Referee #1 (Reza Tabar)

1) By checking the number of minima (k1) of PDF of two variables and number of attractive fixed points (k2) of drift terms, rule out that dynamics may have noise-induced phase transition. For k1 \neq k2 (k1 > k2) dynamics will have noise-induced transition.

We believe that this comment relies on a misunderstanding with respect to the term 'noise induced transition'. In our manuscript, the term refers to a transition between two stable states in response to a strong noise pulse. However, the same term is used to describe the transition from a unimodal to bimodal pdf in response to an increase of the noise amplitude in a monostable system. We believe that the referee had this second meaning in mind.

To prevent this misunderstanding we explicitly added the term 'noise-induced transition' to the description thereof (l.24):

Second, random perturbations may push the system across a basin boundary (noise-induced transition).

In fact, this concept might be relevant in view of the monostable $\delta^{18}O$ drift and the two-regime character of the record. Also, the reference (Majda, 2006) provided by referee #3 relates to this point to some extent. The combination of two apparent regimes and a single stable fixed point in the $\delta^{18}O$ time series and the role of multiplicative noise in this context will be discussed in upcoming research already underway. Both, this comment and the one by referee #3 are very helpful in this regard.

2) In lines 150, checking the short range correlations of increments will not show the Markovianity of time series. For linear Ornstein-Uhlenbeck process, despite of the process is Markov, the increments have negative correlations, see Ex. 21.1 (Tabar2019).

Remove this statement. Use a \chi^2 test (similar to sec. 16.4) or state that this is your assumption.

We agree with the referee, that the short range correlation of the time series' increments does not prove Markovianity. However, it does rule out important long-term memory effects. In the revised manuscripts, we emphasize that we do not provide a sufficient criterion for Markovianity.

Original:

The autocorrelation functions of the increments of both proxies shown in Fig. 2 exhibit weak anti-correlation at a shift of one time step and exhibit negligible correlations beyond this. Such small level of correlation certainly speaks against memory effects to have played a major role in the emergence of the given time series and hence in favour of considering the data Markovian.

Revised:

The autocorrelation functions of the increments of both proxies shown in Fig. 2 exhibit weak anti-correlation at a shift of one time step, while correlations beyond this are negligible. Such a small level of correlation certainly rules out long-term memory effects to have played a major role in the emergence of the given time series. Bear in mind that this is a necessary yet not sufficient criterion to consider the data Markovian. For practical reasons we refrained from further Markovianity tests.

3) Before Eq. 3, state combined \delta^{18}O and dust, … by two dimensional Langevin

equation (using the It\^o description)…

We thank the referee for the comment. In the revised manuscript, we specify that the equation (3)

$dx = F(x)dt + d\xi$

must be understood in the Ito sense. We considered referring to the equation as a Langevin equation, but decided not to do so for the reason given also in the revised manuscript (l. 177):

Notice that we could formulate our method equally well in terms of the simpler Fokker–Planck equation. However, operating with the Fokker–Planck equation implicitly assumes that the stochastic process under investigation follows a Langevin equation in a strict sense, i.e. the noise term in Eq. (3) would be restricted to the case of Brownian motion. However, in ongoing research we find indications that the description of the driving noise ξ(t) as Brownian motion might be overly simplistic (Rydin Gorjão et al., 2023). The use of the KM instead of the Fokker–Planck equation in this work aims at emphasizing that ξ(t) might be more complex than Brownian motion and contain for example discontinuous elements.

4) In 195 the Silverman rule of thumb is valid for estimations of PDFs, there is not such rule for KM coefficients. Authors can use a gaussian kernel with bandwidth h=0.3.

We agree with the reviewer. Indeed Silverman's rule is only valid for the the estimation of PDFs, yet currently there are no known ideal bandwidth rules for KM coefficients. We believe it is best to use 1) the same "ideal" bandwidth and "ideal" kernel that we employ for the estimation of the PDF for the estimation of the KM coefficients; 2) we believe a kernel with bounded support is better since it necessarily only captures local effects in the estimation and does not require truncation in computational estimation.

