the Northern Hemisphere. They are the prime demonstration that large-scale climate change can happen at decadal time scales, and thus have high relevance to the understanding of current anthropogenic climate change.

As of now, no comprehensive picture to explain these climate changes has emerged from analyzing different available proxy observations of glacial climate parameters. Likewise, detailed climate models under glacial boundary conditions typically lack the necessary abruptness to realistically reproduce DO-type variability. Identifying which modifications need to be made in the climate models in order for them to reproduce the observed abrupt climate changes of the past would greatly increase our confidence in their future projections of anthropogenic climate change. About 30 abrupt DO climate changes have been identified in the last glacial period, and they constitute a very complex temporal pattern. Apart from questions regarding the

physical mechanism underlying DO events, i.e., an identification of the components of the climate system that are involved, there are fundamental questions regarding the nature of the events. Are they part of a self-sustained oscillation in the climate system? Or are they a manifestation of switches in between two alternative stable states of the climate system, driven by large fluctuations in fast components of the climate? Answers to questions like these can subsequently be used to distinguish between hypothesized physical mechanisms. This approach is pursued in this thesis.

In the remainder of this Chapter, we will give an introduction to the phenomenon of abrupt glacial climate change, starting with evidence from paleoclimatic archives in Sec. 1.1. Thereafter, in Sec. 1.2 we review leading physical hypotheses to explain DO events, followed by a review of previous research on the dynamics of DO events in Sec. 1.3. The methodology and scientific questions that form the basis of this thesis are introduced in Sec. 1.4.

1.1 Evidence of abrupt glacial climate change

The first evidence that the glacial climate might have been more unstable than the current interglacial (the Holocene) came from the ice core record at the Camp Century military base on the Greenland ice sheet, which showed persistent millennial-scale fluctuations of water stable isotopes [Joh+72]. Water stable isotopes in precipitation have been known to be correlated with atmospheric temperature. This is due to isotopic fractionation, i.e., a depletion of the heavy isotopes ^{18}O and deuterium ^{2}H relative to the lighter isotopes ^{16}O and hydrogen H. This temperature dependent process occurs at the different phase transitions of water in an air mass as it moves from the site of evaporation to the accumulation site, where its isotopic composition is measured [Dan64]. It was conjectured that this relationship could be used to infer past temperatures from old ice in glaciers [Dan+73]. As a consequence, the isotope variations measured in the Camp Century ice core were interpreted as swings in between very cold (*stadial*) and mild (*interstadial*) conditions over Greenland. The signal from Camp Century was found to be robust and not just local, as the same isotope fluctuations were discovered in the second deep Greenland ice core DYE-3, which was drilled about 1400 km away from Camp Century [Dan+82]. The positions of the two cores on the ice sheet were not optimal with respect to the flow of the ice, but were a compromise for logistical and other reasons. Subsequently, these two records were confirmed by the GRIP ice core [Dan+93] and the close-by GISP2 ice core [Gro+93], drilled at the optimal location of the ice sheet, namely

the summit where ice flow is slow and vertical. In both GRIP and GISP2 ice cores, the stratigraphic sequence in the bottom of the core was disturbed, thus corrupting the record from the oldest parts of the last glacial period. The newer NGRIP ice core presents an undisturbed sequence, which spans about 5 kyr into the previous interglacial period (the Eemian) at roughly 123 kyr BP [NGR04]. Because it allows for an analysis of the entire last glacial period, most of the analysis in this thesis is based on the oxygen isotope record of this ice core.

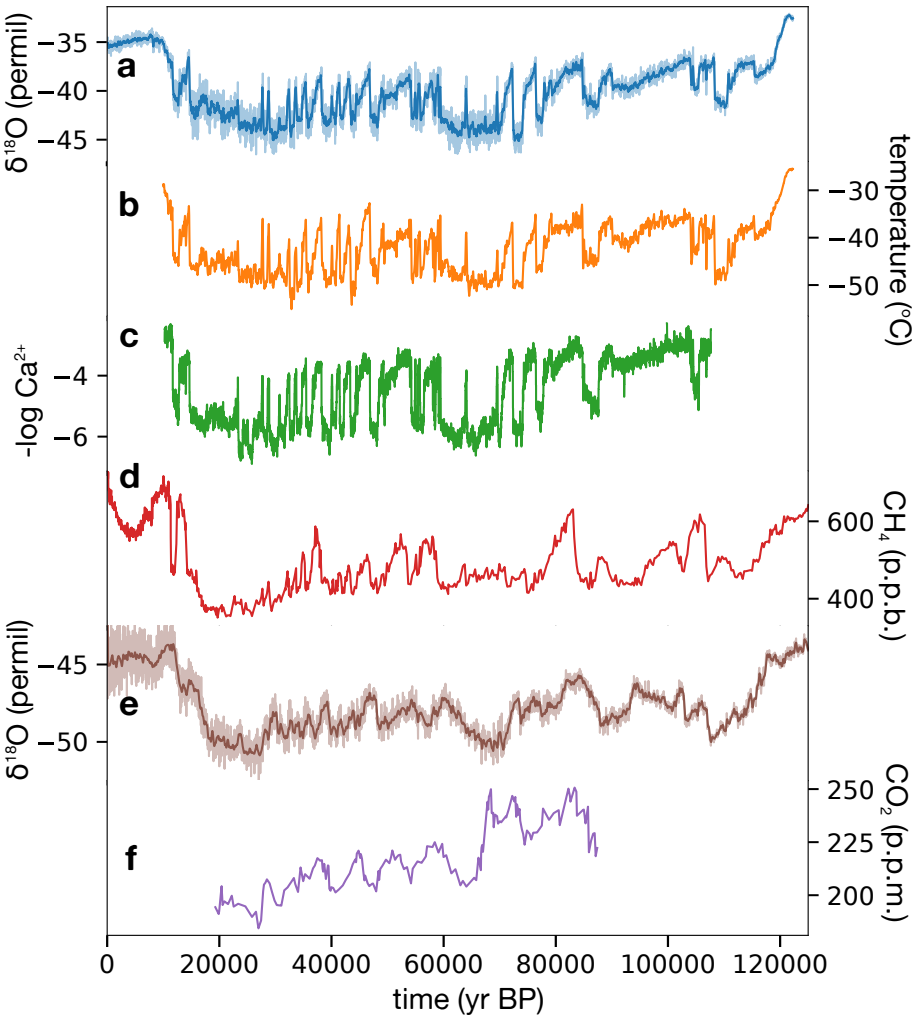

Fig. 1.1.: Several proxy records and paleoclimatic reconstructions of the last glacial period. **a**: NGRIP δ^{18}O record [NGR04]. **b**: Greenland temperature reconstruction by Kindler *et al.* [Kin+14]. **c**: Logarithm of the calcium concentration in the NGRIP ice core [Ras+14]. **d**: Methane record from the Antarctic EPICA Dome C ice core [Lou+08]. **e**: Antarctic δ^{18}O record from the EDML ice core [EPI10]. **f**: CO_2 record from the Antarctic Byrd ice core [AB08].

Most other parameters measured in Greenland ice cores were also found to vary dramatically during the last glacial period. Already in the DYE-3 ice core it was observed that CO_2, trapped in small air bubbles in the ice, co-varies with the water stable isotopes [Sta+84]. However, it was later found that the measured record was not reliable and the actual global atmospheric CO_2 rather co-varies with the Antarctic

isotope record and not with Greenland. More reliable are records of deposited terrestrial dust and other chemical impurities that vary strongly and in synchrony with the water stable isotopes, implying significant atmospheric circulation pattern changes [May+97]. A number of important paleoclimatic reconstructions are shown in Fig. 1.1, for a time interval spanning parts of the previous interglacial period (119-130 kyr BP), the last glacial period (12-119 kyr BP) and the current interglacial period (0-12 kyr BP).

The evidence for abrupt climate changes found in Greenland ice cores has subsequently been supported by similar findings in other paleoclimatic archives. Reconstructions of North Atlantic sea surface temperatures from marine sediment cores have shown strong resemblance with the variations recorded in Greenland [Bon+93]. Furthermore, Northern Hemisphere terrestrial archives established close analogues to the abrupt events over Greenland, and thus revealed their large spatial extent [Wan+01; Gen+03]. With an independent temperature reconstruction based on nitrogen isotopes in air bubbles, the qualitative validity of the paleothermometer based on ice core water stable isotopes has been confirmed [Kin+14], as shown in Fig. 1.1b. From this reconstruction it has been found that the warming transitions from stadial to interstadial conditions corresponded to a Greenland surface warming of +5 K to +16.5 K. The transitions are furthermore very rapid, with temperatures rising to their maximum within a few decades [Wol+10]. Changes in some climate parameters indicating shifts in atmospheric circulation have even been shown to occur within 1-3 years [Ste+08]. This combination of evidence forms a convincing picture of large-scale and abrupt climate changes in the Northern Hemisphere, which are now known as Dansgaard-Oeschger (DO) events. Because of their magnitude and abruptness, these climate changes had even more widespread imprints, beyond the climatological signal. Dramatic changes in vegetation have been observed from paleosoil records in Loess sequences [Rou+07; Rou+17]. Even evidence for an influence of DO events on the extinction of megafauna populations is beginning to be obtained [Coo+15].

Even though the climatic fluctuations of the Southern Hemisphere during the last glacial period, as inferred from Antarctic ice cores, are more subdued and less abrupt, the DO events are in fact of global scale. All events recorded in Greenland have analogues in Antarctic isotope records, referred to as Antarctic Isotope Maxima (AIM) [Blu+98; Jou+07; Cap+10]. These events are closely tied to DO events and a study based on recent efforts in synchronizing Greenland and Antarctic records suggests that DO events on average start 200 years earlier than the corresponding Antarctic events [WAI15]. For a comparison of the structure of the DO and AIM events, see Fig. 1.2. The millennial-scale variability also manifests itself in reconstructed

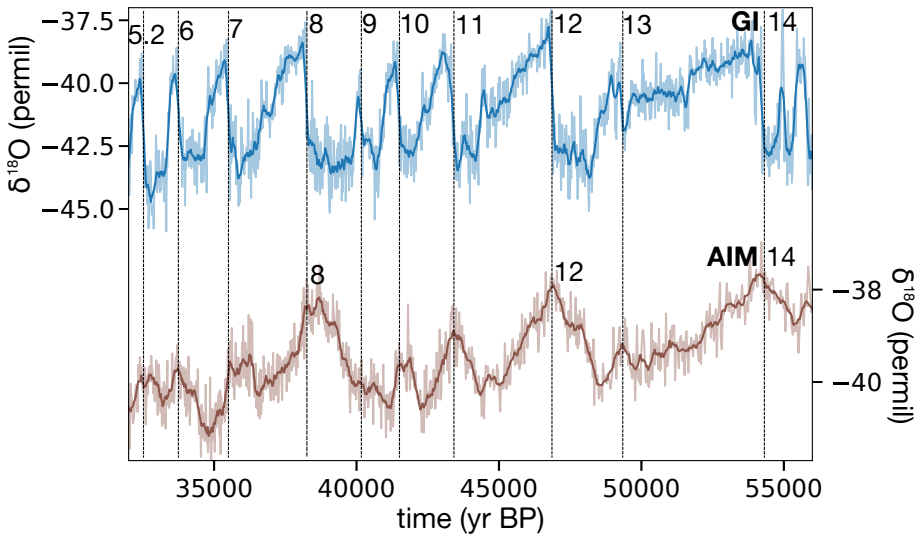

Fig. 1.2.: Comparison of the Greenland NGRIP δ^{18}O (*top*) and Antarctic WAIS δ^{18}O (*bottom*) [WAI15] records for a section of the last glacial period. The canonical numbering of the Greenland interstadials (GI) is given, as well as the numbering of some corresponding Antarctic Isotope Maxima (AIM) events.

Greenhouse gas concentrations, with CO_2 resembling the Antarctic [AB08], and NH_4 the Greenland isotopic record [Flü+04].

Lastly, even though the Greenland ice core records only cover the most recent glacial period, different evidence suggests that there have been DO events in all previous glacial periods, at least in the last 800 kyr since the glacial-interglacial cyclicity elongated to 100 kyr. First, AIM events have been identified in earlier glacial periods [Jou+07], and their close correspondence to DO events thus suggests DO variability in the Northern Hemisphere. In fact, the relation of timing and shape in between AIM and DO events ca be used to create a synthetic Greenland isotope record for earlier glacial periods [Bar+11]. A similar suggestion comes from the Antarctic methane record, which closely resembles the Greenland water stable isotopes for the last glacial, and shows consistent variability in the previous 8 glacial periods [Lou+08]. Other independent records, which have corresponding events in the last glacial, show similar dynamics in older glacial periods, such as a sea surface temperature record from the Iberian margin [Mar+07] or speleothem proxy records of the Asian monsoon [Che+09; Che+16].

1.2 Hypotheses to explain abrupt glacial climate change

In many leading hypotheses to explain DO events, the Atlantic Meriodional Overturning Circulation (AMOC) plays a major role. It is an ocean circulation largely driven by buoyancy and meridional gradients in heat and salt, and is responsible for a large part of the heat transport from lower to higher (polar) latitudes. Thus a change in the circulation can account for large temperature changes at Northern latitudes. It was suggested that the AMOC can switch between an 'on' and an 'off' state [Sto61], where the switching is controlled by freshwater input into the North Atlantic, e.g. from ice sheet melt, which dilutes the heavy salt water, thus turning off the sinking of heavy water (deep water formation) [Bro+85]. Indeed, globally coupled climate models verified the existence of two stable equilibria of the AMOC [MS88]. It has been subsequently reported for various models that a sudden input of freshwater into the North Atlantic can fully or partially shut down the AMOC, leading to an abrupt cooling of higher Northern latitudes. However, multi-model studies show that the amount of freshwater needed in order to significantly impact the AMOC, or, in other words, the stability of the corresponding climate state, is very model dependent [Kag+13]. Furthermore, the sudden input of freshwater explains an abrupt cooling, but not an abrupt warming, since the AMOC resurgence after removal of freshwater input is typically more gradual. These types of studies also do not explain the actual trigger of an event, i.e., the freshwater input dynamics. Still, AMOC variability remains an appealing candidate to cause DO events, not least because it can explain the bi-polar relationship expressed by Greenland and Antarctic temperature reconstructions.

There have been studies introducing models, which do not rely on a freshwater forcing to obtain transitions in the AMOC, but which possess other control parameters that cause feedback mechanisms to result in a change of AMOC circulation regime. For instance, a varying height of Northern Hemisphere ice sheets can trigger changes in atmospheric circulation and subsequently sea ice, which ultimately leads to large transitions in AMOC strength [Zha+14]. Here, there is a hysteresis of AMOC strength with respect to the ice sheet height, which can explain both abrupt warming and cooling transitions. A similar hysteresis with respect to atmospheric CO_2 as control parameter has been found, where abrupt climate transitions can be triggered by changes in CO_2, which are of similar magnitude compared to the observed fluctuations in the last glacial [Zha+17].

A different dynamical mechanism has been identified in a model of intermediate complexity [GR02]. Here, only the weak AMOC state is stable and the strong mode

is marginally unstable. Excursions from the weak to the strong AMOC state are induced by stochastic freshwater forcing. This phenomenon is often referred to as excitability. Furthermore, the possibility of self-sustained oscillations of the AMOC has been postulated, driven by alternating phases of gradual salinity build-up and freshening through meltwater in the Atlantic Ocean [Bro+90]. Salt-driven oscillatory behavior of the AMOC, which however does not require meltwater input, has been subsequently found in a realistic climate model [PV14; VP16]. In this model, changes in North Atlantic sea ice cover are crucial to obtain a climate response similar to DO events. Indeed, in many modeling studies, a change in the AMOC circulation regime is not enough to explain the amplitude and abruptness observed in the Greenland ice core signal. It has thus been suggested, that rapid retreat or expansion of sea ice can act as an amplifier of climate variability, via the positive ice-albedo feedback and because of its insulating effect, which decouples the atmosphere from the heat reservoir of the ocean [Li+05; Li+10; Dok+13]. If the fast time scale provided by sea ice changes is combined with a slower time scale, such as regrowth of an ice shelf after collapse [Pet+13], one might obtain fast-slow dynamics resembling abrupt warming and gradual cooling observed for the Greenland interstadials.

Apart from hypotheses that involve self-sustained oscillations, slow changes in a control parameter or a direct external trigger, there have been a few instances of completely unforced abrupt climate transitions in climate models. Such seemingly random transitions can be the result of extremes in the chaotic atmospheric dynamics, which then trigger large-scale climate regime shifts through different feedback processes [Dri+13; Kle+15].

This overview of proposed hypotheses is not exhaustive, but should illustrate the wide range of different dynamics that arise in the respective scenarios. Note that we did not discuss the connection of DO events with massive iceberg discharges occurring in the same period during the last glacial, called Heinrich events. The connection in between these two types of events is not fully understood, but recent studies propose promising mechanisms to explain the interaction of the two [Bas+17].

1.3 Dynamics of Dansgaard-Oeschger events

In the previous section, it was seen that there exist numerous hypotheses to explain glacial climate variability, supported by experiments with detailed climate models and qualitative analysis of different paleoclimatic archives. From these manifold hypotheses a breadth of qualitatively different dynamics of DO events arises, rang-

ing from periodic oscillations to presumably unpredictable occurrences of events triggered by the fast and chaotic atmospheric dynamics. Here, two approaches to gain insight into the dynamics of DO events are highlighted, which both lie at the heart of this thesis. First, a quantitative look on the statistical properties of the proxy records, such as the Greenland water stable isotopes, allows one to test hypotheses regarding the dynamics of DO events. Second, because it is mathematically and computationally very challenging to understand and identify the dynamics of a realistic climate model with millions of degrees of freedom, low-order dynamical systems models of important climate components can be a very useful tool to understand glacial climate variability.

Early spectral analysis of the GISP2 ice core revealed a spectral peak corresponding to a periodicity of approximately 1500 years [GS97; Sch+99]. Subsequently, by considering the times elapsed in between rapid warming events, i.e., so-called waiting times, it has been suggested that the warming events are paced by a fundamental period of 1470 years and are spaced at multiples of this period [Sch02a]. This was hypothesized to be a result of a phenomenon called stochastic resonance, given a periodic input [All+01]. Consequently, it has been attempted to explain the origin of the fundamental periodicity by a combination of two solar cycles [Bra+05]. However, by statistical hypothesis testing of a stochastic resonance model and the isotope record, it has been found that only a very small periodic component is consistent with the data [Dit+05]. Furthermore, the significance of the spectral peak at 1470 years has been challenged [Bra+10], and the pacing of events by a fundamental period has been shown to be not robust and instead the events are consistent with a memoryless, random process [Dit+07].

There have been additional insights into glacial climate dynamics by observing dynamical behaviors that arise when interacting components of the climate system are idealized as dynamical systems. To appreciate this, note that it was in fact a paper by Stommel [Sto61], which in 1961 first formulated the idea of multiple equilibria in a density-driven circulation, such as the AMOC, by analyzing a simple dynamical systems model. Stommel's model is a so-called box model, where the climate system is highly discretized and simplified by, for instance, describing the Atlantic ocean with two variables (average temperature and salinity) both at equatorial and polar latitudes, and a fixed atmospheric boundary condition. From Stommel's model, different dynamical scenarios may be deduced, such as noise-induced jumps in between two asymptotically stable equilibria, when including a stochastic freshwater forcing [Ces94]. This concept can explain a number of qualitative features of the ice core record [Dit99]. On the other hand, a different extension of Stommel's model allows for the occurrence of relaxation oscillations [RS16b]. The appeal of these relaxation oscillations to describe DO events lies in their fast-slow characteristics,

i.e., a fast time scale to mimic rapid warmings and a slow, relaxation time scale corresponding to the gradual interstadial coolings. Similarly, other box models of the Atlantic oceanic circulation display relaxation oscillations [SP99; Sch02b; Sah15]. Finally, also excitable dynamics have been found in ocean box models [Ces96] and simplified coupled climate models [Tim+03].

With different conceptual models that claim to display essential features of DO cycles, as well as more realistic climate models covering roughly the same set of dynamical behavior, it is necessary to find frameworks that can objectively compare these models against the data and assess their quality. Here, one could either pick specific conceptual models that have been proposed, or more generally consider abstractions of similar conceptual models and replace them with generic stochastic dynamical systems that represent a specific dynamical paradigm. This has been the motivation for recent efforts to compare different dynamical paradigms to Greenland ice core data [Kwa13; MC17], and we will follow up on this approach in this thesis.

Another line of research uses ideas from dynamical systems theory to infer dynamical mechanisms that might underlie certain abrupt changes of behavior in a complex system, such as the climate. These abrupt changes, also called critical transitions, occur when the system crosses a so-called tipping point [Sch+12]. Different generic scenarios in the way such a tipping point is crossed have been identified [Ash+12], namely tipping via slow passage of a bifurcation (B-tipping), tipping by noise-induced escape from an asymptotically stable state (N-tipping), and tipping by failure of tracking a moving, quasi-stable equilibrium when changing a control parameter faster than a critical rate (R-tipping). An important distinction between the different tipping scenarios is the occurrence of early-warning signals due to critical slowing down before the tipping point is reached, which happens for B-tipping but not for N-tipping. The potential occurrence of early-warning signals before R-tipping is currently being explored [RS16a]. Both N-tipping and B-tipping have been considered as potential scenarios for the abrupt warmings associated with DO events. The possibility of an R-tipping mechanism to underlie abrupt climate change is part of our current research effort (see Outlook). Based on these ideas, early-warning imprints have been searched for in Greenland ice core data. While first it has been shown that there is no critical slowing down before DO events in the NGRIP record [DJ10], it was later reported that weak early-warning signals exist when considering the ensemble average of DO events [Cim+13]. However, it is not clear whether this result is statistically robust. More recently, evidence for critical slowing down in the NGRIP record before DO events has been reported, when only considering the higher frequencies in the record [Ryp16], and it has been hypothesized that this is an imprint of a passage through a bifurcation in a fast subsystem of the climate, such as sea ice [Boe18].

1.4 Scientific questions and methodology

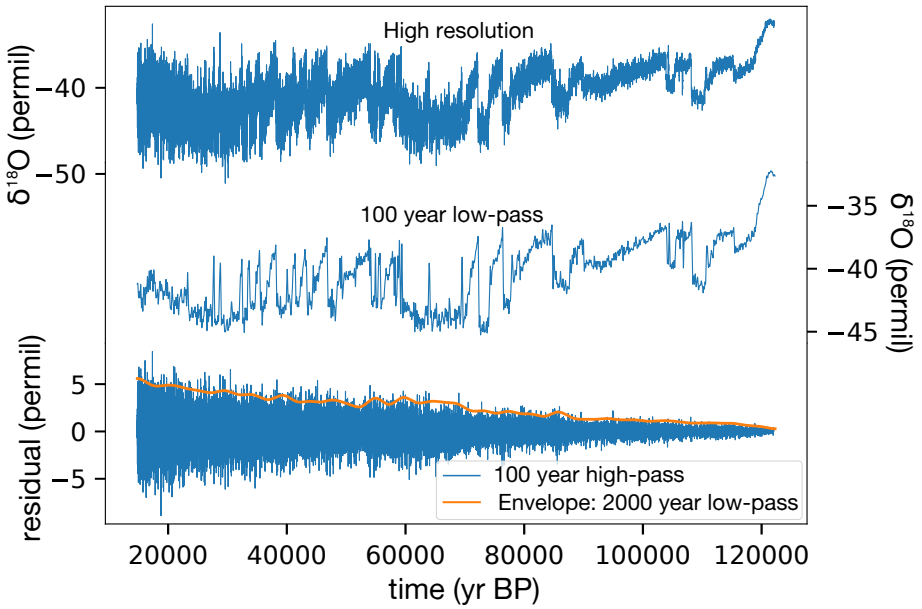

Fig. 1.3.: The high-resolution NGRIP δ^{18}O record of the last glacial (*top*), corresponding to isotope measurements at 5cm intervals on the GICC05 time scale [NGR04; Gki+14; Sve+06]. For comparison, low-pass filtered (*middle*) and high-pass filtered (*bottom*) versions of the record are shown, together with the envelope of the high-pass filtered time series. Note the increasing amplitude of the high-frequency fluctuations towards the end of the glacial.

The aim of this thesis is to further the insight into mechanisms that underlie DO events by extracting new information out of the proxy record, and by processing new and existing information in novel ways. We believe that a comprehensive statistical description of the features of DO events, as inferred from the ice core record, is still missing. Investigations into the governing mechanisms should be guided by the data, and there is more quantitative information available from the ice core records, as has been used previously. Interpretation of proxy records, such as the high-resolution NGRIP δ^{18}O record, should be done with care. As seen in Fig. 1.3, the amplitude of high-frequency variations increases over time in the glacial period. However, it is not obvious to which degree the high frequencies carry a true climatic signal, since there are other processes affecting the record at these frequencies, including variations in the accumulation rate, and post-depositional effects such as wind-driven mixing of deposited snow, diffusion of the isotopes in firn and ice, and a thinning of the layers due to ice flow. Still, there are robust features of DO events that can be extracted from the record. A complete analysis of these features should serve to benchmark DO-type variability that is currently beginning to be obtained in different realistic climate models. In climate model experiments, results are often compared only qualitatively with the data. This is true both for studies with detailed climate models, but also conceptual ones, with the exception of a handful of recent studies, which have a quantitative approach similar to our work [Kwa13; Kru+15; MC17; Boe+17].

Additionally, often only specific parts of the data are being picked that suit best to support the hypotheses, such as the rather regular DO cycles in the period 32-48 kyr BP. The approach taken in this thesis is to formulate clear hypotheses based on the ice core data and test them with simple statistical or dynamical models, which can be readily interpreted in terms of dynamical mechanisms. With this approach, we try to answer the following scientific questions:

1. Can the high temporal variability of the properties of DO events, such as residence times in warm and cold climate states, be expected due to chance by randomly occurring events?

2. Are the DO cycles, and more generally the millennial-scale glacial climate variability, modulated by external forcing, such as insolation?

3. Can we infer from the data whether the DO cycles are a result of noise-induced jumps in between two meta-stable states, or of noisy, self-sustained oscillations?

4. Is there evidence in the statistical properties of the data in favor of certain physical mechanisms to explain DO events?

It is clear that not all questions one might have concerning the cause of DO events can be answered from the Greenland isotope records alone. Other archives will need to be consulted to make statements about the more detailed nature of DO event dynamics, and glacial climate dynamics in general. Still, there is more to be learned from single isotope records, and for some purposes they can stand for themselves, such as answering questions concerning the timing and durations of DO events. We focus on identifying dynamical mechanisms that generate abrupt climate change as observed in Northern Hemisphere paleoclimate proxies. As seen from Antarctic ice cores, the more subdued millennial-scale variability in the Southern Hemisphere is closely related to DO events. A connection of the signals observed in both Hemispheres can be reasonably approximated by simple conceptual models, such as the thermal bipolar seesaw model [SJ03] or a simple integration of the Northern Hemisphere signal to obtain the Southern Hemisphere signal [Bar+11]. Since the Antarctic records obtained from ice cores at different locations have slightly different characteristics across AIM events, it is an open question which model is most appropriate to describe the signals. The details of the bipolar teleconnection mechanisms are still a matter of scientific debate and are beyond the scope of this thesis.

Apart from the above-stated scientific questions, our research interest goes beyond climate science. How can models be compared in a consistent way when focusing strictly on key statistical properties of the data? What is the most sensible way to compare models to data, when the data is one short realization of a stochastic process? What is the influence of different noise structure on the qualitative dynamical behavior in stochastic dynamical systems, and how can such structure be inferred from data? More broadly, we are interested in methods to understand, classify and detect critical transitions.

Our methodology consists of an initial, broad statistical analysis of the data, followed by the identification of key statistical features and a search of dynamical mechanisms that could be in line with these features. These mechanisms are then tested by classical statistical hypothesis tests and Bayesian model comparison. The scientific content of this thesis consists of three separate manuscripts. The first manuscript in Chapter 2, referred to as Paper 1 hereafter, is a statistical analysis of the NGRIP δ^{18}O record with the objective to extract prominent features of DO events, such as the gradual interstadial cooling rates, in order to analyze the distributions of these features, and establish to which degree they correlate with each other and are influenced by external forcing. The second manuscript in Chapter 3 (Paper 2) discusses the time-varying structure of residence times in stadials and interstadials, and tests whether this structure could arise by chance from random processes governing warming and cooling events. Furthermore, a random process model with time-varying parameters is fitted to the data, driven by a combination of external forcings. The third manuscript in Chapter 4 (Paper 3) is a Bayesian model comparison study, which uses summary statistics to compare different stochastic dynamical systems to the NGRIP record. This is achieved with a method called Approximate Bayesian Computation. The manuscripts have been included un-altered from their published or submitted version, except for reformatting to match layout of the thesis. The results of the three manuscripts are synthesized in the conclusion in Chapter 5, and an outlook to current and future work related to this thesis is given in Chapter 6.

2 Objective extraction and analysis of statistical features of Dansgaard-Oeschger events

Abstract

The strongest mode of centennial to millennial climate variability of the last glacial period are the Dansgaard-Oeschger events. They are best recorded in proxies from Greenland ice cores, such as the NGRIP $\delta^{18}O$ record. These proxy records are of very high resolution, but contain a lot of noise, the cause of which is partly due to glaciological effects unrelated to climate. Furthermore, the properties of this noise change throughout the record. This hampers the applicability and interpretability of classical time series analysis techniques. To overcome this, we remove the high-frequency noise and extract the most robust features in the data by fitting a consistent piecewise-linear model to the whole time series. This enables us to investigate new quantitative features of DO events in a statistical manner. Apart from stationary distributions, we search for causalities hidden in between different features and for modulations of them in time via external climate factors.

2.1 Introduction

The physical mechanism(s) and cause of the Dansgaard-Oeschger (DO) events are unknown and debated. Modeling and simulations of the events are guided by the proxy records, among which the stable water isotope records from Greenland ice cores are most prominent. The records are noisy, and since we are not guided by theories about how they should evolve, there is no obvious filter to extract the climate signal from the record. A common characteristics for the DO events seems to be an abrupt temperature increase from the cold stadial conditions to a maximum temperature in the warm interstadial state followed by a gradual cooling until there

is another abrupt jump back into the stadial state. This is referred to as the saw-tooth shape of the events.

Due to the high noise level in the record it is however difficult to discern this specific structure in all of the events. Some events do not seem to follow the generic shape. Furthermore, there are very short events so that it is difficult to speak of a gradual cooling episode. Even other events are interrupted by shorter cooling episodes, referred to as sub-events [Cap+10]. As interpretation of noisy time series are often biased, subjective and one is prone to recognize patterns that can arise by chance, we seek a quantitative evaluation of the record. Assuming the saw-tooth shape of the events, we develop an algorithm for fitting the saw-tooth shape to the entire NGRIP δ^{18}O record of the last glacial, similar to ramp-fitting a jump in a record.

Firstly, our method will give an objective basis of the validity of the generic saw-tooth description of the DO-events and identify which individual DO-events fall outside this description. Secondly, with the simple piecewise-linear fit, we obtain estimates for the stadial and interstadial levels, the abruptness of the transitions and the gradual cooling rate in the interstadial periods. Furthermore, by bootstrapping, we obtain both an estimate of the uncertainty in extracting these parameters from the noisy background and the distribution of parameter values across the DO-events of the last glacial period. This could potentially be used for identifying or excluding proposed mechanisms and for bench-marking model results.

Previous efforts to extract robust DO event features from the record were mostly conducted on only part of the record and were focused on single or very few features. In [Sch02b], linear fits to the interstadials were used to infer the cooling rates starting with Greenland interstadial (GI) 14. Estimates for abruptness of warming transitions and durations of interstadials have been derived in [Rou+17], starting at GI-17.1. A comprehensive survey of onset times of all interstadial and stadial periods is given in [Ras+14]. Our work is different in that we derive all features that can be extracted from a saw-tooth shaped fit to all events at once, by using a fit that is consistent and continuous throughout the record. We thus do not have to rely on any subjective choice of stadial and interstadial onsets or levels. We do, however, not attempt to define the DO events themselves from the record, but instead use the fixed set of all previously classified events [Ras+14].

A naive approach to obtain a piecewise-linear fit of the whole record could proceed in the following way. Considering the stadial episode as constant levels, first cut the time series at a predefined beginning and end of two consecutive stadials episode. Then fit a saw-tooth shape to the event in the middle. The end of the fit to this event then determines the start of the next stadial episode, to be used to fit the

following event. However, the point at which one initially cut the stadial periods determines the levels of the two stadials that have been used to determine the fit to the event. For a consistent fit, the start and end points of a stadial are determined by the fits to two neighboring events. In this way, the fit to each event depends on both its neighboring events before and after, and we cannot simply fit the events sequentially. A solution would be to fit the whole time series at once to a piecewise linear model with approximately 180 parameters, corresponding to 6 times the number of DO events. However, due to high noise and abundance of sub-event features we believe that such a fit will be difficult to achieve without invoking very complicated constraints. Instead, we propose an iterative fitting routine that converges to a consistent fit of the whole time series. We start with a guess for the stadial onset and end times, which determine the constant stadial levels, and fit a saw-tooth shape to each event. Thereafter, we update the stadial onset and end times according to the fit and repeat fitting. When after some iterations the onset and end times do not change significantly anymore the fit has converged and is consistent.

The paper is structured in the following way. In Sec. 2.2 we introduce the data used in the study and its pre-processing, the iterative fitting algorithm, the features we extract from the saw-tooth shape fit to the events and the statistical tools to analyze these features. In Sec. 2.3 we report the results of the fit and the subsequent data analysis. First, in Sec. 2.3.1, we demonstrate the convergence of our fit and discuss the appropriateness of the saw-tooth fit to the events. Then, in Sec. 2.3.2, we give an overview of the features derived from the fit. Section 2.3.3 discusses the uncertainty in estimating the fit parameters and the derived features. In Sec. 2.3.4 we analyze in detail the features characterizing the stadial, interstadial and abrupt warming periods. The results of the fit and the implications of the subsequent data analysis are discussed in Sec. 2.4.

2.2 Methods and Materials

2.2.1 Data

The basis of our study is the $\delta^{18}O$ record of the last glacial period, which lasted from about 120 kyr BP (kyr BP = one thousand years before present) to 12 kyr BP. $\delta^{18}O$ in the NGRIP ice core has been measured in 5 cm samples along the core [NGR04; Gki+14; Ras+14]. These raw depth measurements are then transferred to the GICC05 time scale [Sve+06]. This results in an unevenly spaced time series with a resolution of 3 years at the end to 10 or more years at the beginning of

the last glacial period. Because it greatly simplifies our analysis, we transfer this to an evenly-spaced time series by oversampling it to 1 year resolution using nearest-neighbor interpolation. This means we do not alter the actually measured values and thus adding or removing any variability. For comparison, the high-resolution $\delta^{18}O$ record from the GRIP ice core on the GICC05 time scale has been used [Joh+97; Ras+14], and processed in the same way.

