# Orbital Insolation Variations, Intrinsic Climate Variability, and Quaternary Glaciations

We would like to thank again all three referees whose careful and constructive reviews have allowed us to substantially improve our manuscript. Most changes implemented at this revision stage are in direct response to a referee's comment. However, we have taken advantage of the opportunity and implemented some additional changes as a result of the repeated intensive dealing with the manuscript. This document first lists these changes and continues with the point by point answers to the referee comments. Whenever we preferred to leave the original version of the manuscript unchanged, where a referee has proposed a change, we have made an effort to justify our view thoroughly. We are convinced that the changes substantially improve the quality and clarity of our manuscript and that they address the referees' objections, questions and suggestions.

**Color coding:**

**Comment by the referee.**

**Reply from the authors.**

**Text from the original version of the manuscript.**

**Improved text.**

## Changes not in direct response to a referee comment

l. 13 We added the following paragraph at the very beginning that shall help to set the scene and prepare the reader for the topic:

*'In the early 20th century, Milutin Milankovitch presented his theory of ice ages (Milankovitch, 1920). Based on his own calculations and on insightful suggestions from Wladimir Köppen and Alfred Wegener (Imbrie and Imbrie, 1986), he proposed that the transitions between glacial and interglacial climate conditions were primarily caused by variations of incoming solar radiation, which by that time was known to vary in a quasi-periodic manner on slow time scales of tens to hundreds of thousands of years (Poincaré, 1892–1899). These variations of insolation, which arise as a consequence of the gravitational interaction of the Earth with the other planets and with its own Moon, are typically referred to as orbital forcing.*
*The orbital forcing comprises variations in (i) the eccentricity of the Earth's orbit around the sun with dominant spectral peaks around 400 kyr and 100 kyr; (ii) the obliquity, or axial tilt, i.e., the angle between the Earth's rotational and its orbital axis, with dominant periodicity around 41 kyr; and (iii) the climatic precession, which determines the phase of the summer*

*solstice along the Earth's orbit and has its most pronounced spectral power around 23 kyr and 19 kyr (Berger, 1978).'*

l.41 The unexplained presence of the 100 kyr peak, found the spectra of glacial-cycle proxy records has mostly been highlighted in boxes and figures in the original manuscript. The following paragraph, introduces this topic on the level of the main text.

*'The spectral peaks near 20 kyr and 40 kyr have been widely interpreted within the geological community as evidence for a linear response of the climate system to the orbital forcing (Imbrie and Imbrie, 1986). A third spectral peak at 100 kyr was, however, the most pronounced, but much more difficult to reconcile with the orbital theory of Quarternary glaciations. Since no sufficiently pronounced counterpart can be found in the spectra of the seasonal insolation forcing, Hays et al. (1976) hypothesized a nonlinear response of the climate system in order to explain this dominant periodicity of the late-Pleistocene glacial–interglacial cycles.'*

Much of the text from section 3.3. has been rewritten – partly in response to referee comments , partly as a result of a repeated internal review. The most prominent changes are:

Fig. 9 We have – inspired by a comment by referee #3 – added an illustration of the nullclines of the FHN model. A new paragraph starting at line 435 explains the dynamical features of the model based on the intersect of the x and y nullclines.

*'This behavior can be better understood by considering the nullclines of (23b) and (23a) in the (x, y)-plane, as shown in Fig. 9. If the branches of the y-nullcline that correspond to y_l and y_r , and thus to stable fixed points of (23b) for a given value of x, intersect with the x-nullcline given by y = γ, then this intersection constitutes a stable fixed point for the entire system.*
*If they do not, the system first relaxes along the fast direction toward the y-nullcline. Only then the adjustment of the slow component starts to drag the system along the y-nullcline in the direction where the distance to the x-nullcline decreases. At the point where the y-nullcline reverses, the fast component is immediately attracted by the other branch of the fast nullcline and the same process starts all over again.'*

Fig. 10 As referee 2 correctly noticed, in the case of slow forcing – substantially slower than any internal time scale - the FHN model as presented in the original version of the manuscript does not inevitably entail the need for an NDS treatment. Hence, we supplemented our discussion with a second case, where we increased the external forcing's frequency, such that the relevant time scales are higher entangled. The resulting dynamics can only be understand in the NDS framework. The paragraph starting at line 478 is dedicated to this faster external forcing example.

*'For panels (f)–(h) of Fig. 10, the time scale separation between the forcing and the internal dynamics is reduced, resulting in a qualitatively different behavior of the nonautonomous system. The frequency of occurrence of B-tipping points is much higher, and hence the trajectories do not even execute a full oscillation during a single time interval that permits oscillations. As a result, two stable patterns of trajectories are formed. These two patterns can be brought into agreement by switching the sign of one pattern and shifting it in time by τ f /2.*

*This symmetry reflects the symmetry of the stable nullcline of the fast system component as shown in Fig. 9.*
*Again, the PBA of this nonautonomous system can be thought of as an infinite repetition of the common trajectory structure that can be observed in Figs. 10(g,h) between −5 000 and 15 000 time units. In contrast to the slow-forcing case, each snapshot $A(t)$ now is comprised of two points in the $(x, y)$-plane. This example illustrates how the action of an external force on an autonomous system can give rise to considerably richer dynamics, which crucially depends on both the system's internal variability and the nature of the forcing.'*

Fig. 11 This new figure illustrates the Random Attractor of a the FHN model in a particular parameter setting. We added this figure and the paragraph starting from line 489 to improve the readers intuition for the Random Attractor concept.