We added a notice to the text to warn the reader about this particular issue. It reads (line 212):

We note that the above formula for the ideal bandwidth has been developed for the estimation of the probability density function. As there is currently no consensus on the optimal kernel and bandwidth for the estimation of the KM coefficients, we will employ an Epanechnikov kernel with bandwidth h s throughout our work.

5) Appendix A has not any content. The title of the appendix is an open problem.

We thank the referee for highlighting this oversight from our side. This was an artifact from a previous iteration of the manuscript that we overlooked. It has now been removed.

We thank the referee for the various suggestions and corrections and for highlighting some issues within our analysis. We hope that our revised manuscript and replies comprise an improved version of our work.

Referee #3

This paper employs a non-parametric kernel-density estimation of the drift coefficient of a two-dimensional stochastic process involving the \deltaO18 and dust NGRIP records to shed some light on DO events.

The author’s findings are consistent with the view that atmospheric dynamics (represented here by the dust) controls DO events, in terms of stabilising the respective stadial and interstadial states and in triggering transitions a la Kleppin et al. Their findings, as the authors nicely elaborate, corroborate such a perspective.

I believe that the results are useful for the community and warrant publication. I have only a few minor issues the authors may want to consider.

I think the emphasis on Kramers-Moyal is a little too strong. The drift coefficients (6) and (7) follow directly from their equation (3) for the underlying dynamics. A discussion on how this is part of a Kramers-Moyal equation should be included, but maybe with a little less emphasis in the overall presentation.

We thank the referee for this constructive comment. The analysis could indeed also have been conducted in terms of the simpler and more common Fokker-Planck equation. The starting point of our investigation was the estimation of KM coefficients of the individual $\delta^{18}O$ and dust time series, which revealed contributions from non-Gaussian noise in the $\delta^{18}O$ record. However, there remain some discrepancies to be reconciled in that analysis. To maintain consistency across publications and to highlight that the noise is not necessarily Gaussian, we stick to the KM formulation also in this paper. Certainly, this merits explanation within the manuscript itself which we included as follows (l. 177).

Notice that we could formulate our method equally well in terms of the simpler Fokker–Planck equation. However, operating with the Fokker–Planck equation implicitly assumes that the stochastic process under investigation follows a Langevin equation in a strict sense, i.e. the noise term in Eq. (3) would be restricted to the case of Brownian motion. This conflicts with findings from ongoing research which indicate that the description of the driving noise ξ(t) as Brownian motion might not be applicable (Rydin Gorjão et al., 2022). The use of the KM instead of the Fokker–Planck equation in this work aims at emphasizing that ξ(t) might be more complex than Brownian motion and contain for example discontinuous elements.

We get back to this point in the discussion section of the revised manuscript (l.284):

As mentioned previously, a univariat estimation of the individual $\delta^{18}O$ and dust KM coefficients indicates that at least the $\delta^{18}O$ noise comprises non-Gaussian and potentially discontinuous components which could play a central role with respect to the transition between the two identified stable states of the drift (Rydin Gorjão et al., 2022). However, there remain discrepancies to be reconciled in the analysis of the higher-order KM coefficients of the individual $\delta^{18}O$ and dust time series and until then, arguments about the role of non-Gaussian noise in the state transitions remain speculative. Ideally, higher-order KM coefficients should be computed for the two-dimensional record, however, this is prevented by the low data resolution.

When I read the title, I was slightly taken aback that only the drift coefficient is considered.

Yes, we fully agree that the original title of the manuscript was misleading. It was inherited from an even earlier version of the manuscript. We replaced the title by

Stable stadial and interstadial states of the last glacial's climate identified in a combined stable water isotope and dust record from Greenland

I understand the difficulties in the estimation and interpretability of the higher-order coefficients, and the authors have clearly outlined this, and this is not a criticism of the work done and the results obtained, but simply a matter of presentation/refocus.