Our study does not aim to present an algorithm that detects or classifies warming events from the time series. Instead, we fit a waveform to previously classified events from Greenland, which have been reported in [Ras+14] together with their time stamps. We do not treat sub-events, which are small dips to colder conditions during a warm period, as separate events, but instead fit them as part of the interstadial periods. The published time stamps will be used to initialize our iterative fitting procedure, and are subsequently refined during the process.

Finally, we use several other data sets that are not derived from Greenland ice core, and investigate to what extent they might correlate with or explain the DO event features. These are loosely referred to as external forcings, although not all are truly external to the climate system, but rather obtained from independent data sources. As proxy for global ice volume, we use the LR04 ocean sediment record stack [RL05]. To represent Antarctic temperatures, we choose the $\delta^{18}O$ measurements from the EDML ice core on the AICC12 time scale [EPI10]. This data was processed by an interpolation to an equidistant 20 year grid and subsequently a smoothing by convolution with a 600 year Hamming window to reduce noise. Furthermore, we consider incoming solar radiation curves at high northern latitudes. Firstly, we use incoming solar radiation at 65 degree North integrated over the summer (referred to as 65Nint hereafter), which we defined as the annual sum of the insolation on days exceeding an average of 350 W/m^2 [Huy06]. Secondly, we use incoming solar radiation at 65 degree North at summer solstice (referred to as 65Nsolst hereafter) [Las+04]. In addition, we also consider the raw orbital parameters of obliquity, eccentricity, and precession index [Las+04].

2.2.2 Fitting routine

Starting from an initial guess, based on the timings reported in [Ras+14], we define beginnings and ends of the stadial periods and keep them fixed throughout an iteration of our algorithm. The time series is then divided in segments at these times. For each event i, we take a segment of the time series consisting of a stadial and interstadial period plus the following stadial period. These segments are then fitted individually to the a piecewise-linear model, as shown in Fig. 2.1 and explained in

the following. The fit starts with a constant line at the beginning of a stadial period. The constant is fixed to the mean level of the stadial period l_i^s, where the duration of this period is determined in the previous iteration. A first break point (parameter b_1) of this constant line is determined, which is followed by a linear up-slope (parameter s_1). The slope ends at the second break point (parameter b_2). After this break point there is a linear down-slope (parameter s_2), which ends at a break point (parameter b_3). After this break point there is a steeper down-slope until a last break point (parameter b_4), which is at a fixed level of the next stadial l_{i+1}^s, where this level is determined from the previous iteration. Note that also the rapid down-slope of the interstadial to the stadial state as well as the interstadial maximum are determined from the six parameters $b_{1,2,3,4}$ and $s_{1,2}$.

After all transitions have been fit, the parameters b_4 and b_1 update the beginnings and ends of the stadial periods. The updated stadial periods yield a new segmentation of the time series and new stadial levels, which are then used as the constant segments in the next iteration of the fitting routine. The idea of this approach is that if the problem is well behaved, the beginnings and ends of the stadial periods do not change significantly anymore after a certain number of iterations, meaning that we found a consistent fit of the time series as a whole. An algorithm for this routine is given in Appendix 2.A.

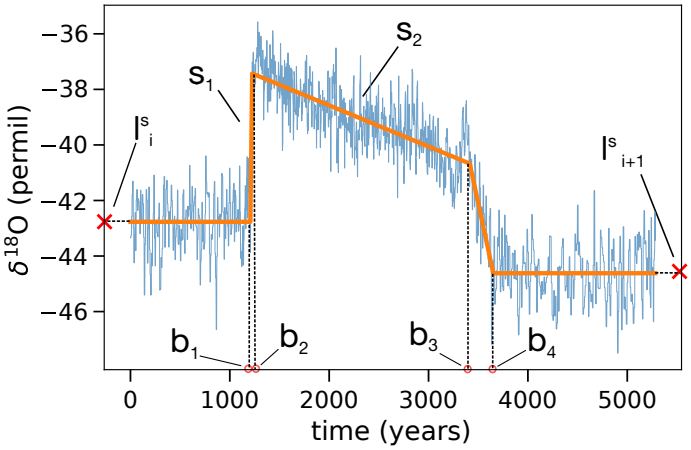

Fig. 2.1.: Generic piecewise-linear model fit to DO event 20, where the time series consists of GS-21.1, GI-20 and GS-20. The parameters of the piecewise-linear model are the four break-points $b_{1,2,3,4}$, the up-slope s_1 and the down-slope s_2. The constant levels l_i^s and l_{i+1}^s of GS-21.1 and GS-20 are constant during an iteration of the fitting routine, and are updated when after each iteration all breakpoints have been determined.

To determine the 6 parameters at each transition, we minimize the root mean squared deviation of the fit from the time series segment. Due to the high noise level, there are many local minima in this optimization problem. Thus, either a brute-force parameter search on a grid or an advanced algorithm is needed to find a global minimum.We chose an algorithm called basin-hopping, which is described

in [Ols+12] and is included in the Scientific Python package scipy.optimize, where it can also be customized. The basic idea of the algorithm is the following: Given a initial coordinates in terms of the parameter vector θ_0, one searches for a local minimum of the goal function $f(\theta)$, e.g., with a Newton, quasi-Newton or other method. The argument to this local minimum θ_n is then randomly perturbed by a Kernel to yield new coordinates $\theta*$, which are the starting point of a new local minimization. Next, there is a Metropolis accept or reject step: We accept the argument of the local minimization θ_{n+1} as new coordinates if the local minimum is deeper than the previous one $f(\theta_{n+1}) < f(\theta_n)$, or else with probability $e^{-(f(\theta_{n+1})-f(\theta_n))/T}$, where T is a parameter relating to the typical difference in depth of adjacent local minima. Now we go back to the perturbation step either with old coordinates θ_n or, if accepted, with new coordinates θ_{n+1}, and repeat. The iterative procedure is repeated for a large number of iterations and the result is the argument to the lowest function value found.

Within basin-hopping, one has the freedom of choosing any local minimizer as well as perturbation Kernel. These have to be adapted to our optimization problem. We have several constraints on the parameters that need to be satisfied by the optimization. For instance, we demand that all segments of the fit are present and do not overlap ($b_1 < b_2 < b_3 < b_4$). Other constraints ensure that the characteristic shape of DO events is fit as good as possible for all events. Among other things, we thus demand the gradual slope to be significantly longer and less steep than the fast cooling transition at the end of an interstadial. An overview of all the constraints we used is given in Appendix 2.A. To satisfy them, we chose a multivariate Gaussian perturbation Kernel, which is truncated at the respective parameter constraints. The local minimizer choice requires further consideration. Our goal function landscape is very rough and not differentiable. Thus, methods like gradient descent give very poor results in our case. A method that does not depend on derivatives and can handle constraints is called Constrained Optimization by Linear Approximation (COBYLA), and we found it to work well in our case.

Two hyper-parameters have to be specified in the basin-hopping algorithm: The variance of the perturbation Kernel, and the parameter T used in the Metropolis criterion. These should both be comparable to typical differences in goal function (temperature) and arguments (perturbation width) of neighboring local minima in the minimization problem. We chose these parameters empirically by observing how the goal function changes as we slightly change the fit. Although this varies significantly from transition to transition, we determined single values as a compromise for all transitions. For the Kernel variance in the directions of $b_{1,2,3,4}$ we chose a value of 15, and for s_1 and s_2 we chose 0.004 and 0.0015, respectively.

The fitting procedure outlined above yields one best fit that we hope to be close to the absolute global minimum of the optimization problem and furthermore as consistent as possible, meaning that the stadial sections that were used for the fit in the last iteration are identical to the stadial sections defined by the resulting fit. Additionally to this best fit we would like to estimate the uncertainty in each of the parameters that arise due to noise in the record. We cannot estimate this from the output of our fitting procedure in a straightforward way. Instead, we use bootstrapping to repeatedly generate synthetic data for each transition and optimize the parameters with basin-hopping. Like this, we yield a distribution on each parameter. Due to computational demands, we do not combine this with our iterative procedure, but rather resample and fit every transition independently. Thus, we neglect the co-variance structure of the errors in the parameters of neighboring transitions. However, we still consider it to be a very good estimate of the uncertainty due to the noise in the record. The detailed procedure is given in Appendix 2.C.

2.2.3 DO event features and data analysis

For each DO event, we yield best fit parameters and from these a variety of features follow immediately. These features are the core of our analysis and will be explained in the following. For each warming period, gradual interstadial cooling period, as well as rapid cooling period at the end of an interstadial, we consider the duration, rate of change and the amplitude. Furthermore, several absolute levels are of interest, including the constant stadial levels, the interstadial levels after the abrupt warming and the interstadial level before the rapid cooling. As a level relative to each event, we consider the level before the rapid cooling above the previous stadial level, which is given by the rapid warming amplitude minus the gradual cooling amplitude. Finally, the gradual cooling amplitude divided by the rapid warming amplitude measures the position of the point of rapid cooling within the event amplitude. In total, we consider 15 features, which are listed in Tab. 2.1.

We use several tools to search for relations in between different features, as well as in between features and external climate factors. Because of the large number of features, we first automatically pre-select potentially relevant relationships and thereafter manually investigate, whether the results are robust to outliers, among other things. First of all we consider Pearson and Spearman correlation coefficients of pairs of features and external climate factors. We pre-select all combinations with p-values smaller than 0.1, assuming independence of the samples. In reality, the p-values are often higher due to autocorrelation. We subsequently investigate this individually for the pre-selected combinations, together with other potential artifacts. In order to find relations of more than two variables, we additionally

Tab. 2.1.: List of DO event features obtained from the fit that are analyzed in this study.

Feature	Definition
Warming duration	$b_2 - b_1$
Warming rate	s_1
Warming amplitude	$s_1(b_2 - b_1)$
Gradual cooling dur.	$b_3 - b_2$
Gradual cooling rate	$-s_2$
Gradual cooling ampl.	$s_2(b_2 - b_3)$
Fast cooling dur.	$b_4 - b_3$
Fast cooling rate	$(b_4 - b_3)^{-1} \cdot [s_1(b_2 - b_1) + s_2(b_3 - b_2) - (l^s_{i+1} - l^s_i)]$
Fast cooling ampl.	$s_1(b_2 - b_1) + s_2(b_3 - b_2) - (l^s_{i+1} - l^s_i)$
Stadial duration	b_1
Stadial level	l^s
Interstadial level	$s_2(b_2 - b_3) + l^s$
Interstadial end level	$s_1(b_2 - b_1) + s_2(b_3 - b_2) + l^s$
Relative Int. end level	$s_1(b_2 - b_1) + s_2(b_3 - b_2)$
Cooling/warming ampl.	$s_2(b_2 - b_3) \cdot [s_1(b_2 - b_1)]^{-1}$

search for multiple linear regression models to explain selected features of the data. For this, we typically use logarithmic quantities because with our features it is otherwise unlikely to find a linear relationship that is not due to outliers. For a given feature as response variable, we fit linear regression models of combinations of two other features or forcings and pre-select models with the largest coefficients of determination, in order to further analyze the fit. We furthermore, in order to find subsets of events that have different properties than other subsets, we perform a clustering analysis on the data. We try two different algorithms, namely K-means and Agglomerative Hierarchical Clustering. As desired number of clusters we choose 2 or 3, since a higher number does not give good results given our sample size of 31. We find clusters for all combinations of 2 and 3 features and forcings and select potentially relevant clusterings by the so-called mean Silhouette coefficient, which is a distance-based measure for the validity of clusters. We perform an analysis with the abovementioned tools on the entire set of features and forcings. From the results obtained, we only report selected findings, which are most robust and relevant to us, in Sec. 2.3.4 of this paper.

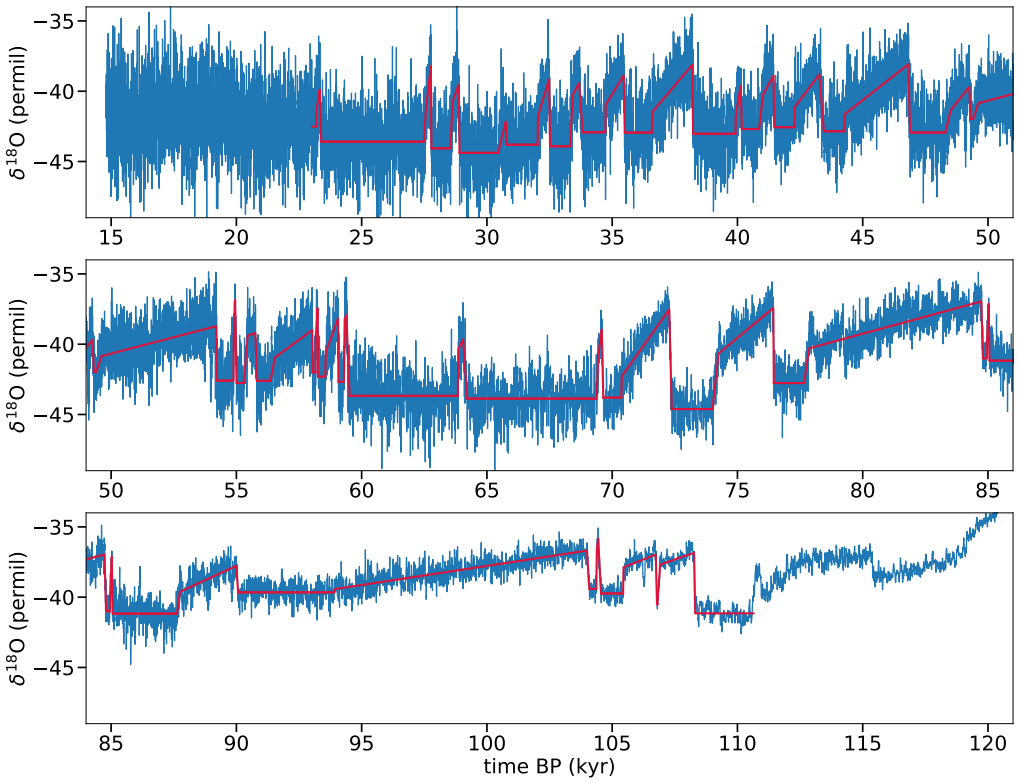

Fig. 2.2.: High-resolution NGRIP δ^{18}O time series and the piecewise-linear fit obtained by our method.

2.3 Results

2.3.1 Piecewise-linear fit

We run the iterative fitting routine for 40 iterations, so that the initial fluctuations in the parameters have died out and converged to a consistent fit, as detailed in Appendix 2.B. In Fig. 2.2 we superimpose the resulting fit onto the high-resolution NGRIP time series. We fit 31 DO events in total, starting with DO 24.2 and ending at DO 2.2, excluding the two outermost events of the last glacial, because they have non-stationary properties in their stadial parts. Table 2.2 shows all parameters obtained from the fit. Instead of $b_{1,2,3,4}$ for each transition, we show the corresponding times of stadial end, interstadial onset, interstadial end and stadial onset.

In our fit, all transitions follow the characteristic saw-tooth shape. For some events, this is because of the constraints we use in the fitting algorithm. Typically, the constraints do not strictly bound the best fit parameters, but they force the fit into another local minimum that is consistent with the saw-tooth shape, which often yields parameters that are still clearly within the constraints. There are, however,

Tab. 2.2.: Parameters resulting from the fitting routine on the NGRIP data.

Event	Stadial End (yr BP)	Interstadial Onset (yr BP)	Interstadial End (yr BP)	Stadial Onset (yr BP)	Warming Rate (permil/yr)	Cooling Rate (permil/yr)
24.2	108313	108270	106914	106810	0.0992	-0.00062
24.1	106790	106743	105452	105439	0.0744	-0.00069
23.2	104556	104441	104387	104366	0.0340	-0.01555
23.1	104090	103996	93916	93898	0.0290	-0.00027
22	90069	89999	87743	87631	0.0270	-0.00082
21.2	85060	85027	84964	84952	0.1230	-0.03992
21.1	84799	84737	77866	77659	0.0655	-0.00049
20	76452	76434	74245	74009	0.2935	-0.00148
19.2	72377	72280	70385	70365	0.0730	-0.00251
19.1	69646	69587	69443	69381	0.0831	-0.01262
18	64212	64051	63858	63846	0.0262	-0.00367
17.2	59520	59390	59323	59294	0.0446	-0.00503
17.1	59076	59061	58571	58549	0.2951	-0.00491
16.2	58266	58245	58168	58162	0.2340	-0.03107
16.1	58051	58023	56536	56364	0.1059	-0.00131
15.2	55821	55759	55449	55296	0.0554	-0.00062
15.1	55011	54950	54892	54887	0.0981	-0.05104
14	54228	54193	49617	49410	0.1092	-0.00046
13	49315	49253	48517	48301	0.0367	-0.00223
12	46890	46826	44286	44277	0.0761	-0.00140
11	43450	43271	42285	42278	0.0225	-0.00236
10	41479	41439	41024	40864	0.0910	-0.00326
9	40175	40131	39933	39928	0.0699	-0.01096
8	38231	38199	36602	36583	0.1549	-0.00210
7	35508	35461	34741	34735	0.0859	-0.00289
6	33822	33681	33434	33314	0.0250	-0.00334
5.2	32521	32485	32039	32028	0.1324	-0.00583
5.1	30794	30752	30514	30473	0.0394	-0.00695
4	28908	28871	28635	28544	0.1302	-0.00485
3	27786	27765	27572	27492	0.2762	-0.01529
2.2	23389	23328	23196	23191	0.0607	-0.01098

four events where the best fit parameters actually lie very close to the bounds set by the constraints. This happens for events 5.1 and 3, which both have ratios of rapid to gradual cooling rates very close to the constraint value of 2.0. Similarly, for the events 15.2 and 6 the ratio of gradual to rapid cooling duration is close to 2.0. Detailed pictures of each transition and the corresponding fit are shown in Fig. S2 in the supplementary material.

The fact that some constraints are needed in order to ensure that the fit of each event follows a saw-tooth shape can be used to classify which events fall outside of this description. To this end, we perform another run of the iterative fitting routine without using constraints 3, 4, 6 and 7 listed in Appendix 2.A. From the resulting fit we then analyze, which of the events are not consistent with the saw-tooth shape.

For this, we use 4 criteria: 1. The abrupt cooling rate is at least twice as large as the gradual cooling rate. 2. The gradual cooling lasts at least twice as long as the abrupt cooling. 3. There is gradual cooling after the rapid warming, i.e., the gradual cooling rate is negative. 4. The abrupt cooling amplitude is larger than 0.5 permil. Criterion 1 is not met by events 23.1, 19.2, 15.1, 11, 5.1, 3 and 2.2, criterion 2 by events 21.2, 19.2, 17.2, 15.2, 15.1, 11, 10, 9, 8, 6, 5.1, 3 and 2.2, criterion 3 by event 11, and criterion 4 by events 23.1 and 15.1. By demanding that all of these criteria are met, we thus conclude that the following 14 out of 31 events fall outside of the saw-tooth description: 23.1, 21.2, 19.2, 17.2, 15.2, 15.1, 11, 10, 9, 8, 6, 5.1, 3 and 2.2.

2.3.2 DO event features

In Fig. 2.3 we show histograms of all the DO event features derived from the fit parameters that we consider in this study, as defined in Sec. 2.2.3. From the histograms we can see that the features have different types of distributions. We discuss which distributions are most likely for selected features in Sec. 2.3.4. Whether the features of single events should be considered as independent samples

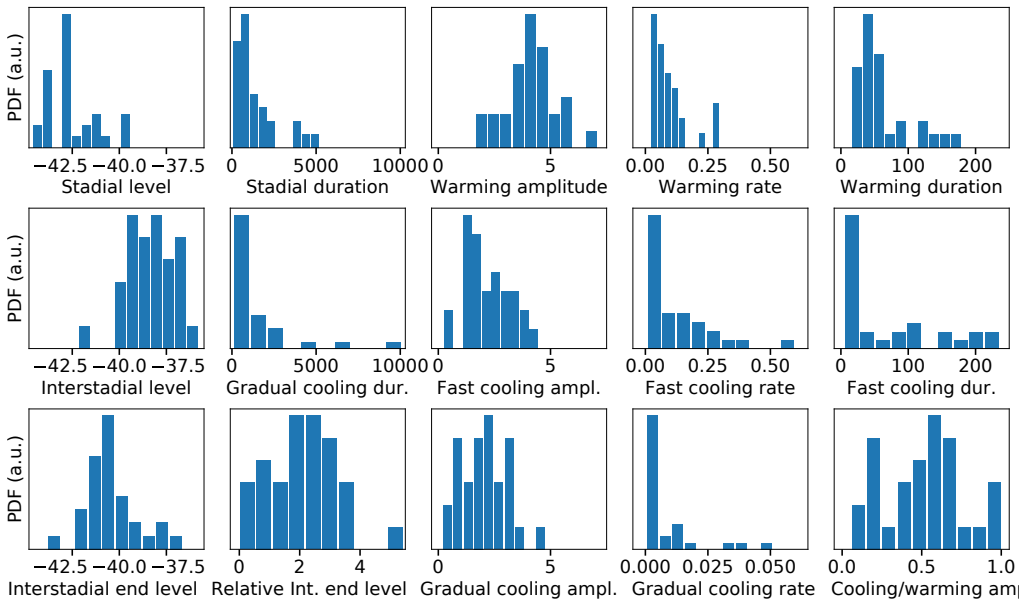

Fig. 2.3.: Histograms of our sample of 31 events for all features considered in this study, as defined in Tab. 2.1.

from a distribution depends on whether they have a significant trend over time. If we consider the event-wise sequence of one feature as an evenly spaced time series we can calculate the autocorrelation until a certain lag and determine by a permutation test, whether the value at a given lag is significantly larger than what could be expected in an uncorrelated sample for a given confidence. By

considering an autocorrelation up to lag 5, we find that the three different levels (stadial, interstadial and level before rapid cooling) show significant autocorrelation at 95% until a lag of 3. We also find significant autocorrelation for four other features at only one specific lag value each, which we consider as false positives. In fact, when independently testing the hypothesis of significant autocorrelation at 95% confidence for 15 different time series (features) at 5 lags, there is an expected value of 3.75 false positives. The corresponding data is shown in Fig. S3 in the supplementary material. As a result, in most cases we can consider the features as independent samples and, e.g., interpret correlation coefficients accordingly. In critical cases we still check whether results are robust by detrending the data.

2.3.3 Uncertainty of fit parameters and features

From the best fit, we estimate the uncertainty of each parameter via bootstrapping, as explained in Appendix 2.C. As an example, we show histograms of the parameters for DO event 20 in Fig. 2.4. In this case the distributions are quite symmetric, but this is not always the case.

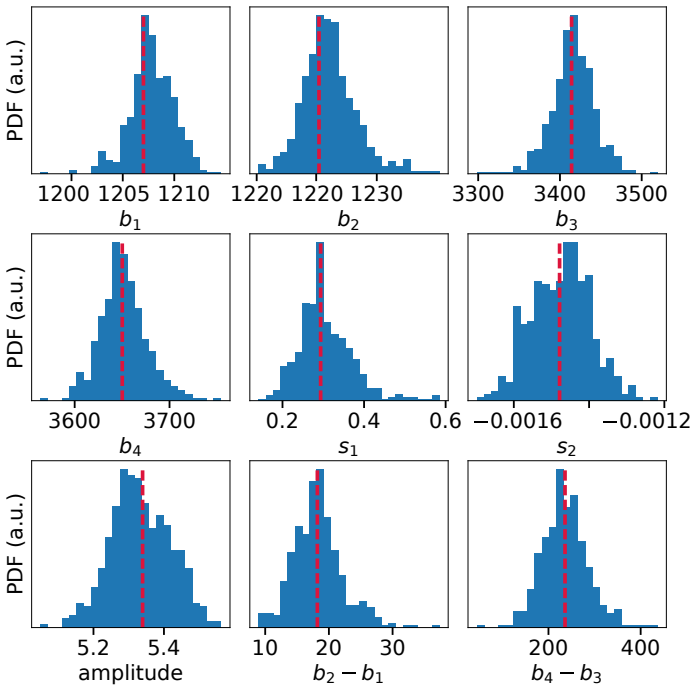

Fig. 2.4.: Histogram of model parameters and some derived quantities for the DO event 20 after 500 iterations of the bootstrap resampling procedure. The parameter values for the best fit, as reported in Sec. 2.3.1, are indicated with *red dashed* lines. The amplitude feature is given by $s_1(b_2 - b_1)$.

In Tab. 2.3 we show the durations and amplitudes of the rapid warmings for each event along with a bootstrap confidence interval consisting of the 16- and 84-percentiles, which would correspond to the $\pm\sigma$ range if the distributions were

Gaussian. The actual distributions are often skewed, so that the best fit values lie close to the edges of the confidence intervals, or even outside of the intervals. In these cases, the $\pm\sigma$ confidence intervals are not the best indicator for the uncertainty, because they barely include the mode of the very skewed distributions. The magnitude of the uncertainties vary from event to event. In the case of the warming durations, the average bootstrap standard deviation is 20.0 years, with a minimum of 3.4 years for GI-16.2 and a maximum of 57.4 years for GI-18. We observe that shorter warmings typically also have smaller uncertainties. As comparison, the durations of the rapid coolings at the end of an interstadial have a larger uncertainty with an average bootstrap standard deviation of 53.6 years. This is expected, because the rapid cooling is typically less well pronounced in the record compared to the rapid warming. The coolings also have a larger spread in the bootstrap standard deviations with a minimum of 4.6 years for GI-16.2 and a maximum of 209.9 years for GI-23.1, because some events have a very clearly defined rapid cooling, while others do not. Similarly, the onset times of the rapid warmings have an average bootstrap standard deviation of 11.4 years, whereas the onset of the stadial periods have a corresponding average uncertainty of 31.7 years.

As complementary approach to assess the uncertainties of the features, we compare them to those derived in the same way from another Greenland ice core. We chose the δ^{18}O record of the GRIP ice core, which is measured at a similar resolution to the NGRIP record and has been transferred to the GICC05 time scale starting at the onset of GI-23-1. We thus started fitting the record from GS-22 with 40 iterations of our algorithm, using the same constraints and hyperparameters. Again, the algorithm converges to a consistent fit, where each of the events is well approximated by a saw-tooth shape. We now describe how well the features of NGRIP and GRIP correspond for the 26 mutual events.

For the gradual cooling rates, the Pearson (Spearman's rank) correlation coefficient is $r_p = 0.64$ ($r_s = 0.65$). Here, the discrepancy in between the two records is only due to a handful of short very events, for which in one record a short but clearly visible linear cooling slope is discernable, whereas this is not the case in the other record, where the interstadials are more like a plateau. This happens for the interstadials 18, 16.2 and 5.1, which don't show a slope in GRIP, and 17.2, which doesn't show a strong slope in NGRIP. If we remove these events, the correlation is $r_p = 0.97$ and $r_s = 0.98$. The warming durations show a correlation of $r_p = 0.55$ and $r_s = 0.63$. There are no outliers, but a rather large spread, indicating that the warming duration is a less robust feature compared to the cooling rate. With 69 years on average, the GRIP warmings are 8 years shorter than the NGRIP average. The average absolute deviation of warming durations in the two cores is 31 years, with a maximum discrepancy of 103 years for GI-10, where we find a warming of 40

Tab. 2.3.: Durations and amplitudes of the rapid warmings inferred from the fit, together with a confidence interval obtained by bootstrapping.

Event	**Warming duration** (yr)			**Amplitude** (δ^{18}O permil)		
	Best fit	16-p	84-p	Best fit	16-p	84-p
24.2	43.4	36.3	47.8	4.30	4.23	4.40
24.1	47.4	34.9	45.0	3.53	3.42	3.61
23.2	115.2	96.1	126.1	3.92	3.72	4.11
23.1	94.1	78.9	127.3	2.73	2.69	2.75
22	70.0	70.3	91.8	1.89	1.78	1.95
21.2	33.0	25.7	39.9	4.06	3.61	4.10
21.1	61.7	53.5	79.6	4.05	3.98	4.09
20	18.2	14.7	21.6	5.34	5.25	5.42
19.2	97.2	74.3	98.1	7.09	6.93	7.19
19.1	58.5	37.7	60.0	4.86	4.45	4.97
18	161.0	74.6	194.0	4.21	3.99	4.51
17.2	129.7	83.7	158.0	5.79	5.47	6.20
17.1	15.3	13.9	27.0	4.53	4.14	4.72
16.2	21.0	18.6	24.0	4.92	4.59	5.19
16.1	28.4	28.9	84.0	3.01	2.88	3.16
15.2	61.6	39.0	100.0	3.41	3.38	3.67
15.1	60.6	56.4	69.2	5.94	5.68	6.12
14	35.5	38.0	79.0	3.87	3.78	3.95
13	62.4	63.4	101.2	2.29	2.07	2.60
12	63.9	45.7	73.8	4.86	4.71	4.94
11	179.5	143.0	201.0	4.05	3.86	4.17
10	40.3	41.3	80.2	3.67	3.40	3.97
9	44.2	37.5	86.4	3.09	2.66	3.22
8	31.7	29.8	53.0	4.91	4.78	4.98
7	47.4	45.3	90.2	4.07	3.86	4.27
6	140.4	110.6	172.1	3.51	3.41	3.92
5.2	36.0	31.1	54.6	4.76	4.45	4.93
5.1	41.8	41.4	82.0	1.65	1.51	1.89
4	37.2	27.1	41.8	4.84	4.47	5.11
3	21.3	18.0	25.0	5.88	5.40	5.92
2.2	61.2	42.0	91.6	3.72	3.21	3.75

years for NGRIP and 143 years for GRIP. Such deviations can arise if there is a slight step in the record before the most rapid warming and the algorithm includes this in the fit.

The warming amplitudes are very well correlated with $r_p = 0.87$ and $r_s = 0.83$. The average amplitude of 3.87 permil in GRIP is 0.45 permil lower than the NGRIP average. The stadial levels are also well correlated with $r_p = 0.78$ and $r_s = 0.66$. There is a quite consistent offset in between the cores of 1.84 permil due to differences in altitude and latitude of the GRIP and NGRIP sites. However, there are

also some clear differences in between the two cores. These include GS-21.1, which does not obey the offset but is at a very similar level in both GRIP and NGRIP, and GS-14, which is difficult to define and thus vulnerable to give different results due to different noise in the cores.

The rapid cooling durations, i.e. $b_4 - b_3$, are less well correlated with $r_p = 0.46$ and $r_s = 0.53$. This is expected because for many transitions this feature is less well defined than the rapid warmings. In these cases, the best fit determined by our algorithm is very susceptible to noise and can give qualitatively different results for different cores, i.e. a very abrupt cooling in one core and a much less abrupt cooling in the other. In this way the abrupt cooling of GI-19.2 lasts 208 years in GRIP and only 20 years in NGRIP, and for GI-12 294 years in GRIP and only 9 years in NGRIP. Conversely, the abrupt coolings of GI-19.1, GI-10 and GI-6 last much longer in NGRIP, with 62, 160 and 120 years in NGRIP versus 2, 5 and 2 years in GRIP, respectively. The average absolute deviation in between the two cores is 59 years.

The stadial and interstadial durations are very well correlated with $r_s = 0.99$ and $r_s = 0.97$, respectively. The average absolute deviation is 59 years for interstadials and 73 years for stadials, which is small compared to the average durations. The biggest discrepancies in between the two cores come from the indeterminacy in the rapid coolings of certain events, as described above.

In summary, the uncertainties obtained by bootstrapping and by comparison with the GRIP ice core are compatible. The average bootstrap standard deviation of rapid warming and cooling durations is 20 and 54 years, respectively. This compares well to the average absolute deviation in between GRIP and NGRIP of warming and cooling durations of 31 and 59 years, respectively. The discrepancy of 31 years for warming durations also includes a systematic bias of on average 8 year longer warmings in GRIP. Thus the unbiased uncertainty is likely even closer to the one obtained by bootstrapping. Shorter time scale features like rapid warming durations are not fully representative for every single event in one core. However, the overall trends are consistent, as seen by significant correlation. Features on longer time scale, such as most of the cooling slopes and stadial levels, are clearly representative. The same holds for stadial and interstadial durations.

2.3.4 Statistical analysis of DO event features

2.3.4.1 Interstadial periods

By our definition, the interstadial periods start at the maximum of the fit after the abrupt warming and ends at the start of the rapid cooling that leads back to stadial conditions. The interstadial durations are thus given by $b_3 - b_2$. In this analysis, we focus on the factors influencing the interstadial durations. As has been noted previously for the younger half of the last glacial in the GISP2 ice core, there is a strong correlation of interstadial durations D_I and their respective cooling rates $\lambda_c = AD_I^{-1}$, where A is the amplitude of the cooling [Sch02b]. If for every interstadial the gradual cooling would be perfectly linear and the jump back to stadial conditions would always occur after the same magnitude of cooling $\bar{A}$, the interstadial duration would be inversely proportional to the cooling rate $D_I = \bar{A}\lambda_c^{-1}$. If on the other hand the interstadials would have a fixed cooling rate $\bar{\lambda}_c$ and the jump back to stadial conditions would happen at a variable threshold, the interstadial durations would be proportional to the cooling amplitudes $D_I = A\bar{\lambda}_c^{-1}$.

We test which of the two scenarios is better supported by the data. In essence, this depends on whether either the cooling amplitudes or the cooling rates have a much larger spread than the other. The coefficient of variation for the amplitudes is $CV = 0.51$, whereas for the rates we find $CV = 1.49$. Furthermore, the Spearman correlation of interstadial durations and cooling rates is $r_s = -0.89$, which is clearly significant given the samples size of 31 events and weak autocorrelation of the sequence of interstadial durations and rates. Similarly, the logarithm of both quantities show a Pearson correlation of $r_p = -0.88$. For interstadial durations and cooling amplitudes we find $r_s = 0.40$, and for the logarithms $r_p = 0.41$. This correlation is largely due to three outliers, namely the two longest interstadials GI-23.1 and GI-21.1, and the event with largest amplitude GI-19.2. Removing these reduces the correlation to $r_s = 0.24$ and $r_p = 0.27$, which is not significant at 95% as determined by a permutation test. Thus, there does not seem to be a relation between durations and amplitudes that goes beyond outlier events, as opposed to durations and cooling rates. Furthermore, the correlation of cooling amplitudes and rates is not significant. Thus there is indeed a strong control by the cooling rates on the interstadial durations over the entire glacial, in agreement with the findings for the younger half of the glacial in [Sch02b].

