HAL
open science

Automatic detection of abrupt transitions in paleoclimate records

W. Bagniewski, M. Ghil, Denis-Didier Rousseau

To cite this version:

W. Bagniewski, M. Ghil, Denis-Didier Rousseau. Automatic detection of abrupt transitions in paleoclimate records. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2021, 31 (11), pp.113129. 10.1063/5.0062543 . insu-03462705

HAL Id: insu-03462705

https://insu.hal.science/insu-03462705

Submitted on 5 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Chaos

Automatic detection of abrupt transitions in paleoclimate records

Cite as: Chaos **31**, 113129 (2021); https://doi.org/10.1063/5.0062543
Submitted: 06 July 2021 • Accepted: 08 October 2021 • Published Online: 16 November 2021

W. Bagniewski, M. Ghil and D. D. Rousseau

COLLECTIONS

This paper was selected as Featured

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Noise-driven topological changes in chaotic dynamics
Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 103115 (2021); https://doi.org/10.1063/5.0059461

Delayed epidemic peak caused by infection and recovery rate fluctuations
Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 101107 (2021); https://doi.org/10.1063/5.0067625

Fast-slow analysis of a stochastic mechanism for electrical bursting
Chaos: An Interdisciplinary Journal of Nonlinear Science **31**, 103128 (2021); https://doi.org/10.1063/5.0059338

© 2021 Author(s).

Chaos ARTICLE scitation.org/journal/cha

Automatic detection of abrupt transitions in paleoclimate records

Cite as: Chaos **31**, 113129 (2021); doi: 10.1063/5.0062543
Submitted: 6 July 2021 · Accepted: 8 October 2021 ·
Published Online: 16 November 2021

View Online Export Citation CrossMark

W. Bagniewski,[1,a)] M. Ghil,[1,2] and D. D. Rousseau[3,4]

AFFILIATIONS

[1]Department of Geosciences and Laboratoire de Météorologie Dynamique (CNRS and IPSL), École Normale Supérieure and PSL University, 75132 Paris Cedex 05, France
[2]Department of Atmospheric and Oceanic Science, University of California at Los Angeles, Los Angeles, California 90095, USA
[3]Geosciences Montpellier, University of Montpellier, CNRS, 34095 Montpellier, France
[4]Lamont Doherty Earth Observatory, Columbia University, New York, New York 10964, USA

[a)]**Author to whom correspondence should be addressed:** wbagniewski@lmd.ipsl.fr

ABSTRACT

Bifurcations and tipping points (TPs) are an important part of the Earth system's behavior. These critical points represent thresholds at which small changes in the system's parameters or in the forcing abruptly switch it from one state or type of behavior to another. Current concern with TPs is largely due to the potential of slow anthropogenic forcing to bring about abrupt, and possibly irreversible, change to the physical climate system and impacted ecosystems. Paleoclimate proxy records have been shown to contain abrupt transitions, or "jumps," which may represent former instances of such dramatic climate change events. These transitions can provide valuable information for identifying critical TPs in current and future climate evolution. Here, we present a robust methodology for detecting abrupt transitions in proxy records that is applied to ice core and speleothem records of the last climate cycle. This methodology is based on the nonparametric Kolmogorov–Smirnov (KS) test for the equality, or not, of the probability distributions associated with two samples drawn from a time series, before and after any potential jump. To improve the detection of abrupt transitions in proxy records, the KS test is augmented by several other criteria and it is compared with recurrence analysis. The augmented KS test results show substantial skill when compared with more subjective criteria for jump detection. This test can also usefully complement recurrence analysis and improve upon certain aspects of its results.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0062543

Paleoclimate proxy records are used to reconstruct past states of the Earth's climate. These records sometimes contain "jumps," or abrupt transitions, which may serve as evidence of dramatic shifts in the Earth's past climates. Such jumps are likely to be caused by the climate system as a whole—or by the subsystem associated with the proxy record under study—crossing a tipping point (TP), at which self-reinforcing feedbacks push the system out of its stable state. Present-day global warming and other impacts of human activity have led to concerns about potential TPs, which, when crossed, could bring irreversible change to the physical climate system and impacted ecosystems.[40,56,61] Therefore, identifying and describing jumps associated with TPs in paleoclimate is essential to properly understanding the Earth system's underlying bifurcation mechanisms and may allow us to make more robust predictions for future climate. As paleoclimate records vary in their origin, time spans, and periodicities, an objective, automated methodology like the one proposed and demonstrated herein is crucial for identifying and comparing TPs. To determine the nature of the tipping mechanism requires theoretical and modeling work.[2,28,39] For the sake of brevity, we assume for the moment that each jump in the time series identified by the methodology described herein is equivalent to an abrupt climate transition and use jump and abrupt transition interchangeably.

I. INTRODUCTION

Proxy records show that climate during the last glacial period of about 115 000–11 700 yr ago was highly variable and characterized by rapid changes in temperature and precipitation occurring on centennial to millennial time scales. Among the most remarkable climatic events of this time period were Heinrich events,[34] associated with iceberg surges in the North Atlantic, as well as Dansgaard–Oeschger events,[17] which were associated

Published under an exclusive license by AIP Publishing

with a warming of several degrees Celsius over as little as a few decades.[36,64] Such transitions have been reproduced in modeling studies[3,7,27,50,79] and are likely to have been driven by major shifts in ice–ocean–atmosphere interactions;[9,43] however, the exact mechanisms driving these changes remain uncertain.[5,10,13,63,75]

Establishing precise dates for critical transitions in Earth's climatic history from jumps in the proxy records is a major challenge in paleoclimate studies.[41,42,55] For the last glacial cycle, reliable dating estimates may be obtained for ice core records, such as the one produced by the North Greenland Ice Core Project (NGRIP).[58] These records contain several types of proxies found in distinct layers of annual snow accumulation, which allow a relatively high sampling resolution. Still, the uncertainties in their dating have been shown to progressively increase for older layers.[8]

Speleothems also contain distinct layers, which may be accurately and precisely dated, mainly by U-Th dating. The descriptions and classifications of climatic events are typically performed by visually inspecting the record, which is both time consuming and rather subjective. In records where independent dating is not available, the age estimates are typically established by synchronization, or "wiggle matching," with well-dated records such as the NGRIP record or speleothem records. With improvements in sampling and dating methodologies, the dates for recognized climatic events have frequently been revised over the last few decades,[16,17,37,58] thus making age estimates in older studies unreliable. Therefore, given the large number of climate records in use and the need to frequently revise the findings from earlier studies, an automated statistical test is the preferred way of robustly analyzing the records.

Early examples of abrupt transition detection in climate time series include analysis of Nile River data.[11,14] Transition detection and change point detection algorithms have since increased in popularity within a variety of disciplines where nonlinear processes are involved, including signal processing, bioinformatics, and finance. The spreading of such algorithms has also contributed to a growing interest in applying such methods to climatic time series.[6,26,31,44,59,71] Truong *et al.*[70] have recently compiled a review of various change point detection methods.

In recent years, recurrence analysis[19,45] has been successfully applied for identifying transitions in paleoclimate data.[18,22,23,31,47,54,60,69,74] Still, climate variability includes both deterministic and stochastic processes, while paleoclimate time series are characterized by high levels of noise and a nonuniform resolution, as well as by irregular periodicities. These combined sources of uncertainty highlight the need for the development of alternative methods as well.

Here, we present a methodology for automatically detecting abrupt transitions in paleoclimatic proxy records of many types and on several time scales. This statistical method is based on the Kolmogorov–Smirnov test and it can be applied to different types of records extending over distinct time spans, thus allowing one to objectively compare them. The jumps so identified in climatic time series may then be further explored with a full hierarchy of models to improve our understanding of the Earth's bifurcation mechanisms[29,30] and identify possible TPs for future climates.

The paper is organized as follows. In Sec. II, we present an augmented Kolmogorov–Smirnov (KS) test and perform receiver operating characteristic (ROC) analysis for varying parameter values. In Sec. III, we apply the method to two paleoclimate records of the last climate cycle, a Greenland ice core,[58] and a speleothem composite record from China[12,73] and compare these results with those of recurrence analysis. Conclusions are drawn in Sec. IV.

II. METHODOLOGY

A. Kolmogorov–Smirnov (KS) test

Our methodology is based on the nonparametric KS test, a goodness-of-fit test for a time series $\{x_k : 1 \le k \le K\}$ named after Andrey Kolmogorov[38] and Nikolai Smirnov.[66] The two-sample KS test compares the empirical distribution functions of two samples, just before and just after a potential change point, in order to test the hypothesis that the two samples come from the same continuous distribution. This test is applied here to a proxy time series, to compare two samples drawn from before and after a potential jump in the paleoclimate record.

The equality of the two samples is quantified using the KS statistic D_{KS},[15,49] defined as the greatest vertical difference at any point between the cumulative distributions F_1 and F_2 of the two samples,

$$D_{\mathrm{KS}} = \sup_k |F_1(x_k) - F_2(x_\ell)|, \tag{1}$$

with $\ell = k + w$. Here, F_1 and F_2 are the empirical distribution functions of the samples drawn from the two intervals $\mathscr{I}_1 = \{t - w \le k \le t\}$ and $\mathscr{I}_2 = \{t \le \ell \le t + w\}$ within a window length w from before and after time t, while $\sup(\cdot)$ is the supremum of the distance between the two. Since both F_1 and F_2 lie between 0 and 1, one also has that $0 \le D_{\mathrm{KS}} \le 1$.

Figure 1 illustrates the quantification of the KS statistic for two different pairs of samples from the NGRIP δ^{18}O time series,[58] with a window of $w = 2000$ yr. Following Eq. (1), the KS statistic is equal to the greatest vertical distance between the two curves in Figs. 1(b) and 1(c), respectively, and clearly the larger distance corresponds to a greater degree of discontinuity.