The authors are careful not to overstate their results and point to several limitations of their approach, which is much appreciated and puts their interesting findings in a broader context.

I would like to add one more point of caution: the authors' focus on the bimodality of the probability density function may be too simplistic to study regimes. Multimodal probability density functions are not necessary for the existence of regimes. For example, in cyclostationary systems (see Wirth (2001)) and chaotic systems with intermittent dynamics (see Majda et al (2006)) the probability density function may be unimodal whereas, if restricted to time windows focusing on the regimes, ''hidden'' regimes may be identified. This is an issue in atmospheric data (see also the discussions in Franzke et al. 2008, 2009). If the authors agree, it may be worthwhile to add this discussion (again, I don't think that the data allow additional analysis along these lines and I am not asking the authors to do this; simply, if they see fit, adding a discussion in the text).

We thank the referee for this remark. Indeed, the concepts and methods presented in the references could be relevant as well for the study of the NGRIP record. In particular, when studied in isolation the $\delta^{18}O$ record appears unimodal, while clearing exhibiting the signature of two-regimes. We will shortly present additional analysis wherein we assess the KM coefficients of the univariate dust and $\delta^{18}O$ time series. We think that this comment might be particularly helpful for this upcoming research. With respect to this manuscript, we believe that we already made an effort to carefully disentangle bistability, bimodality and the existence of two regimes. Still, we find the reference to Majda, 2006 important and helpful and therefore added the following comment (l.114):

Notice that the somewhat counterintuitive combination of meta-stable distinct dynamical regimes and unimodal distributions of the associated variables has been discussed also in the context of atmospheric dynamics (Majda et al., 2006).

Wirth, V., 2001: Detection of hidden regimes in stochastic cyclostationary time series. Phys. Rev. E, 64, 016136,

A. Majda C. Franzke, A. Fischer, and D. T. Crommelin, 2006: Distinct metastable atmospheric regimes despite nearly Gaussian statistics: A paradigm model. Proc. Natl. Acad. Sci. USA, 103, 8309–8314

C. Franzke, D. T. Crommelin, A. Fischer, and A. J. Majda, 2008: A hidden Markov model perspective on regimes and metastability in atmospheric flows. J. Climate, 21, 1740–1757.

C. Franzke , I. Horenko, A. J. Majda, and R. Klein, 2009: Systematic metastable atmospheric regime identification in an AGCM. J. Atmos. Sci., 66, 1997–2012.

Typos etc:

Page 4: All ages are according to —> The dating was performed according to ?

We merged the two sentences

This translates into non-equidistant temporal resolution ranging from sub-annual resolution at the beginning to ~ 5 years at the end of the period 59944.5 – 10276.4 yr b2k. All ages are according to the Greenland Ice Core Chronology 2005 (GICC05), the common age-depth model for both proxies (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2008).

into one sentence (l.85)

This translates into non-equidistant temporal resolution ranging from sub-annual resolution at the beginning to ~ 5 years at the end of the period 59944.5 – 10276.4 yr b2k according to the Greenland Ice Core Chronology 2005 (GICC05), the common age-depth model for both proxies (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2008).

Eqns (6) and (7): I would delete the n! and m! since only the cases n,m=0,1 are considered (and n,m do not appear on the left-hand side anyway).

Agreed, our need to maintain generality is unnecessary, as we only consider n and m either 0 or 1. Thank you, we changed this accordingly.

Page 8: first line: normalisable —> normalised?

Thank you, corrected.

Page 8: s-shape —> S-shape?

Thank you, corrected, and corrected as well in the caption of Fig. 3.

Figure 3: the figure captions for (c) and (d) should be D_{1,0} and D_{0,1} (as in the respective figure labels).?

Thank you for spotting this, the figure's annotations were changed around. We have corrected this.

We thank the referee for the various remarks and corrections, which have helped improve the manuscript. We submit our revised manuscript once more for a renewed appreciation.