In Fig. 2.5a we show a scatterplot of $\log \lambda_c$ and $\log D_I$ along with a linear regression yielding a slope of -0.94. The 95% confidence interval of this slope obtained via bootstrapping is [-1.12, -0.75]. Because we do not account for errors in the rates estimated from the data the regressed slope is biased towards 0 due to attenuation

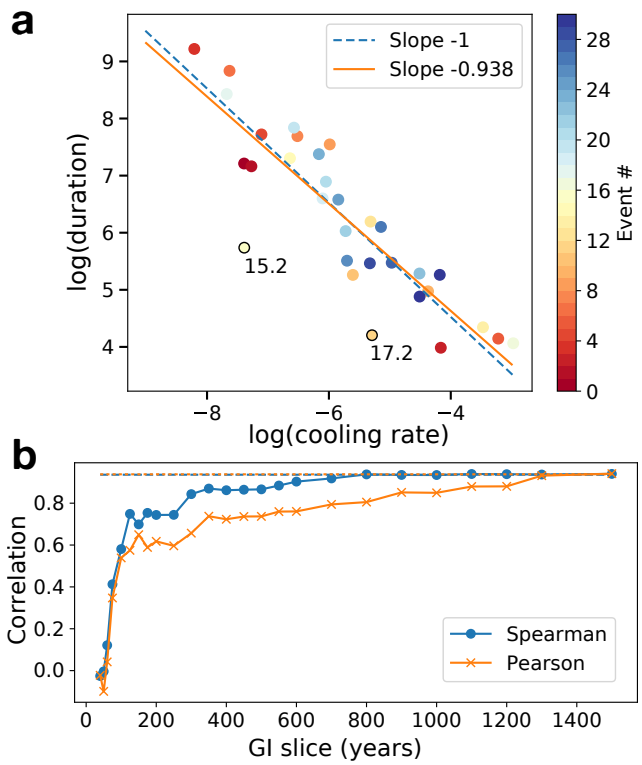

Fig. 2.5.: **a)** Scatterplot of the logarithms of interstadial durations and cooling rates. Two linear fits obtained by ordinary least squares are shown. For one of them we fixed the slope to -1 and varied only the intercept. **b)** Correlation coefficients of the logarithms of interstadial duration and the linear slope fitted to a slice of the beginning of the interstadial as a function of the length of that slice. The values of the Spearman (Pearson) correlation coefficients using slopes obtained from the full interstadials is marked with a *dashed* (*dotted*) line.

and the true slope will be closer to -1. Even though we in fact have estimates for the errors in the rates, we do not attempt to correct this bias at this point. Instead we can conclude that the model $D_I \propto \lambda_c^{-1}$ is consistent with the data, where the spread is caused by the fact that the jump back to stadial conditions happens after varying cooling amplitudes, which have a mean of 2.04 and standard deviation of 1.04.

The relation between interstadial durations and cooling rates also manifests itself in the respective distributions. As seen in Fig. 2.3, both durations and rates have strongly skewed histograms. They are both consistent with log-normal distributions, as shown by Anderson-Darling tests with $p = 0.47$ and $p = 0.89$ for durations and rates, respectively. Because of the relation of the two features, the fact that one is log-normally distributed implies that the other is, too. Consider D_I and λ_c^{-1} as random variables with $D_I = \bar{A} \cdot \lambda_c^{-1}$. If D_I is distributed log-normally with parameters μ and σ, then λ_c^{-1} also follows a log-normal distribution with parameters $-\mu + \ln(\bar{A})$ and σ. We show that in our case this relation holds as follows: We estimate μ and σ from the data D_I and use the observed average amplitude $\bar{A} = 2$. It follows that the data

λ_c^{-1} is consistent with a log-normal distribution with $-\mu + \ln(2)$ and σ , as seen by an Anderson-Darling test with $p = 0.33$.

While other skewed distributions like the exponential, gumbel and power law are rejected by statistical tests at high confidence, both durations and cooling rates are also consistent with an inverse Gaussian distribution. In fact, as indicated by relative likelihoods of 2.6 and 1.7 for durations and rates, respectively, the inverse Gaussian fits better than the log-normal. Given our small sample size, this might not be significant, however. The observation that the durations and rates are inversely related and both are well fit by the inverse Gaussian comes from the fact that the reciprocal inverse Gaussian distribution has a very similar shape. More specifically, if the true distribution of a variable is inverse Gaussian $X \sim IG(x)$, then the distribution of $Y = \frac{\bar{A}}{X}$ is reciprocal inverse Gaussian $Y \sim \frac{\bar{A}}{x^2} IG(\bar{A}/x)$. It is still expected that a moderately sized sample of Y is likely to be also consistent with an inverse Gaussian distribution, due to the similarity of the two. The inverse Gaussian could make an appealing model for the interstadial durations, since it arises as distribution of first hitting times of a constant level for Brownian motion with a constant drift. However, the interstadial time series look qualitatively much different than what is expected from this model, because they are quite linear, but have strongly varying slopes. In order for the model to produce roughly linear time series, the drift has to be high, which results in very similar slopes of the time series. In this parameter regime, the resulting inverse Gaussian distribution of first hitting times converges to a Gaussian. We leave a further discussion on which mechanism could yield log-normal or inverse Gaussian distributions of durations or cooling rates for upcoming studies. Instead, in the following we focus on implications of the approximate linearity of the interstadial time series.

The simple relation of interstadial durations and cooling rates might have some implications on the understanding of DO event dynamics. If the cooling rates imply the interstadial durations, then one must conclude that the durations are already determined as soon as the cooling rate is established, which might happen early on in the interstadial. This is different from the idea that the transition from interstadial to stadial might be a noise-induced escape from one metastable state to another. To test this, we take small slices of the beginnings of each interstadial, fit a linear slope s to them and then calculate how strongly these slopes determine the durations of the full interglacials as we increase the length of the slices. Due to noise in the beginning of the interstadials, for some interstadials a small positive slope is being detected. We set these slopes instead to $s = -0.0001$, because in our analysis we will use the logarithms of cooling rates and durations. For the relatively short events 15.2 and 17.2, no negative slope is obtained when fitting the whole interstadial part independently, as opposed to the slopes obtained in the fit of the

entire time series. We thus have to exclude these two outliers in the following. In Fig. 2.5b we show how the correlation in between the slopes s of these slices and the interstadial durations D_I evolves as we increase the length of the slices. For better visualization, we report correlations of $\log D_I$ and $-\log(-s)$. The correlation of the slopes estimated from the full interstadials and the durations when excluding events 15.2 and 17.2 is $r_s = 0.94$ ($r_p = 0.94$), and is indicated by a dashed (dotted) line in the figure. We can see that the correlation coefficient rapidly increases up to a length of 150 years. Thereafter the correlation stabilizes until another more rapid increase at a length of 350 years. The rapid increase in correlation is partly due to a non-negligible number of events already being at full length (6 events at 150 years and 12 events at 350 years). Still, also the slopes of the remainder does already correlate well with the final durations. At 350 years, the durations are almost as well determined by the slopes estimated from the slices as they are from the full interstadials. The remaining indeterminacy comes from a handful of longer events that did not settle to a clear negative slope after 350 years, namely interstadials 23.2, 22, 14, 11. For the latter three events, this is due to sub-events that occur shortly after the interstadial onset. Although we can see that there are exceptions, because of the rapid increase of the correlation coefficient towards -1 already for short slice lengths, we conclude that for most events the interstadial duration is determined at a relatively early stage within a few hundred years after the rapid warming.

Having established control of the interstadial durations by the cooling rates, we investigate whether the variability in the cooling rates can be explained by other features in the DO event series, or by external climate forcings. Although under initial inspection we find significant correlation of the cooling rates with several other features, none of them are relevant, either because they are caused by few outliers or else directly due to their definition. As an example, the cooling rates are (anti-)correlated with the rates (durations) of the steep drop to stadial conditions, because of the fact that very short interstadials (with high cooling rates) only support steep and short drops at the end.

Among the external climate factors, we find a correlation of $r_s = 0.40$ and $r_p = 0.35$ with the ice volume proxy LR04. This correlation is largely due to a common linear trend of the two quantities. The correlation is not significant anymore at 95% when removing a linear trend, as determined by a permutation test. On the other hand, there is a large sub-set of events which appears to be linearly related. As shown in Fig. 2.6, the furthest outliers from an approximate linear relationship are the interstadials 23.2, 21.2, 16.2 and 15.1. When removing these outliers, the correlation is $r_p = 0.79$, which clearly goes beyond a common linear trend with a correlation of $r_p = 0.63$ after linearly detrending. When only considering the younger half of the record starting with GI-14, which does not contain outliers, the

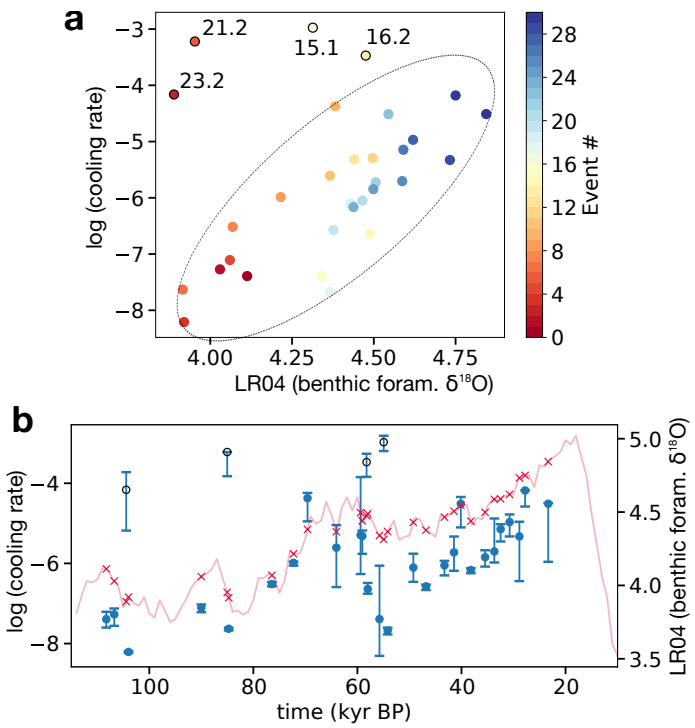

Fig. 2.6.: **a)** Scatterplot of the logarithm of the interstadial cooling rates and the LR04 values at time points corresponding to the interstadial onsets. **b)** Time series of the cooling rates (*dots*) and the LR04 stack (*crosses*). The error bars on the cooling rates are given by the 16- to 84-percentile obtained by bootstrapping.

correlation is $r_p = 0.84$. This corresponds to the finding in [Sch02b], who report that the interstadial cooling rates starting from GI-14 are forced by global sea level. We note, however, that the correlation is mostly due to the common trend of the two quantities, as we find $r_p = 0.37$ after linear detrending, which is not significant at 95%. Nevertheless, as shown above, when discarding outliers there is evidence for significant correlation as we include older parts of the record.

Similarly, in a subset of the events, there are indications of a linear relationship of the logarithm of the cooling rates and the Antarctic EDML record at the interstadial onsets. While the correlation of the entire data set is not significant at 90% with $r_p = -0.19$ and $r_s = -0.23$, when removing the events 24.2, 23.2, 23.1, 21.2, 16.2 and 15.1, the remaining events appear to have an approximate linear relationship, as indicated in Fig. 2.7. The correlation then becomes $r_p = -0.81$ and $r_s = -0.78$, and furthermore $r_p = -0.72$ and $r_s = -0.61$ after linearly detrending. Assuming independence, these values are significant with $p < 0.01$. Thus, in this subset there is evidence for anti-correlation beyond a simple linear trend. The linear relationship is strongest for the younger half of the record, which starts at GI-14 and does not have outliers. Here, we find $r_p = -0.89$, and $r_p = -0.70$ after linearly detrending.

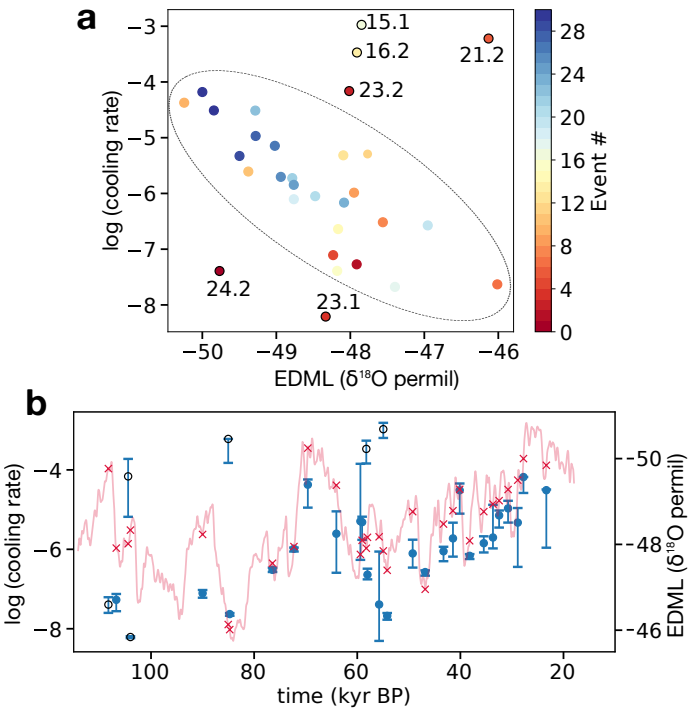

Fig. 2.7.: **a)** Scatterplot of the logarithm of the interstadial cooling rates and the EDML values at time points corresponding to the interstadial onsets. **b)** Time series of the cooling rates (*dots*) and the EDML stack (*crosses*). The error bars on the cooling rates are given by the 16- to 84-percentile obtained by bootstrapping. Note the inverted axis for EDML.

A corresponding linear relation of the logarithms of interstadial durations and Antarctic temperature has been noted before in [BS15], for different Antarctic ice cores. Importantly, they lump together into one event the two events comprising each of the interstadials 24, 23, 21, 17, 16, 15 and 2. With our entire data set we obtain correlations of $r_p = 0.29$ and $r_s = 0.27$ which are not significant at 95%. If we remove the abovementioned outliers plus GI-17.2, we find a strong linear relationship of $r_p = 0.93$, comparable to the findings in [BS15]. It is robust to linear detrending with $r_p = 0.87$. With the exception of GI-24.2, these outliers are all very short events. Removing the outliers, just like lumping them together with adjacent longer events as in [BS15], thus removes a lot of the variability in the interstadial durations.

Without any clear reason to treat the outliers as less significant events, we conclude that the relations of cooling rates and both LR04 and EDML records rather hold only on average, such as when averaging over multiple consecutive events. Because of the large variability in the interstadial features neither record sufficiently predicts all interstadial durations or cooling rates.

2.3.4.2 Stadial periods

The stadial periods are defined to start after the rapid cooling and end at the onset of the rapid warming, and are thus simply given by the values b_1. They have highly variable durations, ranging from the 20 years of GS-24.2 to 5169 years for GS-19.1, with an average of 1328 years. Our definition of the stadials gives rise to the exceptionally short duration of GS-24.2, where the proxy does not stabilize on a constant level, but rapidly warms again right after the rapid cooling. The histogram in Fig. 2.3 shows that the stadial duration distribution is clearly skewed. The data is consistent with an exponential ($p = 0.79$ with Anderson-Darling test) and a log-normal distribution ($p = 0.18$). The exponential distribution is 16 times more likely compared to the log-normal, as determined by the relative likelihood. Exponential distributions arise in the low noise limit of noise-induced escape times from asymptotically stable equilibria in dynamical systems [Day87].

In the following we discuss whether the stadial duration variability is to some degree influenced by other features in the data, or a result of external factors. Among external factors, the stadial durations are best correlated with 65Nsolst ($r_s = -0.64$). Among NGRIP features, the only one that is significantly and robustly correlated with the stadial durations are the stadial levels with $r_s = -0.43$. In Fig. 2.8a we show a scatterplot of the stadial levels and the logarithms of the corresponding durations. If one discards the first 6 events of the record, there is a clear linear correlation of $r_p = -0.80$. After linearly detrending both features, we obtain $r_p = -0.76$, and thus an anti-correlation that goes beyond a common linear trend. The stadial levels themselves are very well explained by external forcing. As a result, the correlation of stadial levels and durations could be either due to common forcing or due to the stadial levels directly controlling the durations. Although these scenarios are difficult to distinguish within our simplistic framework, we start by exploring how well the stadial levels can be explained by external forcing.

The stadial levels correlate well with ice volume and the Antarctic record, where much of the calculated correlation is due to a common linear trend. They are even better explained by insolation forcing. The linear correlation with 65Nsolst is $r_p = 0.60$ with two outliers (GS-24.2 and GS-22). When removing these, the correlation is $r_p = 0.82$, and does not change when linearly detrending because the insolation does not have a linear trend. To see whether this insolation forcing explains most of the correlation of durations and levels, we remove a linear fit to 65Nsolst from each variable and find a remaining linear correlation of $r_p = -0.38$. Even though the significance of this correlation is hard to assess due to the large autocorrelation of the stadial level series, this could imply that there is more information in the stadial levels about the durations than simply common insolation forcing. However, there

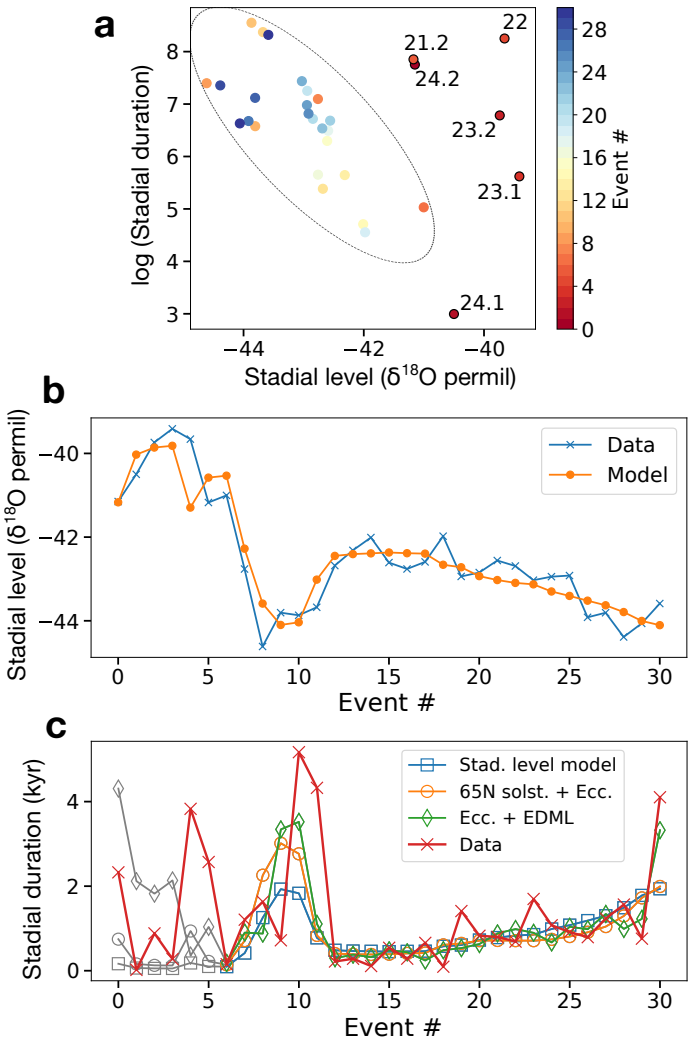

Fig. 2.8.: **a)** Scatterplot of stadial levels and logarithmic durations. Outliers from an approximate linear relationship are labeled. **b)** Event series of observed stadial levels and those modeled by $Y = 3.52 \cdot X_1 + 98.84 \cdot X_2 - 57.96$, where X_1 is 65Nint and X_2 the eccentricity. **c)** Models predicting the observed stadial durations (*crosses*). The first 6 events, indicated by *gray* markers, were discarded when fitting the models. The model based on predicted stadial levels from insolation (*squares*) is explained in the main text. The second model (*circles*) is given by $\log Y = -0.037 \cdot X_1 - 27.11 \cdot X_2 + 25.24$, where X_1 is 65Nsolst and X_2 eccentricity. The third model (*diamonds*) is given by $\log Y = -0.90 \cdot X_1 + 75.39 \cdot X_2 + 38.71$, where X_1 is the Antarctic EDML record and X_2 eccentricity.

could be additional components to the insolation forcing that might explain more of the observed variability.

We investigate whether multiple linear regression models with two predictors explain the stadial levels and durations significantly better. With a model comprised of 65Nint and eccentricity, the stadial levels are very well determined ($R^2 = 0.86$), as shown in Fig. 2.8b. The model is given by $L_{mod} = 3.52 \cdot X_1 + 98.84 \cdot X_2 - 57.96$, where X_1 is 65Nint and X_2 the eccentricity, both at stadial onset. The modeled levels also correlate reasonably well with the logarithm of the stadial durations ($r_p = -0.64$

when excluding the first 6 events). We check whether this is a good model for the durations in the original scale by regressing a linear model of the modeled stadial levels to the logarithm of the stadial durations, which yields $\log(D_{mod}) = -0.90 \cdot L_{mod} - 32.18$, and then exponentiating the predicted values. The result is shown in Fig. 2.8c, where we compare this model to two other multiple linear regression models, that directly regress the external forcings on the logarithm of the durations and then exponentiate the predicted values. After the first 6 events, for which none of the models fits adequately, all three models produce a similar trend. The model based on predicted stadial levels, and a model with direct forcing by 65N summer solstice insolation and eccentricity show similar skill with $R^2 = 0.29$ and $R^2 = 0.30$, respectively. The third model based on eccentricity and the EDML record is slightly better with $R^2 = 0.46$, mainly because it fits two of the longest stadials better. Still, all of the models fit only the overall trend and leave a large variability on top of the trend unexplained. A linear correlation of the logarithm still leaves a lot of room for variability in the original scale, unless the correlation is perfect.

To demonstrate that the exponential tail in the variability of the stadial durations is not a result of the modulation by external forcing, we remove the trend due to forcing by fitting a linear model of one or more forcings to the logarithm of the stadial durations. We obtain detrended data in logarithmic scale by adding the mean of the logarithmic data to the residuals of the fit. Finally we exponentiate to obtain detrended data in the original scale. When using 65Nsolst as forcing, we find $p = 0.15$ in an Anderson-Darling test on the exponential distribution. With the slightly better linear model of both eccentricity and 65Nsolst, as introduced above, we find $p = 0.29$. Thus, the distribution of the detrended data is still long-tailed and consistent with an exponential distribution.

2.3.4.3 Abrupt warming periods

The rapid warming transitions in NGRIP as determined by our piecewise-linear fit have an average duration of 63.2 years. There is quite a large spread with a minimum duration of 15.3 years for GI-17.1 and a maximum of 179.5 years for GI-11. There does not seem to be a trend, as we find both short and long warmings in early and later parts of the record. The distribution is skewed as seen from the histogram in Fig. 2.3. We find 5 transitions that last for more than a hundred years (interstadials 6, 11, 17.2, 18, 23.2). For all of them there is not only a single abrupt warming, but also a systematic departure from stadial to warmer values before, as can be seen in Fig. S1 of the supplemental material. Our algorithm includes these early warming trends into the warming transition. It is difficult to argue whether this is appropriate. Clearly, other methods to define the abrupt warmings might give different results in these cases. In [Rou+17], the transition onsets are defined by the

derivative of the signal and consequently the warming transitions into interstadials 6 and 11 are reported to be much shorter. Given our definition of abrupt warmings, we can at least argue that the longest warming transitions are not a result of noise, because in our fit of the GRIP record the same transitions are also among the longest and are clearly above average.

To investigate which distributions are consistent with the observed data we use Anderson-Darling, Cramer-von Mises and Kolmogorov-Smirnov tests. The Anderson-Darling test shows that the log-normal (p=0.63), Gumbel (p=0.053) and inverse Gaussian (p=0.95) distributions cannot be rejected at 95% confidence by the data. The other tests give equivalent results. By computing the relative likelihood from the Akaike information criterion, we find that the inverse Gaussian distribution is 9.7 times more likely than the Gumbel distribution, and the log-normal distribution is 7.6 times more likely than the Gumbel distribution. We cannot choose in between log-normal and inverse Gaussian with any confidence.

In the following we investigate whether we can infer anything about the mechanism of the warming transitions from the distribution of their durations. To begin with, it is clear that the durations of the warmings are much shorter than the time spent in the stadial state. If we consider the stadial-interstadial transition as a noise-induced transition from one metastable state to another starting at the stadial onset, most of the time is spent in the vicinity of the stadial state. The part of the trajectory that leaves this vicinity for the last time and then moves towards the other state (interstadial) is referred to in the literature as the reactive trajectory. The rapid warming durations of our piecewise-linear fit are estimates for the duration of these reactive trajectories and can be compared to what is expected in the framework of noise-induced transitions in multi-stable systems. For one-dimensional stochastic dynamical systems it has been proven that the distribution of the reactive trajectory durations converges to a Gumbel distribution in the zero noise limit [Cér+13]. Similarly, there is numerical evidence for the Gumbel distribution arising in one-dimensional spatially extended systems for low noise amplitudes [Rol+16]. Because in our data we cannot separate true climatic noise that might drive the observed large-scale climate transitions from other types of noise (non-climatic or regional climatic), it is hard to say whether a low-noise condition is indeed met and a Gumbel distribution should be expected.

To clarify what should be expected for finite noise amplitudes and small sample sizes, we conduct a numerical experiment with a stochastic double well model. It is given by the stochastic differential equation $dX_t = \left(-\frac{dV(X_t)}{dx}\right) dt + \sigma dW_t$, with the potential $V(x) = x^4 - x^2$ and the Wiener process W_t. There are two fixed points at $x = -1$ and $x = 1$. We initialize the system at $x = -1$ and repeatedly collect reactive

trajectories, which start when they last leave $x < -0.9$ and end as they enter $x > 0.9$. We find that small samples of 31 of these reactive trajectory durations are indeed typically consistent with a Gumbel distribution, but can be consistent with other distributions, too. To illustrate this, we perform Anderson-Darling tests on many of these small samples and record the p-values. This is done for a range of different noise levels. For the Gumbel distribution at $\sigma = 0.5$ the p-value of 90% of the samples lies in the interval (0.005, 0.85), which is given by the 5- and 95-percentile. For $\sigma = 0.00045$ the interval is (0.07, 0.92). Thus, in this case, very rarely a sample of 31 reactive trajectory durations would be rejected by a hypothesis test on the Gumbel distribution. However, the log-normal distribution fits equally well with 90% of the p-values in (0.05, 0.94). The distribution that most reliably fits the data is the inverse Gaussian distribution with 90% of the p-values in (0.50, 0.99). The fact that the inverse Gaussian distribution fits well also for large sample sizes has been already noted in [Cér+11]. Even non-skewed distribution can be consistent with the samples, as seen for the Gaussian distribution, which yields 90% of the p-values in (0.005, 0.50). Similar values are obtained for the logistic distribution. The results do not change qualitatively when we use KS or CVM tests instead, and also the noise level does not seem to have a large influence. These results imply that from this small sample size we cannot reliably identify the true distribution and thus an underlying mechanism. Still, the data are at least consistent with the expected behavior of noise-induced escape from a metastable state. Clearly, there are also other simple mechanisms that are consistent with data. For example, as already mentioned above, the inverse Gaussian arises as distribution of time elapsed for a Brownian motion with drift to reach a fixed level.

The variability in the warmings cannot be explained by other features of the record, as there is no significant and robust correlation, other than with the warming rates and amplitudes, which are trivially related. Multiple linear models with two features also cannot predict the warming durations. Similarly, we do not find any external forcing, or a combination of two forcings, that can explain a significant part of the variability.

Finally, we discuss the amplitudes of the rapid warmings, which have an average of 4.2 permil, with most events clustering around this value. Still, there is considerable variability, and the most extreme values are 7.1 permil for GI-19.2 and 1.7 permil for GI-5.1. The latter is not surprising, because GI-5.1 is almost not visually discernible as en event in the δ^{18}O series. When considering whether the variability can be explained by external forcing, we do not find a relation of the DO event amplitudes and global ice volume (LR04), as has been proposed before [McM+99; Sch+99]. It should be noted, however, that these earlier studies have a completely different notion of the amplitudes of DO events and millennial-scale variability. We assess

the actual amplitude of the rapid transitions, for which our approach based on the high resolution data seems ideal, because any low-pass filtering would reduce the amplitude of shorter events. Instead of ice volume, we find a correlation with 65Nint of $r_p = -0.36$ and $r_s = -0.31$, which is significant at 95% as determined by permutation tests. However, the correlation is visually not striking, and is largely determined due to the fact that GI-19.2 happens close to an insolation minimum. Removing GI-19.2 yields correlation coefficients that are not significant at 90%. Finally, note that the NGRIP record and thus the inferred amplitudes are not directly interpretable as temperature, since the relationship of δ^{18}O and temperature is likely to have changed over time. To illustrate, in the temperature reconstruction by Kindler et al. [Kin+14], the largest amplitude occurs for GI-11 with an estimated 16.5 K warming, and not for GI-19.2, which is estimated by 14 K.

2.4 Discussion and Conclusion

We developed a consistent fitting routine that allows us to extract robust features of DO events from the noisy, high-resolution NGRIP record. We demonstrated that the algorithm converges to a continuous piecewise-linear fit of the whole time series, where each DO event cycle is given by a constant stadial period, an abrupt warming period, a gradually cooling interstadial period and an abrupt cooling period. The resulting fit is satisfactory in the sense that each event receives a reasonable saw-tooth shaped fit. Not for all events this is necessarily the overall best piecewise-linear fit. For example, there are transitions that do not really have a significant rapid cooling period at the end of the interstadials, but rather cool gradually until reaching roughly constant stadial values. In fact, in Sec. 2.3.1 we showed that 14 out of the 31 DO events analyzed in this study do not follow the saw-tooth shape that is often reported as being generic for all DO events.

We assess the uncertainties of the fit parameters by using a bootstrap resampling technique and alternatively by comparing with a fit obtained from a different ice core (GRIP). The average absolute deviations of the GRIP and NGRIP features are of very similar magnitude compared to the average standard deviations obtained by bootstrapping. This gives us confidence in the validity of the uncertainty estimates of the latter method. From the uncertainties it follows that some of the shorter time scale features derived from the fit have to be taken with care, such as the rapid warming durations. Here, not all individual values might be reliable. However, the comparison with GRIP shows that we have reason to believe that the overall trends and distributions, also of the shorter time scale features, are robust. The results also depend on the method or model to define the features. As an example, our piecewise-linear method yields quite different estimates of the abrupt warming

durations as compared to the estimates given in [Rou+17], where abrupt warmings are defined by an estimated derivative of the signal. Our results have an average absolute deviation of 25 years (26 years) compared to their algorithmically (visually) determined warming durations starting at GI-17.1.

We subsequently analyzed different features that describe each DO event and that can be derived from the fit parameters. These features include the proxy levels in the stadials, at the interstadial maxima and ends, the durations and rates of warming and cooling periods, as well as stadial and interstadial durations. In general, all features except for the proxy levels develop rather irregularly from event to event, as shown by the absence of significant autocorrelation, and many of them are very broadly distributed. With statistical hypothesis tests we evaluate which distributions describe the individual features best. Identifying the distributions of the underlying processes can give some insight into the nature of the mechanisms giving rise to the abrupt climate changes. Furthermore, we investigated whether the variability in some features can be explained by other features or combinations of them. In order to do this, we searched for significant correlations of pairs of features and multiple linear regression models of two features that determine another feature well. This yields a long list of significant correlations in between features, many of which are not surprising because they occur in pairs of features that are closely related by definition. Other correlations are the result of a few outliers and are thus not robust. Among the remaining relevant relationships in between features, we test if these are due to a common underlying trend caused by external forcing. To do this, we correlate all features with a selection of external climate factors.

We synthesized some of the findings of the data analysis and reported the results in terms of different aspects of DO events, namely the interstadial, stadial, and abrupt warming periods. Except for a common forcing envelope of stadial and interstadial levels, there is hardly any inter-relation in between the features of the different periods. Additionally, they have different statistical properties and appear to be influenced by different external factors, as detailed in the following.

The interstadial periods have highly variable durations and are characterized by a roughly linear cooling, which is given by the gradual cooling rate. We find that the cooling rates are also highly variable and clearly determine the interstadial durations, as opposed to the cooling amplitudes, which cannot robustly explain the variability of the durations. Interstadial durations and cooling rates are consistent with a simple inverse relationship. Because the interstadial cooling is approximately linear, the interstadial durations are determined to a good approximation as soon as the cooling rates have stabilized. We estimate from the data that for most transitions this happens within the first 150 to 350 years of the interstadial. Next, we assess

whether the large variability in the cooling rates can be explained by other features of the record or external factors, and compare our results to two previously proposed external forcing mechanisms to control the interstadials. Based on the GISP2 ice core record of the younger half of the last glacial it has been proposed, that the interstadial cooling rates are controlled by global sea level [Sch02b]. While we can confirm this finding by observing a linear correlation of the logarithm of the cooling rates with the LR04 record for global ice volume, the relation seems to be weaker in the older half of the glacial. Here, there are a handful of outliers. Similarly, a control of Antarctic temperature on interstadial durations has been reported [BS15], but is only valid if certain outliers are discarded. Assuming that our classification of outliers is robust, it could indicate that not all of the transitions are caused by the same trigger.

The stadials have different properties compared to the interstadials, going beyond the fact that the temperature within stadials is approximately constant. The duration distribution closely resembles an exponential, and is thus consistent with noise-induced escape from a metastable state to another. The large dispersion of this distribution cannot be explained by external forcing alone. Instead, the distribution is still consistent with an exponential after detrending with the best fit to insolation forcing. Although there are different possibilities to obtain a good fit of the trend of the stadial durations, a forcing by insolation seems most plausible. We additionally find indications for a control by the stadial levels on the durations, but it is difficult to conclude from our data whether there is a true causal link, or merely common insolation forcing on both variables.

The piecewise-linear fit furthermore gives estimates for the amplitudes, rates and durations of the rapid DO warmings. We find no evidence for the rapid warmings to be influenced by any other features of the record or by external factors. Still, there is considerable variability of the warming durations, and we find the distribution to be consistent with the durations of so-called reactive trajectories in systems with noise-induced escapes in between multiple metastable equilibria [Cér+13; Rol+16].

Thus, our analysis suggests that both the stadial period durations and subsequent warming durations are consistent with the stadial to interstadial transition as a noise-induced escape from a metastable state. This is different from the interstadial to stadial transition, which seems to occur in a more predictable fashion, because our analysis shows that the linear cooling rate anticipates the interstadial durations. Additionally, the interstadial durations are not consistent with an exponential distribution. It has, however, recently been suggested that there is a bifurcation before DO events in a fast sub-system of the climate, which was based on evidence for critical

slowing down in the high-frequencies of the ice core record prior to a significant number of DO warming transitions [Ryp16; Boe18]. If this is the case it would mean that there is some predictability of the warming transitions, too. Finally, we find that the influence of external forcing is different for stadial and interstadial periods, with more evidence for insolation forcing on stadials and ice volume on interstadials, which is equivalent to the findings in [LD18].

In conclusion, we developed an iterative method to fit a continuous piecewise-linear waveform to the whole last glacial record, which converges well. By using parameter constraints, we can fit a characteristic saw-tooth shape to every DO event. However, we find that for many of the transitions this is ad-hoc. Almost half of the events do not show a distinct and significant rapid cooling episode after the more gradual interstadial cooling. An analysis of the DO event features that we derive from the fit confirms the irregularity and randomness that is evident from visual inspection of the record. There is hardly any evidence for relationships linking the features that describe the stadial, interstadial and abrupt warming periods, except for a common envelope that governs the stadial and interstadial levels via insolation forcing. A statistical analysis hints at different mechanisms underlying warming and cooling transitions. This manifests itself in different distributions and external influences of the stadials and interstadial durations, as well as the fact that the interstadial durations can be predicted to some degree by the interstadial linear cooling rates.