The null hypothesis that the samples are drawn from the same distribution is rejected if D_{KS} is greater than a critical value,[67]

$$D_{\mathrm{KS}} > c(\alpha)\sqrt{\frac{n_1 + n_2}{n_1 n_2}}. \tag{2}$$

Here, n_1 and n_2 are the sizes of the two samples, and $c(\alpha)$ is the inverse of the Kolmogorov distribution at a chosen significance level α,

$$c(\alpha) = \sqrt{-\frac{1}{2}\ln\left(\frac{\alpha}{2}\right)}. \tag{3}$$

B. Application of the KS test to paleoclimate data

Since paleoclimate records vary in their length, resolution, dominant periodicity, and amplitude, the notion of "abruptness" requires a more precise definition than what is evaluated in Eq. (2). The KS test can give very different results depending on the window length being used. As may be seen in Fig. 1, the window length is inadequate when it is significantly shorter or significantly longer than the time interval between two discontinuities. Thus, to adapt

Published under an exclusive license by AIP Publishing

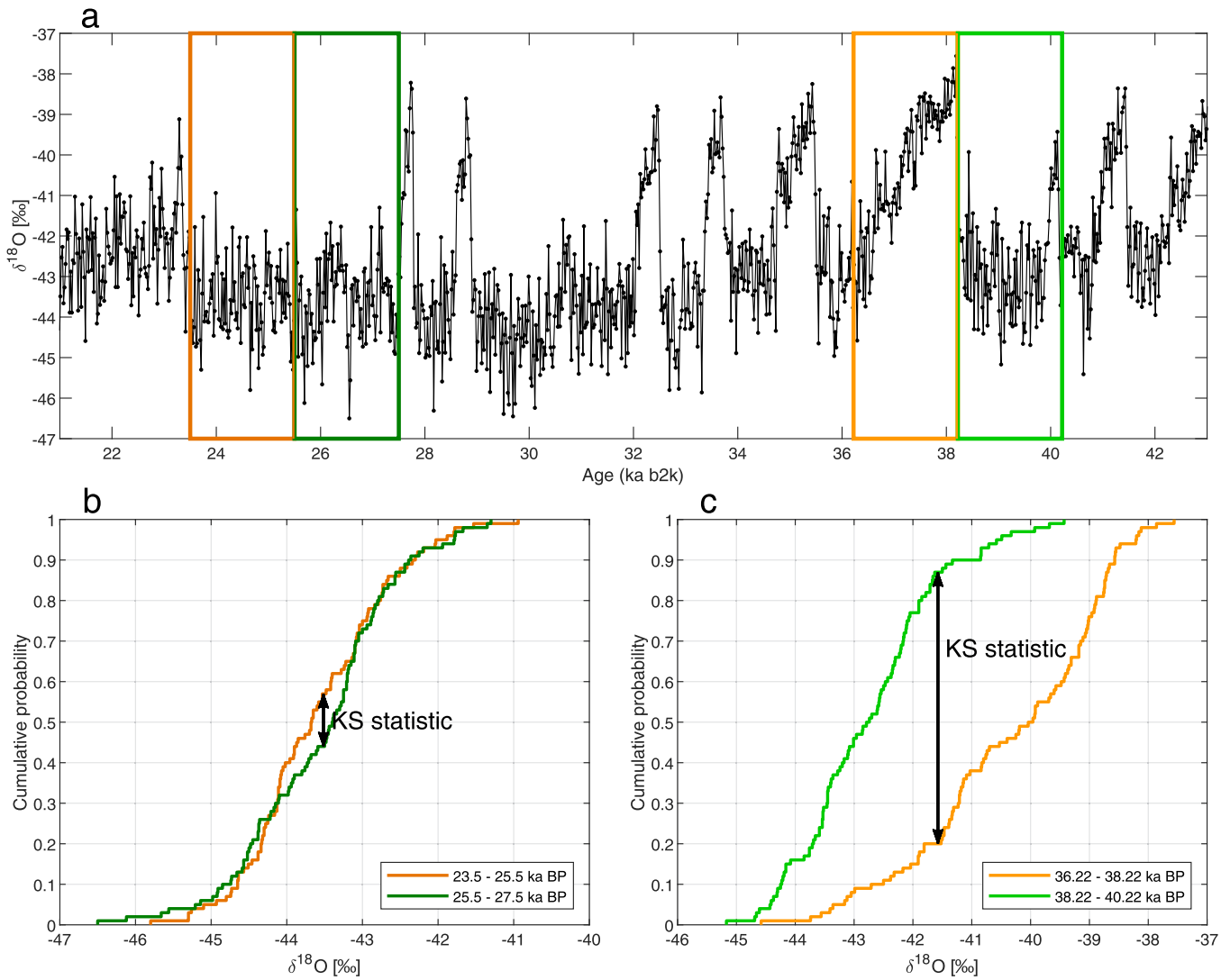

FIG. 1. Example of the KS test applied to the NGRIP δ^{18}O time series.[58] (a) Snapshot of the record at 21–43 ka b2k. The green and orange rectangles correspond to the sample windows of equal width w used for evaluating the KS statistic, green before and orange after the potential jump. (b) and (c) Empirical distribution functions of the two pairs of samples. The length of the black double arrow is equal to the KS statistic D_{KS}.

the KS test to detect transitions that may occur at distinct recurrence times, we introduce a varying window length w_i, with a range between $w_{\min}$ and $w_{\max}$,

$$w_i = w_{\min}\left(\frac{w_{\max}}{w_{\min}}\right)^{(i-1)/(N_w-1)}, \tag{4}$$

where N_w is the number of distinct window lengths used in the analysis.

For each window w_i, the KS statistic is calculated following Eq. (1). As climatic variability occurs at a spectrum of different time scales that may be as short as one hour or as long as the age of the Earth,[52] the frequency at which transitions are detected is subject to the values of $w_{\min}$ and $w_{\max}$. Thus, the values of these two parameters should correspond to the time scale at which a given paleorecord is investigated.

By calculating the KS statistic of Eq. (1), as visualized in Fig. 1, for every time step in a paleoclimate record, one may obtain a time series of D_{KS} with higher values that correspond to a higher degree of discontinuity in the record. Following Eq. (2), one wishes to define a critical value above which the discontinuity between two samples compared with the KS test is classified as an abrupt transition. As noted by Conover,[15] however, in almost any goodness-of-fit test, the null hypothesis will be rejected if a large enough sample

Published under an exclusive license by AIP Publishing

size is employed. Therefore, we replace the critical value defined in Eqs. (2) and (3) by a cut-off threshold D_c for the KS statistic, with $0 \leq D_c \leq 1$, and scale D_{KS} by the sizes of the two samples,

$$1 - \frac{1 - D_{KS}}{1 - \sqrt{\frac{n_1+n_2}{n_1 \cdot n_2}}} > D_c. \tag{5}$$

Furthermore, as the KS test by itself does not depend on the amplitude of the "jump" in the time series, it is desirable to discard any smaller jumps that may be the result of an error in the observed data, e.g., measuring error or small-scale variability that occurs over time intervals shorter than the sampling resolution of the proxy record. Thus, for each of the two adjacent samples drawn from $\mathscr{I}_1$ and $\mathscr{I}_2$, respectively, we compute the mean change $|\bar{x}_1 - \bar{x}_2|$ in magnitude between them, as well as the standard deviation for the two samples, σ_1 and σ_2. For a transition to be considered significant, the change in magnitude between the two samples in proportion to their standard deviation needs to exceed a threshold σ_c,

$$\frac{|\bar{x}_1 - \bar{x}_2|}{\sigma_1} > \sigma_c \quad \text{and} \quad \frac{|\bar{x}_1 - \bar{x}_2|}{\sigma_2} > \sigma_c. \tag{6}$$

Last, as the KS test requires a large enough sample size to be significant, we specify a minimum sample size threshold, n_c. For any window w_i, if either of the two samples has a size n smaller than n_c, the result of the KS test is deemed insignificant,

$$n_i \geq n_c. \tag{7}$$

An abrupt transition at time t is detected when all of the conditions in Eqs. (5)–(7) are satisfied. The dates for which these conditions are satisfied are obtained as time intervals of varying lengths, as shown in Fig. 2 for two different window lengths. The precise date for a transition within such a time interval is then determined by the maximum D_{KS} value found within that time interval. When several time steps share the same value D_{KS} that equals the maximum over the given interval, the time step corresponding to the maximum change in magnitude $|\bar{x}_1 - \bar{x}_2|$ is used; moreover, if there is more than one such time step, then the one with the earlier date is used.

Since the transition dates so identified will vary depending on the window length, the priority assigned to the results obtained with different windows will depend on that length. Given that the statistical significance of the KS test improves with the sample size, we first identify the transitions detected with the longest window, namely, w_N. These results are then supplemented with additional transitions detected for the next-longest window, w_{N-1}, and eventually to all other window lengths, until w_1.

Finally, as the same transition may be found at slightly different dates depending on the window length that is used, we want to ensure that each of the transitions thus identified is a separate event and thus avoid having the same transitions identified multiple times at neighboring dates. Hence, for window w_i we discard transitions identified at time t if the interval $\{t - w_i \leq t \leq t + w_i\}$ contains transitions that were previously identified with greater window lengths. These steps are visualized in Fig. 3.

