A First Step Toward Quantifying the Climate's Information Production over the Last 68,000 Years

Joshua Garland[1,3](✉), Tyler R. Jones[2], Elizabeth Bradley[1,3], Ryan G. James[4], and James W.C. White[2]

[1] Department of Computer Science, University of Colorado, Boulder, CO, USA
garland.joshua@gmail.com
[2] University of Colorado, INSTAAR, Boulder, CO, USA
[3] Santa Fe Institute, Santa Fe, NM, USA
[4] Department of Physics, University of California, Davis, CA, USA

Abstract. Paleoclimate records are extremely rich sources of information about the past history of the Earth system. We take an information-theoretic approach to analyzing data from the WAIS Divide ice core, the longest continuous and highest-resolution water isotope record yet recovered from Antarctica. We use weighted permutation entropy to calculate the Shannon entropy rate from these isotope measurements, which are proxies for a number of different climate variables, including the temperature at the time of deposition of the corresponding layer of the core. We find that the rate of information production in these measurements reveals issues with analysis instruments, even when those issues leave no visible traces in the raw data. These entropy calculations also allow us to identify a number of intervals in the data that may be of direct relevance to paleoclimate interpretation, and to form new conjectures about what is happening in those intervals—including periods of abrupt climate change.

1 Introduction

The Earth system contains a vast archive of geochemical information that can be utilized to understand past climate change. Using continually improving analytical techniques, records of change have emerged from corals, marine and lake sediments, tree rings, cave formations, pollen distribution, and the ice sheets. These heterogeneous data sets paint an intricate history of climate change on Earth, often being linked in time by common features, but also containing distinct information about local, regional, and global processes.

To our knowledge, no one has applied information-theoretic techniques to these data—an approach that holds promise for improving climatic interpretations. Knowledge as to where information is created in the climate system, and how it propagates through that system, could reveal and elucidate triggers, amplifiers, sources of persistence, and globalizers of climate change [2,23].

H. Boström et al. (Eds.): IDA 2016, LNCS 9897, pp. 343–355, 2016.
DOI: 10.1007/978-3-319-46349-0_30

For example, ice cores provide high-resolution proxies for hydrologic cycle variability, greenhouse gases, temperature, and dust distribution, among others. The spatiotemporal information captured in these records of climate change could reveal intricacies about the Earth climate system.

This paper is about one piece of that question: what the Shannon entropy rate of the water isotope signals in a specific Antarctic ice core tells us—about that data, about the past conditions at the core site, and about the overall climate. In an ice core, layers capture information about the local conditions at the time of deposition. A depth-wise series of measurements of some chemical or physical property of the ice, then, is effectively a time-series trace of those conditions. Water isotopes are a particularly useful property to study because they are good proxies for temperature and atmospheric circulation that result from variability in the hydrologic cycle. The time scale is unknown, though, and understanding the specific form of the relationship between the measured quantity and different aspects of the climate system requires forensic reasoning. These issues are discussed further in Sect. 2.

The Shannon entropy rate is a potentially useful way to carry out forensic reasoning about the climate system. It measures the average rate at which new information—unrelated to anything in the past—is produced by the system that generated the time series. If that rate is very low, the current observation contains a lot of information about the past and the signal is perfectly predictable. If that rate is very high, all of the information in the observation is completely new: i.e., the past tells you nothing about the future. Calculated over time-series data from ice cores, this quantity—described in Sect. 3—allows one to explore temporal correlations in the climate, which are critically important in understanding the underlying spatiotemporal mechanisms of this complex dynamical system. The results of these calculations, described in Sect. 4, are quite promising; they not only corroborate known facts, but also suggest new and sometimes surprising geoscience, and pave the way towards more-advanced interhemispheric entropy comparisons that could elucidate some of the deeper questions posed above about the larger climate system.

2 Paleoclimate: Dynamics and Data

At long time scales, the climate alternates between glacial and interglacial periods. A few of these cycles are shown in Fig. 1: the warm Holocene in which we live, which began $\approx 12,000$ years before present (12 ka), then the last glacial period from 110–12 ka and the Eemian interglacial period from 135–110 ka. *(NB: time runs backwards in most paleoclimate data analysis: the plots in this paper start at the current era and move into the past, from left to right.)* There is finer-grained structure in the record as well: the Younger Dryas “cold snap” between 12.8–11.5 ka, for instance, which interrupted the slow temperature rise into the Holocene [11]. There are also meaningful differences between records in different parts of the world.