2.A Iterative algorithm to fit piecewise-linear model

Algorithm 1 Pseudocode for fitting algorithm

while $j < J$ **do**
  if $j = 0$ **then**
    $\{t_i^b\} = \{t_i^{b0}\}$; $\{t_i^e\} = \{t_i^{e0}\}$
  else
    Set Stadial durations: $\{D_i^{St}\} = \{b_{1,i} + t_i^b - t_{i-1}^b - b_{4,i-1}\}$
    Set Interstadial durations: $\{D_i^{Is}\} = \{b_{4,i} - b_{1,i}\}$
    Set Stadial beginning times: $\{t_i^b\} = \left\{\sum_{n=0}^{i-1}(D_n^{St} + D_n^{Is})\right\}$
    Set Stadial end times: $\{t_i^e\} = \left\{\sum_{n=0}^{i} D_n^{St} + \sum_{n=0}^{i-1} D_n^{Is}\right\}$
  end if
  Define Stadial levels: $\{l_i^s\} = \left\{\left\langle X_{t_i^b,\ldots,t_i^e}\right\rangle\right\}$
  Cut into segments: $\{s_i\} = \left\{X_{t_i^b,\ldots,t_{i+1}^e}\right\}$
  while $i < N$ **do**
    if j=0 **then**
      Initial conditions: $\theta_i^* = \theta_i^0$
    else
      Initial conditions: $\theta_i^* = \theta_i$
    end if
    Find optimal θ_i^{new} of segment s_i with θ_i^*, l_i^s and l_{i+1}^s
    $i = i + 1$
  end while
  Update parameters $\{\theta_i\} = \{\theta_i^{new}\}$
  $j = j + 1$
end while

The following list contains all constraints used in the optimization problem in order to ensure convergence of the algorithm to a fit within the qualitative limits of the desired characteristic waveform. Specifically, constraints 3 and 4 shall guarantee that there is a distinction in between gradual cooling and rapid cooling at the end of an interstadial. With these constraints we can prevent that our algorithm splits an interstadial in half with two very similar slopes, which can easily happen because there are interstadials which arguably have a rather gradual cooling all the way down to the next stadial with no easily discernable steep cooling at the end. The lower limit of constraint 6 shall help to only fit to the steep part of warming transitions,

which might have a slight warming prior to it. The upper limit of constraint 7 is needed in order to force a small negative slope on very short transitions which otherwise could also be viewed as plateaus.

1. No overlap of segments:
   $b_2 > b_1$, $b_3 > b_2$ and $b_4 > b_3$

2. Gradual slope cannot go below following stadial level l^s_{i+1}:
   $s_1(b_2 - b_1) + s_2(b_4 - b_3) > l^s_{i+1}$

3. Gradual slope must be twice as long as steep drop:
   $b_3 - b_2 > 2 \cdot (b_4 - b_3)$

4. Drop at end of interstadial must be at least twice as steep as gradual slope:
   $2 \cdot s_2 < \frac{s_1(b_2-b_1)+s_2(b_3-b_2)-l^s_{i+1}+l^s_i}{b_4-b_3}$

5. Stadial period not shorter than 20 years:
   $b_1 > 20$, $b_2 > 20$, $b_3 < (D^{St} + D^{Is} - 20)$
   and $b_4 < (D^{St} + D^{Is} - 20)$

6. Limit steepness of up-slope (permil y^{-1}):
   $0.02 < s_1 < 1.5$

7. Limit steepness of down-slope (permil y^{-1}):
   $-0.3 < s_2 < -0.0001$

For the basin-hopping algorithm we use a multivariate Gaussian Kernel of fixed variance with $\sigma_{b_1} = 15$, $\sigma_{b_2} = 15$, $\sigma_{b_3} = 15$, $\sigma_{b_4} = 15$, $\sigma_{s_1} = 0.004$ and $\sigma_{s_2} = 0.0015$.

2.B Convergence of iterative fitting routine

We repeatedly run our iterative fitting routine and monitor whether the individual parameters converge, so that a consistent fit is obtained in the end. Critical for obtaining a consistent fit is that the stadial levels do not change substantially, as explained in the Methods section. In Fig. 2.9a we show the evolution over 40 iterations of the incremental deviations of the stadial levels compared to the previous

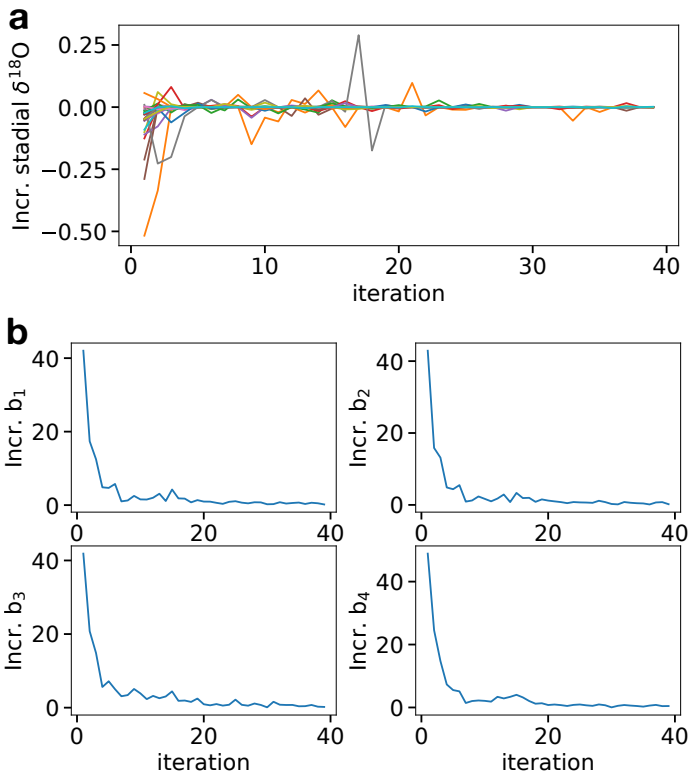

Fig. 2.9.: **a)** Evolution of the incremental change of all stadial levels compared to the previous iteration for all 40 iterations of the fitting routine. **b)** Average over all transitions of the incremental change (absolute value) of the break point parameters b_1, b_2, b_3 and b_4.

iteration. Most stadial levels converge rapidly so that their increments stay below 0.05 permil. Two short stadials keep fluctuating until around iteration 20 before they converge. Because of the convergence of stadial levels, we consider our fit to be consistent. Furthermore, the best fit parameters are robust, which can be seen in Fig. 2.9b. Here, we show the average absolute incremental deviations to the break point parameters at each iteration. After 15 iterations the procedure is stable, with average incremental deviations of roughly 0.4 years for b_1 and b_2 to 0.5 years for b_3 and b_4, which result from the stochastic fitting algorithm. Note that these values are already well below the smallest sample spacing of the original unevenly spaced time series.

2.C Uncertainty estimation of fitting parameters

Because of the nature of the data, care has to be taken when generating the synthetic data. The properties of the data changes throughout the record and are also quite different in between adjacent stadials and interstadials. Stadials have both a larger

variance and a larger effective sample spacing in time than the interstadials. For this reason, synthetic data will be created for each stadial and interstadial period individually. The original data is unevenly spaced, which would provide difficulties on its own, while our data is nearest-neighbor interpolated and oversampled to a 1-year resolution. This means that there typically are multiple neighboring point with the same value, making it difficult to find a good autoregressive or ARMA model for the residuals to generate synthetic data. Instead, we use a block bootstrap resampling technique to keep all relevant structure in the data. We chose a simple block bootstrap where non-overlapping blocks of fixed length of the time series are randomly ordered, because it preserves the correct mean of the individual stadial and interstadial residuals. Alternatively, a more involved method called stationary bootstrap could be applied, but we do not think it will change any of the conclusions.

In the following, we present the procedure for uncertainty estimation. We denote the original data time series of a given transition as $\{X_t\}$, the fit obtained by the data as $\{Y_t\}$ and the residuals to the fit as $\{R_t\} = \{X_t - Y_t\}$. We furthermore use the break points $b_{1,2,3,4}$ obtained in the fit of this transition.

1. Divide the residuals into four segment R_t^i at the breakpoints:
   $\{R_t^i\} = \{R_t\}_{t=b_{i-1}...b_i}$ for $i = 1...4$, where $b_0 = 0$.
   Denote the length of $\{R_t^i\}$ as n_i.

2. For each segment: Divide into n_i/l blocks of length l.
   Append remaining data points to last block if n_i/l non-integer.
   The block length l is determined by the length of the segment, as explained below.

3. For each segment: Randomly sample blocks without replacement and concatenate until all blocks have been used. This yields resampled segments $\{\bar{R}_t^i\}$.

4. Concatenate the four resampled segments and add the fit to get synthetic data
   $\{X_t^*\} = \{Y_t\} + \{\{\bar{R}_t^1\}, \{\bar{R}_t^2\}, \{\bar{R}_t^3\}, \{\bar{R}_t^4\}\}$

5. Fit $\{X_t^*\}$ to a piecewise-linear model with the basin-hopping algorithm.

6. Repeat from step 2.

In order to also be able to resample the shortest segments, while also preserving the autocorrelative structure in all but the shortest segments, we choose the following

scheme for the block length l: If the segment length n_i is larger than 40 years, choose $l = 20$. If $40 > n_i \geq 20$ choose $l = 10$. If $20 > n_i \geq 10$ choose $l = 5$. If $n_i < 10$ do not resample and simply return original segment. The scheme has been determined by looking at the residuals of each segments in all transitions and observing that the autocorrelation drops to non-significant values for all segments after 10-15 years. It thus seems reasonable to use the same block length rule for all transitions and segments.

3 Random and externally controlled occurrence of Dansgaard-Oeschger events

Abstract

Dansgaard-Oeschger (DO) events constitute the most pronounced mode of centennial to millennial climate variability of the last glacial period. Since their discovery, many decades of research have been devoted to understand the origin and nature of these rapid climate shifts. In recent years, a number of studies have appeared that report emergence of DO-type variability in fully coupled general circulation models via different mechanisms. These mechanisms result in the occurrence of DO events at varying degrees of regularity, ranging from periodic to random. When examining the full sequence of DO events as captured in the NGRIP ice core record, one can observe high irregularity in the timing of individual events at any stage within the last glacial period. In addition to the prevailing irregularity, certain properties of the DO event sequence, such as the average event frequency or the relative distribution of cold versus warm periods, appear to be changing throughout the glacial.
By using statistical hypothesis tests on simple event models, we investigate whether the observed event sequence may have been generated by stationary random processes or rather has been strongly modulated by external factors. We find that the sequence of DO warming events is consistent with a stationary random process, whereas dividing the event sequence into warming and cooling events leads to inconsistency with two independent event processes. As we include external forcing, we find a particularly good fit to the observed DO sequence in a model where the average residence time in warm periods are controlled by global ice volume and cold periods by boreal summer insolation.

3.1 Introduction

During the last glacial period, lasting from approximately 120 kya BP to 12 kya BP (thousands of years before present), a large number of abrupt large-scale climate

changes have been recorded in Greenland ice cores and other Northern Hemisphere climate proxies. These so-called Dansgaard-Oeschger (DO) events [Dan+93] are characterized by an abrupt warming of 10-15 K from cold conditions (stadials) to warmer conditions (interstadials) within a few decades. This is typically followed by gradual cooling, lasting centuries to thousands of years, until a more abrupt jump back to cold conditions is observed. The warming events are not regularly spaced over the glacial, but rather distributed in a complex temporal pattern, as can be seen in the NGRIP ice core record in Fig. 3.1. This raises questions about the causes of these recurring climate changes. Could an internal oscillation of large components of the climate system under strongly varying conditions give rise to this pattern? Are the climate changes in contrast manifestations of highly sensitive, multistable climate system components, where jumps in between different states are triggered in an unpredictable way by one or possibly many different other chaotic components?

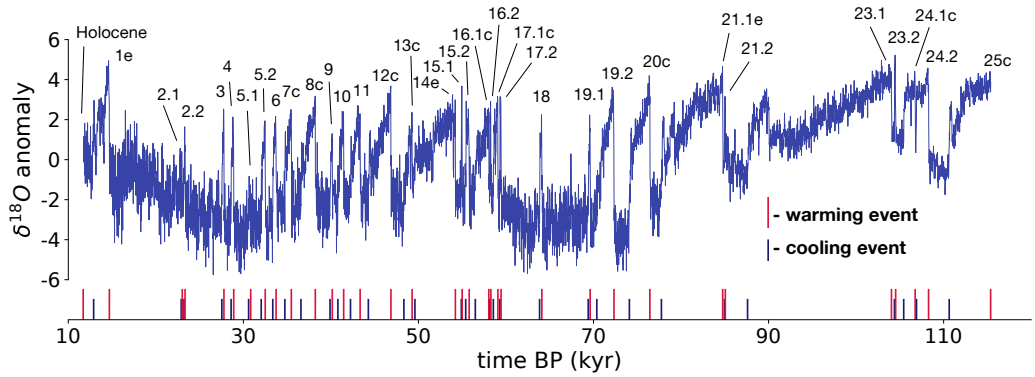

Fig. 3.1.: NGRIP oxygen isotope ice core record in 20-year binned resolution and associated Dansgaard-Oeschger warming and cooling events. The numbers above the time series indicate the warming transitions into the respective Greenland interstadials. The nomenclature is adopted from [Ras+14] and only events considered in this study are marked. On the time axis we marked the timing of warming (red) and cooling (blue) events.

Since the discovery of these unexpected climate events with no known cause, questions of this kind have been addressed. Whereas high-resolution coupled climate models under glacial conditions typically lack DO-type variability, models of intermediate complexity and simpler conceptual models have been proposed to explain qualitative features of the sequence of last glacial climate changes. Starting from the discovery of an approximate 1500 year spectral signature in the GISP2 ice core record [GS97] and an apparent in-phase pacing of individual events by multiples of this time period [All+01; Sch02b; Rah03], a number of competing hypotheses have been compared to the data. Among these were studies aiming to establish a mechanism for this periodicity, including direct triggering by periodic forcing [Bra+05], stochastic resonance [All+01], ghost resonance [Bra+07] and coherence

resonance [Tim+03]. On the other hand, it has been shown that there is limited significance to the periodic spectral signature [Bra+10] and pacing of individual events [Dit+07]. When including data reaching further back in time than 50 kyr BP it is found that only very weak periodic contributions to modeled switching sequences are compatible with the data and that instead it is more likely that the observed sequence of events is a realization of a purely noise-driven process [Dit+05].

In this work, we want to expand on this idea by testing whether the observed sequence of events is indeed consistent with one or more random, stationary processes, or whether the changes over time of the properties of the observed event sequence require a modulation of parameters of the governing process over time. To this end, we consider the whole glacial period, as opposed to previous efforts focusing on a rather regular period in the middle of the glacial. We investigate two different levels in detail of description by first only regarding the sequence of warming events and second the combined sequence of alternating transitions in between cold and warm conditions. We proceed by testing two null hypotheses: 1. The sequence of DO warming events is a realization of a Poisson process with fixed rate parameter. 2. The sequence of stadials and interstadials is a realization of two independent Poisson processes with fixed rate parameters giving rise to transitions in between stadials and interstadials. In order to test the hypotheses, we consider the evolution of the number of warming events in a moving window of 20 kyr. This quantity measures how variable the average event frequency is over time, a property which we denote as irregularity, and in the DO sequence it deviates strongly from a constant occurrence frequency of events over time. We test whether samples from the above mentioned stationary processes show a similar irregularity.

In addition to the evolution of the frequency of warming events we look at the evolution of the abundance of the stadial over the interstadial condition, which changes significantly over time in the DO sequence. This additional non-stationary structure in the data is the basis for another hypothesis test we perform. Finally, we test how the models' support with respect to the data is improved as we force the rate parameters with a combination of a global climate proxy and orbital variations of insolation, to incorporate changing background climate conditions. The main findings of this study are: 1. A Poisson process with fixed rate parameter, modeling warming transitions only, is consistent with the time variations in the NGRIP DO warming event sequence. 2. A model composed of two independent stationary Poisson processes governing transitions in between stadials and interstadials is not consistent with the time variations in the observed DO event sequence. 3. Forcing the aforementioned models with a combination of a global ice volume proxy and a summer insolation curve leads to good statistical agreement with the observed sequence. Specifically, we find good agreement for a model with two individual

processes, where the average transition rate from interstadial to stadial is controlled by global ice volume forcing, obtained from independent ocean core isotope records, and the average transition rate from stadial to interstadial is controlled by boreal summer insolation.

The paper is structured in the following way. In Section 2 we introduce in more detail the data used in this study, the summary statistics used to investigate irregularity in the event series, the models used to explain the data and the hypothesis test procedure. In Section 3 we present the results of the hypothesis tests on the different models. We discuss and interpret the results in Section 4.

3.2 Methods and Models

Our study of the sequence of DO events is based on the refined dating represented by the GICC05 time scale [Sve+06], the classification of Greenland stadials (GS) and Greenland interstadials (GI) given in Rasmussen et al. [Ras+14] and the timings reported therein. We consider all stadials and interstadials and corresponding transitions, starting with GI-25c at 115370 kyr BP and ending with the transition from GS-1 to the Holocene at 11703 kyr BP. We do not include events classified as subevents, i.e., drops in the middle interstadials to colder, but not fully stadial conditions, with the exception of GS-14. This yields a total number of 34 warming events and 33 cooling events. This increase in number from the 25 originally reported warming events is due to refined subdivision [Ras+14].

Given sequence and timing of transitions in between stadials and interstadials, we construct time-varying indicators of irregularity in the sequence of event timings, which are shown in Fig. 3.2a,b. To this end, we calculate the number of warming transitions within a moving window of 20 kyr at midpoint in time t, which we denote as $E(t)$. The window size of 20 kyr is chosen as trade-off between resolution and statistical robustness of longer-term features in the event sequence given the characteristic time scale of event occurrence of 3.1 kyr. The window is furthermore of comparable size to dominant variations in global background climate and insolation forcing, which will be investigated below. We obtain a time series indicating the deviation in the occurrence frequency of warming events from an evenly spaced (regular) event occurrence. We summarize this in a scalar test statistic E_S defined as the root mean squared deviation of the time series $E(t_n)$ from the expectation value $\bar{E}$

$$E_S = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(E(t_n) - \bar{E})^2}, \tag{3.1}$$

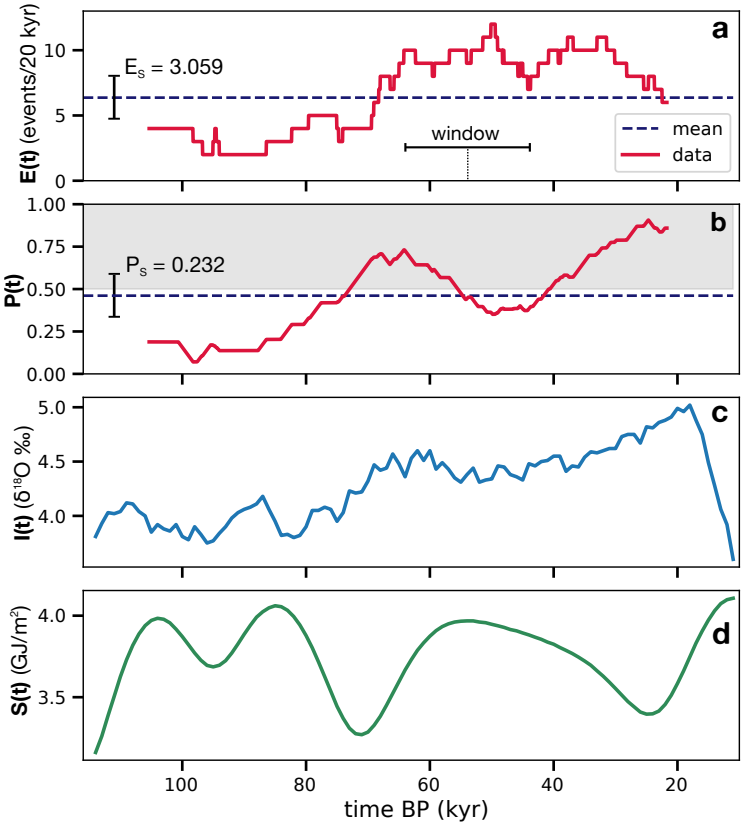

Fig. 3.2.: Time-varying irregularity indicators calculated from the NGRIP DO sequence, and climate forcings. (a): The number of warming events in a running 20 kyr window $E(t)$ (red) and the mean value (dashed blue). (b): The abundance of stadials in a 20 kyr window $P(t)$. For values greater than 0.5 (indicated by gray shading) the portion of stadials is larger than the portion of interstadials within the window. (c) Ocean sediment proxy record for global ice volume $I(t)$. (d) Integrated summer insolation at 65 degree North $S(t)$.

where $\bar{E} = 6.367$ the average number of events per 20 kyr of the whole DO sequence. For events occurring periodically with a period significantly smaller that the window size, the test statistic E_S is close to zero. For completely randomly occurring events the test statistic will show a finite value, which depends on the moving window size relative to the average waiting time in between events. The same time-varying indicator has previously been used to complement a model comparison study aiming to quantify the influence of external forcing to conceptual models of the NGRIP ice core record [MC17]. With this statistic we test whether the observed DO sequence departs significantly further from regularity as compared to what is expected by a random, uncorrelated event sequence. If this is the case, it would hint either at non-stationarity of the underlying process or a super-exponential event waiting time distribution. We consider the latter scenario to be less likely since no clear motivation for such a process exists.

While no significant correlation between duration of individual stadials and preceding or subsequent interstadial is observed (Pearson's $r = 0.04$ and $r = -0.15$,

respectively), the data suggests long-term variations in stadial and interstadial duration distributions. If these variations are systematic for stadials and interstadials (i.e. correlated or anti-correlated) they should be detectable in the correlation of individual neighboring stadial/interstadial durations given a large enough sample size. However, due to the small sample size of events in this study and the broad distribution of event waiting times a correlation due to long-term trends is not observed in practice. It is thus necessary to devise another time-varying indicator in order to capture additional detail in the structure of the DO sequence. When observing a given number of events in a time window, this may be either comprised of a combination of long stadials and short interstadials, or short stadials and long interstadials. This is not resolved in the statistic E_S. To capture this structure, we investigate the total portion of stadials within a moving window. Given the sum of the duration of all stadials $T_{st}(t)$ in a time window around a given midpoint in time t, the indicator is defined as $P(t) = T_{st}(t) \cdot (20 \text{ kyr})^{-1}$. We summarize this indicator by a scalar test statistic P_S. It is defined as the root mean squared deviation from the average value $\bar{P}$

$$P_S = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(P(t_n) - \bar{P})^2}, \tag{3.2}$$

where $\bar{P} = 0.461$ is the sum of all stadial durations divided by the total duration of the last glacial period.

We now describe the models which are used to evaluate our hypotheses on the data using the test statistics described above. The first model used in our study models the process generating the sequence of warming events as a Poisson process with fixed rate parameter λ, i.e., we disregard the cooling transitions in between warming events. It is denoted as 'one-process model' hereafter. The inverse of the rate parameter corresponds to the average waiting time in between warming events. The Poisson process corresponds to a situation where there is no memory of the past and thus the probability for a transition is determined by λ and is independent of time. All information on climate stability is represented in the parameter λ. We set its value equal to the inverse of the empirically observed average waiting time over the entire glacial record. This yields $\lambda = (3.141 \text{ kyr})^{-1}$.

As a second model, labeled 'two-process model' hereafter, we propose two individual processes for generating warming transitions from stadials to interstadials and cooling transitions from interstadials back to stadials. Each is represented by a Poisson process with a fixed rate λ_1 and λ_2, for warming and cooling, respectively. Again, the parameters are derived from the data by considering the empirical average residence times in stadials and interstadials, yielding $\lambda_1 = (1.477 \text{ kyr})^{-1}$ and $\lambda_2 = (1.663 \text{ kyr})^{-1}$. The model is different from the one previously introduced in that the sequence of warming transitions is not a Poisson process, but a more regular

one that is obtained from the sum of two independent processes. The probability distribution of waiting times T in between warming events is not exponential, but can be evaluated, yielding

$$P(t > T) = (\lambda_1 - \lambda_2)^{-1} \cdot (\lambda_1 e^{-\lambda_2 T} - \lambda_2 e^{-\lambda_1 T}). \tag{3.3}$$

The average interstadial and stadial durations of the data seem to behave differently over the course of the glacial, as captured by our second test statistic. This motivates us to study whether this behavior is likely to be encountered by chance assuming randomness and independence of both warming and cooling transitions.

As comparison to our hypothesis of stationary random processes, we consider the same models with time-varying rate parameters, which are given by a linear combination of two external climate factors: $\lambda = \tilde{\lambda} + aS(t) + bI(t)$. Firstly, we use a measure of incoming solar radiation at 65 deg North integrated over the summer $S(t)$ [Huy06]. It is defined as the annual sum of the insolation on days exceeding an average of 350 W/m^2. Secondly, we use the LR04 ocean sediment record stack as proxy for global ice volume $I(t)$ [RL05]. We note that, in contrast to insolation, global ice volume is not an external factor in the strict sense. However, its dominant variability is on longer timescales than DO events and most importantly it is obtained from an independent data source. Time series of these forcings are shown in Fig. 3.2c,d. The models' parameters are chosen such that the time-varying indicators are on average closest to those of the data. Specifically, by Monte Carlo simulation we generate many realizations for a fixed model parameter, compute time-varying indicators for each realization and then construct an average curve. Finally, the root-mean-square deviation (RMSD) from this curve with respect to the time-varying data statistic is computed. For best fit, we search for the least RMSD on a grid in parameter space. This corresponds to a numerical calculation of the maximum likelihood fit to the observed data. The two-process model is fitted to both statistics $E(t)$ and $P(t)$ simultaneously, by minimizing the normalized sum of the errors RMSD$_{E,P}$ to each of the statistics, defined as RMSD$_{sum}$ = RMSD$_P/E_S$ + RMSD$_P/P_S$.

The hypothesis tests are performed in the following way. For a given model we simulate a large number of realizations, which are collections of subsequent events with the same total duration as the record (104 kyr). For each realization we calculate the time-varying indicator of interest and the corresponding scalar test statistic. We then use the distribution of test statistics for a one-sided hypothesis test. The test simply counts how many test statistics in the ensemble are as large as or larger than the test statistic obtained from the data. Divided by the sample size, this yields a p-value, which estimates the probability of generating a random realization under the null hypothesis model that is at least as extreme as the observed data. We

can reject the null hypothesis at a confidence level α if the p-value is smaller than $1 - \alpha$.

3.3 Results

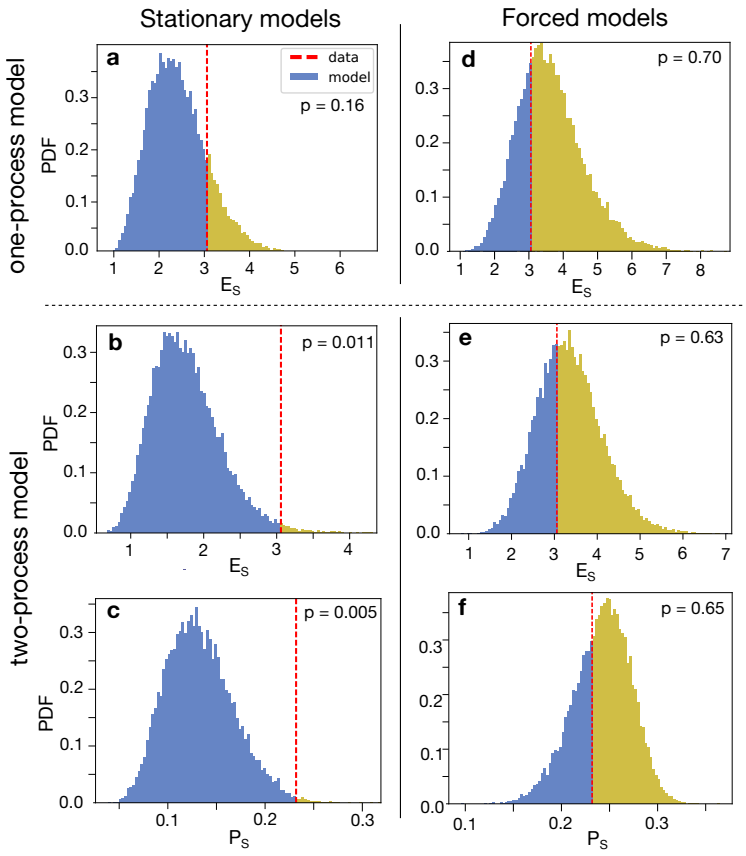

Fig. 3.3.: Empirical distributions from Monte Carlo simulation of the test statistics for the stationary (a)-(c) and the forced null model (d)-(f). The test statistics E_S of the one-process models are shown in panels (a) and (d). The test statistics E_S and P_S of the two-process models are shown in panels (b) and (e), and (c) and (f), respectively. The position of the data test statistic within the distribution is marked in red and determines the p-value of the hypothesis tests.

The results of the hypothesis test on the stationary one- and two-process models are shown in Fig. 3.3a-c. The plots show test statistic distributions of the respective null models and the corresponding test statistic value of the data. For the one-process model, the data test statistic lies well within the distribution, yielding a p-value of $p_E = 0.16$, as seen in Fig. 3.3a. Thus, we cannot rejected the null hypothesis at a level >85%. This indicates that the variations in the timing of warming transitions are consistent with a stationary Poisson process, i.e., without invoking variations in the rate parameter. Fig.'s 3.3b,c show the hypothesis tests of the two-process model, yielding low p-values p_E and p_P for both test statistics. The stationary two-

process model is thus rejected by the hypothesis tests with both test statistics at high confidence >98%.

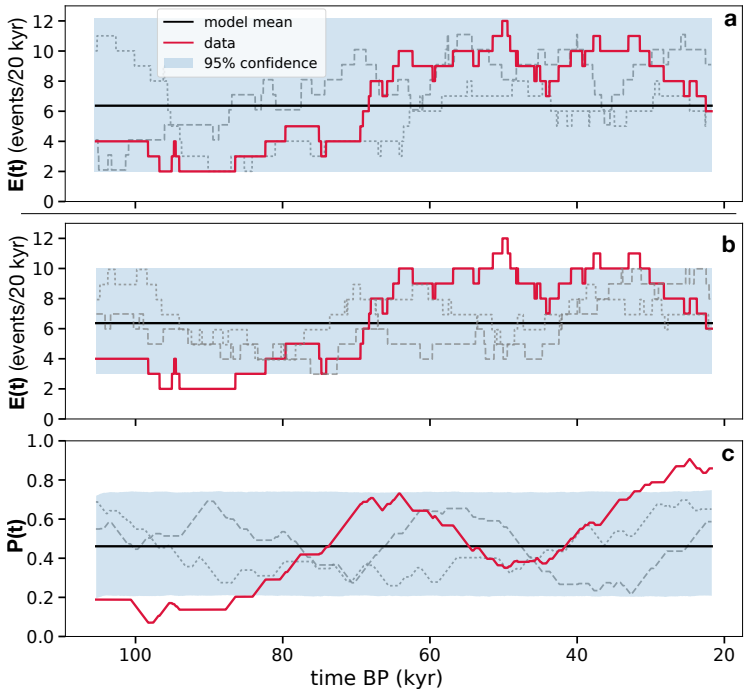

Fig. 3.4.: Point-wise 95% confidence bands and model mean (black line) for the time-varying indicators $E(t)$ and $P(t)$ from Monte Carlo simulations. (a) $E(t)$ for the stationary one-process model. (b) and (c), $E(t)$ and $P(t)$ for the stationary two-process model. The indicators for the data are shown in red and two typical model realizations are shown in gray.

To better visualize the outcomes of the hypothesis tests, we show confidence bands for the time-varying indicators from our Monte Carlo simulations in Fig. 3.4. The indicator $E(t)$ of the data lies within the 95% point-wise confidence band of the one-process model. Moreover, this band can be calculated analytically from the fact that the probability distribution of observing k events in a time period T is given by the Poisson distribution $P(k,T) = \frac{(\lambda T)^k}{k!} e^{-\lambda T}$. The cumulative distribution thereof allows us to calculate the probabilities of observing the minimal and maximal number of events per 20 kyr found in the data indicator $E(t)$. We find the probability to observe 2 or less events is $P = 0.047$ and to observe 12 or more events $P = 0.030$. This confirms that we cannot exclude the possibility of observing only 2 events and as much as 12 events during 20 kyr of the record at 95% confidence. The 95% confidence band of $E(t)$ for the two-process model in Fig. 3.4b is narrower and does not include the most extreme parts of the data curve. The same holds for the indicator $P(t)$, thus confirming that the two-process model can be ruled out with high confidence as null model for the observed sequence of events.

In the following we present the hypothesis tests performed on the one- and two-process models forced with insolation $S(t)$ and ice volume $I(t)$, which are both

Model	Parameters	p-Value	Goodness-of-Fit
Stationary one-process	$\lambda = 0.32$	$p_E = 0.16$	RMSD $= 3.05$
Stationary two-process	$\lambda_1 = 0.68$ $\lambda_2 = 0.60$	$p_E = 0.011$ $p_P = 0.005$	$\text{RMSD}_{sum} = 2.0$
Non-stat. one-process	$\tilde{\lambda} = 0.32$ $a = 0.43$, $b = 0.82$	$p_E = 0.70$	RMSD $= 1.42$
Non-stat. two-process	$\tilde{\lambda}_1 = 0.97$, $\tilde{\lambda}_2 = 0.97$ $a_1 = 1.60$, $b_1 = -0.57$ $a_2 = -1.96$, $b_2 = 2.56$	$p_E = 0.63$ $p_P = 0.65$	$\text{RMSD}_{sum} = 0.59$

Fig. 3.5.: Summary of model parameters, hypothesis test results and Goodness-of-Fit of the mean model time-varying indicators with respect to the data.

scaled to zero mean and range 1. Fig. 3.6 shows the time dependent transition rates as obtained from the parameter fit. For the one-process model we obtain

$$\lambda(t) = 0.32 + 0.43 \cdot S(t) + 0.82 \cdot I(t). \tag{3.4}$$

The best-fit two-process model has warming transition rate $\lambda_1(t)$ and cooling transition rate $\lambda_2(t)$

$$\begin{aligned} \lambda_1(t) &= 0.97 + 1.60 \cdot S(t) - 0.57 \cdot I(t) \\ \lambda_2(t) &= 0.97 - 1.96 \cdot S(t) + 2.56 \cdot I(t). \end{aligned} \tag{3.5}$$

The hypothesis tests for the fitted models are shown in Fig. 3.3d-f and yield high p-values, where the data statistic lies near the mode of the distributions. Note that we only measure the deviation of the time-varying statistics from a constant average value. Thus, the statistical test is not targeted at evaluating the fit to the data, but merely at probing whether the fluctuations over time of the indicators are of the right magnitude. Goodness-of-fit can be seen by means of the confidence bands and mean of the time-varying indicators, as shown in Fig. 3.7. For both models, the mean indicators lie close to the data curves, which consequently lie within 95% confidence bands. We summarize all model parameters, hypothesis test results and goodness-of-fit values in the table in Fig. 3.5.