C. Characterizing Greenland interstadials and stadials

Many of the abrupt transitions in the NGRIP δ^{18}O record correspond to shifts between a warmer climate during Greenland Interstadials (GIs) and a colder climate during Greenland Stadials (GSs). In order to determine whether a transition in the proxy record marks the start of a GS or a GI mode, we computed two additional time series for the proxy record and plot them in Fig. 4: they correspond to the stadial and interstadial values in the record.

To transcribe the bistable behavior, observed at varying characteristic time scales in climate records, we employ a running window w_i of varying length, according to Eq. (4), to extract samples from the w_i-filtered time series. These samples are then divided into two sets, below and above the mean value of the series, for each w_i. The GS values are then calculated from the 25th percentile of the set below the mean, while GI values are calculated from the 75th percentile of the set above the mean.

These GS and GI values calculated with each w_i window are then averaged, yielding the blue ("stadial") and the red ("interstadial") curves in Fig. 4. These two curves are then averaged at each time t, yielding the purple curve in Fig. 4. These curves help visually

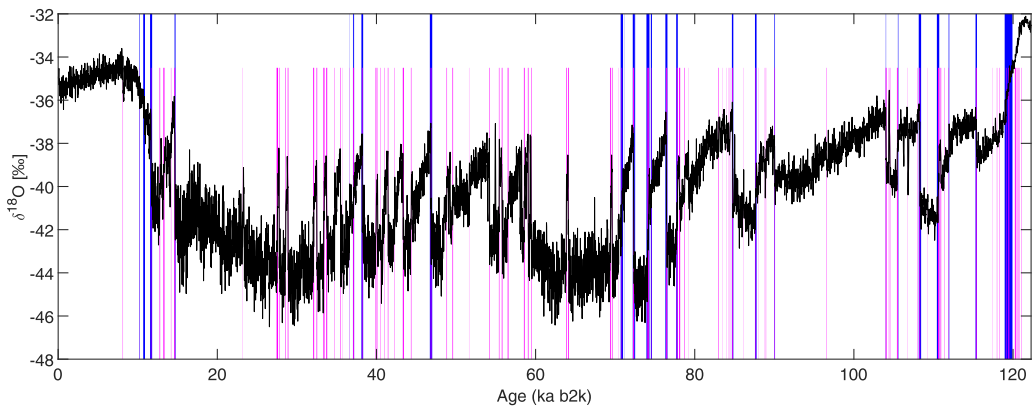

FIG. 2. Main transitions detected for the NGRIP δ^{18}O record[58] (black line) that satisfy the conditions in Eqs. (5)–(7) using windows of length $w = 1000$ yr (long, blue bars) and 200 yr (short, magenta bars).

Published under an exclusive license by AIP Publishing

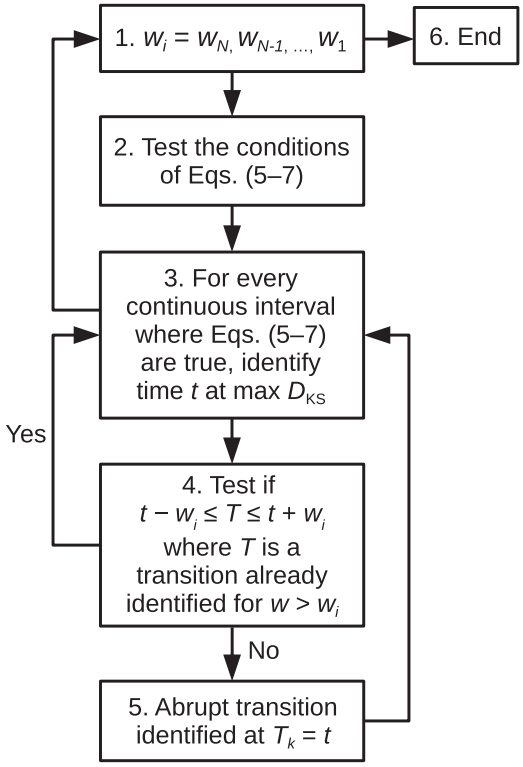

FIG. 3. Schematic of the steps involved in identifying abrupt transitions with a varying window length w_i. Steps 2–5 are repeated for every window length defined in step 1, starting with the longest window w_N and ending with w_1. Steps 3–5 are repeated for every continuous interval within which the conditions in Eqs. (5)–(7) are satisfied. Step 4 is used to discard transitions that match those already detected with a longer window. Abrupt transitions are identified in step 5.

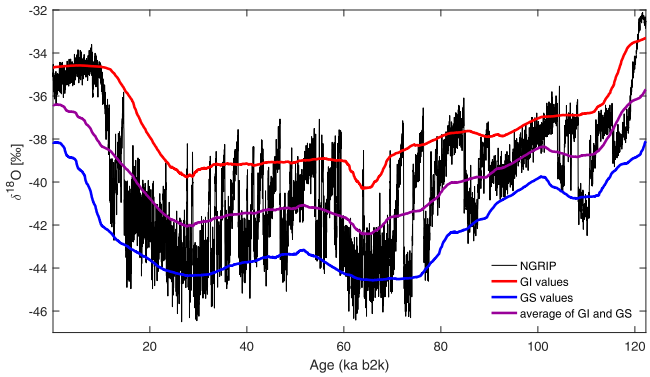

FIG. 4. NGRIP δ^{18}O record[58] (black line). The red line corresponds to the calculated interstadial values (GI), the blue line corresponds to the calculated stadial values (GS), and the purple line is the average between the red and blue lines. The purple line represents the threshold that is used to identify shifts between stadials and interstadials.

identify dominant GS and GI regimes within the time series and are compared with the proxy values at the detected transitions. When a detected transition coincides with the proxy record crossing the middle curve (purple), it is recognized as corresponding to a shift between a stadial and an interstadial.

D. ROC analysis of jump identification in paleodata

A receiver operating characteristic (ROC) curve illustrates the diagnostic capabilities of a binary classifier.[24,33] ROC curves were first used during World War II to measure the ability of a radar receiver operator to correctly detect enemy aircraft from radar signals.[53,57] A binary classifier classifies the elements of a given set into two groups, such as "pass" or "fail."[62] Here, the ROC curves are used to compare the efficacy of different criteria used in our augmented KS methodology for detecting abrupt transitions, as described in Sec. II B above, and to optimize its parameters.

The conditions in Eqs. (5) and (6) are binary classifiers, whose parameters D_c and σ_c we wish to tune to bring our results closer to a desired goal. In the case at hand, the goal is provided by the heuristically determined change points of Rasmussen *et al.*,[58] which we compare with the abrupt transitions identified in Fig. 2 of Sec. II B. The two ROC curves for D_c (red) and σ_c (blue) are plotted in Fig. 5.

In order to plot such curves, we must first define the following concepts and numerical definitions:

- True positive or "hit": a positive classifier value that corresponds to a known positive; for example, if the condition in Eq. (5) is "pass" at $t = 28\,900$ yr, a time that corresponds to the Start of GI-4 given by Rasmussen *et al.*[58]

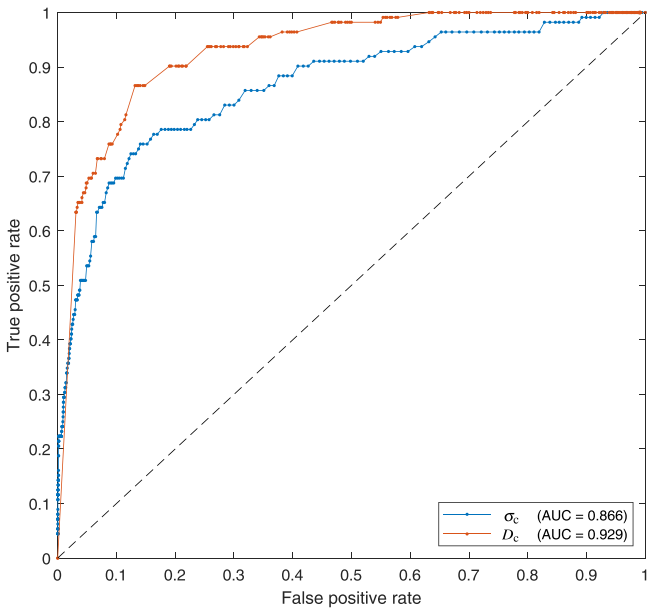

FIG. 5. ROC curves for the parameters D_c (red) and σ_c (blue), which control minimum jump size. Please see the text for details.

Published under an exclusive license by AIP Publishing

- True negative or "correct rejection": a negative classifier value that corresponds to a known negative; for example, if the condition in Eq. (5) is "fail" at $t = 20\,000$ yr, a time that does not correspond to any event given by Rasmussen *et al.*[58]
- False positive or "false alarm": a positive classifier value that corresponds to a known negative; for example, if the condition in Eq. (5) is "pass" at $t = 20\,000$ yr, a time that does not correspond to any event of Rasmussen *et al.*[58]
- False negative or "miss": a negative classifier value that corresponds to a known positive; for example, if the condition in Eq. (5) is "fail" at $t = 28\,900$ yr, it means that it does not agree with the identification of the Start of GI-4 given by Rasmussen *et al.*[58]
- True positive rate: $\mathrm{TPR} = \frac{\text{sum of true positives}}{\text{all known positives}}$.
- False positive rate: $\mathrm{FPR} = \frac{\text{sum of false positives}}{\text{all known negatives}}$.

A good classifier is one that finds as many true positives and true negatives as possible, and as few false positives and false negatives as possible. Hence, we aim for a TPR that is close to 1 and an FPR that is close to 0.

Note that, given the known imperfections in the NGRIP record's dating, we allow for a difference of up to 30 yr when comparing the jumps identified by our augmented KS methodology with those determined by the subjective expert methods used by Rasmussen *et al.*[58] Some similar tolerance has to be used in any comparison of change point analysis methods for time series with dating uncertainties.