Greenland cores like the NGRIP one in Fig. 1, for example, preserve strong signatures of Dansgaard-Oeschger (DO) events [9], where the temperature rises

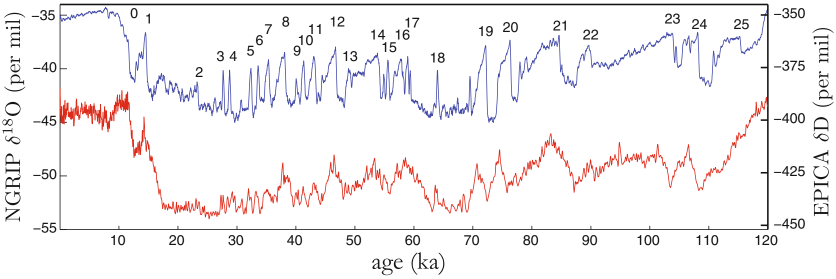

Fig. 1. Climate records from Greenland (NGRIP; top, in blue) and Antarctica (EPICA; bottom, in red). The horizontal axis is time before present in thousands of years. The quantities δ^{18}O and δD, as explained at more length in the text, are temperature indicators. The numbers identify Dansgaard-Oeschger events. (Color figure online)

9–16 °C over a span of decades or even years, then slowly falls over the course of ≈ centuries, and finally decays rapidly back to the baseline. These events, shown with superimposed numbers in Fig. 1, are thought to involve large-scale redistribution of oceanic heat. The trigger is rapid warming in the north Atlantic, perhaps because of injection of fresh water from ice-sheet melting; this disturbs the deep "conveyor belt" currents, causing Antarctica to cool some 200 years later, which is evident in ice cores from that region in the form of an AIM event (Antarctic isotope maxima) [22]. There are many other meaningful features in these data, as well; see [9,16] for good reviews.

Modern ice cores, from which data sets like the one in Fig. 1 are derived, cover timespans of up to 800,000 years. These can reach over 3 Km in length and are typically analyzed on a scale of cm—and, for some properties, mm. Each sample may involve dozens of measurements: different kinds of ions and isotopes, dust levels, conductivity, and so on. Some of the more useful of these are the stable and radiogenic isotopes, the amount and type of dust (which are correlated to the energy and humidity of the atmosphere), and the conductivity. The dynamic ranges of these measurements can be huge: sulfate levels go up by a factor of 1000 when a volcano erupts, for instance. Noise levels vary greatly across the different measurements, but those levels are not well established—and indeed are the subject of some important arguments about how to distinguish signal from noise. And of course the analysis equipment affects the data, sometimes without leaving any visually obvious trace in that data. That issue will return later in this paper.

The study reported here involves data from the 3405 m long West Antarctic Ice Sheet Divide core (WDC), which was gathered and analyzed by a team involving authors Jones and White [21,22]. This core, which covers a period of roughly 68 ka, is the highest-resolution and longest continuously measured record of its kind ever recovered from Antarctica. The high accumulation rate at the WAIS Divide—about 23 cm/yr in recent times—results in annual isotopic signals that persist for the last ≈ 16 thousand years, as well as signals at three years and greater that persist throughout the entire 68 thousand year record. These high-frequency signals have never before been interpreted across the last glacial-interglacial transition in Antarctica. In this paper, we focus on the water isotope measurements in this record: specifically δD, the ratio of ^{2}H (deuterium, D) to ^{1}H, and δ^{18}O, the ratio of ^{18}O to ^{16}O. Their values are reported in mille (parts per thousand, or "per mil"), relative to a calibrated standard of the isotopic composition of fresh water [1], and are generally negative for glacier ice. A δD value of -250 mille, for instance, means that that water sample is depleted in deuterium by 250 parts per thousand, relative to that standard.

Both δD and δ^{18}O are good proxies for temperature at the time of deposition of the associated section of the core [10], but the underlying mechanisms are not straightforward. The initial values of these ratios are known from ocean chemistry. The heavier isotopes (deuterium, ^{18}O) precipitate out preferentially, at known rates, as the water is carried to the polar regions in the form of vapor and the air mass cools. Other factors also affect that air mass along the way, however, so δD and δ^{18}O are not simple functions of temperature. And some of those effects—as well as some of the post-depositional processes that affect these two quantities once they are embedded in an ice core—are different for the two isotopes because of their differing molecular weights [8]. One method for understanding these molecular differences, for example, is to study the secondary measure $dxs = \delta\text{D} - 8 \times \delta^{18}\text{O}$, which is considered to be an effective proxy for kinetic effects during evaporation at the moisture source (largely a function of sea surface temperature), or a reflection of changing moisture sources over time.