We additionally report how the goodness-of-fit of the forced models changes when using only partial forcing and thus a reduced number of parameters. When forcing the one-process model with both ice volume and insolation, we yield a RMSD of the model mean $E(t)$ from the data curve of 1.42. Forcing with ice volume (insolation) only yields a best-fit RMSD of 1.64 (3.00). As baseline comparison, the RMSD from the unforced model to the data curve is equal to E_S, i.e. has a RMSD of 3.05. The model forced with ice volume fits the data only marginally worse than the model with both forcings and for comparison we show the mean time-varying indicator $E(t)$ for this model in Fig. 3.7a with a green dashed curve. For the two-process model,

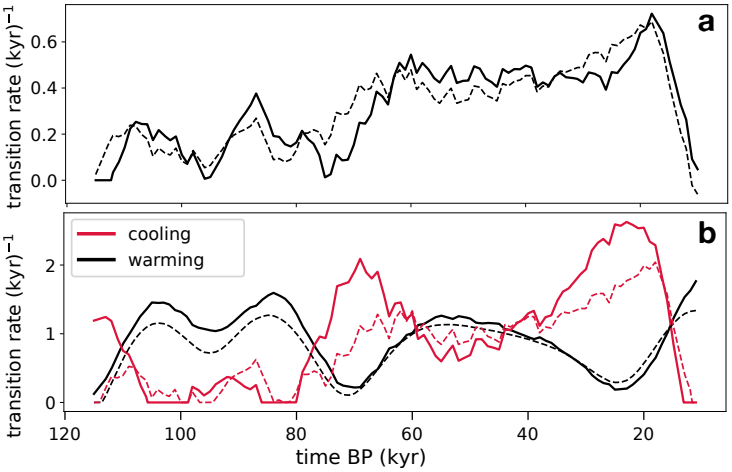

Fig. 3.6.: Time-varying transition rate parameters of the best-fit one-process $\lambda(t)$ (a) and two-process (b) models $\lambda_1(t)$ and $\lambda_2(t)$, as well as of the reduced models (dashed lines).

we considered all combinations where both warming and cooling processes are only forced by either ice volume or insolation. Goodness-of-fit in the two-process model is given by the sum of errors of both indicators RMSD_{sum}. For the best-fit model with full forcing we find $\text{RMSD}_{sum} = 0.59$, whereas the baseline of an unforced model gives $\text{RMSD}_{sum} = 2.0$. Forcing both warming and cooling processes with ice volume (insolation) only yields a best-fit of $\text{RMSD}_{sum} = 0.90$ (1.60). Using insolation for the warming transitions and ice volume forcing for the cooling transitions yields $\text{RMSD}_{sum} = 0.68$, while the converse choice yields $\text{RMSD}_{sum} = 1.68$. Thus, the only reduced two-process model yielding a comparable goodness-of-fit compared to the model with full forcing is the model with insolation forcing on warming transitions and ice volume forcing on cooling transitions. It is defined by

$$\begin{aligned} \lambda_1(t) &= 0.81 + 1.54 \cdot S(t) \\ \lambda_2(t) &= 0.80 + 2.39 \cdot I(t). \end{aligned} \tag{3.6}$$

We show the mean time-varying indicators for this model in Fig. 3.7b,c with a green dashed curve.

3.4 Discussion and Conclusions

Our first result considers only the warming events. While the distribution of waiting times in between warming events is well modeled by an exponential distribution (not shown here), we show here that the number of events in a moving window of 20 kyr (and thus the mean waiting time) is clearly changing over time, but no more

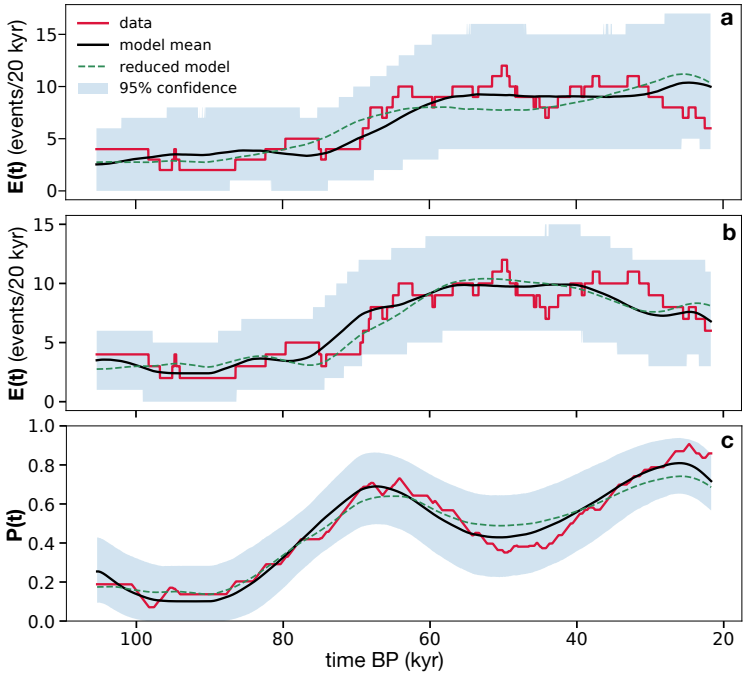

Fig. 3.7.: Point-wise 95% confidence bands and model mean (black curve) for the time-varying indicators $E(t)$ and $P(t)$ from Monte Carlo simulations. (a) $E(t)$ for the best-fit non-stationary one-process model with full forcing. (b) and (c), $E(t)$ and $P(t)$ for the best-fit non-stationary two-process model with full forcing. The indicators for the data are shown in red. The green dashed line indicates the best-fit model mean curves for the respective models with reduced forcing, as described in the main text.

than would be expected from a realization of a stationary Poisson process. Thus, if there is a unique process giving rise to the warming transitions, it need not be changing over time due to external factors. Although the description of DO events solely by the timing of the abrupt warming is very simplistic, we still think it is a useful result since the abrupt warmings are the most robust feature in ice core and other proxy records and are commonly used to assess synchronicity and pacing of abrupt climate change in the last glacial.

The second result indicates, however, limits to the stationarity in the sequence of events as we increase the detail of description. Assuming two independent processes giving rise to transitions from stadials to interstadials and vice versa, the null hypothesis of stationarity can be rejected with both our statistics. Specifically, both the variations over time of the number of warming events and the relative durations of stadials and interstadials are too large to be consistent with our two-process model using constant parameters. This model gives rise to a more regular sequence of warming events, compared to the one-process model. This is because one DO cycle is the sum of two independent processes and thus its duration does not follow an exponential distribution (coefficient of variation CV $= 1.0$), but Eq. 3.3, which is less dispersed (CV $= 0.708$). In the limiting case of a DO cycle comprised of a very

large sequence of N independent and stationary processes, one finds a Gaussian distribution of waiting times with decreasing variance as N grows. This would then correspond to an almost evenly-spaced sequence of events, which is not supported by the observations.

Next, we investigated improvements of the consistency of the models with the data by allowing their parameters to vary over time as linear combination of two climate forcings. Choosing the best fit linear combination of forcings, we found the average time varying indicators of both models to match very well to the data curve. Thus, whereas the data was seen as a rather out-lying realization that is consistent with a one-process model but not with a two-process model, when introducing forcings it becomes the expected behavior of the models. The goodness-of-fit follows from the correlation of the time-varying indicators and the forcings, which can be seen in Fig. 3.2. For the ice volume proxy we find a Pearson correlation of $r^2 = 0.78$ with $P(t)$ and $r^2 = 0.58$ with $E(t)$. We assess significance of this correlation by fitting an AR(1) process to the linearly detrended ice volume and conduct a hypothesis test yielding a correlation with $P(t)$ of 0.33 with a p-value of $p = 0.035$ and thus significance at 95% confidence. In contrast, the correlation of the indicator $E(t)$ and ice volume does not go beyond the linear trend. We do not assess the significance of correlations of insolation with the time-varying indicators, since it is difficult to find a good null model in this case.

Finally, we discuss the importance of ice volume and insolation in the best-fit one- and two-process models. In the one-process model, Eq. 3.4 shows that both increased insolation and ice volume lead to higher occurrence rates of events, with the contribution of ice volume approximately twice as large as that of insolation. We note that the decrease in DO activity towards the Last Glacial Maximum is not captured by the model because the ice volume forcing is dominating. The individual contributions in the best-fit model do not completely capture the importance of the two forcings. There are directions in the likelihood landscape of parameters which are very flat. As a result, we found that the best-fit model with only ice volume forcing yields a fit only marginally worse than the best-fit model with both forcings. We thus conclude that ice volume is clearly the more important control on the sequence of warming events. This is consistent with the findings in Mitsui and Crucifix [MC17], where Bayesian model selection criteria show that global ice volume is a more important forcing than insolation in stochastic dynamical systems as models for Greenland ice core records.

In the two-process model the warming and cooling transition rates are influenced by the forcing in opposite ways, as can be seen from Eq. 3.5: Warming transitions from stadials to interstadials become more likely for higher insolation and lower

ice volume, and vice versa for cooling transitions from interstadials to stadials. For cooling transitions, the contribution of ice volume is slightly larger than that of insolation. The warming transition rate is dominated by insolation, which contributes three times more than ice volume. With this model, the overall trend of mean waiting times in between warming events and of the stadial abundance is well captured, including the decrease in activity towards the LGM. Similar to the one-process model, we found a more parsimonious model which fits the data almost as well as the best-fit model with full forcing. This model uses only insolation forcing for stadials and ice volume forcing for interstadials, which complements the analysis of the individual contributions in the fully forced model. We thus hypothesize that based on our study there is evidence for insolation control on average stadial duration and ice volume control on average interstadial duration. This finding could hint at two distinct mechanisms responsible for transitions in between regimes.

An exhaustive investigation of whether our model description and subsequent findings are consistent with governing mechanisms for DO-type variability inferred from detailed data and realistic model studies is beyond the scope of this paper. Nevertheless we conclude the discussion with some interpretations which are more speculative in nature. We begin with insolation control on stadial duration. Boreal summer insolation might influence the occurrence frequency of warming transitions by modulating the ice-ocean albedo feedback, which amplifies break-up or export of larger areas of sea ice. Sea ice decrease could subsequently cause rapid warming through subsurface ocean heat release [Dok+13]. Initial openings of the sea ice cover might be created by wind stress. Evidence for stochastic wind stress forcing and subsequent sea ice changes have been reported in unforced model studies of rapid climate transitions [Dri+13; Kle+15]. To explain global ice volume control on interstadial duration we invoke different influences on the strength and stability of the interstadial (strong) mode of the Atlantic Meridional Overturning Circulation (AMOC). If we consider global ice volume as an indicator of mean global climate, we find consistency with coupled climate simulations that show correlation of the stability of the strong AMOC branch to freshwater hosing and mean climate state [Kaw+17]. We furthermore note the study in Buizert and Schmittner [BS15], where a correlation of individual interstadial duration and Antarctic temperatures from ice cores is established and explained by influences of Southern Ocean processes on the strength and stability of the AMOC. Given the strong similarity of the global ice volume record and Antarctic ice core records on longer time scales, this is closely related to our findings. We finally note that in our model description, the trigger for warmings and coolings is stochastic and thus different from near-periodic DO cycles [PV14].

In conclusion, we show that the long-term variations in DO warming event frequency, often described as millennial climate activity, is consistent with a memory-less stationary random process. From the data at hand we cannot exclude the possibility that the long-term variations have occurred by chance. If we however divide a DO cycle into two independent processes governing warming and cooling, this is not true anymore and significant time-varying structure is detected. We thus propose a model that incorporates long-term variations through forcing of the parameters with external climate factors. We find good agreement with the data in a model where the mean duration of interstadial phases of the DO cycle are controlled by global ice volume and the stadial phases by boreal summer insolation. This finding can help to distinguish in between different mechanisms that have been proposed to cause DO events.

4

A consistent model selection analysis of abrupt glacial climate changes

Abstract

The most pronounced mode of climate variability during the last glacial period are the so-called Dansgaard-Oeschger (DO) events. The underlying dynamical mechanism of these abrupt climate changes remains unknown and they are elusive in most simulations of state-of-the-art coupled climate models. There has been significant debate over whether the climate system is exhibiting self-sustained oscillations with vastly varying periods across these events, or rather noise-induced jumps in between two quasi-stable regimes. In previous studies, statistical model comparison has been employed to the NGRIP ice core record from Greenland in order to compare different classes of stochastic dynamical systems, representing different dynamical paradigms. Such model comparison studies typically rely on accurately reproducing the observed records. We aim to avoid this due to the large amount of stochasticity and uncertainty both on long and short time scales in the record. Instead, we focus on the most important qualitative features of the data, as captured by summary statistics. We perform Bayesian inference and model comparison experiments based solely on these summary statistics via Approximate Bayesian Computation. This yields an alternative approach to existing studies that helps to reconcile and synthesize insights from Bayesian model comparison and qualitative statistical analysis.

4.1 Introduction

The last glacial period, lasting from roughly 120 kyr to 12 kyr before present (1 kyr = 1 thousand years), has seen around 30 very abrupt changes in climate conditions of the Northern Hemisphere, known as Dansgaard-Oeschger (DO) events [Dan+93]. These events are the most pronounced climate variability on the sub-orbital timescales, i.e., below $\approx$ 20 kyr. In Greenland, they are marked by rapid

warmings from cold conditions (*stadials*) to approximately 10 K warmer conditions (*interstadials*) within a few decades. This is usually followed by a more gradual cooling, which precedes a quick jump back to stadial conditions. The spacing and duration of individual events is highly variable and largely uncorrelated in time over the course of the last glacial period. Some interstadials show gradual cooling for thousands of years, while others jump back to stadial conditions within 100-200 years. DO events are the primary evidence that large-scale climate change can happen on centennial and even decadal timescales. It is thus imperative to understand the underlying mechanisms in order to improve predictions of future anthropogenic climate change.

While significant climate change concurrent with DO events is well documented in various climate proxies from marine and terrestrial archives all over the Northern Hemisphere, it is most clearly observed in proxy records from Greenland ice cores. An important proxy is $\delta^{18}O$, which measures the ratio of the heavy oxygen isotope ^{18}O to the light isotope ^{16}O in the ice. This ratio is widely accepted as a proxy for temperature at the accumulation site. We consider the $\delta^{18}O$ record of the NGRIP ice core, which has been measured in 5 cm samples along the core. This results in an unevenly spaced time series with a resolution of 3 years at the end to 10 or more years at the beginning of the last glacial period. It is a matter of debate whether the highest frequencies in ice core records correspond to a true large-scale climate signal. Studies of ice coring sites with low accumulation rates have shown that the highest frequencies in the record can be dominated by post-depositional disturbances to the snow [Mün+16]. To facilitate analysis and to filter out some of these high frequencies, we will use an evenly spaced time series of 20 year binned and averaged $\delta^{18}O$ measurements. Still, it is unclear to what degree adjacent samples of this time series represent true large-scale climate variability. In our attempt to analyze and model the data, we thus concentrate instead on more characteristic statistical features, which do not concern the highest frequencies in the record.

Even after decades of research following their discovery, there is no consensus on the triggers of DO events, or on whether they are a manifestation of internal climate variability. In simulations of globally coupled climate models, DO-type events are largely elusive, although some recent studies report occurrences thereof, albeit through different mechanisms at play. Development in this area is hampered by very high computational costs of investigating millennial-scale phenomena with high-resolution climate models. Similarly, the paleoclimate data community has not settled on a comprehensive explanation by examining evidence from different proxy variables at different locations. With this work, we want to advance the understanding of mechanisms that could be a likely cause of DO events. We attempt to investigate whether it is possible to establish evidence in favor of one causal mech-

anism above others from the NGRIP $\delta^{18}O$ time series alone. To this end, we compare a suite of simple, stochastic dynamical systems models to each other via Bayesian model comparison. The models represent different dynamical paradigms and arise as conceptual climate models with different underlying physical hypotheses.

The NGRIP data set is characterized by the high amounts of irregularity that is displayed both on the very short time scales (possibly non-climatic noise) and longer time scales, as manifested in the high temporal irregularity of the abrupt events. We thus choose to view the time series at hand as one realization of a highly stochastic process, produced by the complex and chaotic dynamics of the climate system. As a consequence, we want to avoid fitting the models point-wise to the data, but rather demand the models to display similar qualitative, statistical features, such that the observations could be a likely or possible realization of the model. In order to do that in a qualitative way, this leads us to the construction of a set of summary statistics replacing the actual time series. To perform Bayesian parameter inference and model comparison implies the evaluation of a likelihood function of a model given a set of parameters and data. Since the likelihood function of our models is completely intractable, especially in the presence of summary statistics, we have to adopt a likelihood-free method. One method permitting this is called *Approximate Bayesian Computation* (ABC, first developed in [Pri+99], see [Mar+12] for a review). This technique allows us to compute Bayes factors and posterior parameter distribution. Compared to simply estimating maximum likelihood parameters, this is advantageous because we can assess the models' sensitivity in parameter space and see how well constrained individual model parameters are by the data.

The paper is organized in the following way. In Sec. 4.2 we will present the models examined in this study, along with some physical considerations motivating the study of these. In Sec. 4.3, our method is presented, i.e., the construction of summary statistics as well as the parameter inference and model comparison approach. Our results are given in Sec. 4.4, where we first demonstrate the method with a study on synthetic data in Sec. 4.4.1 and then present the study on the NGRIP data set in Sec. 4.4.2. We discuss our results and conclude in Sec. 4.5.

4.2 Models

In this study we restrict our analysis to models defined by stochastic differential equations of two variables, where the variable x will be identified with the climate

proxy. Several well-studied stochastic dynamical systems models are of the following form:

$$\begin{aligned} dx_t &= \left(a_1 x_t - a_3 x_t^3 + a_0 + b y_t\right) dt + \sigma_x dW_{x,t} \\ dy_t &= f(x_t, y_t) dt + \sigma_y dW_{y,t} \end{aligned} \tag{4.1}$$

The individual models investigated by us differ in the choice of $f(x, y)$ and specific parameters:

1. $f = 0$, $b = 0$: Double well potential (DW)

2. $f = -x_t + c$, $a_0 = 0$: Van der Pol oscillator (VDP)

3. $f = \tan(\beta) y_t - x_t + c$, $a_0 = 0$: FitzHugh-Nagumo model (FHN)

The DW model corresponds to stochastic, overdamped motion of a particle in a double well potential. It has been proposed previously as model for glacial climate variability [Dit99; TL00], and displays jumps in between cold and warm states at random times similar to a telegraph process. It can be derived from Stommel's classical model of a bi-stable Northern Atlantic Overturning Circulation (AMOC), which has been one of the most prevalent mechanisms invoked to explain DO events [Sto61]. When including stochastic wind stress forcing in the Stommel model and going to the limit of very fast ocean temperature equilibration, one yields stochastic motion in a double well potential [SY93; Ces94].

Similarly, relaxation oscillators, such as the VDP or FHN models, have been proposed for modeling Greenland ice cores [Kwa13; RS16b; MC17]. At first glance, they seem good candidates for generating DO events, since during a relaxation oscillation cycle one can get a characteristic fast rise and slow decay of the fast variable in a certain parameter regime ($c \neq 0$). We illustrate the most important dynamical regimes in Fig. 4.1. In the VDP model, the oscillatory regime is given if $|c|$ is small compared to the ratio a_1/a_3. On the other hand, as depicted in Fig. 4.1a, if $|c|$ is beyond a certain critical value, the deterministic system has one stable fixed point. Noise perturbations can kick the system out of this fixed point and excite a larger excursion in phase space until the fixed point is reached again. This is often referred to as the excitable regime. If we decrease b in the oscillatory regime, the period of oscillation grows, as the trajectory spends more time close to the stable parts of the nullcline of the x variable, which is also referred to as the slow manifold and is indicated in Fig. 4.1b. In the limit of $b = 0$ in Eq. 4.1, the variables decouple and we are left with a double well potential model for the variable x. Thus, both VDP and FHN models include a symmetric ($a_0 = 0$) DW model as a special case. The general form in Eq. 4.1 permits transitions between the very different models proposed in

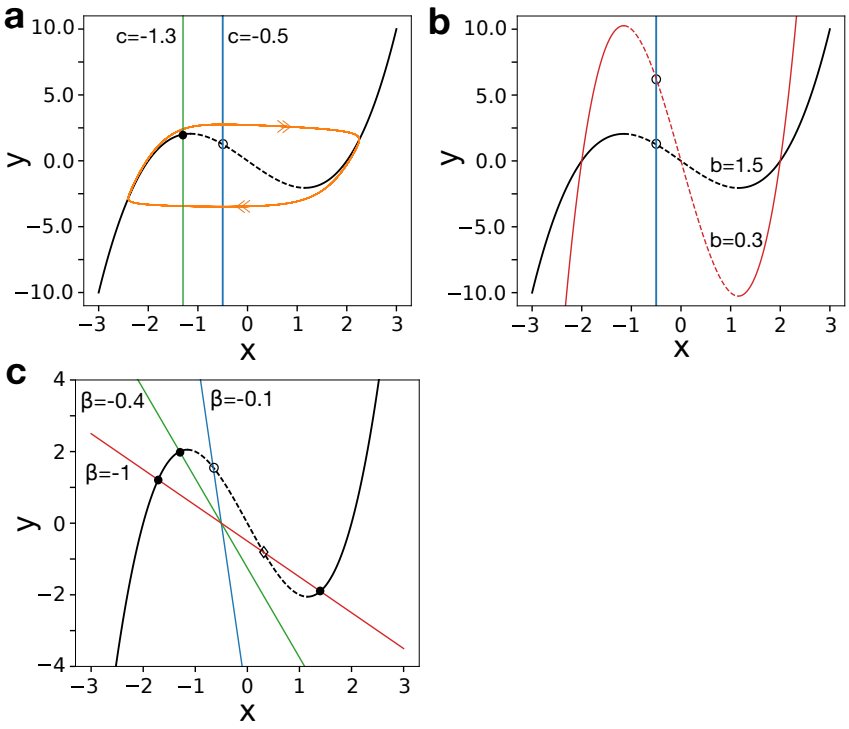

Fig. 4.1.: Phase portrait and nullclines of the VDP and FHN models with $a_1 = 4$ and $a_3 = 1$. The nullclines of the x variables are given by $y = (a_3 x - a_1 x^3)/b$ and are drawn in *black* for $b = 1.5$ in all panels. The *solid* part of that curve is the slow manifold. Stable (unstable) spirals are marked by *solid* (*open*) *circles*, and saddle points by open diamonds. a): Two y-nullclines of the VDP model given by $x = c$, indicating a transition from oscillatory ($c = -0.5$, limit cycle drawn in *orange*) to excitable dynamics ($c = -1.3$). b): x-nullclines of the VDP model for two different values of b indicating a stretching of the slow manifold and thus a lengthening of the period in the oscillatory regime. c): Three y-nullclines of the FHN model given by $y = (x - c)/\tan(x)$, indicating a transition from oscillatory ($\beta = -0.1$) to excitable ($\beta = -0.4$) and bi-stable ($\beta = -1$) dynamics.

the literature by continuous changes of parameter values. Similarly, the oscillator models we consider are nested, as explained in the following.

The VDP model is a special case of the FHN model, obtained by setting $\beta = 0$. Initially developed as simplified model for spiking neurons, the FHN model can display even richer dynamical behaviors including relaxation oscillations, excitability and bi-stability. The latter regime occurs for negative β, where below a certain critical value two stable fixed points emerge. As b decreases, this critical value gets closer to zero. Including additive noise in this regime induces stochastic jumps in between the two states. We indicated a transition from oscillatory to excitable and bi-stable dynamics by changing β and otherwise fixed parameters in Fig. 4.1c. For more details on the dynamics of the VDP and FHN models, and the rich bifurcation structure that appears especially close to the boundaries of the dynamical regimes, we refer the reader to [Roc+00]. Relaxation oscillator models, similar to the ones regarded in this study, can also be derived from Stommel's model, e.g., by including

an additional feedback from the ocean state to the atmosphere [RS16b]. We include noise forcing in the oscillator models, which is crucial in order to obtain the highly irregular oscillatory behavior that is seen in the data.

We simulate all models with a Euler-Maruyama method, a time step of $\Delta t = 0.0005$ and time scaled to units of 1 kyr. The actual model output we consider is given as a binned average of 20 years, i.e. 40 time steps, mirroring the pre-processing of the NGRIP data at hand.

4.3 Materials and methods

4.3.1 Data

Our model comparison study starts by preprocessing the NGRIP data set, as explained in the following. We use the 20 year averaged $\delta^{18}O$ data on the GICC05modelext time scale, as published by [Ras+14]. We remove a 25 kyr running mean, corresponding to a highpass eliminating variations due to orbital forcing on time scales longer than 20 kyr, which are not investigated in this study. With this, we are able to assess the statistical properties of the sub-orbital timescale dynamics in the signal using summary statistics. Finally, we cut the time series starting at 110 kyr, i.e. during GS-25 (GS = Greenland interstadial), and ending at 23 kyr b2k (before AD 2000), i.e. just after GI-2.2 (GI = Greenland interstadial). We do this in order to exclude the high early glacial $\delta^{18}O$ values before GS-25 and the rising $\delta^{18}O$ values in GS-2.1 with very high noise level in order to be able to objectively define warming events, as described below. The resulting time series is shown in Fig. 4.2.

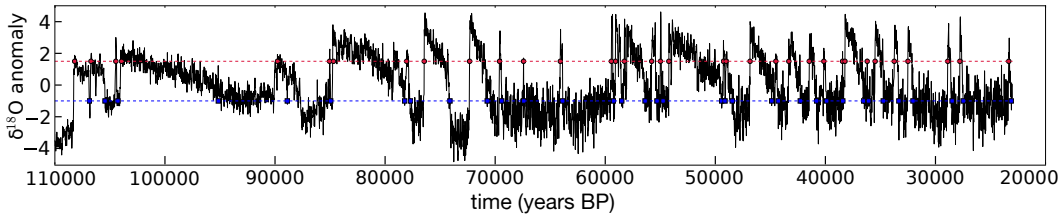

Fig. 4.2.: Slice of the NGRIP $\delta^{18}O$ time series high pass filtered with 25 kyr running mean, on which our study is based. Also shown are thresholds used to define warming (cooling) events, which are marked by *red* (*blue*) *dots*.

4.3.2 Summary statistics

As next prerequisite to perform parameter inference and model comparison one needs to specify a measure to quantify the goodness-of-fit of model output with

respect to data. We do not compare model output time series and data pointwise, e.g., using a root mean squared error. Due to the high stochasticity displayed in the data, it is irrelevant and possibly overfitting to find a model which would be able to produce a time series which is pointwise close to the data. Practically, one can use one-step prediction errors, assuming these are uncorrelated Gaussian. This has been done with the NGRIP record using Kalman filtering [KL09; Kwa13; MC17]. However, due to the high noise level and uncertainty in the interpretation of high-frequencies in the ice core data, our strategy is to replace the time series with a set of summary statistics and assess goodness-of-fit by comparing summary statistics of model and data time series. The summary statistics are described in the following.

We choose summary statistics which contain as much information as possible about the qualitative aspects of the NGRIP data that we want our models to reproduce. First of all, the models should show DO-type events, i.e. switching in between higher and lower proxy values. To define events, we introduce one lower and one upper threshold at $x = -1$ and $x = 1.5$, respectively. An up-switching event is defined by the first up-crossing of the upper threshold after up-crossing the lower threshold [Dit+07]. In the same way, the first down-crossing of the lower threshold after down-crossing the upper threshold defines a down-switching event. The result of this procedure applied to the detrended NGRIP data can be seen in Fig. 4.2. Periods in between up- and down-switching events (and vice versa) are denoted as interstadials (stadials). The thresholds are defined such that when applied to the detrended NGRIP data, the original classification of DO events and Greenland stadials/interstadials is reasonably well preserved [Ras+14]. Our classification differs such that GI-5.1 is not detected and GI-16.2 and 16.1 are detected as one single interstadial. Additionally, three very short spikes, which are not classified DO events, are identified as warming events (in GS-8, GS-9 and GS 19.1). We furthermore detect some of the most pronounced climate changes typically classified as DO sub-events, yielding 35 warming events in total.

With events defined as above we construct three summary statistics in the following way: Since one notable characteristic of the data is a broad distribution of durations in between events, we compare models and data using empirical cumulative distribution functions (ECDFs) of these durations. Specifically, given a time series, ECDFs are constructed for durations of stadials, interstadials, and for the waiting times in between two warming events. Two time series are then compared by computing the Kolmogorov-Smirnov distance of the respective ECDFs, which yields a scalar measure of goodness-of-fit for each of the statistical properties. These are denoted as s_1, s_2 and s_3, for stadial durations, interstadial durations and waiting times in between warming events, respectively. We visualize this construction in Fig. 4.3a-c.

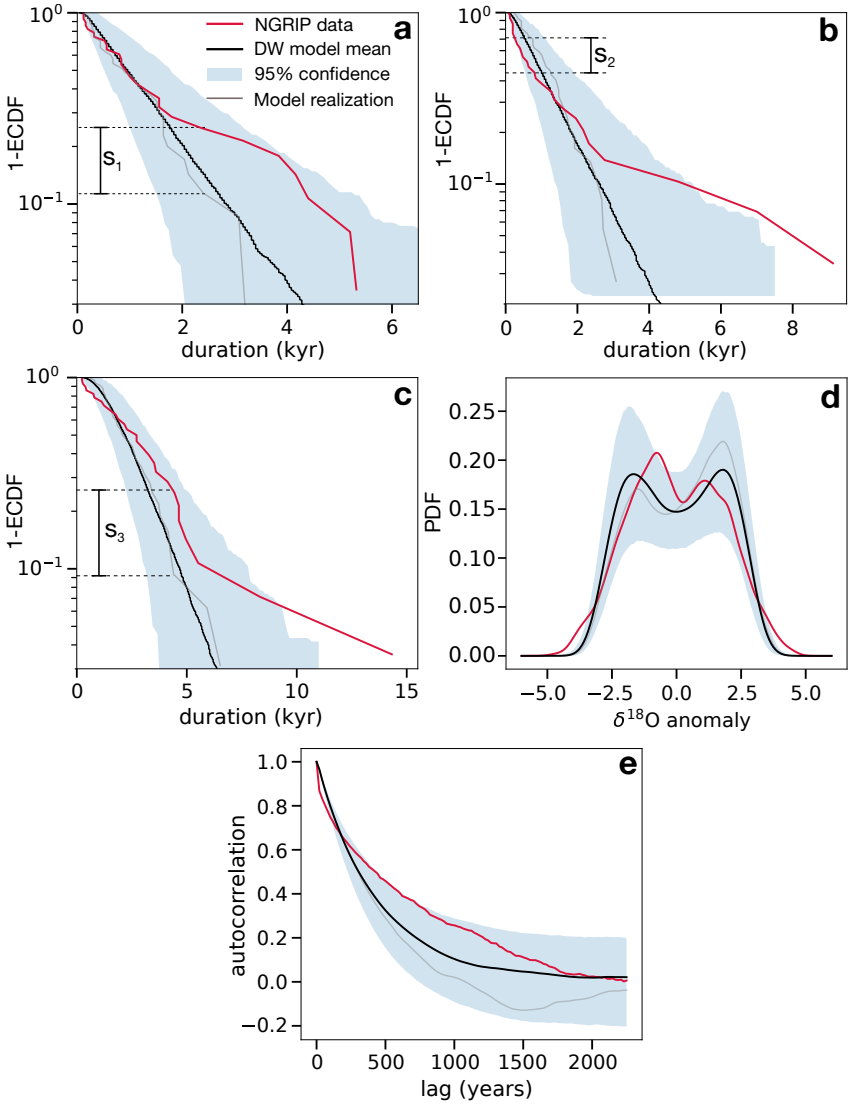

Fig. 4.3.: Statistical properties investigated in this study. a), b) and c) Complement of empirical cumulative distribution function $1 - \text{ECDF} = P(X > x)$ of stadial durations, interstadial durations and waiting times, respectively. The NGRIP data statistics are shown in *red*, the asymptotic statistics for a DW model with $a_0 = 0.16$, $a_1 = 2.86$, $a_3 = 0.93$ and $\sigma = 4.17$ is shown in *black*, corresponding 95% simultaneous confidence bands are shown with *blue shading* and an example realization is shown in *gray*. The maximal vertical distances of data and model realization are illustrated with *dashed* lines and correspond to our summary statistics s_1, s_2 and s_3. d) Probability density function (PDF) of the time series, used to compute s_4. e) Autocorrelation function up to a lag of 2250 years, which underlies s_5.

We introduce a fourth summary statistic in order to capture the bi-modal structure of the NGRIP time series, which is best observed from the stationary density shown in Fig. 4.3d. We compute the ECDF of the whole time series and make a pairwise comparison by computing the Kolmogorov-Smirnov distance, which we denote as s_4.

Finally, to capture the persistence properties of the detrended climate record, we base another summary statistic on the autocorrelation function up to a lag of 2250 years, as shown in Fig. 4.3e for both NGRIP data and a DW model. Two time series are compared by computing the root mean squared deviation (RMSD), which will be denoted as s_5. This yields a total of 5 scalar quantities to assess the fit of model output to data, which we summarize in a vector $\underline{s} = (s_1, s_2, s_3, s_4, s_5)^T$. For a good fit, we require all individual components to be sufficiently small, as will be discussed in more detail below.

An important qualitative feature of the NGRIP record so far missing from this description with summary statistics is the characteristic saw-tooth shape of the DO events. This behavior can also be captured with summary statistics, but with the models considered here it turns out to be hardly compatible with the other summary statistics introduced above. We discuss this statistical feature separately in Sec. 4.4.2.3.

4.3.3 Inference and model comparison

The measures for goodness-of-fit as defined above enable us to perform parameter inference and model comparison in an approximate Bayesian approach. Specifically, we aim to approximate two entities. First, we want to sample from the posterior distribution of model parameters θ given data D

$$p(\theta|D) = \frac{p(D|\theta)\, p(\theta)}{p(D)}, \tag{4.2}$$

where $p(\theta)$ is a prior distribution of the parameters. Second, we wish to compute the relative probabilities of each model given the data $p(\mathcal{M}_i|D)$, which is evaluated using Bayes' theorem:

$$p(\mathcal{M}_i|D) = \frac{p(D|\mathcal{M}_i)\, p(\mathcal{M}_i)}{p(D)}. \tag{4.3}$$

Here, $p(\mathcal{M}_i)$ is the prior probability of model $\mathcal{M}_i$. Thus, the relative posterior probability of two models is

$$\frac{p(\mathcal{M}_1|D)}{p(\mathcal{M}_2|D)} = \frac{p(D|\mathcal{M}_1)}{p(D|M_2)} \cdot \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)} = \mathcal{B}_{1,2} \frac{p(\mathcal{M}_1)}{p(\mathcal{M}_2)}, \tag{4.4}$$

where $\mathcal{B}_{1,2}$ is called the Bayes' factor and $p(D|\mathcal{M}_i)$ is referred to as the marginal likelihood or model evidence. The latter is an integral over parameter space of the product of likelihood and prior:

$$p(D|\mathcal{M}_i) = \int p(D|\theta, \mathcal{M}_i)\, p(\theta, \mathcal{M}_i)\, d\theta. \tag{4.5}$$

The computation of both the posterior parameter distribution and the model evidence require the likelihood $p(D|\theta)$, for which no explicit form is available for our models and summary statistics. We thus adopt a likelihood-free method called Approximate Bayesian Computation (ABC). An explicit expression for the likelihood is replaced by the following approximation

$$p(D|\theta) = \int p(x|\theta)\,\delta(x, D)dx \approx \int p(x|\theta)\,\pi_\epsilon(x, D)dx, \tag{4.6}$$

where x denotes model output, $\delta(x, D)$ is the Dirac function and $\pi_\epsilon(x, D)$ is a Kernel function which is strongly peaked where model output x and data D are similar according to some metric. Our choice of metric is the vector of summary statistics $\underline{s}(x, D)$ introduced earlier. We choose a rectangular Kernel function that is $\pi_\epsilon(x, D) = 1$ if all components $s_k(x, D)$ for given model simulation x satisfy $s_k(x, D) < \epsilon_k$, and $\pi_\epsilon(x, D) = 0$ otherwise. The vector $\underline{\epsilon} = (\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4, \epsilon_5)^T$ denotes tolerances to be specified. The only requirement for using this approach is the ability to produce model simulations x for given values of parameter θ, or in other words to sample from $p(x|\theta)$. The integral in Eq. 4.6 is replaced by a Monte Carlo sampling approximation.