Every point on the two ROC curves in Fig. 5 corresponds to a different parameter value: for example, the values of D_c range from 0 to 1. Point (0, 0) on the graph corresponds to parameter values that always yield a negative result, i.e., such a value does not identify any jumps in the record. Point (1, 1) corresponds to parameter values that always yield a positive result, i.e., such a value identifies a jump at every single step in the record. One seeks an optimal balance between the two extremes by finding the point on the ROC curve that is closest to the error-free point (0, 1), calculated as the maximum value of $J = \mathrm{TPR} - \mathrm{FPR}$, where J is the Youden[76] J-statistic.

The area under the curve (AUC) is an important measure of the performance of a classifier.[33,48] The closer the ROC curve is to point (0, 1) on the graph, the higher its AUC and the better the classifier's diagnostic ability. Thus:

- If AUC = 1, the classifier is "perfect"; e.g., in our case, it finds the exact same jumps as in Rasmussen *et al.*[58]
- If AUC is close to 1, the classifier is "good."
- If AUC is only a little above 0.5, the classifier is "poor."
- If AUC = 0.5, i.e., the area under the curve is the same as the area under the diagonal in Fig. 5, the classifier is considered to be "useless," i.e., only as good as a random guess.
- If AUC $<$ 0.5, the classifier is doing the opposite of what was intended.

In Fig. 5, the ROC curves for the parameters D_c and σ_c have AUCs of 0.929 and 0.866, respectively. Thus, both classifiers are quite good, but Eq. (5) is the better one of the two.

The optimal values of D_c and σ_c are 0.735 and 1.92, respectively. However, since these values are obtained independently of one another, they do not equal necessarily the optimal values that can be computed through multi-dimensional optimization. In the meantime, given that the separate AUCs are quite high, we use in practice parameter values that are slightly different from the values above.

III. APPLICATIONS

A. Greenland ice core record

As discussed in Sec. II, our augmented KS test has been tuned to identify jumps that match the stadial–interstadial boundaries identified heuristically for the NGRIP ice core record.[58] Following the optimization via ROC analysis, we chose the parameter values $D_c = 0.77$ and $\sigma_c = 1.9$ for the criteria given by Eqs. (5) and (6). The range of window lengths w_i in Eq. (4) corresponds to the centennial-to-millennial time scale of glacial–interglacial transitions, and thus we used $w_{\min} = 0.1$ kyr and $w_{\max} = 2.5$ kyr to calculate D_{KS}. The minimum sample size used is $n_c = 3$.

The GS and GI boundaries in Rasmussen *et al.*[58] were established using both δ^{18}O and Ca^{2+} proxies from the NGRIP record. In the present analysis, though, for the sake of simplicity and to allow a more direct comparison between ice core and speleothem records, we only used the δ^{18}O proxy values. The results of this augmented KS test for the NGRIP δ^{18}O record are shown in Fig. 6. Note that the interstadials associated with the gray bars in the figure are not necessarily delimited by the most abrupt transitions: they were identified by the stadial–interstadial boundaries determined following the procedure described in Sec. II C, by relying on the additional information contained in Fig. 4.

Our method properly identifies most of the abrupt transitions described by Rasmussen *et al.*,[58] including Dansgaard–Oeschger events and glacial–interglacial transitions. In addition to the stadial–interstadial transitions, the KS method also identifies abrupt transitions that are not associated with these climatic shifts. The total number of transitions detected is, therefore, higher than in Rasmussen *et al.*[58]

For example, while Rasmussen *et al.*[58] show a firm boundary between the Eemian interglacial and GS-26 at 119 140 yr b2k, we find this transition to be occurring over several steps, between 120 760 and 118 320 yr b2k. We also find several such events during GI-21.1, between 84 760 yr b2k and 77 760 yr b2k and during GI-23.1, between 104 020 yr b2k and 91 740 yr b2k, as well as distinct events during stadials at 109 300 yr b2k, 86 160 yr b2k, and 67 260 yr b2k.

Furthermore, several notable differences in timing of the stadial–interstadial transitions are observed. Here, GI-23.1 is found to end at 91 740 yr b2k, an earlier time than the 90 140 yr b2k given by Rasmussen *et al.*,[58] where the following GS-23.1 is identified as a "quasi-stadial" and a full GS only begins at 87 600 yr b2k. Furthermore, the GI-5.1 interstadial, at 30 840–30 600 yr b2k, is not identified here at all, while in the place of GI-2.2 and GI-2.1, we only identify a single warming event at 23 360 yr b2k. The most likely reason is that δ^{18}O is the only proxy used in the present analysis, while Rasmussen *et al.*[58] used Ca^{2+} as well.

Published under an exclusive license by AIP Publishing

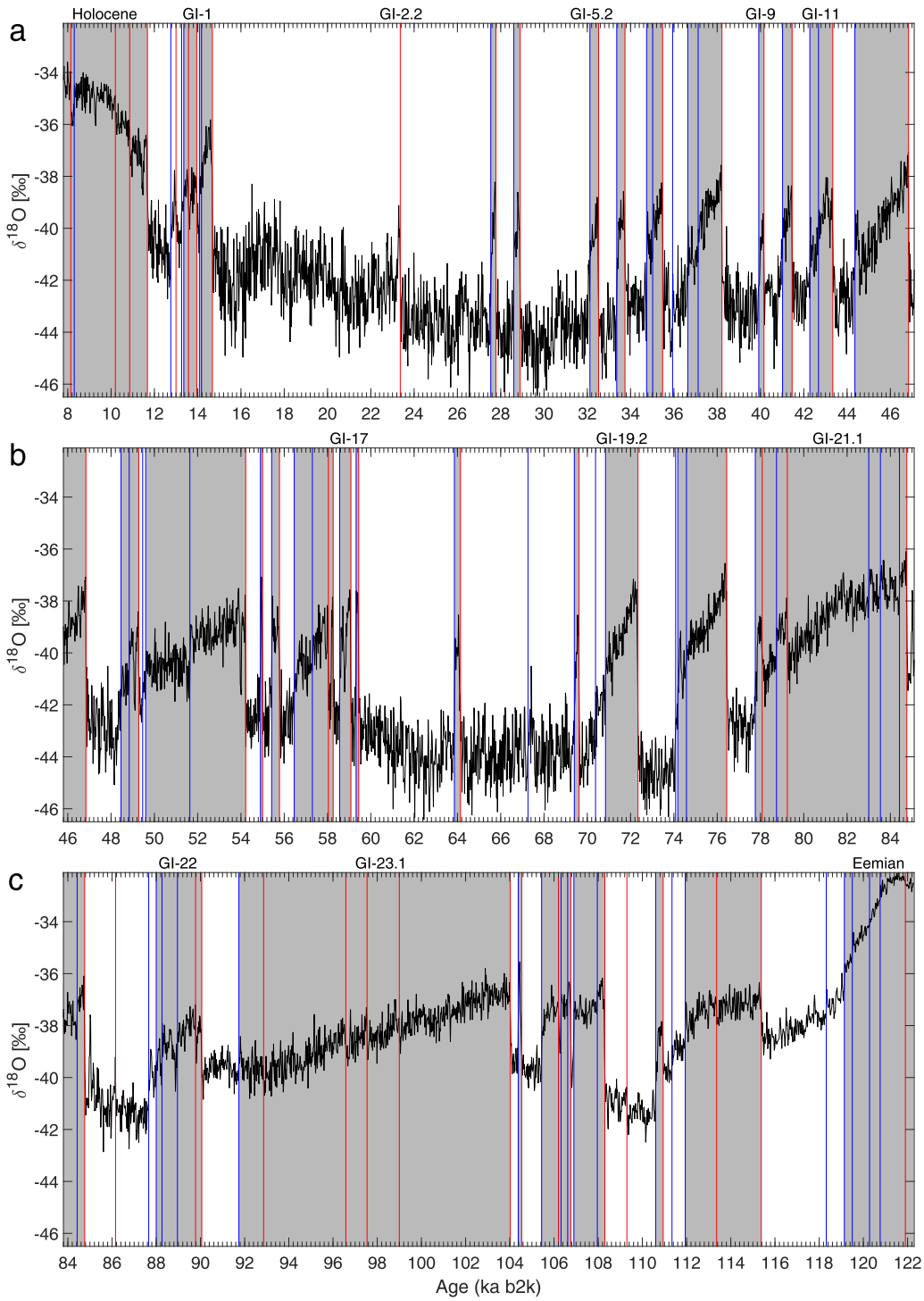

FIG. 6. NGRIP δ^{18}O record[58] (black line): (a) 7.8–47.1 ka b2k; (b) 45.8–85.1 ka b2k; and (c) 83.8–122.3 ka b2k. Vertical lines represent detected transitions, with red lines for warming transitions and blue lines for cooling ones. Grey bars represent interstadials.

Published under an exclusive license by AIP Publishing

B. China speleothem record

In order to determine whether our augmented KS test may be successfully applied to other types of records with different time scales and resolutions, and subject to different nonlinear trends, we applied the test to a composite speleothem record.[12] This composite δ^{18}O record was constructed using multiple speleothem records from three caves in China, the Hulu, Sanbao, and Dongge caves,[72,73,77] and it is reproduced here in Fig. 7.

While this speleothem record extends back to 641 ka BP, we apply our methodology only to the time interval covered by the NGRIP record, i.e., the past 122 ka. To obtain a clean comparison with the results of Sec. III A, we run the method with the same settings and using the same parameter values as there. The abrupt transitions identified therewith, including stadial–interstadial changes, are shown in Fig. 7.