All of those measurements are on a depth scale; to do any kind of time-series analysis, one must convert them to an age scale. This requires an "age model" for the core: a mapping of depth to age. Constructing this mapping requires a subtle, complicated combination of data analysis and scientific reasoning. Layers can be counted, for instance, but only to a maximum of 40–50 ka because the upper layers compress the ice underneath, thinning the layers to the point that they are unrecognizable. The measurements in the core play a key role in age-model construction: the astronomically based "Milankovitch" theory of ice ages predicts how δ^{18}O should vary through time, for instance. But ocean δ^{18}O also depends on the total volume of land ice on Earth[1], so this quantity is also a useful climate proxy. And near the base of the ice sheet, the ice often melts and/or deforms, making dating—or *any* kind of data analysis—very difficult. For the WAIS Divide core, the construction of the age model required several person-years of effort. The top 31.2 ka of the core was dated by four different individuals and one HMM-based software tool [24]; this procedure entailed visual identification of annual fluctuations in several different chemical traces along thousands of meters of core, followed by cross-corroboration between different proxies and different daters [20,21]. From 31.2–67.8 ka, the age scale was based on stratigraphic matching to "gold standard" Greenland ice cores and cross referenced using uranium/thorium ratios from cores drilled from cave features [5]. This represents the state of the art for data analysis in this field.

[1] Since ice sheets preferentially collect ^{16}O, while oceans preferentially collect ^{18}O.

3 Calculating the Rate of Information Production

The rate at which new information appears in a time series has been shown to be an effective method for signaling regime shifts: e.g., epileptic seizure detection in EEG signals [6], bifurcations in the transient logistic map [6], and recognizing voiced sounds in a noisy speech signal [3]. Estimating that quantity from an arbitrary, real-valued time series can be a real challenge, however. Most approaches to this problem use the Shannon entropy rate [15,19] and thus require categorical data: $x_i \in \mathcal{S}$ for some finite or countably infinite *alphabet* $\mathcal{S}$. This is an issue in the analysis of the type of high-resolution data produced by an ice-core lab because symbolization introduces bias and is fragile in the face of noise [4,13].

Permutation entropy (PE) [3] is an elegant solution to this problem. It symbolizes the time series in a manner that follows the intrinsic behavior of the system under examination. This method is quite robust in the face of noise and does not require any knowledge of the underlying mechanisms of the system. Rather than calculating statistics on sequences of values, as is done when computing the Shannon entropy in the standard way, permutation entropy looks at the statistics of the *orderings* of sequences of values using ordinal analysis. Ordinal analysis of a time series is the process of mapping successive elements of a time series to value-ordered permutations of the same size. For example, if $(x_1, x_2, x_3) = (7, 2, 5)$ then its *ordinal pattern*, $\phi(x_1, x_2, x_3)$, is 231 since $x_2 \leq x_3 \leq x_1$. The ordinal pattern of the permutation $(x_1, x_2, x_3) = (7, 5, 2)$ is 321.

Given a time series $\{x_i\}_{i=1,\ldots,N}$, there is a set $\mathcal{S}_\ell$ of all $\ell!$ permutations π of order ℓ. For each $\pi \in \mathcal{S}_\ell$, one defines the relative frequency of that permutation occurring in $\{x_i\}_{i=1,\ldots,N}$:

$$p(\pi) = \frac{|\{i | i \leq N - \ell, \phi(x_{i+1}, \ldots, x_{i+\ell}) = \pi\}|}{N - \ell + 1} \tag{1}$$

where $p(\pi)$ quantifies the probability of an ordinal and $|\cdot|$ is set cardinality. The permutation entropy of order $\ell \geq 2$ is:

$$\mathrm{PE}(\ell) = -\sum_{\pi \in \mathcal{S}_\ell} p(\pi) \log_2 p(\pi) \tag{2}$$

Since $0 \leq PE(\ell) \leq \log_2(\ell!)$ [3], it is common in the literature to normalize permutation entropy as follows: $\frac{PE(\ell)}{\log_2(\ell!)}$. With this convention, "low" PE is close to 0 and "high" PE is close to 1.