The standard ABC approach proceeds in the following way: We sample a parameter from the prior $p(\theta)$, simulate a model output x and accept the parameter value θ as a sample from the posterior distribution with probability $\pi_\epsilon(x, D)$. This yields a sampling estimate of the parameter posterior, which furthermore gives us a Monte Carlo estimate of the integral needed to compute the model evidence in Eq. 4.5:

$$p(D|\mathcal{M}_i) \approx \frac{1}{J}\sum_{j=1}^{J} \pi_\epsilon(x_j, D), \tag{4.7}$$

where J is the number of samples drawn from the prior. In our case of a rectangular Kernel, this expression is equal to the rate at which randomly drawn parameter samples are accepted, i.e., yield summary statistics which are below the given tolerances.

Sampling parameters from the prior distribution in order to obtain the posterior is typically inefficient, since most of the prior parameter space has very small posterior probability. Instead we use an approach known as ABC population Monte Carlo (ABC-PMC) [Bea+09], which uses sequential importance sampling to approximate the posterior distribution through a sequence of intermediary distributions using decreasing values for the tolerances ϵ_k. In this approach, we start at some relatively large tolerance and draw parameter samples from the prior distribution $p(\theta)$ until a desired number of samples have satisfied $s_k(x, D) < \epsilon_k$. This population of parameters is then perturbed by a Gaussian Kernel and sampled from in the next

iteration with slightly lower tolerance. This perturbed distribution is referred to as the proposal distribution $f(\theta)$. From the second iteration on, the population of accepted parameter samples has to be weighted according to importance sampling in order to compensate that it was not drawn from the prior distribution but from the proposal distribution. The weights of a particle j in importance sampling is given by the likelihood ratio of prior and proposal distribution $w_j = \frac{p(\theta_j)}{f(\theta_j)}$.

Furthermore, in ABC-PMC the Gaussian Kernel used to perturb the previous population is adaptive. [Bea+09] use a diagonal multivariate Gaussian Kernel where the diagonal entries are given by the two-fold variance of the previous population samples. Instead, we use a multivariate Gaussian Kernel, where each entry is given by the two-fold co-variance of the previous population. This allows us to sample more efficiently when there is co-variant structure in the parameter posterior, as shown later. We stop the iterative procedure when the tolerances are so low that it is computationally very expensive to get a reasonable amount of posterior samples. The algorithm used in this study to obtain parameter posteriors and Bayes factors is is given and explained in the appendix.

It is important to note that because of the use of summary statistics there is no point to point model output and data comparison and thus we can use a length of model simulations different to the data length. Increasing model simulation length can sometimes increase performance, as will be discussed later.

4.4 Results

In order to demonstrate the method's abilities, we first apply it to synthetic data from within our model ensemble in Sec. 4.4.1. Thereafter, we present the results of the method when applied to the NGRIP data set in Sec. 4.4.2.

4.4.1 Synthetic data study

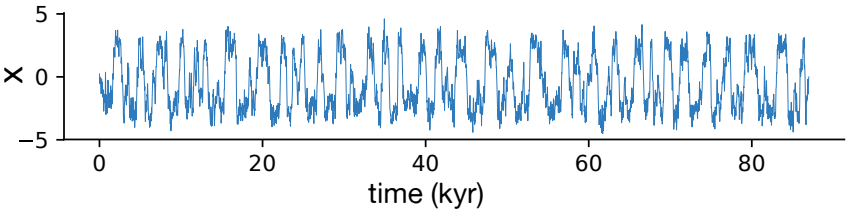

Fig. 4.4.: Time series of VDP model used as synthetic data to test the ABC-SMC method. The model parameters are $b = 6$, $a_1 = 6$, $a_3 = 1$, $c = -0.5$, $\sigma_X = 4.5$ and $\sigma_Y = 0$.

Tab. 4.1.: Sequence of tolerances used in the ABC-SMC experiment with synthetic data.

Iteration	$\epsilon_{1,2,3}$	$\epsilon_{4,5}$	Iteration	$\epsilon_{1,2,3}$	$\epsilon_{4,5}$
1	0.400	0.300	9	0.175	0.085
2	0.325	0.250	10	0.170	0.070
3	0.275	0.225	11	0.165	0.055
4	0.250	0.200	12	0.160	0.045
5	0.225	0.175	13	0.155	0.035
6	0.200	0.150	14	0.150	0.030
7	0.190	0.125	15	0.145	0.025
8	0.180	0.100			

As synthetic data, we choose a 87 kyr simulation output from the VDP model in a dynamical regime of noisy oscillations, which can be seen in Fig. 4.4. With this we demonstrate the following abilities of our method: 1. The correct model parameters are recovered from the posterior parameter distribution of the true model. 2. The true model is selected very strongly over a model which cannot operate in a comparable dynamical regime. 3. A model that can operate in the same dynamical regime as the true model but is of higher complexity is disfavored by the model selection procedure due to the higher number of parameters. 4. The results of model and parameter inference at sufficiently low tolerance are not critically dependent on data length and parameter prior distributions.

The model comparison parameters used in this synthetic data study are as follows. We used 15 ABC-PMC iterations with descending tolerances as specified in Tab. 4.1. Each model simulation output was equally long as the data (87 kyr), and a total number of 500 particles were used at each step. The prior parameter distributions for all models were chosen to be uniform.

4.4.1.1 Parameter inference

We first discuss the results for parameter inference, starting with the true model. In the violin plot of Fig. 4.5, we show the kernel density estimates of the VDP intermediate marginal parameter distributions for each iteration and indicate the bounds of the uniform prior distributions (red). We observe a gradual decrease in dispersion of the distributions as well as a convergence of the means close to the true values. Figure 4.6a shows in more detail the marginal posterior parameter distributions for the VDP model after the last iteration. We can see that there remains both an uncertainty in the parameter estimate as well as a small bias of the distribution mode for some parameters. The uncertainty is mostly due to the

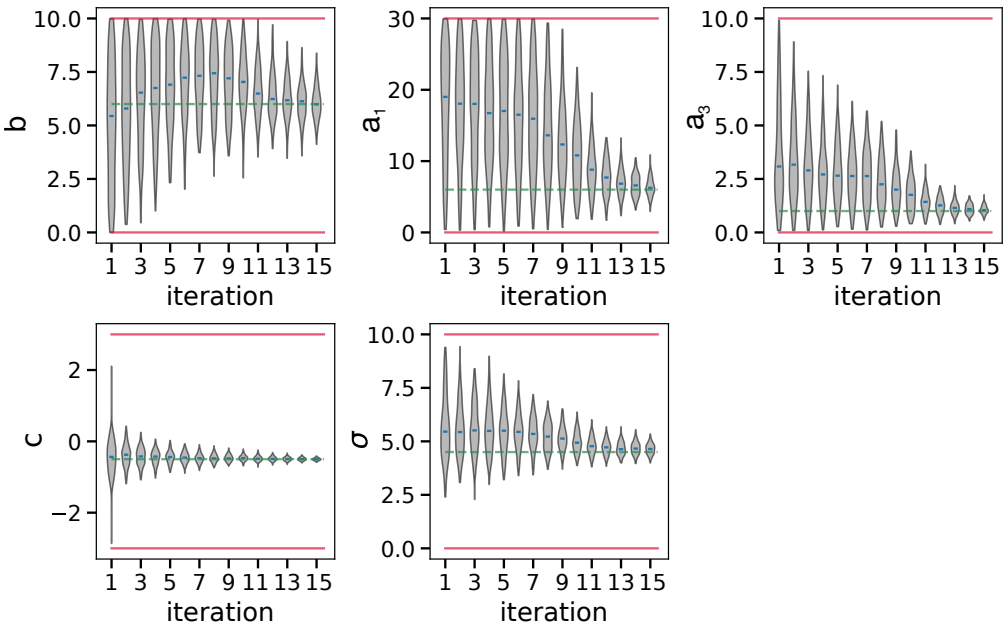

Fig. 4.5.: Violin plot illustrating the convergence of VDP marginal intermediate distributions for increasing iterations of the ABC-PMC algorithm. For each iteration, a Gaussian kernel estimate of the density is shown, together with the median. The true parameter values are indicated with a *green dashed* line. The bounds of the uniform distributions on the prior are indicated with the *red* lines.

non-zero tolerance and short simulation length, while the bias is due to random sampling and shortness of the test data. We conducted experiments with various data and model simulation lengths: When using shorter data length, the summary statistics are always quite different from the mean model statistics. Thus we find a bias in the inferred parameters. Longer data yields statistics closer to the model mean and thus less biased inference. However, the posterior dispersion does not change. If we increase the model simulation length, we can reduce the posterior dispersion because the statistics of model output samples are sharper for a given parameter and thus less wrong parameter samples scatter into the posterior. The Bayes factors are not systematically influenced in either case.

We furthermore observe that some parameters are better constrained by our summary statistics than others and are thus easier to infer, such as can be seen for c in contrast to b. Additionally, while most parameters seem independent of each other, the parameters a_1 and a_3 show a linear dependency in the posterior, which can be seen in the bottom right panel of Fig. 4.6a. This gives rise to most of the uncertainty seen in the marginal posteriors.

The parameter inference results for the FHN model are shown in Fig. 4.6b. We see that parameters, which the FHN model has in common with the VDP model, also converge to the true values. The additional parameter β stays quite uncertain but has most weight in a region close to 0, which would then correspond to the

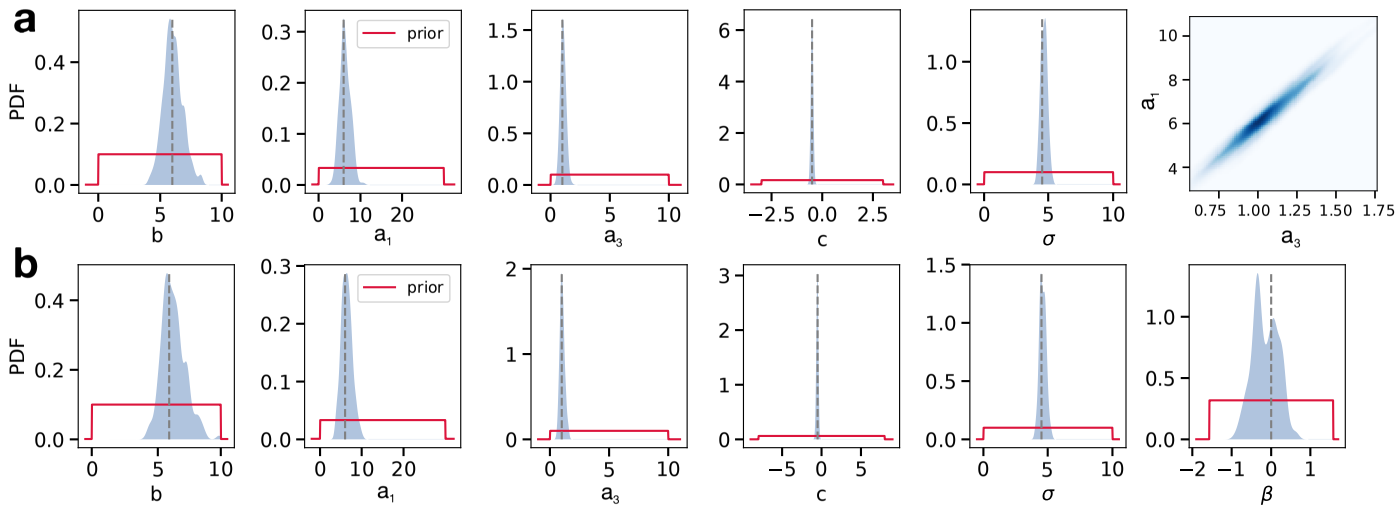

Fig. 4.6.: Gaussian kernel density estimates of the marginal posterior distribution of a) VDP and b) FHN model parameters as obtained after the last iteration of ABC-SMC on synthetic data. The *gray dashed* lines indicated the true parameter values. In the top right panel of a) we show the two-dimensional marginal distribution of parameters a_1 and a_3 for the VDP model.

VDP model. We do not show marginal posterior parameter distributions for the DW model, since it is eliminated by our model selection procedure after iteration 6, as discussed in the following.

4.4.1.2 Model comparison

We now discuss model comparison results. In Fig. 4.7 we show Bayes factors for four different ABC experiments at all iterations. The data from the experiment discussed in the previous section are shown in circle markers. The Bayes factors of the DW model over the VDP model, shown in Fig. 4.7a, drop to zero already at large tolerances, which means that the ABC procedure can efficiently exclude the wrong model. In contrast, Fig. 4.7b shows that the Bayes factors of the FHN model over the VDP model settles after some fluctuations to $B_{FHN,VDP} \approx 0.5$ as the tolerance approaches zero. This is because the two models are nested, i.e. the FHN model includes the VDP model but has an additional parameter. We can thus use this Bayes factor as an estimate of how much an additional parameter is penalized among models explaining the data equally well.

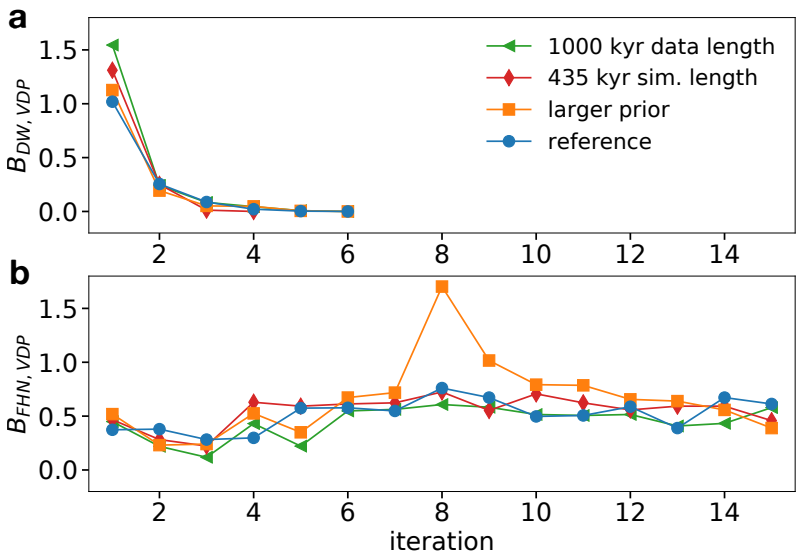

Fig. 4.7.: Bayes factors a) $B_{DW,VDP}$ and b) $B_{FHN,VDP}$ at all iterations of four different ABC-PMC runs using VDP synthetic data.

The squares in Fig. 4.7 show an ABC experiment where we doubled the prior range of the parameters b and a_1 in the VDP model. It is seen that the model comparison results do not depend on the width of the priors, given they are wide enough to contain the full posterior distribution. In the third ABC experiment, marked with triangles, we used a synthetic data length of 1000 kyr. The last experiment is marked with diamonds and shows results for using a longer simulation length of 435 kyr. In both of these experiments the model comparison results do not change qualitatively.

Tab. 4.2.: Sequence of tolerances used in the ABC-PMC experiment with NGRIP data.

Iter.	$\epsilon_{1,2,3}$	ϵ_4	ϵ_5	Iter.	$\epsilon_{1,2,3}$	ϵ_4	ϵ_5
1	0.4	0.3	0.3	7	0.19	0.08	0.11
2	0.3	0.225	0.225	8	0.185	0.075	0.1
3	0.25	0.175	0.175	9	0.18	0.07	0.09
4	0.225	0.15	0.15	10	0.175	0.065	0.08
5	0.2	0.125	0.125	11	0.17	0.06	0.07
6	0.195	0.1	0.115				

4.4.2 NGRIP data study

The ABC-PMC runs with NGRIP data were performed with the sequence of tolerances given in Tab. 4.2. Because none of the models can perfectly reproduce the NGRIP statistics, we had to stop the sequential algorithm due to computational demand at slightly higher tolerances compared to the synthetic data study. As in the synthetic data study, we used uniform priors for all parameters. The ranges of these priors can be seen in the respective figures showing the posterior distributions (Fig.'s 4.8-4.10) and were chosen wide enough to contain the full posterior distribution.

4.4.2.1 Parameter inference

The posterior distributions of the DW model ($b = 0$) are shown in Fig. 4.8 and lie well constrained within the priors. There remains considerable dispersion in the marginals of a_1 and a_3, most of which comes from a linear dependency of the two, as can be seen in the bottom right panel of the figure. With a_0 close to zero, the double well potential inferred from the data is approximately symmetric.

We now discuss the inferred dynamics of the oscillator models when only including noise in the x variable, i.e., $\sigma_Y = 0$. The posterior distributions are shown in Fig. 4.9. Because different regions in parameter space describe different dynamical regimes, we analyze the posterior samples as an ensemble. In the posterior samples of the VDP shown in Fig. 4.9a, we can see that the distribution of b is approaching 0. Thus the dynamics are effectively one-dimensional and approximate a symmetric double well potential, as discussed in Sec. 4.2. Still, 91% of the posterior samples are in a regime of noisy oscillations, because $|c|$ is too small compared to the ratio a_1/a_3. However, the oscillation periods expected from the deterministic system increase as b goes to zero and are much longer than the waiting times of the stochastic dynamics. The median ratio of deterministic period to stochastic waiting time is 38.4, with 10- and 90-percentiles at 8.2 and 165.1. Thus, the dynamics are such that much time

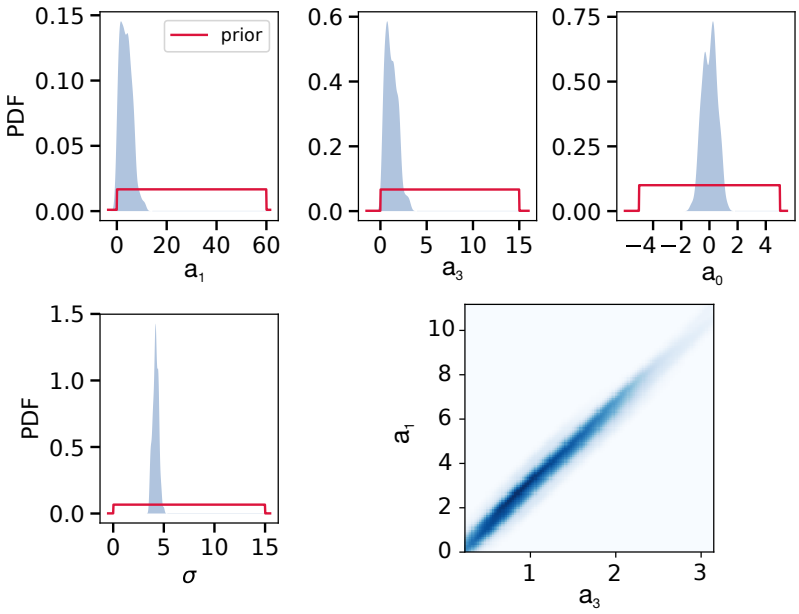

Fig. 4.8.: Marginal posterior distribution of DW model parameters after the last iteration of the ABC-PMC inference using NGRIP data. The bottom right panel shows the joint distribution of parameters a_1 and a_3.

can be spent on each branch of slow manifold, which is then escaped via noise. In effect, the model is noise dominated and the dynamics are closely similar to a double well potential.

Figure 4.9b shows the posterior distributions of the FHN model. We observe that b again becomes close to zero, albeit not as strongly as in the VDP model. Additionally, β approaches its limit of $-\pi/2$. The combination of these two parameter regimes typically yields two stable steady states, as explained in Sec. 4.2. Indeed, 95% of the posterior samples are in a bi-stable regime, whereas the remaining ones are in the excitable regime. The dynamics in the x variable in a bi-stable regime are again effectively very similar to double well potential dynamics. There is a large remaining dispersion in c, since the effect of c on the dynamics becomes negligible as β approaches $-\pi/2$.

As we include noise in the y variables of the oscillator models, the inferred parameter regimes change as seen from the marginal distributions in Fig. 4.10. In the VDP_Y model, the parameter b no longer tends to zero. As a result, the dynamics are no longer quasi one-dimensional. Out of the posterior samples, 83% are in an oscillatory regime, the rest being excitable. Within the oscillatory samples, the median ratio of deterministic period to stochastic waiting time is 1.02 (10- and 90-percentile at 0.73 and 1.75). Due to the parameter a_1 approaching very small values, the amplitude of the deterministic limit cycles is small compared to the amplitude of the noisy signal. Thus, the dynamics are again very noise-driven and apart from the mean

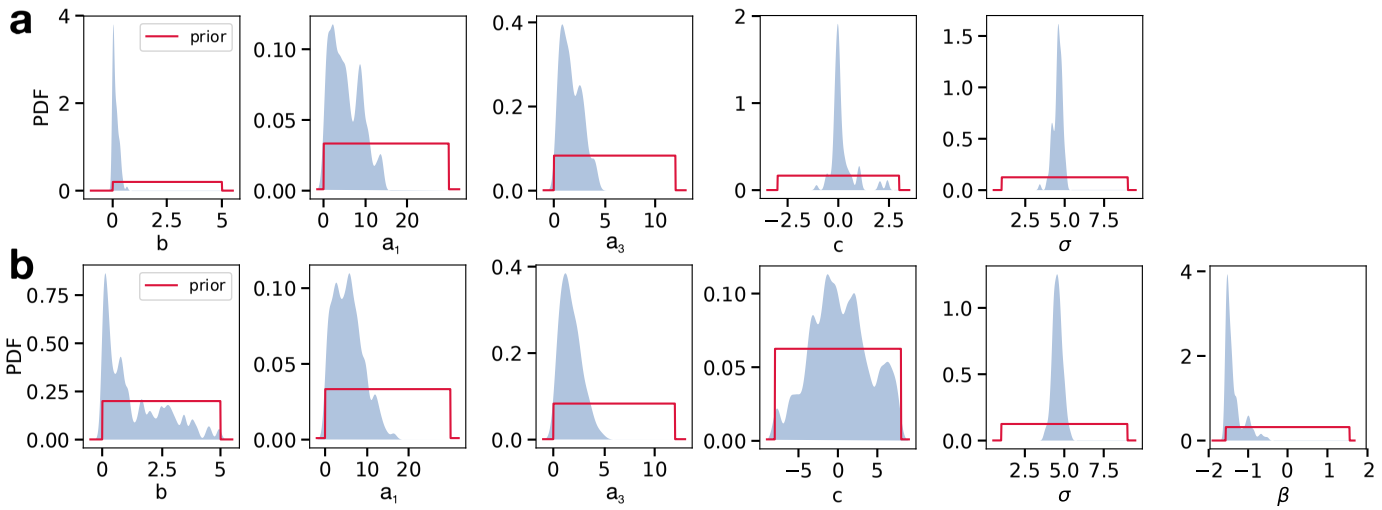

Fig. 4.9.: Marginal posterior distribution of a) VDP and b) FHN model parameters with $\sigma_Y = 0$ after the last iteration of the ABC-PMC inference using NGRIP data.

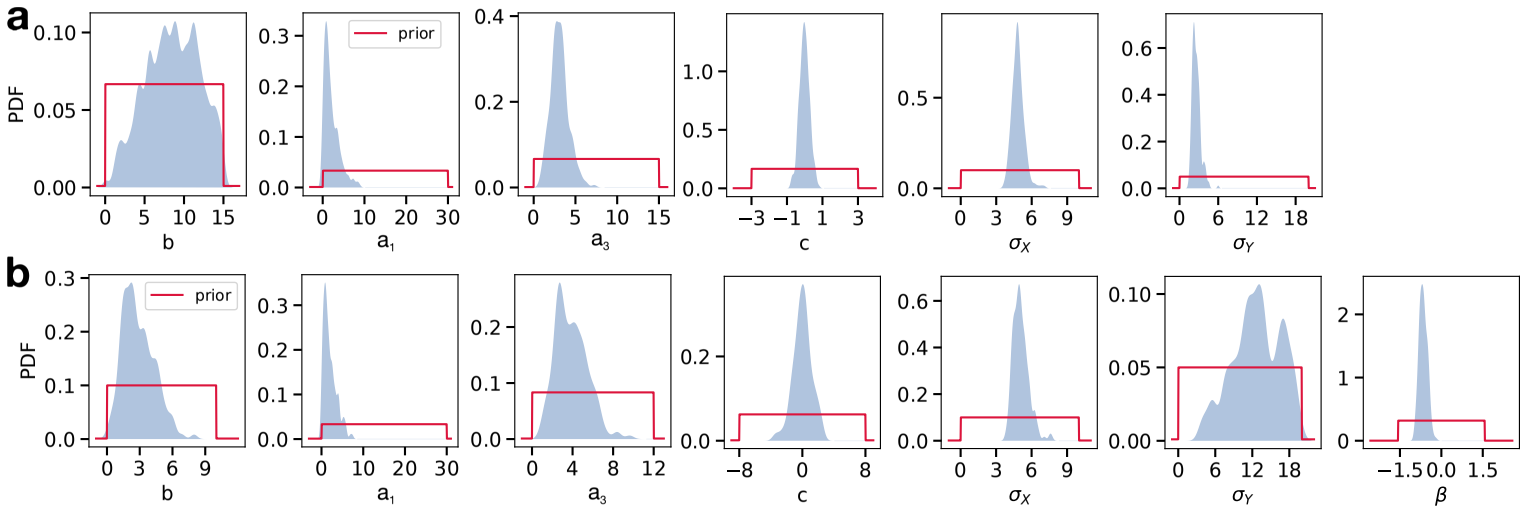

Fig. 4.10.: Marginal posterior distribution of a) VDP_Y and b) FHN_Y model parameters with noise in the y variable after the last iteration of the ABC-PMC inference using NGRIP data.

Tab. 4.3.: Highest probability parameters within the posterior sample estimated by Gaussian Kernel smoothing.

Model	Best parameter estimates						
DW	$a_0 = 0.16$	$a_1 = 2.86$	$a_3 = 0.93$		$\sigma = 4.17$		
VDP	$b = 0.04$	$a_1 = 4.42$	$a_3 = 1.35$	$c = 0.07$	$\sigma = 4.44$		
FHN	$b = 0.04$	$a_1 = 2.23$	$a_3 = 0.82$	$c = -6.98$	$\sigma = 4.46$		$\beta = -1.51$
VDP_Y	$b = 10.23$	$a_1 = 1.43$	$a_3 = 2.89$	$c = 0.01$	$\sigma_X = 4.90$	$\sigma_Y = 2.45$	
FHN_Y	$b = 2.55$	$a_1 = 0.63$	$a_3 = 2.71$	$c = 0.22$	$\sigma_X = 4.80$	$\sigma_Y = 11.08$	$\beta = -0.67$

period do not inherit any features of the deterministic system. For the FHN model, Fig. 4.10b shows that the parameters b and β no longer approach their boundaries of 0 and $\pi/2$, respectively. As a result, the FHN_Y model posterior samples contain 79% mono-stable, 17% oscillatory and 4% bi-stable parameter regimes. Thus, the excitable regime is the most prevalent. It does not seem to matter, whether the single fixed point in the mono-stable samples is in the 'warm' or 'cold' state, as they are roughly equally distributed among the ensemble. Furthermore, as for the VDP model, the parameter a_1 tends to very small values.

To get an idea of the maximum likelihood parameters of our models and to show representative time series, we estimate the parameter sample which lies in the highest density region of the posterior distribution. This is done via Gaussian Kernel smoothing, where the Kernel width is chosen manually. Although the method is typically robust over a wide range of Kernel widths, the result still has to be taken with care because of the relative sparseness of the posterior samples in parameter space. This is especially true if parameter samples tend to accumulate at the edges of their valid domain, as is often the case in our study. The resulting parameter estimates are given in Tab. 4.3 and model realizations are shown together with the data in Fig. 4.11.

4.4.2.2 Model comparison

As detailed in Sec. 4.3.3, the ratio of acceptance rates in ABC-PMC runs of two models at a given tolerance gives our approximation of the Bayes factor $\mathcal{B}_{1,2}$. The results are summarized in the Tab. 4.4. As can be seen in the table's first column, the DW model is slightly preferred over the oscillator models without noise in the y variable, while the converse is true as we add noise to both variables. Thus, the performance of the oscillators clearly improves by adding noise also to the y-variable, which is reflected by Bayes factors of 6.27 and 4.83 for the VDP_Y over VDP and FHN_Y over FHN models, respectively. Comparing the two oscillator models with and without noise in the y variable, we observe that in both cases the FHN model is very slightly preferred with Bayes factors of 1.24 and 1.61, respectively. As a result, the

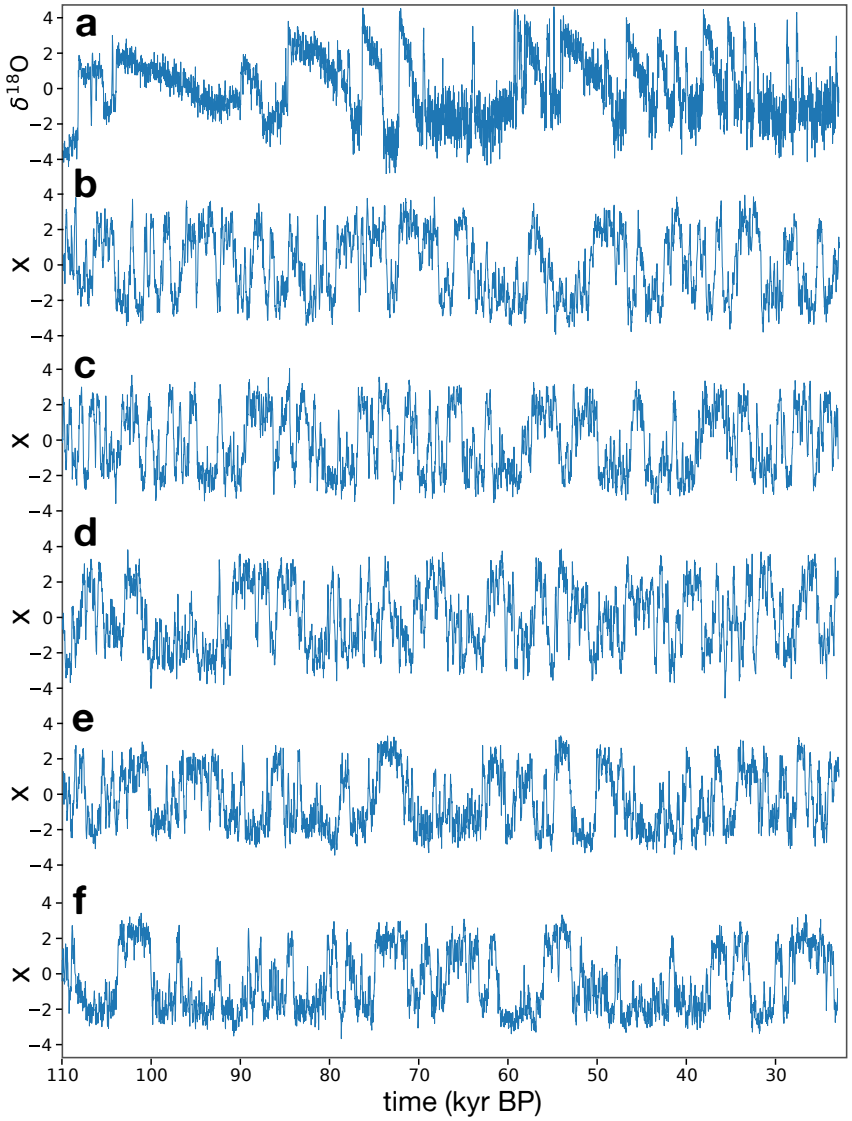

Fig. 4.11.: a) NGRIP data set used as basis of our study. b)-f) Model realizations of all models considered in the study with highest posterior probability parameters from Tab. 4.3. Panels b), c), d), e) and f) correspond to the DW, VDP, FHN, VDP_Y and FHN_Y models, respectively.

model that is most supported by the data in terms of the summary statistics chosen by us is the FHN_Y model with additive noise in both variables.

4.4.2.3 Time reversal asymmetry

We now address the characteristic saw-tooth shape of the DO events, which is not accounted for by the summary statistics used in the model comparison experiments of this study so far. On average, the NGRIP record rises much faster to high values during warming periods as it falls to low values during cooling periods. This feature

Tab. 4.4.: Bayes factors obtained from the ABC-PMC experiment with NGRIP data. The rows and columns are organized such that the value in column i and row j is the Bayes factor $\mathcal{B}_{ij}$ of model i in favor over model j, as defined in Eq. 4.4.

$\mathcal{B}_{ij}$ (j \ i)	**DW**	**VDP**	**FHN**	**VDP**$_Y$	**FHN**$_Y$
DW	-	0.26	0.42	1.62	2.01
VDP	3.87	-	1.61	6.27	7.78
FHN	2.41	0.62	-	3.90	4.83
VDP$_Y$	0.62	0.16	0.26	-	1.24
FHN$_Y$	0.50	0.13	0.21	0.81	-

is often referred to as time-reversal asymmetry and can be measured in a time series $x(t)$ by the skewed difference statistic

$$M(\tau) = \frac{\langle [x(t) - x(t+\tau)]^3 \rangle}{\langle [x(t) - x(t+\tau)]^2 \rangle}, \tag{4.8}$$

where $\langle \cdot \rangle$ denotes the expectation value over the time series and τ corresponds to a characteristic time scale (see e.g. [The+92]). A similar indicator has been used before to analyze the results of the model comparison study by [Kwa13]. In contrast to the DW model, both VDP and FHN models can in principle show such time reversal asymmetry, in a regime of relaxation oscillations. Due to similarity in shape of the DO events and relaxation oscillations, the latter are often invoked as plausible dynamical mechanism.