Our method identifies the start of the Holocene, at 11 510 yr b2k; GS-1 or Younger Dryas, at 12 800 yr b2k; the GI-1, at 14 620 yr b2k; most of the known Dansgaard–Oeschger events; as well as the beginning of the Last Glacial Period, at 120 450 yr b2k. Due to a lower resolution of the speleothem record, as well as to a generally less abrupt nature of the warming events, fewer transitions are identified than in the NGRIP record; those missed include the one attributed to the GI-18 interstadial at 64 100 yr b2k and, most strikingly, the entire GI-25 interstadial at 110 640–115 370 yr b2k.

Some of the transitions are identified at a different time than in the NGRIP record, e.g., the beginning of the Last Glacial Period is identified to occur approximately 1300 yr earlier. Such differences are likely the result of different dating methods used for the two records and are consistent with uncertainties that arise from layer-counting-based dating of ice cores.[8] Some events, however, display a very different pattern, e.g., GI-23.1 ends at 99 200 yr b2k and is followed by an additional cooling transition at 98 070 yr b2k, while in the NGRIP record, a significant cooling transition during GI-23.1 does not take place until 91 740 yr b2k, which marks, therewith, the end of this interstadial. GI-5.1, while not identified in the NGRIP record at all, is found here to end at 30 660 yr b2k and, as this date is not preceded by a warming transition, appears as an extension of GI-5.2.

C. Comparison between the Greenland and China records

The results in Sec. III B demonstrate that the KS detection method, as augmented herein, may be applied successfully to different types of records, of different origin, resolution, and other characteristics. Given the well-established transition dates in Rasmussen *et al.*,[58] the ROC analysis has only been used here to tune the method for the NGRIP record.

One may be able to improve the results in Sec. III B further by additional tuning of the method's parameters to account for the differences between speleothem and ice core records, in particular, the varying resolution of the former. Here, we merely wish to demonstrate the applicability of the methodology to different types of records; hence, a more accurate identification of abrupt transitions is not our primary goal.

Although we have not performed specific tuning for the China cave composite record, it is quite encouraging that the method was nevertheless capable of finding transitions that closely match those in the NGRIP record. The differences in our results between the two records are most likely due to their different age models as well as to actual differences in climate evolution between Greenland and China; they are thus less likely to be caused by inconsistencies in the method itself. Furthermore, note that the NGRIP record comes from a single, continuous ice core, while the Chinese speleothem record is a composite from three caves many hundreds of kilometers apart. Therefore, the differences in geochemical properties of the individual speleothems, as well as regional climate differences, may affect the continuity of this record.

D. Comparison with recurrence analysis

To further evaluate the methodology of this paper, we compare the results from the transition detection method for the NGRIP δ^{18}O record with an analysis based on the recurrence plots (RPs) introduced by Eckmann *et al.*[19] and popularized in the climate sciences by Marwan *et al.*[45,46] The RP for a time series $\{x_k : k = 1, \ldots, K\}$ is constructed as a square matrix in a Cartesian plane with the abscissa and ordinate both corresponding to a time-like axis, with one copy x_i of the series on the abscissa and another copy x_j on the ordinate. A dot is entered into a position (i, j) of the matrix when x_j is sufficiently close to x_i, i.e., $|x_i - x_j| < \varepsilon$ with ε being the recurrence threshold. For the details—such as how close is "sufficiently close"—we refer to Eckmann *et al.*[19] and Marwan *et al.*[46]

Clearly, all the points on the diagonal $i = j$ have dots and, in general, the matrix is rather symmetric, although one does not always define closeness symmetrically; to wit, x_j may be "closer to" x_i than x_i is to x_j.[19] An important advantage of the recurrence method is that it does apply to dynamical systems that are not autonomous, i.e., that may be subject to time-dependent forcing. The latter is certainly the case for the climate system on time scales of 10–100 kyr and longer, which is affected strongly by orbital forcing.

Eckmann *et al.*[19] distinguished between large-scale typology and small-scale texture in the interpretation of the square matrices of dots that are the visual result of RP. Thus, if all the characteristic times of an autonomous dynamical system are short compared to the length of the time series, the RP's typology will be homogeneous and, thus, not very interesting. In the presence of an imposed drift, a more interesting typology will appear.

The most interesting typology in RP applications so far is associated with recurrent patterns that are not exactly periodic but only nearly so. Hence, such patterns are not that easily detectable by purely spectral approaches to time series analysis. Marwan *et al.*[46] discuss how to render the purely visual RP typologies studied up to that point more objectively quantifiable by recurrence quantification analysis (RQA)[45] and bootstrapping.[20,21]

While the selection of an optimal recurrence threshold ε is not straightforward, several rules of thumb have been proposed.[45] For example, it has been suggested that it should not exceed 10% of the maximum phase space diameter[51,78] and that it should be at least five times larger than the standard deviation of the observational noise,[68] a value which is difficult to estimate for ice core δ^{18}O, but likely exceeds 0.2‰. Thus, we chose a ε of 1.3‰, which is close to the maximum value allowed by the first condition. The RP of NGRIP δ^{18}O obtained using this threshold is shown in Fig. 8, middle panel.

Published under an exclusive license by AIP Publishing

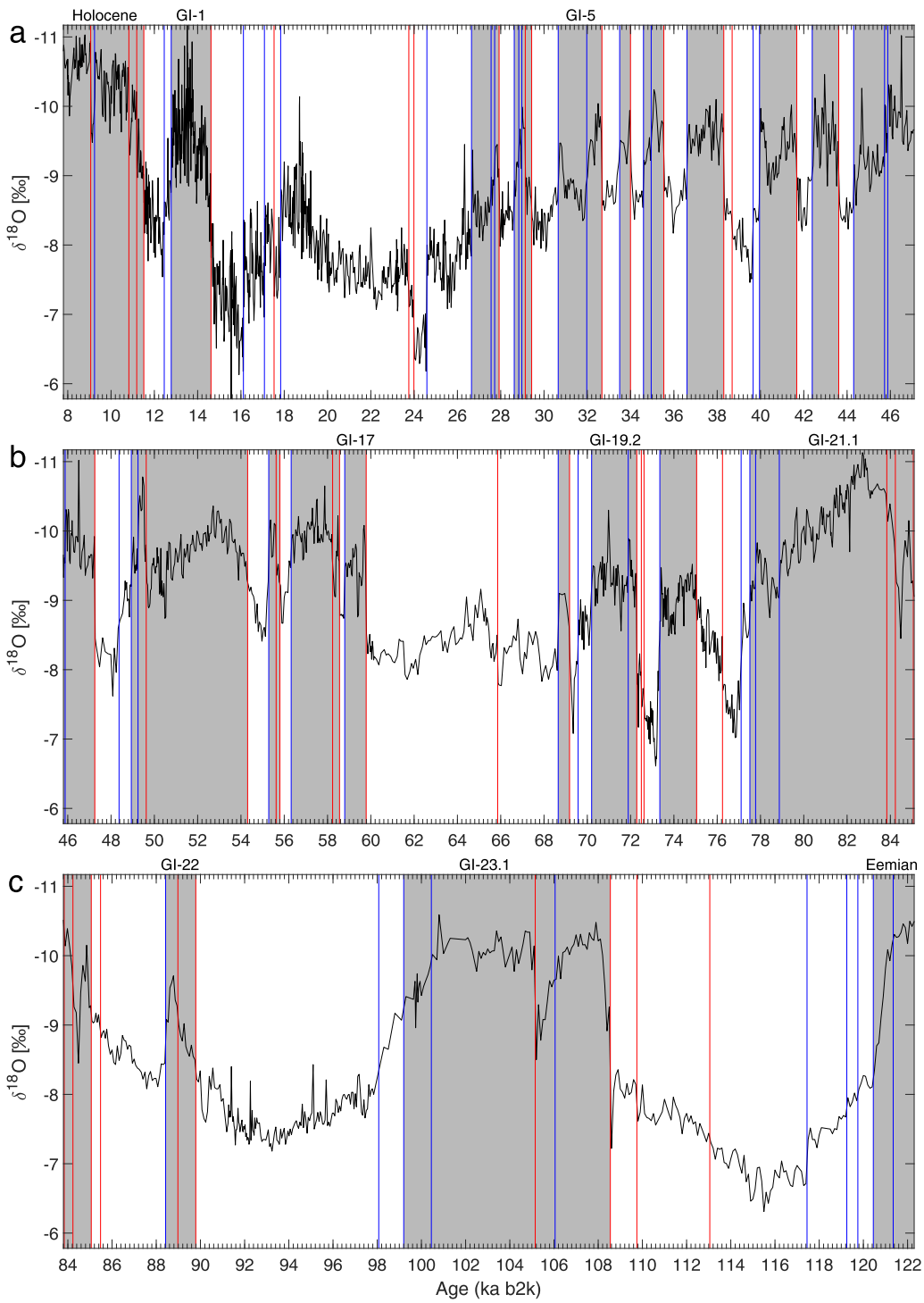

FIG. 7. China cave δ^{18}O composite record[12] showing the last climate cycle (black line) and the abrupt transitions therein: (a) 7.8–47.1 ka b2k; (b) 45.8–85.1 ka b2k; and (c) 83.8–122.3 ka b2k. Same conventions as in Fig. 6.