PE runs into trouble if the observational noise is larger than the trends in the data, but smaller than its larger-scale features. *Weighted permutation entropy* (WPE) [12] addresses this issue by taking the weight of a permutation into account:

$$w(x_{i+1}^{\ell}) = \frac{1}{\ell} \sum_{j=i}^{i+\ell} \left(x_j - \bar{x}_{i+1}^{\ell}\right)^2 \tag{3}$$

where x_{i+1}^{ℓ} is a sequence of values $x_{i+1}, \ldots, x_{i+\ell}$, and $\bar{x}_{i+1}^{\ell}$ is the arithmetic mean of those values. The weighted probability of a permutation is defined as:

$$p_w(\pi) = \frac{\sum\limits_{i \le N-\ell} w(x_{i+1}^{\ell}) \cdot \delta(\phi(x_{i+1}^{\ell}), \pi)}{\sum\limits_{i \le N-\ell} w(x_{i+1}^{\ell})} \tag{4}$$

where $\delta(x, y)$ is 1 if $x = y$ and 0 otherwise. Effectively, this weighted probability emphasizes permutations that are involved in "large" features and de-emphasizes permutations that are small in amplitude, relative to the features of the time series. The standard form of weighted permutation entropy is:

$$\mathrm{WPE}(\ell) = -\sum_{\pi \in \mathcal{S}_\ell} p_w(\pi) \log_2 p_w(\pi), \tag{5}$$

which can also be normalized by dividing by $\log(\ell!)$, to make $0 \le \mathrm{WPE}(\ell) \le 1$.

In practice, calculating permutation entropy and weighted permutation entropy involves choosing a good value for the word length ℓ. The primary consideration in that choice is that the value be large enough to allow the discovery of forbidden ordinals, yet small enough that reasonable statistics over the ordinals can be gathered. If an average of 100 counts per ordinal is considered to be sufficient, for instance, then $\ell = \mathrm{argmax}_{\hat{\ell}}\{N \gtrapprox 100\hat{\ell}!\}$. In the literature, $3 \le \ell \le 6$ is a standard choice—generally without any formal justification. In theory, the permutation entropy should reach an asymptote with increasing ℓ, but that can require an arbitrarily long time series. In practice, the right thing to do is to calculate the *persistent* permutation entropy by increasing ℓ until the result converges, but data length issues can intrude before that convergence is reached. We used that approach to choose $\ell = 4$ for the calculations in this paper. This value represents a good balance between accurate ordinal statistics and finite-data effects.

WPE is a powerful technique, but it is not without issues. The choice of the ℓ value is one; another is the notion of significance. As is the case with many nonlinear measures on data, it is quite difficult to define what qualifies as a significant change in a WPE plot. One way to tell if a particular feature (e.g., jump, spike, valley) is important is by first understanding the time scales of the system and comparing them to the size of the window of data over which the WPE calculation is performed. This can help establish whether a change on that time scale makes sense. It is also important to remember that for a particular system, small-scale fluctuations over short time intervals may indeed signal some small event. To distinguish between signal and noise, it is important to vary the window size (as the data allows) and see if the result persists. We take that approach in the calculations reported in the following section.

4 Results

Ice cores are sampled at evenly spaced intervals in depth, but these measurements are spaced nonlinearly (and unevenly) in time because of the progressive downcore thinning of the ice and differing annual accumulation rates of snow. To create evenly spaced time-series data for δD, δ^{18}O, and dxs, we first used the age model described at the end of Sect. 2 to convert depths to ages, and then re-mapped the data to a constant temporal spacing of 1/20th of a year using linear interpolation. The effective resolution of the data is 0.005 m. In the upper portions of the ice core, annual layer thicknesses are about 20 cm, so there are roughly 40 data points per year. At greater depths in the core, an annual layer may only be 4 cm thick, yielding eight data points per year. The accuracy involved in interpolating these unevenly spaced data to a uniform spacing of 1/20th year varies over the depth of the core; this matter, and its potential effects on the results, are discussed further at the end of this section. The specific age scale spacing of 1/20th per year was chosen because it preserves the structure and amplitude of the data—that is, there are no instances of significantly reduced amplitude in the signal, or losses in spectral power.