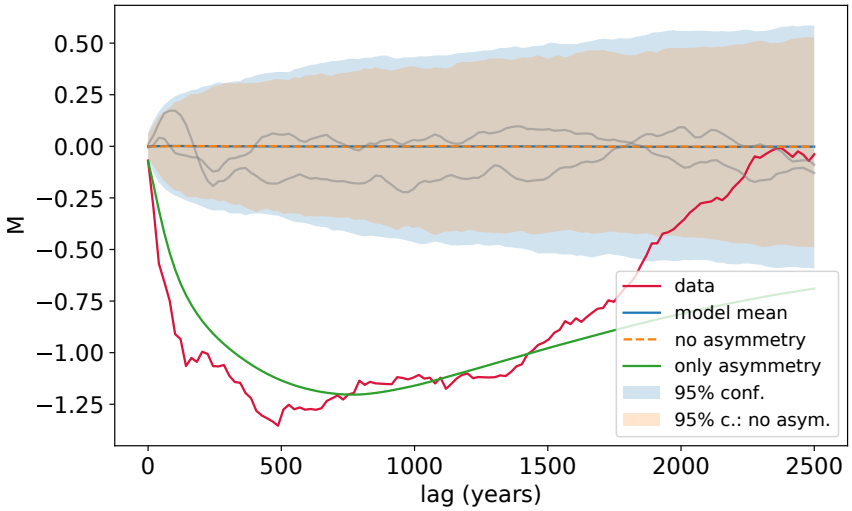

Fig. 4.12.: Time-reversal asymmetry statistic $M(\tau)$ for the NGRIP data (*red*) and model averages over the posterior samples from ABC-PMC runs with following summary statistics: 1. $s_{1,2,3,4,5,6}$ (*blue*). 2. $s_{1,2,3,4,5}$ (*orange dashed*). 3. s_6 and standard deviation (*green*). For the former two runs, 95% simultaneous confidence bands are shown. Two example realizations of the first run are shown in *gray*.

In order to test whether the oscillator models can show asymmetry behavior similar to the NGRIP time series, we include the RMSD of $M(\tau)$ up to a lag of $\tau = 2500$ years for model output and data as an additional summary statistic s_6. The RMSD of the data curve $M(\tau)$ to a straight line, i.e., a model with no asymmetry, is 0.92, which serves as a baseline for our asymmetry summary statistic s_6 and respective tolerance ϵ_6. We restrict our analysis to the FHN_Y model since it has the richest dynamics.

In Fig. 4.12 we compare $M(\tau)$ of data, and FHN_Y posterior samples of different ABC runs. For illustrative purposes, we conducted a ABC-PMC run that only used s_6 and the standard deviation as summary statistics. The average model statistics for posterior samples obtained from this run are shown as a green line in the figure and demonstrate that the FHN_Y model has a dynamical regime with asymmetry of the desired magnitude. Next, we performed a ABC-PMC run with all six summaries $s_{1,2,3,4,5,6}$. We gradually decreased the tolerance of s_6 from $\epsilon_6 = 0.8$ to $\epsilon_6 = 0.575$, while lowering the other tolerances to the rather moderate values of $\epsilon_{1,2,3} = 0.225$ and $\epsilon_{4,5} = 0.15$. At this point it becomes computationally very expensive to continue with lower tolerances, mirroring the fact that the FHN_Y model cannot both display time reversal asymmetry and the statistical behavior discussed earlier in this work. The summaries $s_{1,2,3,4,5}$ force the oscillator model into a regime, where it can only show asymmetry throughout a whole realization by chance, which is very rare. From the figure we can see that on average, the posterior samples of the run that included s_6 show no asymmetry. The same holds for the posterior samples inferred from the ABC-PMC run without s_6. The posterior samples with s_6 are also only marginally more likely to show significant asymmetry compared to the ones without s_6, as can be seen from the confidence bands.

4.5 Discussion and Conclusion

This study presents Bayesian model comparison experiments of stochastic dynamical systems given the NGRIP δ^{18}O record of the last glacial period, and aims to further the knowledge on which dynamical mechanism underlies DO events. The highly stochastic nature of these climate changes, as well as of the underlying data set prompted us to base this model comparison solely on statistical properties of the time series, captured by summary statistics. This approach is different from previous model comparison studies concerned with Greenland ice core data and stochastic dynamical systems [Kwa13; Kru+15; MC17; Boe+17]. Even though these studies also aim to compare different models in terms of their statistical properties, such as stationary densities and mean waiting times, they first estimate maximum-likelihood parameters from the 1-step prediction error with various techniques and subse-

quently use the Bayesian Information Criterion for model selection. Afterwards, they qualitatively compare the statistical properties of the best fit models. However, it is unclear how the statistical properties of the models emerge in the fitting procedure. As a consequence, there might arise a mismatch in between the models or parameter regimes preferred by the model comparison procedure and by qualitative analysis of the statistical properties, as reported by [Boe+17]. This motivates us to base the entire parameter inference and model comparison on summary statistics. Additionally, our approach is different in that we are able to show full parameter posterior distributions, which allows the assessment of parameter sensitivity and uncertainty. This becomes especially important in models with physically motivated parameters.

As prerequisite result, using synthetic data, we demonstrated in Sec. 4.4.1 how parameter inference and model comparison can be successfully done with a set of summary statistics and the ensemble of models considered. For this purpose, we adopted a version of ABC to our needs, and showed that given a model realization from within the model ensemble, the true model and its parameters can be inferred in a robust way. Furthermore, we estimated the penalty on the Bayes factor arising if one model of our ensemble has a superfluous parameter. In the case of two models being both correct, we yield $\mathcal{B} \approx 2$ in favor of the model without the superfluous parameter, which we consider to be only a small penalty.

We subsequently applied the model comparison framework to the NGRIP data set, mainly aiming to establish evidence in favor or against the DW model over one or both of the oscillator models. We found that the results depend on whether one includes noise only in the observed x variable of the oscillators, or in both. There is evidence that the DW model is better supported by the data than the oscillator models without noise in the y variable. Our estimate of the Bayes factor in favor of the DW model over the VDP and FHN models is 3.87 and 2.41, respectively. By looking at the posterior parameter distributions, we can see that the oscillator models in fact operate in regimes where they approximate dynamics similar to the DW model. Specifically, the VDP oscillator dynamics can be characterized by deterministic oscillations with very long residences in either of the two branches of the slow manifold, which are however abandoned prematurely by a stochastic jump to the other branch. The FHN model, on the other hand, operates in a bi-stable regime, where transitions are noise-induced. As a consequence, we believe that in the case of $\sigma_Y = 0$ there is a large contribution to the Bayes factors by a penalty on the additional parameters of the oscillator models. Even though the Bayes factors are not very high, we can thus conclude that the double well potential paradigm is clearly favored over oscillator models with additive noise only in the x variable.

As we add noise to the y variable of the oscillator models, we saw clear improvement over the case with $\sigma_Y = 0$, as inferred from Bayes factors of 6.27 and 4.83 for the VDP and FHN model, respectively. We inferred from the posterior parameter distributions that the oscillator models now operate in dynamical regimes different from the case $\sigma_Y = 0$. While the VDP model still is in an oscillatory regime, albeit with different properties, the FHN model now prefers an excitable regime with one fixed point either in a 'warm' or 'cold' state. From the Bayes factors of 1.62 and 2.01 we now find slight evidence in favor of the VDP_Y and FHN_Y models over the DW model. As a consequence, our results agree in principle with previous model comparison studies that also compare a DW potential model with a VDP oscillator including additive noise in both variables [Kwa13; MC17]. However, while these studies find quantitatively overwhelming evidence in favor of the oscillator, we only find very mild evidence. To complement our quantitative analysis via Bayes factors, one can qualitatively observe the models' statistical properties underlying our summary statistics. We show this in the Electronic supplementary material for both best fit parameter estimates and posterior parameter ensemble averages. We conclude that none of the models can fit all statistics at the same time in a robust way. Furthermore, the different models don't fit the individual statistics equally well. Although there might be a slight overall advantage for the FHN_Y model, our analysis does not suggest that either one of the DW, VDP_Y and FHN_Y models is much worse than the others in describing the statistical properties of the record.

Finally, we considered an additional summary statistic in our Bayesian model comparison experiment, which measures the time reversal asymmetry of a time series and captures the characteristic saw-tooth shape of the DO events in the NGRIP record. We used this additional statistic in a ABC-PMC run on the FHN model, which can show time reversal asymmetry in a regime of relaxation oscillations. We observe that it is not possible to yield time reversal asymmetry comparable to the data in the model when also obeying constraints posed by the other statistical properties $s_{1,2,3,4,5}$, in particular the temporal irregularity of events captured by the long-tailed distributions of waiting times. This is consistent with the results of the study by [Kwa13], where it is observed that the best-fit VDP model inferred from the NGRIP data also does not show time-reversal asymmetry. We thus conclude that the time-asymmetry of the record cannot be explained by chance. It is a real feature of the data, which is not captured by the simple class of models investigated here. More complex models are necessary, such as models including time delays, which were shown to yield time-reversal asymmetry to a certain degree when inferred from the NGRIP data [Boe+17].

Our study does not address external forcing directly, since we use summary statistics based on stationary properties only. This can however readily be done by including

summary statistics of time-varying properties in the data, such as the summary statistics used in [LD18]. Even though there is evidence for a contribution of external modulation to the statistical properties in the record [MC17; LD18], we still find it useful to analyze a class of models that can approximate the observed statistics without a forced modulation of parameters. We believe that the observed statistical properties are largely due to stationary variability and not external modulation.

In conclusion, this study investigates the ability of a class of models to explain the statistical properties of the glacial climate. This class of models incorporates different dynamical paradigms, which can be interpolated by continuous changes of parameters. We conducted model comparison experiments using only key statistical properties of the data. Although we find that relaxation oscillator models with noise in both variables have a slight advantage over stochastic motion in a double well potential, the Bayes factors are not very conclusive. None of the models can accurately fit all data statistics and all models have to rely heavily on chance for a realization to fit closely. This means that the dynamics inferred from the glacial climate record must be noise-dominated and the deterministic backbone is less well-defined. As a result, different deterministic regimes from the spectrum in between double well potential and relaxation oscillations can be equally consistent with the data.

4.A ABC-PMC algorithm

In this appendix we present our adaption of the ABC-PMC algorithm first presented in [Bea+09]. It is an iterative procedure over subsequent populations t of N parameter samples θ_t^j, called particles in the following. Each population is weighted by importance sampling weights w_t^j, which are the likelihood ratios of the prior parameter distribution $p(\theta_t^j)$ and the proposal distribution. The proposal distribution is a perturbation of the previous population by a Gaussian Kernel $K_t(\cdot|\theta)$, whose Kernel width adapts after every population t. As discussed in Sec. 4.3, $\underline{s}(D', D)$ is a vector of summary statistics, and $\underline{\epsilon}_t$ is a vector of tolerances, whose entries decrease for increasing population t.

1. Set population indicator $t = 0$

2. Set particle indicator $j = 1$

3. If $t = 0$ sample θ^{**} from $p(\cdot)$.
   If $t > 0$ sample θ^* from previous population with weights $\{w_{t-1}^j\}$ and perturb

particle to obtain $\theta^{**} \sim K_t(\cdot|\theta^*)$, where K_t is a Gaussian Kernel with covariance Σ_{t-1}.
If $p(\theta^{**}) = 0$ return to 3.
Simulate data D' from $p(\cdot|\theta^{**})$.
If $\underline{s}(D', D) > \underline{\epsilon}_t$ return to 3.

4. Set $\theta_t^j = \theta^{**}$ and calculate the particle weight $w_t^j = \frac{p(\theta_t^j)}{\sum_{i=1}^N w_{t-1}^i K_t(\theta_{t-1}^i)|\theta_t^j))}$
If $j < N$, set $j = j + 1$ and go to 3.

5. Normalize weights and set Σ_t to twice the covariance of $\{\theta_t^j\}$
If $t < T$ set $t = t + 1$ and go to 2.

5 Conclusion

This thesis investigates the causes and mechanisms that underlie DO events by statistical analysis and modeling of Greenland ice core records. With a combination of time series analysis, statistical hypothesis testing and Bayesian model comparison, we gather evidence in order to distinguish in between different hypotheses concerning DO events. In the following, the main findings are summarized.

1. The stadial and interstadial periods have different statistical properties. Whereas the stadial durations are very well described by an exponential distribution, the interstadials are consistent with a log-normal or inverse Gaussian distribution.

2. The interstadial durations are strongly correlated with the respective cooling rates. In fact, the cooling rates anticipate the durations. This is because the interstadial coolings are roughly linear and the variability of the cooling rates is higher than the variability of the levels at which the interstadials terminate. We show that within 150-350 years after the interstadial onset, the interstadial durations become well determined. This does not hold for every single interstadial, but for most.

3. With statistical hypothesis tests based on Poisson processes we show that the observed variability in the timings of the DO warming and cooling event sequence is too high to have been the result of two stationary, independent random processes that trigger warming and cooling transition.

4. There is evidence for external modulation of the properties of DO events, and the influencing factors are different for stadial and interstadial durations. The stadials are controlled by insolation, and the interstadials by global ice volume.

5. External forcing is, however, not enough to explain all variability, but just an underlying trend. When considering the stadial durations, we still find exponentially distributed variability, as for a memoryless process, after a removal of the insolation component from the data.

6. Based on key statistical properties of the NGRIP δ^{18}O record, we compare stochastic dynamical systems that can display bi-stable, excitable and oscillatory dynamics. In order to reproduce the data's statistics, all models have

to be driven by high intensity noise, which dominates over the specific deterministic dynamics of the individual models. As a result, double well potential, relaxation oscillator and excitable models may all be equally consistent with the data.

7. The saw-tooth shape of DO events is a robust aspect of the data, and indicates some determinism. We find that simple relaxation oscillator models driven by additive noise cannot display this behavior when they are also constrained to fulfill the more stochastic aspects in the data, such as waiting and residence time distributions.

We now discuss these findings in terms of the main scientific questions we have posed in the introduction and compare them with related, previously published results.

1. *Can the high temporal variability of the properties of DO events, such as residence times in warm and cold climate states, be expected due to chance by randomly occurring events?*

There is a visible trend in the properties, as captured by the time-varying indicators defined in Paper 2 (Sec. 3.2). Specifically, the frequency of DO warming events increases from 2-4 events per 20 kyr at the beginning of the last glacial to up to 12 events per 20 kyr around 50 kyr BP, after which there is a decline to around 5 events per 20 kyr towards the end of the last glacial. At the same time, the portion of stadial compared to interstadial periods changes from around 20% stadials to 80% stadials. We find that the long-term variations in DO warming event frequency is consistent with a memory-less stationary random process. We thus cannot exclude the possibility that the long-term variations have occurred by chance. If we however divide a DO cycle into two independent processes governing warming and cooling, this is not true anymore and significant time-varying structure is detected, which might be due to external forcing. Since our statistical analysis in Paper 1 (Sec. 2.3.4) suggests that the processes of warming and cooling transitions are indeed independent, the latter result seems more relevant. On the other hand, if one allows for processes that give rise to super-exponential variability, stationarity might still be consistent with the data. Although we have not tested it explicitly, we cannot exclude this possibility, given that the interstadial durations are consistent with a log-normal distribution, as reported in Paper 1 (Sec. 2.3.4.1). Still, an externally forced scenario seems likely as discussed in the following paragraph.

2. *Are the DO cycles, and more generally the millennial-scale glacial climate variability, modulated by external forcing, such as insolation?*

Following up on the evidence for non-stationarity, in Paper 2 (Sec. 3.2) we propose a model that incorporates long-term variations through linear forcing of the parameters with external climate factors. We find good agreement with the data in a model where the average duration of interstadial phases of the DO cycle are controlled by global ice volume and the stadial phases by boreal summer insolation. Similarly, we find that several DO event features derived from the piecewise-linear fit in Paper 1 are correlated with external forcing (Sec. 2.3.4). For stadial and interstadial durations, the findings are very similar to the model proposed in Paper 2. Importantly, there are always outliers and the relationships of forcings and durations are very non-linear, which leaves a lot of variability unexplained even after accounting for external forcing.

Time variations of millennial climate activity, i.e. the occurrence frequency of DO events, have been noted before. Through analysis of both ocean sediment cores and Antarctic ice cores it has been established that the occurrence of millennial-scale events is most frequent for intermediate glacial climate, and less frequent for warmer climate and fully glacial climate, as measured by global ice volume [McM+99] or mean Antarctic temperature [Kaw+17]. The model presented in Paper 2, which takes into account both stadial and interstadial periods with individual ice volume and insolation forcings, can reproduce this behavior for the last glacial period. Additionally, a forcing of the individual interstadial durations by Antarctic temperature [BS15] and global ice volume [Sch02b] has been reported. As discussed more in detail in Paper 1 (Sec. 2.3.4.1), these results are not fully robust when also including the shortest DO events into the analysis, especially in the older part of the last glacial. Based on our analysis, we thus propose either a) a control by external forcing of the mean stadial and interstadial durations with a very large variability around the mean (as modeled in Paper 1), or b) a more direct forcing of the majority events with less variability, but with some outlier events that are independent from the forcing, and which might be triggered via a different mechanisms.

As a side note, we mention that the expression *external forcing* is used rather loosely in this thesis. It can indicate both a time varying quantity that is truly external to the climate system, such as boreal summer insolation or orbital parameters, or another paleoclimatic record that is representative of large scale climate on long time scales obtained from a data source that is independent from the Greenland isotope record, such as global ice volume records.

3. *Can we infer from the data whether the DO cycles are a result of noise-induced jumps in between two meta-stable states, or of noisy, self-sustained oscillations?*

Before discussing this question, note that our studies do not directly address periodicity of the events. It is clear from the record that the oscillations have no simple periodicity, if one regards the whole glacial. The distributions of waiting times in between events and residence times in cold and warm states are long-tailed, as discussed in Paper 1 (Sec. 2.3.4). Furthermore, we have shown that the variability in these features cannot be fully explained when accounting for external forcing driving the instantaneous period of the DO cycles. This indicates that there is a strong stochastic component to the occurrences of DO events. We also do not discuss the hypothesis that events are triggered by integer multiples of a fundamental period [All+01; Sch02a], because we think it has been sufficiently covered in the literature [Dit+07; PF10]. Even though, when allowing for a vast range of these multiples, we cannot exclude that some results in this thesis might be consistent with this mechanism, there is generally not enough evidence in the data in favor of this hypothesis [Dit+07]. Instead, we compare different noise-driven dynamical systems to the statistical properties of the data. In Paper 3, a model comparison via Approximate Bayesian Computation shows that one cannot conclusively decide whether the underlying deterministic dynamical regime is that of a relaxation oscillator, an excitable system or a bi-stable, one-dimensional potential system (Sec. 4.4.2). There is a small advantage of both the relaxation oscillator and the excitable system over the potential model, which we attribute mostly due to the possibility of unobserved dynamical noise in the hidden variable. Our ongoing work shows that using additional observational noise for a double well potential model significantly improves the model's performance (not presented in the thesis). From the perspective of this low-dimensional model description, we conclude that the stochastic influences dominate and it is not so crucial where the underlying deterministic dynamics lie in the continuum in between the dynamical regimes outlined above.

Note that our interpretation of this model comparison is different from previous studies [Kwa13; MC17], which report strong evidence in favor of relaxations oscillator models. A main difference is their use of the time series increments for a one-step prediction likelihood, whereas we use statistical properties that do not depend much on these increments, which we justify by the difficulty in interpreting the highest frequencies of the record. What generally motivates the use of relaxation oscillator models is their ability to show saw-tooth shaped oscillations, which break time reversal symmetry, and are similar to the characteristic shape of DO events. However, we show that the models cannot display significant time reversal asymmetry if they also have to obey the constraints of the other statistical properties of the record, such as the temporal irregularity captured by the distributions of residence and waiting times.

4. *Is there evidence in the statistical properties of the data in favor of certain physical mechanisms to explain DO events?*

In Paper 1 (Sec. 2.3.4) we find that both the stadial durations and the durations of the rapid warming transitions are consistent with what is expected for noise-induced escapes from a meta-stable state. This suggests that mechanisms with a random trigger could underlie the transitions, such as those of studies that report noise-induced, unforced climate transitions in detailed climate models [Dri+13; Kle+15]. This is different for the interstadial periods, whose linear cooling indicate more deterministic dynamics, given by a cooling time scale, which itself has a very large variability. Thus, the cooling transitions at the end of interstadials don't seem to be compatible with a random and memoryless trigger. However, what kind of process could set the highly variable cooling rates is unclear.

The different forcing influences on interstadial and stadial durations (Paper 1 and 2) let us speculate that there are different mechanisms governing the two transitions (warming and cooling), where the specific form of the external forcings can give hints into the mechanisms. As one potential mechanism that seems consistent with the forcing influence we inferred, we hypothesize a modulation of stadial durations by insolation via the ice-ocean albedo feedback (Sec. 3.4). Global ice volume control on interstadial duration might be explained by different influences on the strength and stability of the interstadial (strong) mode of the Atlantic Meridional Overturning Circulation (AMOC). However, contrary to our finding, the influence of ice volume on AMOC stability reported by studies with globally coupled models is often such that increases of Northern Hemisphere ice sheets actually enhance the stability of the strong AMOC state, which would intuitively result in longer interstadials [Zha+14]. This has been addressed in [BS15], where Southern Ocean processes are invoked to control interstadial durations. The fact that there are distinct outlier events which do not follow the trend by external forcing that is present in the majority of events might suggest that there are types of events that are triggered by a different mechanism.

As a final note, we remark that from our model comparison study in Paper 3 we cannot exclude the possibility of an underlying self-sustained oscillation, such as reported by Peltier and Vettoretti [PV14]. However, our analysis suggests that such an oscillation needs to be heavily influenced by noise, in order to fulfill the observed temporal irregularity.

6 Outlook

We conclude with some brief thoughts on possible directions to use and extend the results obtained in this thesis. In the model comparison study in Paper 3, it was clear that none of the models could accurately reproduce all the statistical features of the data in a reliable way. Thus, better stochastic dynamic models are needed, which fulfill both stochastic and deterministic aspect of DO events. One direction that we currently pursue is the inclusion of other types of noise, such as non-Gaussian and multiplicative noise, or a combination of both. Initial tests lead us to believe that this permits the use of additive noise with only small amplitude and as a consequence can preserve deterministic dynamical features, while still producing an irregular sequence of events.

Based on the forcing influences we inferred for stadial and interstadial durations, it would be interesting to check whether detailed climate models that produce a DO-like cycle show a response to variations in boundary conditions that is compatible with our findings. Similarly, by combining all statistical properties of DO events that we have collected in the three manuscripts, one could look for a conceptual, low-order box model, which can fulfill as many constraints set by our findings as possible. Such a conceptual model should furthermore be extended to qualitatively reproduce the bipolar dynamics seen by comparison of Greenland and Antarctic ice cores, which was not a focus of this thesis.

Concerning the statistical analysis in Paper 1, it seems straightforward to go into more depth with some of the aspects covered, such as the interstadial cooling rates, and construct simple dynamical models to test against the data. Additionally, the data analysis presented in the paper is not exhaustive. Some features have not been analyzed extensively, the nature of outlier events should be investigated, and in general more advanced statistical methods could be tried to analyze the data generated by the piecewise-linear fit.

Finally, in addition to the dynamical paradigms investigated in this thesis, new mechanisms should be considered. One such mechanism that we currently investigate concerns rate-induced tipping. The fact that the simple Stommel box model actually supports this phenomenon could hint at a potential relevance in relation to DO events. Rate-dependent transitions of the AMOC have also been observed in a detailed climate model [SS97]. In conclusion, we have gathered some evidence to give partial answers to the scientific questions that we posed in the introduction.

However, it is clear that in order to get more definite answers more work is needed, not least in order to overcome the issue of the limited number of DO events available to conduct statistical tests. A satisfactory reproduction of the glacial climate record with tractable stochastic dynamical systems models remains a challenge.

A

Supplemental material to 'Objective extraction and analysis of statistical features of Dansgaard-Oeschger events'

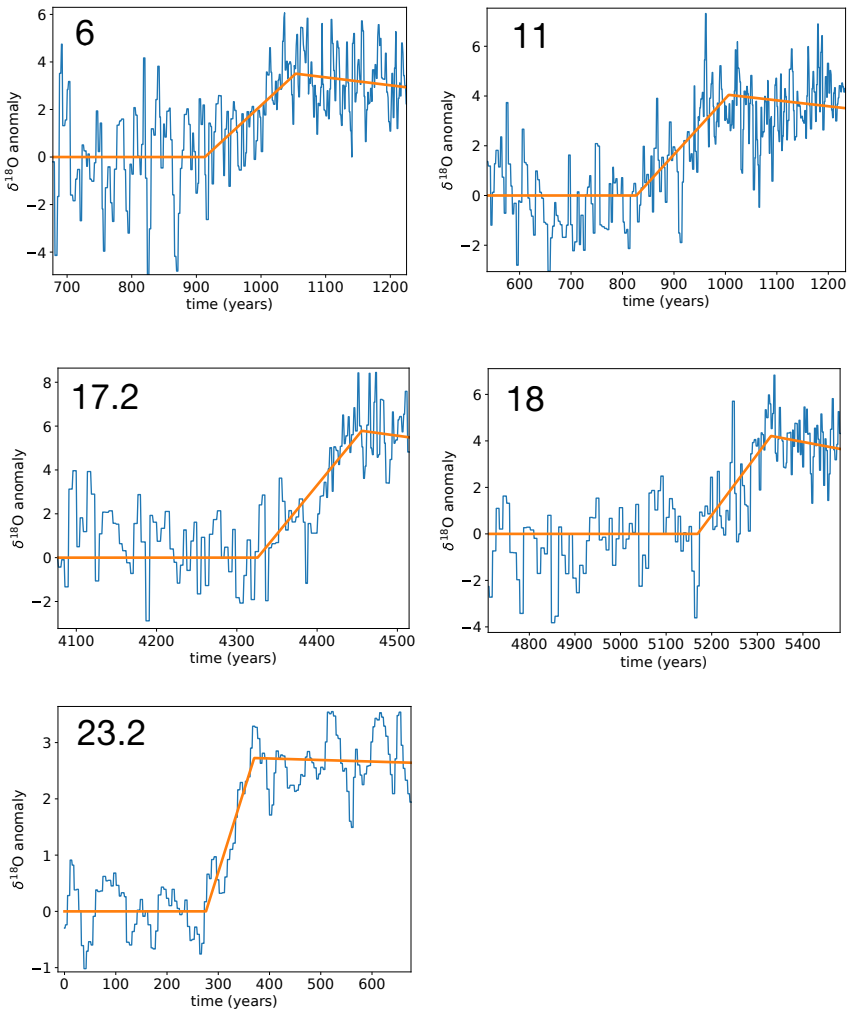

Fig. A.S1.: Zoom into time series and piecewise-linear fit of the 5 longest rapid warmings leading up to interstadials 6 (warming of 140.4 years), 11 (179.5 years), 17.2 (129.7 years), 18 (161.0 years) and 23.2 (115.2 years).

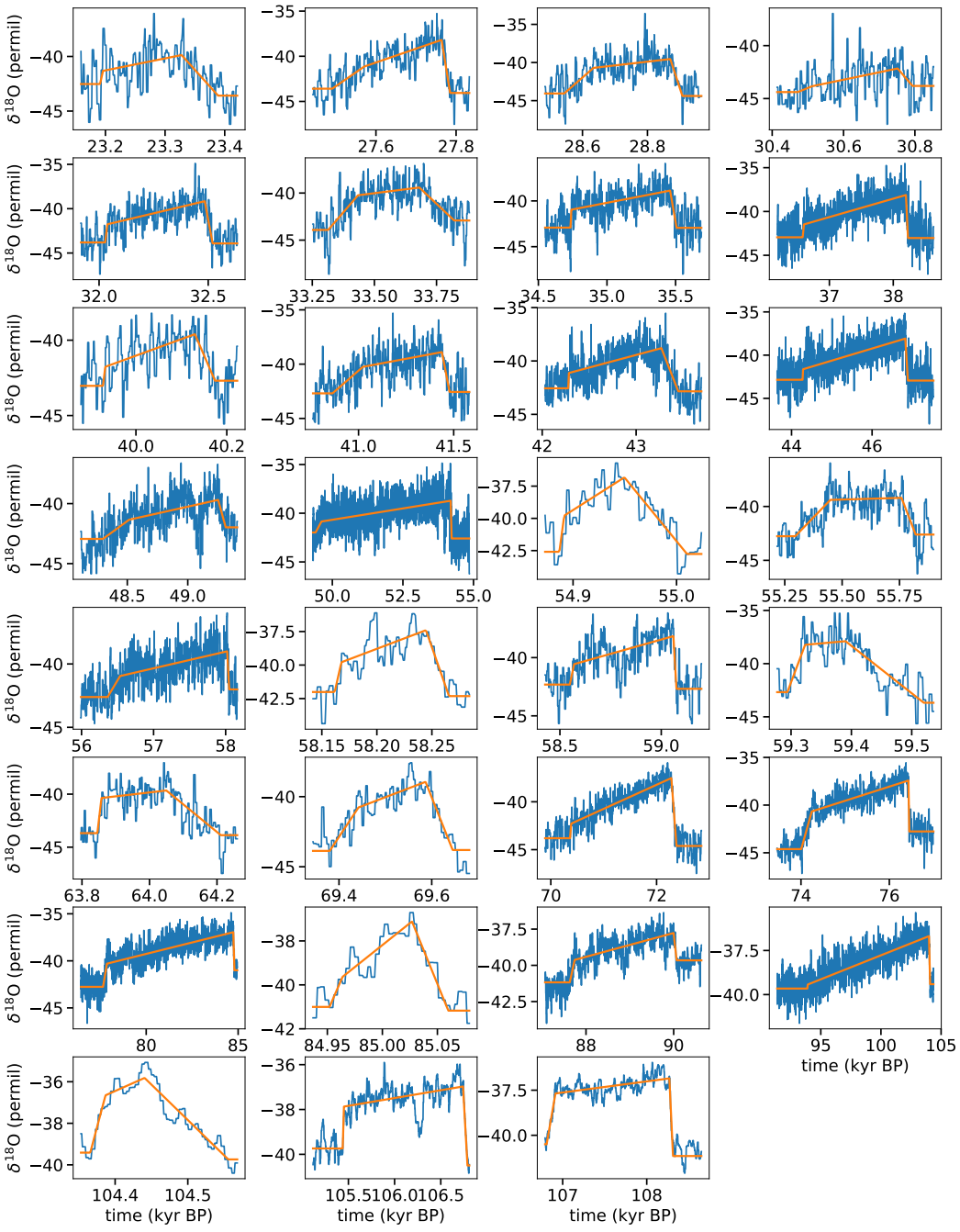

Fig. A.S2.: All interstadial transitions fitted with a saw-tooth shape by our algorithm. The panels start with GI-2.2 on the top left and end with GI-24.2 on the bottom right.

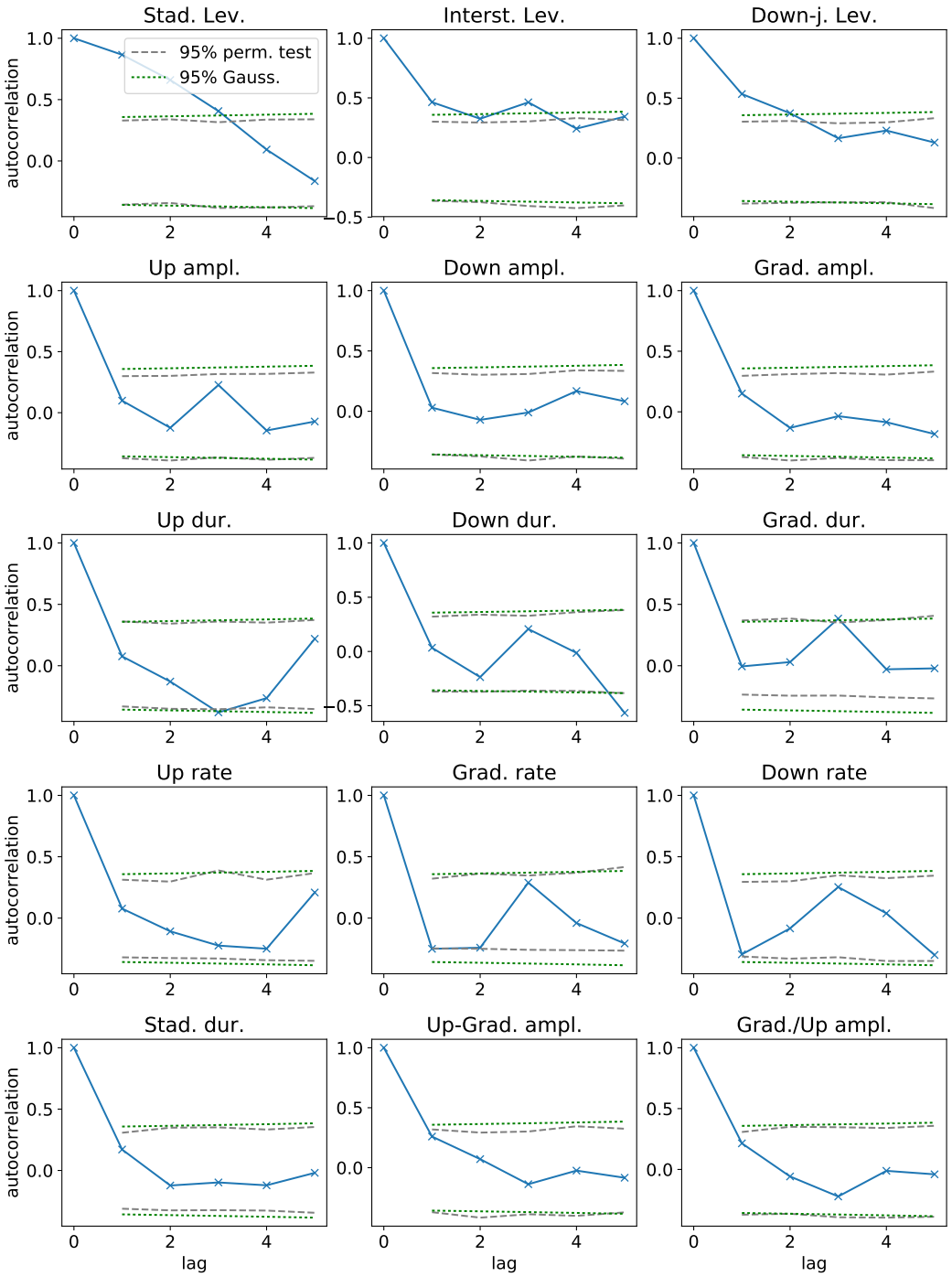

Fig. A.S3.: Autocorrelation functions of all 15 features considered in the study, which are defined in the Methods and Materials section of the main article. We give 95% confidence bands, outside of which an autocorrelation is considered significant. The theoretical confidence band assuming Gaussianity and large sample size is shown with a *dotted* line, while a confidence band obtained from the respective samples and a permutation test is shown with a *dashed* line.