Published under an exclusive license by AIP Publishing

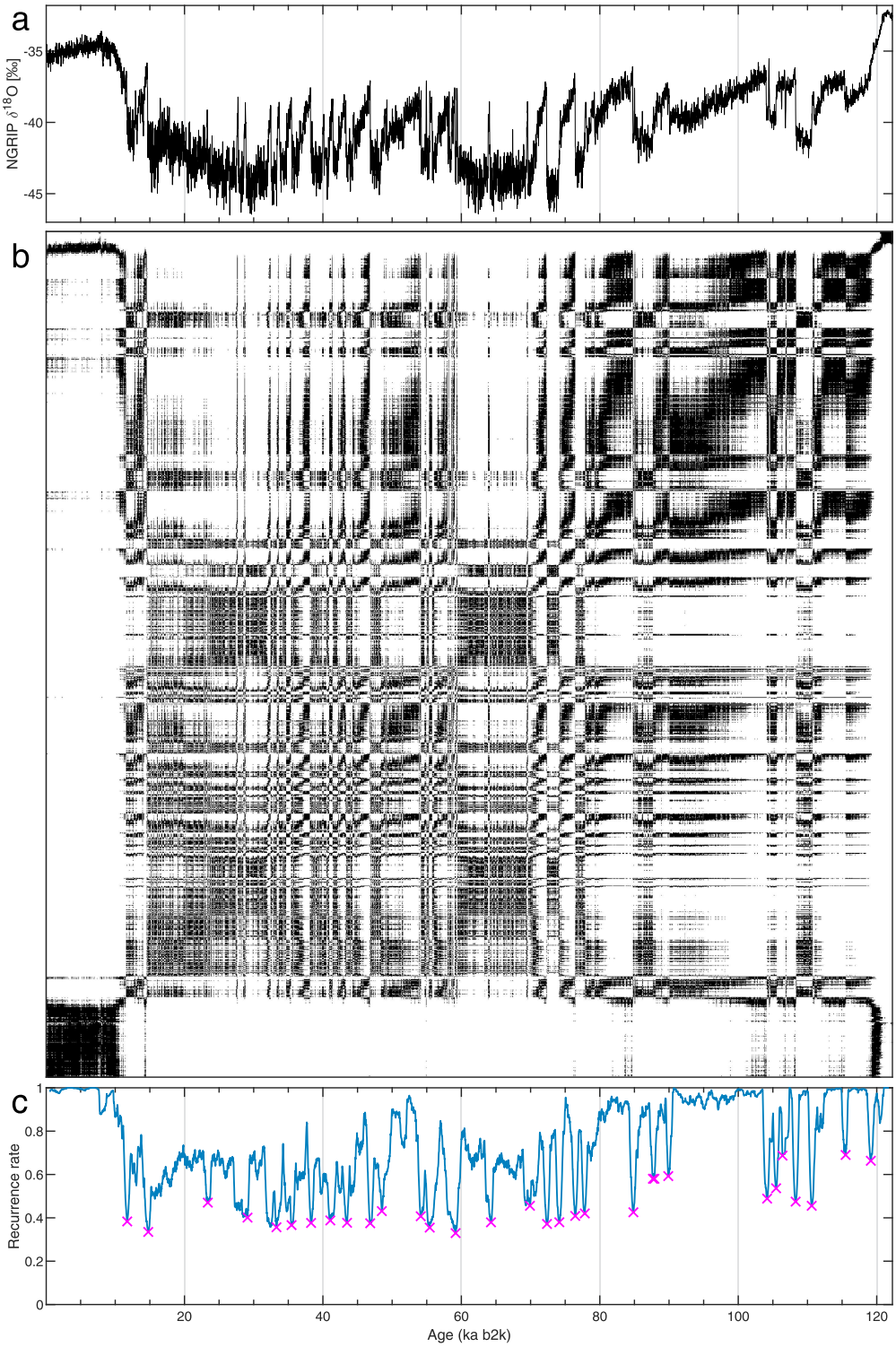

FIG. 8. Comparison of visual and quantitative identification of recurrences in the NGRIP record:[58] (a) time series; (b) recurrence plot (RP); and (c) recurrence rate (RR) measures. Magenta crosses represent local RR minima and they correspond to abrupt transitions.

Published under an exclusive license by AIP Publishing

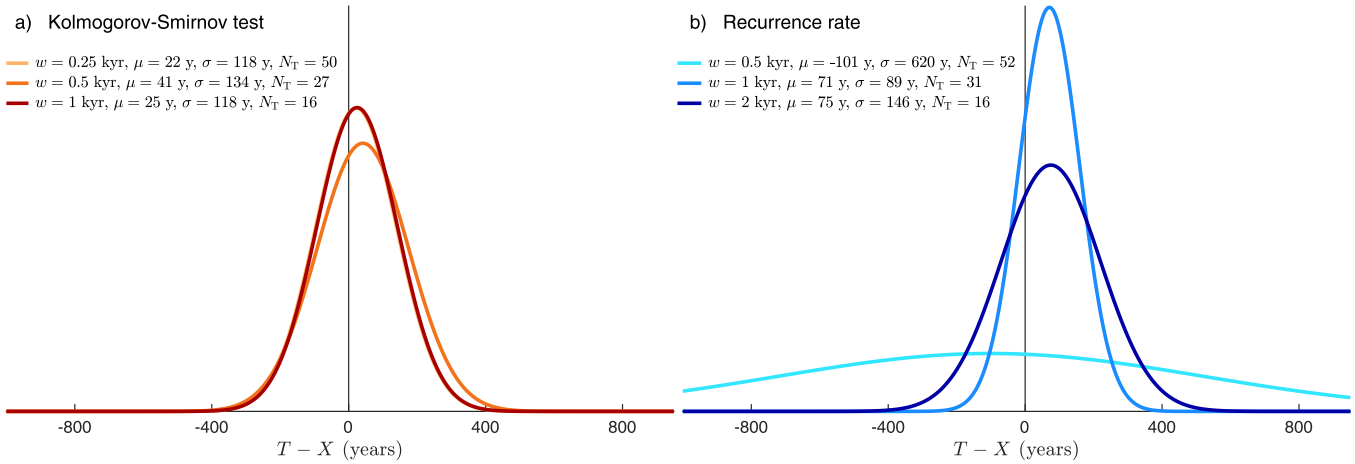

FIG. 9. Difference between abrupt transitions detected in the NGRIP δ^{18}O record with the KS and RQA methods (T_i) and their closest neighboring transitions in Rasmussen *et al.*[58] (X_i); normal distributions for different window lengths w were calculated from the means and standard deviations of $T_i - X_i$. (a) Results of the augmented KS test; and (b) RQA results using the RR. The Y-axes have the same scale in both panels. As the distributions for $w = 0.25$ kyr and $w = 1$ kyr in panel (a) are nearly identical, the lighter curve is hidden behind the darker curve.

Following Ref. 45, visual inspection of RPs may be complemented by RQA, which quantifies selected recurrence characteristics (see also Ref. 69). One of the simplest RQA measures is the recurrence rate (RR), namely, the density of dots within the recurrence plot, which describes the probability of states of the system recurring within a particular time interval. By evaluating RQA measures such as RR in a sliding window, it is possible to identify changes in the time series. The bottom panel of Fig. 8 shows the RR computed over a 1000-yr sliding window; this window was found to have the most effective length, among those evaluated, for identifying abrupt transitions (Fig. 9).

Low RR values correspond to an unstable behavior of the system. Thus, abrupt transitions in a time series may be identified by the local RR minima. As it is often unclear whether neighboring minima represent several distinct events or a single event, we select local minima with the greatest prominence, a measure originally used in mountaineering and geography to distinguish independent mountains from sub-peaks[25,35] and, in recent years, successfully applied in other fields.[1,32,65] Prominence is defined as the shortest vertical distance between a peak and a saddle that connects it to any higher peak; if no higher peak exists, the peak's prominence is equal to its elevation. In finding minima, the sign of RR is reversed. RR minima with a prominence greater than the standard deviation of the RR time series mark the dates of the abrupt transitions identified by windowed RQA (magenta crosses in Fig. 8).

The major transitions identified by our KS method in Fig. 6 are also identified by RQA in Fig. 8, but properly recognizing the transition points by RQA becomes more challenging at time scales shorter than the window length. For example, the transitions between GI-17.2 and GI-16.1c, at 59 440 and 58 040 yr b2k, respectively, are properly identified by the KS method but they are not easy to distinguish on the RP and only one transition, at 59 170 yr b2k, is found by RQA. Likewise, none of the sub-events during GI-1 between 14 075 and 12 896 yr b2k are identified by RQA. Furthermore, for the Dansgaard–Oeschger events of GI-18 to GI-2.1 at 64 100 to 22 900 yr b2k, RQA only identifies either the start or the end of the event.

On the other hand, recurrence analysis appears to be more useful for recognizing changes in periodicity in paleorecords. Such changes are clearly visible in the RP at 71 000 yr b2k (GI-19.2) and at 14 700 yr b2k (start of GI-1), when more frequent interstadials occur and the record is characterized by greater variance than before 71 000 yr b2k or during the Holocene. While these two transitions have also been identified with the KS method, there is no clear distinction between them and the other transitions in Fig. 6.

We have seen that the augmented KS test uses a variable window length to find transitions at different time scales. For comparing it with RQA, though, it is desirable to use equivalent window lengths for both methods. We show, therefore, in Fig. 9 a general comparison of the two methods, including the effect of the window length on their performance.

In the figure, abrupt transitions are identified using three different window lengths. Windows used for the KS test are half the length of those used in RQA because the KS test compares two neighboring samples of a time series, while the RR is calculated for a single sample. For both methods, we calculated the difference in years between each of the detected transitions (T_i) and their closest neighboring transitions in Rasmussen *et al.*[58] (X_i).