We then used the normalized version of Eq. (5) with $\ell = 4$ to calculate WPE in 500-year long windows across each of those time-series traces. Each point in the resulting calculation captures the rate of information production over the previous 500 years. The sliding-window nature of this calculation is intended to bring out the fine-grained details of the information mechanics of the system. Since WPE's statistics are built up over the full span of the data that is passed to it, performing that calculation over a longer segment of the climate data—one that spanned different regimes—would intermingle the mechanics of those different regimes. In the case of the questions that we were asking about the WAIS Divide core, the minimum scale of the interesting events was 100–500 years. We ran all of the calculations reported here for that range of window sizes and observed no change in the results. (All of these window sizes, incidentally, satisfied the theoretical data-length requirements for successful WPE calculations.)

Figure 2 shows the δD data from the WDC, along with the WPE of that trace, calculated as described above. The δD WPE is ≈ 0.2 from 10–60 ka, indicating

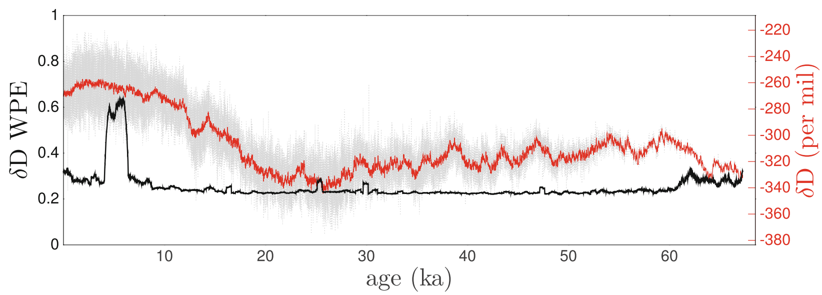

Fig. 2. The deuterium/hydrogen ratio (δD) measured from the WAIS Divide Core. The original data is shown in grey, the smoothed data (500-year moving average) in red, and the weighted permutation entropy (WPE) calculated from the original data in black. (Color figure online)

that δD values depend strongly on their previous values during this period—i.e., that very little new information is produced by the system at each time step. A very interesting feature here is the large jump in WPE between 5–8 ka. As it turns out, an older instrument was used to analyze the ice in this region. The WPE results clearly show that that instrument introduced noise into the data: i.e., every measurement contains completely new information, unrelated to the previous ones. As can be seen from examination of the red and grey traces in the figure, that noise was not visually apparent in the δD data itself, so the instrument issue was not detected immediately by the laboratory team. The fact that WPE brings out the disparity between the two instruments so clearly is a major advantage. (Indeed, that revelation has caused author White's team to re-examine the data in the depth ranges where the blips occur in the WPE results, near 17, 26, and 30 ka.) Another interesting feature of Fig. 2 is the rise in δD WPE from 62–68 ka. This may be due to geothermal heat at the base of the ice sheet, which causes water isotopes to diffuse in that region, thereby injecting new information into the oldest section of the time series. This matter is discussed at more length below.

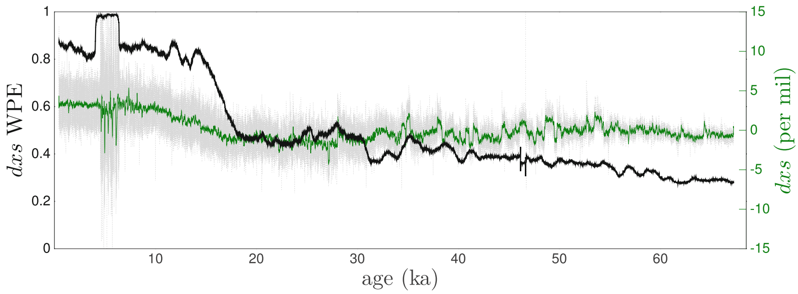

Fig. 3. The secondary measure dxs calculated from the δD and δ^{18}O measurements in the WAIS Divide Core. Original and smoothed values of dxs are shown in green and grey respectively; the WPE of the original signal is shown in black. (Color figure online)

The instrument issue that was invisible in the δD data **was** visible in the dxs trace, as shown in Fig. 3. Recall that $dxs = \delta\mathrm{D} - 8\delta^{18}\mathrm{O}$, and that the intent of performing this weighted sum is to deconvolve the differential scientific effects that are at work in these two quantities and thereby zero in on sea-surface temperature at the moisture source. It may seem that as dxs is an affine transformation of δD and δ^{18}O, the WPE would similarly be an affine transformation of the information production of the individual signals, i.e., $\mathrm{WPE}(\ell; dxs) = a\mathrm{WPE}(\ell; \delta\mathrm{D}) + b\mathrm{WPE}(\ell; \delta^{18}\mathrm{O})$. However, *weighted* permutation entropy is not preserved by affine transformations—and this is useful. Among other things, it allows us to leverage the information production of dxs to gain insight into the second-order dynamics that the calculations of dxs is intended to get at: specifically, the interplay between kinetic fractionation effects and moisture source conditions.