B

Supplemental material to 'A consistent model selection analysis of abrupt glacial climate changes'

In this supplement we give a qualitative overview of the statistical properties of the three models that are best supported by the data, i.e., the DW, VDP_Y and FHN_Y models. In Fig. B.S1, we show the statistics obtained under a single parameter value for each model, which is given by our estimate of the largest posterior probability parameter. Figure B.S2 instead shows statistics averaged over the full posterior parameter ensemble. Note that in both figures there is a bend for higher values in the black lines for $s_{1,2,3}$, i.e., the model averages of the empirical cumulative distribution functions, which is an artifact due to the finite size of each realization. In both figures, the average statistics show the same characteristics and yield the following qualitative picture. It is seen that none of the models can robustly satisfy all statistics of the data. For the autocorrelation (s_5) there is a clear advantage of the FHN_Y over the DW and VDP_Y models, whereas for the stationary density (s_4) DW performs better than VDP_Y and FHN_Y. We do not see a clear advantage of any model for the stadial durations (s_1), whereas for the waiting times (s_3) there might be a slight advantage for the oscillators, since they seem to be closer to an exponential distribution. The same holds for the interstadial durations, where the oscillators show a slightly super-exponential distribution, which is advantageous in terms of the data.

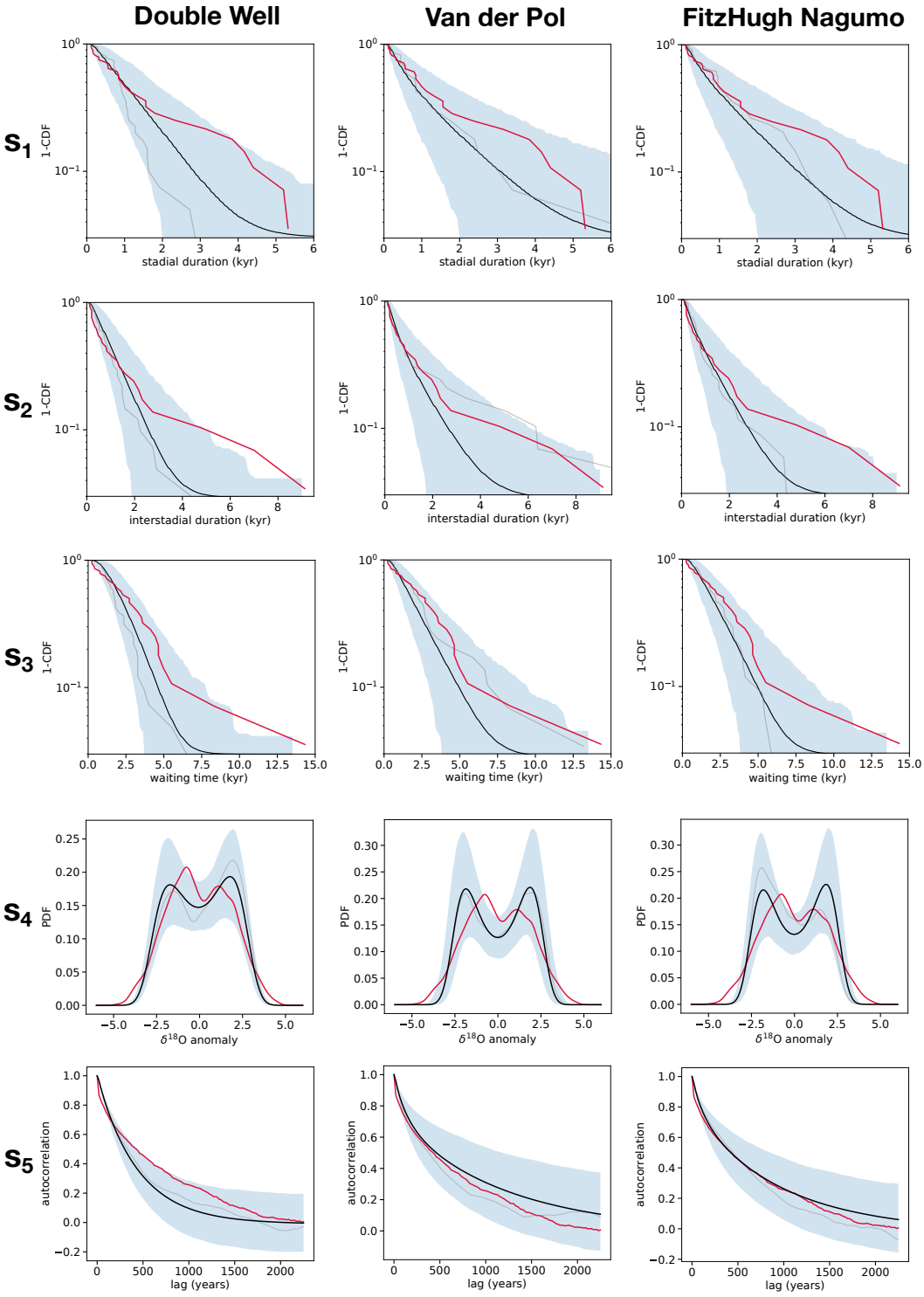

Fig. B.S1.: Statistical properties underlying the summary statistics $s_{1,2,3,4,5}$ for the DW, VDP_Y and FHN_Y models with largest posterior probability parameters, as explained in the main text. Red lines are data statistics, black lines are the model averages over 20.000 realizations and the shading corresponds to 95% simultaneous confidence bands.

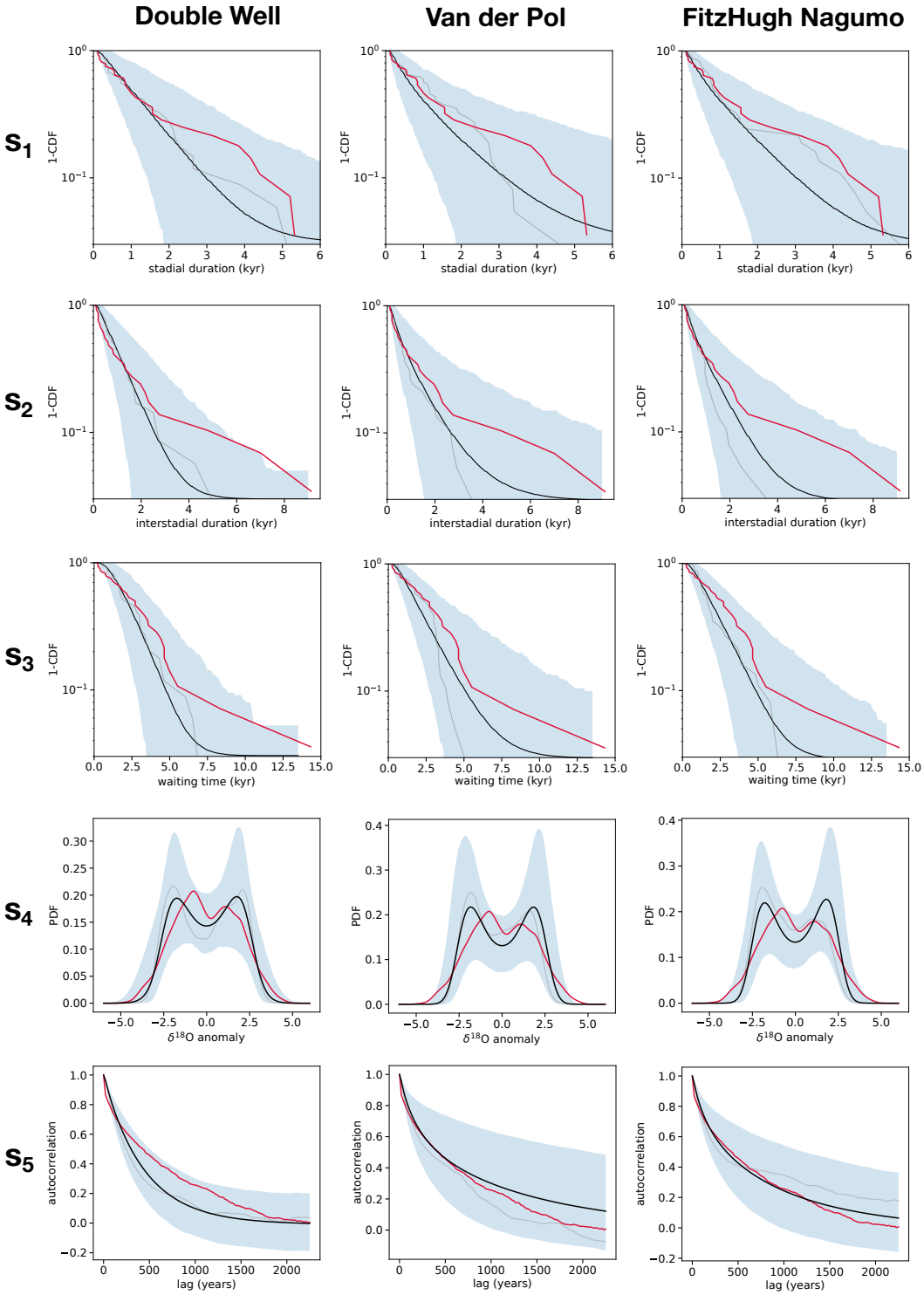

Fig. B.S2.: Statistical properties underlying the summary statistics $s_{1,2,3,4,5}$ for the posterior parameter ensembles of the DW, VDP_Y and FHN_Y models. Red lines are data statistics, black lines are the model averages and the shading corresponds to 95% simultaneous confidence bands. We generated multiple realizations with each parameter in the ensembles, such that the total number of realizations was roughly 20.000.

Bibliography

[AB08] J. Ahn and E. J. Brook. „Atmospheric CO2 and Climate on Millennial Time Scales During the Last Glacial Period“. In: *Science* 322 (2008), pp. 83–85 (cit. on pp. 3, 5).

[All+01] R. B. Alley, S. Anandakrishnan, and P. Jung. „Stochastic Resonance in the North Atlantic“. In: *Paleoceanography* 16.2 (2001), pp. 190–198 (cit. on pp. 8, 50, 94).

[Ash+12] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox. „Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system“. In: *Phil. Trans. R. Soc. A* 370 (2012), pp. 1166–1184 (cit. on p. 9).

[Bar+11] S. Barker, G. Knorr, R. L. Edwards, et al. „800,000 Years of Abrupt Climate Variability“. In: *Science* 334 (2011), pp. 347–351 (cit. on pp. 5, 11).

[Bas+17] J. N. Bassis, S. V. Petersen, and L. Mac Cathles. „Heinrich events triggered by ocean forcing and modulated by isostatic adjustment“. In: *Nature* 542 (2017), pp. 332–334 (cit. on p. 7).

[Bea+09] M. A. Beaumont, J.-M. Cornuet, J.-M. Marin, and C. P. Robert. „Adaptive approximate Bayesian computation“. In: *Biometrika* 96.4 (2009), pp. 983–990 (cit. on pp. 74, 75, 89).

[Blu+98] T. Blunier, J. Chappellaz, J. Schwander, et al. „Asynchrony of Antarctic and Greenland climate change during the last glacial period“. In: *Nature* 394 (1998), pp. 739–743 (cit. on p. 4).

[Boe+17] N. Boers, M. D. Chekroun, H. Liu, et al. „Inverse stochastic–dynamic models for high-resolution Greenland ice core records“. In: *Earth Syst. Dynam.* 8 (2017), pp. 1171–1190 (cit. on pp. 10, 86–88).

[Boe18] N. Boers. „Early-warning signals for Dansgaard-Oeschger events in a high-resolution ice core record“. In: *Nature Comm.* 9 (2018), p. 2556 (cit. on pp. 9, 42).

[Bon+93] G. Bond, W. Broecker, S. J. Johnsen, et al. „Correlations between climate records from North Atlantic sediments and Greenland ice“. In: *Nature* 365 (1993), pp. 143–147 (cit. on p. 4).

[Bra+05] H. Braun, M. Christl, S. Rahmstorf, et al. „Possible solar origin of the 1,470-year glacial climate cycle demonstrated in a coupled model“. In: *Nature* 438 (2005), pp. 208–211 (cit. on pp. 8, 50).

[Bra+07] H. Braun, A. Ganopolski, M. Christl, and D. R. Chialvo. „A simple conceptual model of abrupt glacial climate events“. In: *Nonlin. Processes Geophys.* 14 (2007), pp. 709–721 (cit. on p. 50).

[Bra+10] H. Braun, P. D. Ditlevsen, J. Kurths, and M. Mudelsee. „Limitations of red noise in analysing Dansgaard-Oeschger events“. In: *Clim. Past* 6 (2010), pp. 85–92 (cit. on pp. 8, 51).

[Bro+85] W. S. Broecker, D. M. Peteet, and D. Rind. „Does the ocean-atmosphere system have more than one stable mode of operation?“ In: *Nature* 315 (1985), pp. 21–26 (cit. on p. 6).

[Bro+90] W. S. Broecker, G. Bond, and M. Klas. „A salt oscillator in the glacial Atlantic? 1. The concept“. In: *Paleoceanography* 5.4 (1990), pp. 469–477 (cit. on p. 7).

[BS15] C. Buizert and A. Schmittner. „Southern Ocean control of glacial AMOC stability and Dansgaard-Oeschger interstadial duration“. In: *Paleoceanography* 30 (2015), pp. 1595–1612 (cit. on pp. 33, 41, 62, 93, 95).

[Cap+10] E. Capron, A. Landais, J. Chappellaz, et al. „Millennial and sub-millennial scale climatic variations recorded in polar ice cores over the last glacial period“. In: *Clim. Past* 6 (2010), pp. 345–365 (cit. on pp. 4, 14).

[Ces94] P. Cessi. „A simple box model of stochastically forced thermohaline flow“. In: *Phys. Oceanography* 24 (1994), p. 1911 (cit. on pp. 8, 68).

[Ces96] P. Cessi. „Convective Adjustment and Thermohaline Excitability“. In: *J. Phys. Oceanography* 26 (1996), pp. 481–491 (cit. on p. 9).

[Che+09] H. Cheng, R. L. Edwards, W. Broecker, et al. „Ice Age Terminations“. In: *Science* 326 (2009), pp. 248–252 (cit. on p. 5).

[Che+16] H. Cheng, R. L. Edwards, A. Sinha, et al. „The Asian monsoon over the past 640,000 years and ice age terminations“. In: *Nature* 534 (2016), pp. 640–646 (cit. on p. 5).

[Cim+13] A. A. Cimatoribus, S. Drijfhout, V. Livina, and G. van der Schrier. „Dansgaard–Oeschger events: bifurcation points in the climate system“. In: *Clim. Past* 9 (2013), pp. 323–333 (cit. on p. 9).

[Coo+15] A. Cooper, C. Turney, K. A. Hughen, et al. „Abrupt warming events drove Late Pleistocene Holarctic megafaunal turnover“. In: *Science* (2015), 10.1126/science.aac4315 (cit. on p. 4).

[Cér+11] F. Cérou, A. Guyader, T. Lelièvre, and D. Pommier. „A multiple replica approach to simulate reactive trajectories“. In: *J. Chem. Phys.* 134 (2011), p. 054108 (cit. on p. 38).

[Cér+13] F. Cérou, A. Guyader, T. Lelièvre, and F. Malrieu. „On the length of one-dimensional reactive paths“. In: *ALEA, Lar. Am. J. Probab. Math. Stat.* 10.1 (2013), pp. 359–389 (cit. on pp. 37, 41).

[Dan+73] W. Dansgaard, S. J. Johnsen, H. B. Clausen, and N. Gundestrup. „Stable Isotope Glaciology“. In: *Meddelelser om Grønland* 197.2 (1973), pp. 1–53 (cit. on p. 2).

[Dan+82] W. Dansgaard, H. B. Clausen, N. Gundestrup, et al. „A New Greenland Deep Ice Core“. In: *Science* 218.4579 (1982), pp. 1273–1277 (cit. on p. 2).

[Dan+93] W. Dansgaard et al. „Evidence for general instability of past climate from a 250-kyr ice-core record“. In: *Nature* 364 (1993), p. 218 (cit. on pp. 2, 50, 65).

[Dan64] W. Dansgaard. „Stable isotopes in precipitation“. In: *Tellus* 16.4 (1964), pp. 436–468 (cit. on p. 2).

[Day87] M. V. Day. „Recent progress on the small parameter exit problem“. In: *Stochastics* 20.2 (1987), pp. 121–150 (cit. on p. 34).

[Dit+05] P. D. Ditlevsen, M. S. Kristensen, and K. K. Andersen. „The Recurrence Time of Dansgaard–Oeschger Events and Limits on the Possible Periodic Component“. In: *J. Climate* 18 (2005), pp. 2594–2603 (cit. on pp. 8, 51).

[Dit+07] P. D. Ditlevsen, K. K. Andersen, and A. Svensson. „The DO-climate events are probably noise induced: statistical investigation of the claimed 1470 years cycle“. In: *Clim. Past* 3 (2007), pp. 129–134 (cit. on pp. 8, 51, 71, 94).

[Dit99] P. D. Ditlevsen. „Observation of α-stable noise induced millenial climate changes from an ice-core record“. In: *Geophys. Res. Lett.* 26.10 (1999), pp. 1441–1444 (cit. on pp. 8, 68).

[DJ10] P. D. Ditlevsen and S. J. Johnsen. „Tipping points: Early warning and wishful thinking“. In: *Geophys. Res. Lett.* 37 (2010), p. L19703 (cit. on p. 9).

[Dok+13] T. M. Dokken, K. H. Nisancioglu, C. Li, D. S. Battisti, and C. Kissel. „Dansgaard-Oeschger cycles: Interactions between ocean and sea ice intrinsic to the Nordic seas“. In: *Paleoceanography* 28 (2013), pp. 491–502 (cit. on pp. 7, 62).

[Dri+13] S. Drijfhout, E. Gleeson, H. A. Dijkstra, and V. Livina. „Spontaneous abrupt climate change due to an atmospheric blocking–sea-ice–ocean feedback in an unforced climate model simulation“. In: *PNAS* 110.49 (2013), pp. 19713–19718 (cit. on pp. 7, 62, 95).

[Flü+04] J. Flückiger, T. Blunier, B. Stauffer, et al. „N2O and CH4 variations during the last glacial epoch: Insight into global processes“. In: *Glob. Bioch. Cycles* 18 (2004), GB1020 (cit. on p. 5).

[Gen+03] D. Genty, D. Blamart, R. Ouahdi, et al. „Precise dating of Dansgaard–Oeschger climate oscillations in western Europe from stalagmite data“. In: *Nature* 421 (2003), pp. 833–837 (cit. on p. 4).

[Gki+14] V. Gkinis, S. B. Simonsen, S. L. Buchardt, J. W. C. White, and B. M. Vinther. „Water isotope diffusion rates from the NorthGRIP ice core for the last 16,000 years – Glaciological and paleoclimatic implications“. In: *Earth and Plan. Sc. Lett.* 405 (2014), pp. 132–141 (cit. on pp. 10, 15).

[GR02] A. Ganopolski and S. Rahmstorf. „Abrupt Glacial Climate Changes due to Stochastic Resonance“. In: *Phys. Rev. Lett.* 88.3 (2002), p. 038501 (cit. on p. 6).

[Gro+93] P. M. Grootes, M. Stuiver, J. W. C. White, S. J. Johnsen, and J. Jouzel. „Comparison of oygen isotope records from the GISP2 and GRIP Greenland ice cores“. In: *Nature* 366 (1993), pp. 552–554 (cit. on p. 2).

[GS97] P. M. Grootes and M. Stuiver. „Oxygen 18/16 variability in Greenland snow and ice with 10^{-3} to 10^{5}-year time resolution“. In: *J. Geoph. Research* 102 (1997), pp. 26,455 (cit. on pp. 8, 50).

[Huy06] P. Huybers. „Early Pleistocene Glacial Cycles and the Integrated Summer Insolation Forcing“. In: *Science* 313 (2006), pp. 508–511 (cit. on pp. 16, 55).

[Joh+72] S. J. Johnsen, W. Dansgaard, H. B. Clausen, and C. C. Langway. „Oxygen Isotope Profiles through the Antarctic and Greenland Ice Sheets“. In: *Nature* 235 (1972), pp. 429–434 (cit. on p. 2).

[Joh+97] S. J. Johnsen, H. B. Clausen, W. Dansgaard, N. Gundestrup, C. U. Hammer, et al. „The δ^{18}O record along the Greenland Ice Core Project deep ice core and the problem of possible Eemian climatic instability“. In: *J. Geoph. Research* 102.C12 (1997), pp. 26397–26410 (cit. on p. 16).

[Jou+07] J. Jouzel, V. Masson-Delmotte, O. Cattani, et al. „Orbital and Millennial Antarctic Climate Variability over the Past 800,000 Years“. In: *Science* 317 (2007), pp. 793–796 (cit. on pp. 4, 5).

[Kag+13] M. Kageyama, U. Merkel, B. Otto-Bliesner, et al. „Climatic impacts of fresh water hosing under Last Glacial Maximum conditions: a multi-model study“. In: *Clim. Past* 9 (2013), pp. 935–953 (cit. on p. 6).

[Kaw+17] K. Kawamura et al. „State dependence of climatic instability over the past 720,000 years from Antarctic ice cores and climate modeling“. In: *Sci. Adv.* 3 (2017), e1600446 (cit. on pp. 62, 93).

[Kin+14] P. Kindler, M. Guillevic, M. Baumgartner, et al. „Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core“. In: *Clim. Past* 10 (2014), pp. 887–902 (cit. on pp. 3, 4, 39).

[KL09] F. Kwasniok and G. Lohmann. „Deriving dynamical models from paleoclimatic records: Application to glacial millennial-scale climate variability“. In: *Phys. Rev. E* 80 (2009), p. 066104 (cit. on p. 71).

[Kle+15] H. Kleppin, M. Jochum, B. Otto-Bliesner, C. A. Shields, and S. Yeager. „Stochastic Atmospheric Forcing as a Cause of Greenland Climate Transitions“. In: *J. Climate* 28 (2015), pp. 7741–7763 (cit. on pp. 7, 62, 95).

[Kru+15] S. Krumscheid, M. Pradas, G. A. Pavliotis, and S. Kalliadasis. „Data-driven coarse graining in action: Modeling and prediction of complex systems“. In: *Phys. Rev. E* 92 (2015), p. 042139 (cit. on pp. 10, 86).

[Kwa13] F. Kwasniok. „Analysis and modelling of glacial climate transitions using simple dynamical systems“. In: *Phil. Trans. R. Soc. A* 371 (2013), p. 20110472 (cit. on pp. 9, 10, 68, 71, 85, 86, 88, 94).

[Las+04] J. Laskar, P. Robutel, F. Joutel, et al. „A long term numerical solution for the insolation quantities of the Earth“. In: *Astronomy and Astrophysics* (2004), pp. 261–285 (cit. on p. 16).

[LD18] J. Lohmann and P. D. Ditlevsen. „Random and externally controlled occurrences of Dansgaard–Oeschger events “. In: *Clim. Past* 14.5 (2018), pp. 609–617 (cit. on pp. 42, 89).

[Li+05] C. Li, D. S. Battisti, D. P. Schrag, and E. Tziperman. „Abrupt climate shifts in Greenland due to displacements of the sea ice edge“. In: *Geophys. Res. Lett.* 32 (2005), p. L19702 (cit. on p. 7).

[Li+10] C. Li, D. S. Battisti, and C. M. Bitz. „Can North Atlantic Sea Ice Anomalies Account for Dansgaard–Oeschger Climate Signals?“ In: *J. Climate* 23 (2010), pp. 5457–5475 (cit. on p. 7).

[Lou+08] L. Loulergue, E. Schilt, R. Spahni, et al. „Orbital and millennial-scale features of atmospheric CH4 over the past 800,000 years“. In: *Nature* 453 (2008), pp. 383–386 (cit. on pp. 3, 5).

[Mar+07] B. Martrat, J. O. Grimalt, N. J. Shackleton, et al. „Four Climate Cycles of Recurring Deep and Surface Water Destabilizations on the Iberian Margin“. In: *Science* 317 (2007), pp. 502–507 (cit. on p. 5).

[Mar+12] J.-M. Marin, P. Pudlo, C. P. Robert, and R. J. Ryder. „Approximate Bayesian computational methods“. In: *Stat. Comput.* 22 (2012), pp. 1167–1180 (cit. on p. 67).

[May+97] P. A. Mayewski, L. D. Meeker, M. S. Twickler, et al. „Major features and forcing of high-latitude northern hemisphere atmospheric circulation using a 110,000-year-long glaciochemical series“. In: *J. Geoph. Research* 102.C12 (1997), pp. 26345–26366 (cit. on p. 4).

[MC17] T. Mitsui and M. Crucifix. „Influence of external forcings on abrupt millennialscale climate changes: a statistical modelling study“. In: *Clim. Dyn.* 48 (2017), p. 2729 (cit. on pp. 9, 10, 53, 61, 68, 71, 86, 88, 89, 94).

[McM+99] J. F. McManus, D. W. Oppo, and J. L. Cullen. „A 0.5-Million-Year Record of Millennial-Scale Climate Variability in the North Atlantic“. In: *Science* 283 (1999), pp. 971–975 (cit. on pp. 38, 93).

[Mil30] M. Milancović. „Mathematische Klimalehre und astronomische Theorie der Klimaschwankungen“. In: *Bd.1: Allgemeine Klimalehre*. Ed. by W. Köppen and R. Geiger. Handbuch der Klimatologie. Berlin: Borntraeger, 1930 (cit. on p. 1).

[MS88] S. Manabe and R. J. Stouffer. „Two Stable Equilibria of a Coupled Ocean-Atmosphere Model“. In: *J. Climate* 1 (1988), pp. 841–866 (cit. on p. 6).

[Mün+16] T. Münch, S. Kipfstuhl, J. Freitag, H. Mayer, and T. Laepple. „Regional climate signal vs. local noise: a two-dimensional view of water isotopes in Antarctic firn at Kohnen Station, Dronning Maud Land“. In: *Clim. Past* 12 (2016), pp. 1565–1581 (cit. on p. 66).

[Ols+12] B. Olson, I. Hashmi, K. Molloy, and A. Shehu. „Basin Hopping as a General and Versatile Optimization Framework for the Characterization of Biological Macromolecules“. In: *Advances in Artificial Intelligence* 2012 (2012), p. 674832 (cit. on p. 18).

[Pet+13] S. V. Petersen, D. P. Schrag, and P. U. Clark. „A new mechanism for Dansgaard-Oeschger cycles“. In: *Paleoceanography* 28 (2013), pp. 24–30 (cit. on p. 7).

[PF10] D. Peavoy and C. Franzke. „Bayesian analysis of rapid climate change during the last glacial using Greenland δ^{18}O data“. In: *Clim. Past* 6 (2010), pp. 787–794 (cit. on p. 94).

[Pri+99] J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman. „Population Growth of Human Y Chromosomes: A Study of Y Chromosome Microsatellites“. In: *Mol. Biol. Evol.* 16.12 (1999), pp. 1791–8 (cit. on p. 67).

[PV14] W. R. Peltier and G. Vettoretti. „Dansgaard-Oeschger oscillations predicted in a comprehensive model of glacial climate: A “kicked” salt oscillator in the Atlantic“. In: *Geophys. Res. Lett.* 41 (2014), pp. 7306–7313 (cit. on pp. 7, 62, 95).

[Rah03] S. Rahmstorf. „Timing of abrupt climate change: A precise clock“. In: *Geophys. Res. Lett.* 30.10 (2003), p. 1510 (cit. on p. 50).

[Ras+14] S. O. Rasmussen et al. „A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy“. In: *Quat. Sc. Rev.* 106 (2014), pp. 14–28 (cit. on pp. 3, 14–16, 50, 52, 70, 71).

[RL05] M. E. Raymo and L. E. Lisiecki. „A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records“. In: *Paleoceanography* 20 (2005), PA1003 (cit. on pp. 1, 16, 55).

[Roc+00] C. Rocsoreanu, A. Georgescu, and N. Giurgiteanu. Mathematical Modelling: Theory and Applications. Volume 10. Springer, 2000 (cit. on p. 69).

[Rol+16] J. Rolland, F. Bouchet, and E. Simonnet. „Computing Transition Rates for the 1-D Stochastic Ginzburg–Landau–Allen–Cahn Equation for Finite-Amplitude Noise with a Rare Event Algorithm“. In: *J. Stat. Phys.* 162 (2016), pp. 277–311 (cit. on pp. 37, 41).

[Rou+07] D.-D. Rousseau, A. Sima, P. Antoine, et al. „Link between European and North Atlantic abrupt climate changes over the last glaciation“. In: *Geophys. Res. Lett.* 34 (2007), p. L22713 (cit. on p. 4).

[Rou+17] D.-D. Rousseau, N. Boers, A. Sima, et al. „(MIS3 & 2) millennial oscillations in Greenland dust and Eurasian aeolian records - A paleosol perspective“. In: *Quat. Sc. Rev.* 169 (2017), pp. 99–113 (cit. on pp. 4, 14, 36, 40).

[RS16a] P. Ritchie and J. Sieber. „Early-warning indicators for rate-induced tipping“. In: *Chaos* 26 (2016), p. 093116 (cit. on p. 9).

[RS16b] A. Roberts and R. Saha. „Relaxation oscillations in an idealized ocean circulation model“. In: *Clim. Dyn.* 48.7 (2016), pp. 2123–2134 (cit. on pp. 8, 68, 70).

[Ryp16] M. Rypdal. „Early-Warning Signals for the Onsets of Greenland Interstadials and the Younger Dryas–Preboreal Transition“. In: *J. Climate* 29 (2016), pp. 4047–4056 (cit. on pp. 9, 42).

[Sah15] R. Saha. „Millennial-scale oscillations between sea ice and convective deep water formation“. In: *Paleoceanography* 30 (2015), pp. 1540–1555 (cit. on p. 9).

[Sch+12] M. Scheffer, S. R. Carpenter, T. M. Lenton, et al. „Anticipating Critical Transitions“. In: *Science* 338 (2012), pp. 344–348 (cit. on p. 9).

[Sch+99] M. Schulz, W. H. Berger, M. Sarntheim, and P. M. Grootes. „Amplitude variations of 1470-year climate oscillations during the last 100,000 years linked to fluctuations of continental ice mass“. In: *Geophys. Res. Lett.* 26.22 (1999), pp. 3385–3388 (cit. on pp. 8, 38).

[Sch02a] M. Schulz. „On the 1470-year pacing of Dansgaard-Oeschger warm events“. In: *Paleoceanography* 17.2 (2002), pp. 4–1–4–9 (cit. on pp. 8, 94).

[Sch02b] M. Schulz. „The tempo of climate change during Dansgaard-Oeschger interstadials and its potential to affect the manifestation of the 1470-year climate cycle“. In: *Geophys. Res. Lett.* 29.1 (2002), p. 1002 (cit. on pp. 9, 14, 28, 32, 41, 50, 93).

[SJ03] T. F. Stocker and S. J. Johnsen. „A minimum thermodynamic model for the bipolar seesaw“. In: *Paleoceanography* 18.4 (2003), p. 1087 (cit. on p. 11).

[SP99] K. Sakai and W. R. Peltier. „A Dynamical Systems Model of the Dansgaard–Oeschger Oscillation and the Origin of the Bond Cycle“. In: *J. Climate* 12 (1999), pp. 2239–2255 (cit. on p. 9).

[SS97] T. F. Stocker and A. Schmittner. „Influence of CO2 emission rates on the stability of the thermohaline circulation“. In: *Nature* 388 (1997), pp. 862–865 (cit. on p. 97).

[Sta+84] B. Stauffer, H. Hofer, H. Oeschger, J. Schwander, and U. Siegenthaler. „Atmospheric CO2 concentration during the last glaciation“. In: *Annals of Glaciology* 5 (1984), pp. 160–164 (cit. on p. 3).

[Ste+08] J. P. Steffensen, K. K. Andersen, M. Bigler, et al. „High-Resolution Greenland Ice Core Data Show Abrupt Climate Change Happens in Few Years“. In: *Science* 321 (2008), pp. 680–684 (cit. on p. 4).

[Sto61] H. Stommel. „Thermohaline Convection with Two Stable Regimes of Flow“. In: *Tellus* 13 (1961), p. 2 (cit. on pp. 6, 8, 68).

[Sve+06] A. Svensson et al. „The Greenland Ice Core Chronology 2005, 15–42 ka. Part 2: comparison to other records“. In: *Quat. Sc. Rev.* 25 (2006), pp. 3258–3267 (cit. on pp. 10, 15, 52).

[SY93] H. M. Stommel and W. R. Young. „The Average T-S Relation of a Stochastically Forced Box Model“. In: *Phys. Oceanography* 23 (1993), pp. 151–158 (cit. on p. 68).

[The+92] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer. „Testing for nonlinearity in time series: the method of surrogate data“. In: *Physica D* 58 (1992), pp. 77–94 (cit. on p. 85).

[Tim+03] A. Timmermann, H. Gildor, M. Schulz, and E. Tziperman. „Coherent Resonant Millennial-Scale Climate Oscillations Triggered by Massive Meltwater Pulses“. In: *J. Climate* 16 (2003), pp. 2569–2585 (cit. on pp. 9, 51).

[TL00] A. Timmermann and G. Lohmann. „Noise-Induced Transitions in a Simplified Model of the Thermohaline Circulation“. In: *Phys. Oceanography* 30 (2000), pp. 1891–1900 (cit. on p. 68).

[VP16] G. Vettoretti and W. R. Peltier. „Thermohaline instability and the formation of glacial North Atlantic super polynyas at the onset of Dansgaard-Oeschger warming events“. In: *Geophys. Res. Lett.* 43 (2016), pp. 5336–5344 (cit. on p. 7).

[Wan+01] Y. J. Wang, H. Cheng, R. L. Edwards, et al. „A High-Resolution Absolute-Dated Late Pleistocene Monsoon Record from Hulu Cave, China“. In: *Science* 294 (2001), pp. 2345–2348 (cit. on p. 4).

[Wol+10] E. W. Wolff, J. Chappellaz, T. Blunier, S. O. Rasmussen, and A. Svensson. „Millennial-scale variability during the last glacial: The ice core record“. In: *Quat. Sc. Rev.* 29 (2010), pp. 2828–2838 (cit. on p. 4).

[Zha+14] X. Zhang, G. Lohmann, G. Knorr, and C. Purcell. „Abrupt glacial climate shifts controlled by ice sheet changes“. In: *Nature* 512 (2014), pp. 290–294 (cit. on pp. 6, 95).

[Zha+17] X. Zhang, G. Knorr, G. Lohmann, and S. Barker. „Abrupt North Atlantic circulation changes in response to gradual CO2 forcing in a glacial climate state“. In: *Nat. Geoscience* 10 (2017), pp. 518–523 (cit. on p. 6).

[EPI10] EPICA Community Members. „Stable oxygen isotopes of ice core EDML“. In: *PANGAEA* (2010), https://doi.org/10.1594/PANGAEA.754444 (cit. on pp. 3, 16).

[NGR04] NGRIP Members. „High-resolution record of Northern Hemisphere climate extending into the last interglacial period“. In: *Nature* 431 (2004), pp. 147–151 (cit. on pp. 3, 10, 15).

[WAI15] WAIS Divide Project Members. „Precise interpolar phasing of abrupt climate change during the last ice age“. In: *Nature* 520 (2015), pp. 661–665 (cit. on pp. 4, 5).