For better visual presentation, the distribution of these differences was modeled using a normal distribution. While varying the window length has a significant impact on the total number N_T of identified transitions in both methods, the dates of these transitions are consistently in good agreement with those of Rasmussen *et al.*[58] only for the KS test; in the latter case, $T_i - X_i$ has means ranging between 22 and 41 yr and standard deviations ranging between 118 and 134 yr. Transitions found by RR are in their best agreement with those of Rasmussen *et al.*[58] for $w = 1$ kyr, where $T_i - X_i$ has

Published under an exclusive license by AIP Publishing

a mean of 71 yr and a standard deviation of 89 yr; however, for both the longer and the shorter windows the standard deviation is much larger.

This analysis suggests that the accuracy of the detected transitions is less sensitive to window length for the augmented KS test than for the RR test of RQA. The KS test is, therefore, the more robust method for establishing precise dates of abrupt transitions, while the transition dates detected by RQA are inconsistent from one window length to another.

IV. CONCLUSIONS

We have shown here that the Kolmogorov–Smirnov (KS) test, which by itself is a powerful tool for quantifying discontinuities in a time series, may be adapted for paleoclimate data analysis when combined with additional statistical tests. Given its statistical robustness and relative simplicity, the methodology described in this paper is very effective in detecting abrupt transitions in paleoproxy records.

We tested the applicability of the methodology on two paleorecords from the last climate cycle. In future studies, we aim to apply this method to various records with a different origin, nature, and length, which originated on different continents and ocean basins. A database of paleoproxy records with automatically detected transitions, currently under development[4] will include a much more extensive application of the augmented KS methodology. Here, the parameters required were conveniently tuned for the NGRIP record, using the transition dates of Rasmussen *et al.*[58]

To evaluate time series with different resolutions, periodicities, noise levels, and strength of the signal, different settings may be used. The D_c parameter specifies the desired level of "abruptness," n_c relates to the record's resolution and to the desired degree of robustness of the KS test, while the σ_c parameter specifies the magnitude of the jumps in relation to background noise. By modifying the window length used for evaluating the KS statistic, the transitions may be found for different time scales of interest.

A good understanding of the record's characteristics is quite helpful in selecting the desired parameter values and it can be aided further by using automated methods for determining them. The formulations of Eqs. (6) and (7) that target the additional criteria of rate-of-change and minimum sample size have improved the method's ability to focus on the most distinct transitions. Still, the introduction of a variable window length in Eq. (4) and the procedure for identifying distinct transitions among those found with different window lengths (cf. Fig. 3) are our two most critical modifications of the KS test: they allow one to apply it to study a record under distinct "magnification lenses."

Comparison of our augmented KS method with recurrence analysis, a frequently used method for transition detection in paleoclimate studies, indicates that our KS method is more useful at determining individual jumps in the record and finding their precise dates, particularly when variable time scales are involved. On the other hand, recurrence analysis may help establish particularly important transitions that correspond to a change in a record's characteristic time scale. The two methodologies appear thus to be both useful and to complement each other.

Since the KS methodology gives precise dates for the abrupt transitions it identifies, its wider application may help reconstruct the chronology of Earth's climatic events and build improved age models for records in which "wiggle matching" is typically used as the dating method. Furthermore, this objective approach to identifying abrupt climate transitions can improve our understanding of the Earth's bifurcation mechanisms and TPs. This, in turn, will allow us to construct better nonlinear and stochastic models of the Earth's climate and its interactions with ecosystems.[28,30]

ACKNOWLEDGMENTS

It is a pleasure to thank two anonymous reviewers for their constructive comments, which have improved the presentation of our results. This paper is Tipping Points in the Earth System (TiPES) Contribution No. 126. This study was funded by the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 820970). This is LDEO contribution.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The NGRIP record is openly available from the Centre for Ice and Climate, University of Copenhagen at https://www.iceandclimate.nbi.ku.dk/data, Ref. 58 and the Chinese cave composite record is in supplementary Table 1 on the Nature website at https://doi.org/10.1038/nature18591, Ref. 12.

REFERENCES

[1] M. Alho, C. S. Wedlund, H. Nilsson, E. Kallio, R. Jarvinen, and T. Pulkkinen, "Hybrid modeling of cometary plasma environments—II. Remote-sensing of a cometary bow shock," Astron. Astrophys. **630**, A45 (2019).
[2] 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," Philos. Trans. R. Soc. A **370**(1962), 1166–1184 (2012).
[3] W. Bagniewski, K. J. Meissner, and L. Menviel, "Exploring the oxygen isotope fingerprint of Dansgaard–Oeschger variability and Heinrich events," Quat. Sci. Rev. **159**, 1–14 (2017).
[4] W. Bagniewski, D.-D. Rousseau, and M. Ghil, "PaleoJump database for research on rapid climate transitions," EGU General Assembly, EGU21-9968, 2021.
[5] S. Barker, J. Chen, X. Gong, L. Jonkers, G. Knorr, and D. Thornalley, "Icebergs not the trigger for North Atlantic cold events," Nature **520**(7547), 333–336 (2015).
[6] C. Beaulieu, J. Chen, and J. L. Sarmiento, "Change-point analysis as a tool to detect abrupt climate variations," Philos. Trans. R. Soc. A **370**(1962), 1228–1249 (2012).
[7] N. Boers, M. Ghil, and D.-D. Rousseau, "Ocean circulation, ice shelf, and sea ice interactions explain Dansgaard–Oeschger cycles," Proc. Natl. Acad. Sci. U.S.A. **115**(47), E11005–E11014 (2018).
[8] N. Boers, B. Goswami, and M. Ghil, "A complete representation of uncertainties in layer-counted paleoclimatic archives," Clim. Past **13**(9), 1169–1180 (2017).
[9] W. S. Broecker, D. M. Peteet, and D. Rind, "Does the ocean–atmosphere system have more than one stable mode of operation?," Nature **315**(6014), 21–26 (1985).
[10] M. A. Cane, "A role for the tropical Pacific," Science **282**(5386), 59–61 (1998).
[11] E. Carlstein *et al.*, "Nonparametric change-point estimation," Ann. Stat. **16**(1), 188–197 (1988).