Careful comparison of the three WPE traces in Fig. 4 illustrates the utility of that reasoning. Recall that both δD and δ^{18}O of a section of the core are, to first approximation, proxies for temperature at the time of the deposition of that material, but that thermodynamics and the difference in atomic weight causes them to behave slightly differently. (The *dxs* calculation, again, is intended to get at these second-order effects.) As expected, the δD and δ^{18}O WPE traces are largely similar, except for a small but fairly consistent vertical offset from 0–17 ka. This offset indicates that the information production of δ^{18}O increased towards the end of the last glacial period, suggesting that, from an information-theoretic perspective, the second-order thermodynamic effects have been playing a greater role in the climate dynamics since 17,000 years ago. On the right-hand side of the figure, however, the similarity fails. δ^{18}O WPE rises slowly from 50–62 ka, at which point δD WPE rises somewhat as well. This could be a thermal diffusion effect due to geothermal heat at the bedrock interface, which will inject noise into the data—and differentially affect the isotopes due to their different molecular masses. It could also be a scaling issue; signal at those depths is attenuated by the effects of time and depth, and also smoothed by the finite resolution of the analysis equipment.

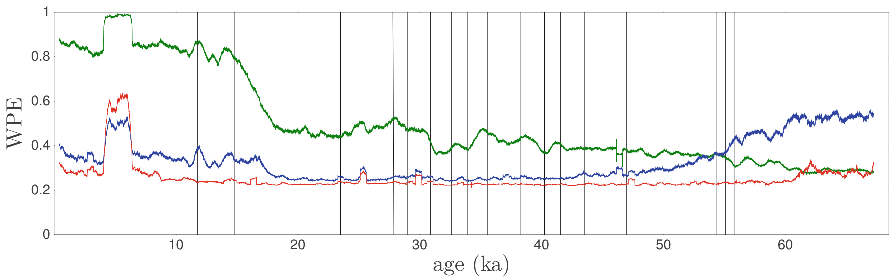

Fig. 4. WPE of δD (red), δ^{18}O (blue), and *dxs* (green); Dansgaard-Oeschger events are marked with black vertical lines. (Color figure online)

The Dansgaard-Oeschger events described in Sect. 2 are climatologically important and scientifically interesting, but their mechanics and dynamics are not completely understood. During the early stages of the collaboration that produced this paper, the geoscientists on the team conjectured that these events would inject new information into the time series. As is clear from Fig. 4, however, that is not the case. That is, while DO events may reflect changes in the dynamics of the climate (cf., recent work on "critical slowing down" [7,14,17]), they are not associated with changes in the *information production* of that system. Rather, they appear to be just part of the normal operating procedure of the climate system. We are currently looking at shorter windows to see whether WPE reveals "triggers" or other early-warning signals for these important events.

The results shown in Fig. 4 catalyzed a number of other new hypotheses about climate science. The WAIS Divide ice core derives moisture mainly from a Pacific Ocean source. During the beginning of the deglaciation at $\approx$ 19 ka, the dxs WPE and δ^{18}O WPE begin to increase, in line with accumulation[2]. At that time, changing climatic conditions in West Antarctica may have produced storm tracks that delivered precipitation from more-diverse locations. For example, increased sea ice extent during the glacial period may have limited storm tracks to those originating primarily in the central Pacific, but upon sea ice decline during deglaciation, more local storm tracks, and possibly storm tracks from the west Pacific and even the Indian Ocean, could have contributed precipitation to the ice core site. Simply put, we suggest that more accumulation means more storms originating from more locations. Of course, from a fractionation standpoint, the increasingly variable location of moisture formation would have to introduce differing kinetic effects that injected more information into the system; we would not expect this as the physics of evaporation should be the same.

It is worth thinking about whether the preprocessing step outlined in the first paragraph of this section—which is the standard approach in this field if one wants an ice-core data set with even temporal sampling—could have disturbed the information mechanics of the data. The ramps introduced by linear interpolation introduce repeating, predictable patterns in the π of Sect. 3, which could skew the distribution of those permutations. For long enough interpolations, this should lower the overall WPE value, but the time scales of this effect are all but impossible to derive.

To explore whether this WPE shrinkage was at work in our results, we carried out the following experiment. We first generated a time series using a random-walk process, which has a theoretical WPE of $\approx$ 0.9405. That only holds, however, if one uses an infinitely long time series; in practice, calculations of WPE on time series like this yield values of $\approx$ 0.85–0.9. We then used the WAIS Divide core age model to invert the time scale of this trace, downsampling it nonlinearly so that the temporal intervals between data points were consistent with an 0.005 m spacing. This is an effective ansatz for a data set from that core: closely spaced points near the beginning of the trace, where the core is less dense and a year's worth of material is thicker, and spreading apart roughly exponentially later in the time series, which corresponds to the highly compressed ice deep in the core. We then subjected that trace to the same preprocessing steps outlined at the beginning of this section and finally computed its WPE. The results showed a correct baseline of 0.82–0.87 early in the time series, where the interpolation interval is small, followed by a slow decrease starting around 12 ka.

That decrease suggests that one should be careful comparing WPE values of a single trace across wide temporal ranges—especially when one is working deep

[2] The accumulation data from the WDC has not yet been released publicly, so we cannot include a plot of it here, but there are some *extremely* interesting correspondences that we hope to be able to include in a few months, when we are allowed to share those data.

in the core. However, the claims offered in this paper concern (a) features high up in the core (viz., the instrument issue) (b) narrow features deeper down, and (c) broad time-scale comparisons of traces produced with identical interpolation processes (Fig. 4). Moreover, we are not completely convinced of the accuracy of our synthetic experiment; note, for example, that the δ^{18}O and δD WPE traces in Fig. 4 actually *rise* at lower depths—the region where interpolation effects should, theoretically, cause a dropoff. It may be the case that the geothermal effects inject much more information into the system than the results presented here suggest, or it may be that we do not completely understand the interpolation effects. We are currently exploring this from several angles: by simply downsampling the WDC data, rather than interpolating it, and by comparing WPE of different WDC traces that were sampled at different intervals. There are other issues as well. Gases diffuse through the material in the core: more readily at the top, where the material is less dense, and more slowly lower down, where the snow has been compressed into solid ice. Since diffusion effectively introduces white noise, it should raise WPE. However, this effect would be quite difficult to deconvolve from the data in an ice core, since a given segment of that core has undergone a continuum of diffusion processes at different temporal and spatial scales during its history.

5 Conclusion and Future Work

Paleoclimatic analyses require complicated forensic reasoning to determine the timing and phasing of past events. In ice-core science, a particularly meaningful paleoclimatic indicator is found in the measurements of water isotopes; these are useful for understanding temperature, atmospheric circulation, and oceanic conditions. Interpreting records of the history of these isotopes requires knowledge of the thermodynamics of the climate system, along with a precise age scale. The former is governed by physics; the latter requires intelligent data analysis performed by multiple individuals and software tools.

In this paper, we used data from the WAIS Divide ice core to analyze information production over the last 68,000 years. Through permutation entropy techniques (WPE), we found that information production in δD, δ^{18}O, and dxs is consistent with thermodynamic expectations. The WPE of δ^{18}O and dxs share common features, likely because these parameters are more sensitive to kinetic fractionation effects in ocean moisture source regions. Conversely, the δD is likely more responsive to temperature-related equilibrium effects. We identify a number of intervals in the data that may be of direct relevance to paleoclimate interpretation. In the deepest sections of the core (> 60 ka), divergence of WPE for δD and δ^{18}O may help identify time periods when geothermal heat flux causes differential solid diffusion of water molecules. Throughout the last glacial period (> 12 ka), rather constant WPE values suggest no information production during large-scale abrupt warming events in Greenland that also appear as isotope maxima in Antarctica: this is an unexpected result. Across the glacial-interglacial transition ($\approx 19 - 11$ ka) and into the Holocene (< 11 ka), increases in WPE for all variables may signal more variability in the hydrologic cycle—for example, receding sea ice may establish more-diverse moisture sources across latitudes and ocean basins. In the last 2000 years, an increase in WPE (the small rise that is visible on the left-hand edge of Fig. 4, in all three traces) may signal anthropogenic effects on the climate system. These comparisons deserve more rigorous treatment, and will be included in future work.