Published under an exclusive license by AIP Publishing

[12]H. Cheng, R. L. Edwards, A. Sinha, C. Spötl, L. Yi, S. Chen, M. Kelly, G. Kathayat, X. Wang, X. Li *et al.*, "The Asian monsoon over the past 640 000 years and ice age terminations," Nature **534**(7609), 640–646 (2016).
[13]A. C. Clement, M. A. Cane, and R. Seager, "An orbitally driven tropical source for abrupt climate change," J. Clim. **14**(11), 2369–2375 (2001).
[14]G. W. Cobb, "The problem of the Nile: Conditional solution to a changepoint problem," Biometrika **65**(2), 243–251 (1978).
[15]W. J. Conover, *Practical Nonparametric Statistics* (John Wiley & Sons, 1999).
[16]W. Dansgaard, H. Clausen, N. Gundestrup, C. Hammer, S. Johnsen, P. Kristinsdottir, and N. Reeh, "A new Greenland deep ice core," Science **218**(4579), 1273–1277 (1982).
[17]W. Dansgaard, S. J. Johnsen, H. B. Clausen, D. Dahl-Jensen, N. Gundestrup, C. Hammer, C. Hvidberg, J. Steffensen, A. Sveinbjörnsdottir, J. Jouzel *et al.*, "Evidence for general instability of past climate from a 250-kyr ice-core record," Nature **364**(6434), 218–220 (1993).
[18]J. F. Donges, R. V. Donner, M. H. Trauth, N. Marwan, H. -J. Schellnhuber, and J. Kurths, "Nonlinear detection of paleoclimate-variability transitions possibly related to human evolution," Proc. Natl. Acad. Sci. U.S.A. **108**(51), 20422–20427 (2011).
[19]J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," Europhys. Lett. **4**(9), 973–977 (1987).
[20]B. Efron, "Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods," Biometrika **68**(3), 589–599 (1981).
[21]B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," Stat. Sci. **1**(1), 54–75 (1986).
[22]D. Eroglu, N. Marwan, S. Prasad, and J. Kurths, "Finding recurrence networks' threshold adaptively for a specific time series," Nonlinear Process. Geophys. **21**(6), 1085–1092 (2014).
[23]D. Eroglu, F. H. McRobie, I. Ozken, T. Stemler, K.-H. Wyrwoll, S. F. Breitenbach, N. Marwan, and J. Kurths, "See-saw relationship of the Holocene East Asian-Australian summer monsoon," Nat. Commun. **7**, 12929 (2016).
[24]T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett. **27**(8), 861–874 (2006).
[25]S. Fry, "Defining and sizing-up mountains," Summit **33**(1), 16–21 (1987).
[26]C. Gallagher, R. Lund, and M. Robbins, "Changepoint detection in climate time series with long-term trends," J. Clim. **26**(14), 4994–5006 (2013).
[27]A. Ganopolski and S. Rahmstorf, "Rapid changes of glacial climate simulated in a coupled climate model," Nature **409**(6817), 153–158 (2001).
[28]M. Ghil, "A century of nonlinearity in the geosciences," Earth Space Sci. **6**, 1007–1042 (2019).
[29]M. Ghil and S. Childress, *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics* (Springer Science+Business Media, Berlin, 1987), Reissued as an eBook in 2012.
[30]M. Ghil and V. Lucarini, "The physics of climate variability and climate change," Rev. Mod. Phys. **92**(3), 035002 (2020).
[31]B. Goswami, N. Boers, A. Rheinwalt, N. Marwan, J. Heitzig, S. F. Breitenbach, and J. Kurths, "Abrupt transitions in time series with uncertainties," Nat. Commun. **9**(1), 1–10 (2018).
[32]J. Griffié, L. Boelen, G. Burn, A. P. Cope, and D. M. Owen, "Topographic prominence as a method for cluster identification in single-molecule localisation data," J. Biophoton. **8**(11-12), 925–934 (2015).
[33]T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer Science+Business Media, 2009).
[34]H. Heinrich, "Origin and consequences of cyclic ice rafting in the Northeast Atlantic Ocean during the past 130 000 years," Quat. Res. **29**(2), 142–152 (1988).
[35]A. Helman, *The Finest Peaks: Prominence and Other Mountain Measures* (Trafford Publishing, 2005).
[36]C. Huber, M. Leuenberger, R. Spahni, J. Flückiger, J. Schwander, T. F. Stocker, S. Johnsen, A. Landais, and J. Jouzel, "Isotope calibrated Greenland temperature record over Marine Isotope stage 3 and its relation to CH_4," Earth Planet. Sci. Lett. **243**(3–4), 504–519 (2006).
[37]S. Johnsen, H. Clausen, W. Dansgaard, K. Fuhrer, N. Gundestrup, C. Hammer, P. Iversen, J. Jouzel, B. Stauffer *et al.*, "Irregular glacial interstadials recorded in a new Greenland ice core," Nature **359**(6393), 311–313 (1992).
[38]A. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione," Giornale IIA **4**, 83–91 (1933).
[39]C. Kuehn, "A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics," Physica D **240**(12), 1020–1035 (2011).
[40]T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, "Tipping elements in the Earth's climate system," Proc. Natl. Acad. Sci. U.S.A. **105**(6), 1786–1793 (2008).
[41]L. E. Lisiecki and M. E. Raymo, "A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records," Paleoceanography **20**(1), PA1003, https://doi.org/10.1029/2004PA001071 (2005).
[42]L. E. Lisiecki and M. E. Raymo, "Plio–Pleistocene climate evolution: Trends and transitions in glacial cycle dynamics," Quat. Sci. Rev. **26**(1-2), 56–69 (2007).
[43]D. MacAyeal, "Binge/purge oscillations of the Laurentide Ice Sheet as a cause of the North Atlantic's Heinrich events," Paleoceanography **8**(6), 775–784, https://doi.org/10.1029/93PA02200 (1993).
[44]B. R. Markle, E. J. Steig, C. Buizert, S. W. Schoenemann, C. M. Bitz, T. Fudge, J. B. Pedro, Q. Ding, T. R. Jones, J. W. White *et al.*, "Global atmospheric teleconnections during Dansgaard–Oeschger events," Nat. Geosci. **10**(1), 36–40 (2017).
[45]N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," Phys. Rep. **438**(5-6), 237–329 (2007).
[46]N. Marwan, S. Schinkel, and J. Kurths, "Recurrence plots 25 years later—Gaining confidence in dynamical transitions," Europhys. Lett. **101**(2), 20007 (2013).
[47]N. Marwan, M. Thiel, and N. R. Nowaczyk, "Cross recurrence plot based synchronization of time series," Nonlinear Process. Geophys. **9**(3/4), 325–331 (2002).
[48]S. J. Mason and N. E. Graham, "Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation," Q. J. R. Meteorol. Soc. **128**(584), 2145–2166 (2002).
[49]F. J. Massey, Jr., "The Kolmogorov-Smirnov test for goodness of fit," J. Am. Stat. Assoc. **46**(253), 68–78 (1951).
[50]L. Menviel, A. Timmermann, T. Friedrich, and M. England, "Hindcasting the continuum of Dansgaard–Oeschger variability: Mechanisms, patterns and timing," Clim. Past **10**(1), 63–77 (2014).
[51]G. M. Mindlin and R. Gilmore, "Topological analysis and synthesis of chaotic time series," Physica D **58**(1-4), 229–242 (1992).
[52]J. M. Mitchell, Jr., "An overview of climatic variability and its causal mechanisms," Quat. Res. **6**(4), 481–493 (1976).
[53]J. M. Murphy, D. M. Berwick, M. C. Weinstein, J. F. Borus, S. H. Budman, and G. L. Klerman, "Performance of screening and diagnostic tests: Application of receiver operating characteristic analysis," Arch. Gen. Psychiatry **44**(6), 550–555 (1987).
[54]I. Ozken, D. Eroglu, T. Stemler, N. Marwan, G. B. Bagci, and J. Kurths, "Transformation-cost time-series method for analyzing irregularly sampled data," Phys. Rev. E **91**(6), 062911 (2015).
[55]F. Parrenin, J.-M. Barnola, J. Beer, T. Blunier, E. Castellano, J. Chappellaz, G. Dreyfus, H. Fischer, S. Fujita, J. Jouzel, K. Kawamura, B. Lemieux-Dudon, L. Loulergue, V. Masson-Delmotte, B. Narcisi, J.-R. Petit, G. Raisbeck, D. Raynaud, U. Ruth, J. Schwander, M. Severi, R. Spahni, J. P. Steffensen, A. Svensson, R. Udisti, C. Waelbroeck, and E. Wolff, "The EDC3 chronology for the EPICA Dome C ice core," Clim. Past **3**(3), 485–497 (2007).
[56]F. Pearce, *With Speed and Violence: Why Scientists Fear Tipping Points in Climate Change* (Beacon Press, 2007).
[57]W. W. T. G. Peterson, T. Birdsall, and W. Fox, "The theory of signal detectability," Trans. IRE Prof. Group Inf. Theory **4**(4), 171–212 (1954).
[58]S. O. Rasmussen, M. Bigler, S. P. Blockley, T. Blunier, S. L. Buchardt, H. B. Clausen, I. Cvijanovic, D. Dahl-Jensen, S. J. Johnsen, H. Fischer *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," Quat. Sci. Rev. **106**, 14–28 (2014).
[59]J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, "A review and comparison of changepoint detection techniques for climate data," J. Appl. Meteorol. Climatol. **46**(6), 900–915 (2007).
[60]D.-D. Rousseau, W. Bagniewski, and M. Ghil, "Abrupt climate changes and the astronomical theory," Clim. Past, 10.5194/cp-2021-103 (2021).

Published under an exclusive license by AIP Publishing

[61] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, "Early-warning signals for critical transitions," Nature **461**(7260), 53–59 (2009).
[62] B. Schölkopf, A. J. Smola, F. Bach *et al.*, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond* (MIT Press, 2002).
[63] R. Seager and D. S. Battisti, "Challenges to our understanding of the general circulation: Abrupt climate change," Global Circ. Atmos. **331**, 371 (2007).
[64] J. P. Severinghaus, T. Sowers, E. J. Brook, R. B. Alley, and M. L. Bender, "Timing of abrupt climate change at the end of the Younger Dryas interval from thermally fractionated gases in polar ice," Nature **391**(6663), 141–146 (1998).
[65] A. Singh, M. A. Papa, H.-B. Eggenstein, and S. Walsh, "Adaptive clustering procedure for continuous gravitational wave searches," Phys. Rev. D **96**(8), 082003 (2017).
[66] N. Smirnov, "Ob uklonenijah empiriceskoi krivoi raspredelenija," Recl. Math. Mat. Sbornik (N.S.) **6**(48), 3–26 (1939).
[67] R. R. Sokal and F. J. Rohlf, *Biometry* (W.H. Freeman, 1995).
[68] M. Thiel, M. C. Romano, J. Kurths, R. Meucci, E. Allaria, and F. T. Arecchi, "Influence of observational noise on the recurrence quantification analysis," Physica D **171**(3), 138–152 (2002).
[69] M. H. Trauth, A. Asrat, W. Duesing, V. Foerster, K. H. Kraemer, N. Marwan, M. A. Maslin, and F. Schaebitz, "Classifying past climate change in the Chew Bahir Basin, Southern Ethiopia, using recurrence quantification analysis," Clim. Dyn. **53**(5), 2557–2572 (2019).
[70] C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," Signal Process. **167**, 107299 (2020).
[71] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," Remote Sens. Environ. **114**(1), 106–115 (2010).
[72] Y. Wang, H. Cheng, R. L. Edwards, X. Kong, X. Shao, S. Chen, J. Wu, X. Jiang, X. Wang, and Z. An, "Millennial-and orbital-scale changes in the East Asian monsoon over the past 224 000 years," Nature **451**(7182), 1090–1093 (2008).
[73] Y.-J. Wang, H. Cheng, R. L. Edwards, Z. An, J. Wu, C.-C. Shen, and J. A. Dorale, "A high-resolution absolute-dated late Pleistocene monsoon record from Hulu Cave, China," Science **294**(5550), 2345–2348 (2001).
[74] T. Westerhold, N. Marwan, A. J. Drury, D. Liebrand, C. Agnini, E. Anagnostou, J. S. Barnet, S. M. Bohaty, D. De Vleeschouwer, F. Florindo *et al.*, "An astronomically dated record of Earth's climate and its predictability over the last 66 million years," Science **369**(6509), 1383–1387 (2020).
[75] C. Wunsch, "Abrupt climate change: An alternative view," Quat. Res. **65**(2), 191–203 (2006).
[76] W. J. Youden, "Index for rating diagnostic tests," Cancer **3**(1), 32–35 (1950).
[77] D. Yuan, H. Cheng, R. L. Edwards, C. A. Dykoski, M. J. Kelly, M. Zhang, J. Qing, Y. Lin, Y. Wang, J. Wu *et al.*, "Timing, duration, and transitions of the last interglacial Asian monsoon," Science **304**(5670), 575–578 (2004).
[78] J. P. Zbilut and C. L. Webber Jr, "Embeddings and delays as derived from quantification of recurrence plots," Phys. Lett. A **171**(3-4), 199–203 (1992).
[79] 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," Nat. Geosci. **10**(7), 518–523 (2017).

Published under an exclusive license by AIP Publishing