The overall results of this study constitute only one kind of information calculation on a few isotopes in a single ice core; there are additional options that could broaden the scope in useful and meaningful ways. For example, comparison of WPE from the WAIS Divide with that of other Antarctic ice cores may elucidate the thermodynamics of varying moisture sources in the Indian, Pacific, Atlantic, and Southern Oceans. It may even provide information about atmospheric processes that affect the isotopic signal. Comparison of WPE of varying WAIS Divide proxies, such as water isotopes and accumulation, may also hold important clues to information production in the climate system. Comparison of Antarctic and Greenland cores could provide interhemispheric viewpoints that can inform us about abrupt climate change events (e.g., trigger mechanisms). Other information-theoretic measures, such as mutual information or transfer entropy [18], may be even more useful than WPE in these kinds of comparisons. None of these sorts of calculations have previously been done on paleoclimatic data sets—let alone multiple ones—and the initial findings presented here hold promise for improving climatic interpretations.

References

1. http://tinyurl.com/hz8xelf
2. Alley, R., et al.: Abrupt climate change. Science **299**, 2005–2010 (2003)
3. Bandt, C., Pompe, B.: Permutation entropy: a natural complexity measure for time series. Phys. Rev. Lett. **88**(17), 174102 (2002)
4. Bollt, E., Stanford, T., Lai, Y.C., Życzkowski, K.: What symbolic dynamics do we get with a misplaced partition?: on the validity of threshold crossings analysis of chaotic time-series. Physica D **154**(3), 259–286 (2001)
5. Buizert, C.: The WAIS divide deep ice core WD2014 chronology–Part 1: methane synchronization (6831 ka BP) and the gas age-ice age difference. Clim. Past **11**, 153–173 (2015)
6. Cao, Y., Tung, W., Gao, J., Protopopescu, V., Hively, L.: Detecting dynamical changes in time series using the permutation entropy. Phys. Rev. E **70**(4), 046217 (2004)
7. Dakos, V., et al.: Slowing down as an early warning signal for abrupt climate change. Proc. Nat. Acad. Sci. **105**(38), 14308–14312 (2008)
8. Dansgaard, W.: Stable isotopes in precipitation. Tellus **16**, 436–447 (1964)
9. Dansgaard, W., et al.: Evidence for general instability of past climate from a 250-kyr ice-core record. Nature **364**, 218–220 (1993)
10. Dansgaard, W., Johnsen, S., Clausen, H., Gundestrup, N.: Stable isotope glaciology. Medelelsser om Gronland **197**, 1–53 (1973)
11. Dansgaard, W., White, J., Johnsen, S.: The abrupt termination of the younger Dryas climate event. Nature **339**, 532–534 (1989)

12. Fadlallah, B., Chen, B., Keil, A., Príncipe, J.: Weighted-permutation entropy: a complexity measure for time series incorporating amplitude information. Phys. Rev. E **87**(2), 022911 (2013)
13. Kraskov, A., Stögbauer, H., Grassberger, P.: Estimating mutual information. Phys. Rev. E **69**(6), 066138 (2004)
14. Lenton, T., et al.: Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness. Philos. Trans. R. Soc. Lond. A: Math. Phy. Eng. Sci. **370**(1962), 1185–1204 (2012)
15. Mantegna, R., Buldyrev, S., Goldberger, A., Havlin, S., Peng, C., Simons, M., Stanley, H.: Linguistic features of noncoding DNA sequences. Phys. Rev. Lett. **73**(23), 3169–3172 (1994)
16. Petit, J., et al.: Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature **399**, 429–436 (1999)
17. Scheffer, M., et al.: Early-warning signals for critical transitions. Nature **461**, 53–59 (2009)
18. Schreiber, T.: Measuring information transfer. Phys. Rev. Lett. **85**(2), 461–464 (2000)
19. Shannon, C.: Prediction and entropy of printed English. Bell Syst. Tech. J. **30**(1), 50–64 (1951)
20. Sigl, M., et al.: The WAIS Divide deep ice core WD2014 chronology–part 2: Annual-layer counting (031 ka BP). Clim. Past **12**, 769–786 (2016)
21. WAIS Divide Project Members: Onset of deglacial warming in west antarctica driven by local orbital forcing. Nature **500**, 440–444 (2013)
22. WAIS Divide Project Members: Precise interpolar phasing of abrupt climate change during the last ice age. Nature **520**, 661–665 (2015)
23. White, J., et al.: Abrupt impacts of climate change: Anticipating surprises. In: EGU General Assembly Conference Abstracts, vol. 16, p. 17028 (2014)
24. Winstrup, M., et al.: An automated approach for annual layer counting in ice cores. Clim. Past **8**, 1881–1895 (2012)