Appendix A: To give the reader a better idea of the extensive research that addresses — more or less explicitly — glacial-interglacial cycles by means of NDS and RDS theory, we added Table A1 of conceptual glacial-cycle models. This addition was stimulated in part by the public comments of István Daruka.

## Comments by Referee 1 and according changes in response

I wonder actually who first-authored the paper. If Prof Ghil “conceived and designed the study”, why would he not write the paper? Or, if he did, why is he not the first author?

We have chosen the order of authors as usual according to how much they contributed to the final manuscript. We agree that the ‘authors contribution’ statement leaves room for interpretation and therefore we have changed the wording from

*‘MG conceived and designed the study. KR and TM carried out the numerical computations. All authors interpreted and discussed the results and wrote the manuscript.’*

to

*‘MG conceived and designed the study. KR and TM carried out the major part of the article's new research. All authors interpreted and discussed the results and wrote the manuscript.’.*

I do sympathise with Dr Daruka Istvan (whom i don’t know) if he has misgivings about any misrepresentation of his work, even if inadvertantly, especially regarding novelty. Although i should say that being completely ignored would be worse. I would like to kindly request from the authors that they do their utmost to be fair. Probably i didn’t even need to say this any more.

In the revised version of the manuscript we are more specific about the differences between the original and the modified version of the model that we use.

In particular, we replaced the original formulation (line 460 – original manuscript)

*‘We deviate from Daruka and Ditlevsen (2016), though, by introducing a slow change in the parameters*
*α(t) and β(t) of Eq. (29b), as follows:*

*\alpha(t) = 2.1 – 1.4 \tanh( [t+1100] / 500 )*
*\beta(t) = 2.5 + 1.4 \tanh( [t+1100] / 500 )’*

by (line 595 – revised manuscript)

*‘In the original DD16 model, MPT-like behavior was produced by a slow sigmoid variation of the parameter κ in Eq. (29b),*
*κ(t) = κ_1 + 0.5(κ_0 − κ_1 ) {1.0 − tanh( [t – t_0] / t_s)}*
*In our M-DD16 model, we introduce instead a slow change in the parameters α(t) and β(t) of Eq. (29b), as follows:*

$$\alpha(t) = 2.1 - 1.4 \tanh((t + 1100)/500) ,$$
$$\beta(t) = 2.5 + 1.4 \tanh((t + 1100)/500) .’$$

The new formulations clearly explain the changes that we applied to the DD16 model, and furthermore clarify, that the original DD16 model certainly was capable to produce an MPT like behavior.

Furthermore, in the revised text we explain better why we actually selected the DD16 model out of the large number of available conceptual glacial cycle models that do reproduce the MPT. To do so, we replaced

(line 447 – original manuscript)

*‘In this section, we illustrate how the PBA concept can help shed more light upon the dynamics of ice age models. For this purpose, we apply the Daruka and Ditlevsen (2016) model of glacial-interglacial cycles with slight modifications. We show first that this model approximates rather well the glacial cycles inferred from the benthic δ18 O proxy reconstruction of Lisiecki and Raymo (2005) and then compute the model’s PBA to investigate the dynamical stability of its glacial cycles.’*

by ( line 574 – revised manuscript)

*‘In this section, we illustrate how the PBA concept can help shed more light upon the dynamics of ice age models. As pointed out in Sect. 3.1 and elsewhere in this paper, there is a long history of modeling the climate of the Quaternary by means of conceptual models, and many nonautonomous models have been proposed to simulate glacial-interglacial cycles of the last 400 kyr to 2.6 Myr, based on the orbital forcing. In Appendix A, we provide a long but still not exhaustive list of glacial-cycle models and specify some of their key characteristics, including the degree of their success at simulating the MPT; see also the discussion in Sect. 2.2.*

*Among these glacial-cycle models, the model of Daruka and Ditlevsen (2016, DD16 hereafter) belongs to the more abstract ones, as it is not derived from detailed physical considerations. Still, its concise form, interesting nonlinear dynamics, and ability to simulate glacial cycles, as well as the MPT, make the DD16 model well suited for our illustrative purposes. We first slightly modify this model from its original formulation. We do so mainly in order for the model to better approximate the benthic δ 18 O proxy reconstruction of glacial–interglacial cycles due to Lisiecki and Raymo (2005), especially the timing of glacial terminations; compare our Fig. 13 with Fig. 1 in DD16. Thereafter, we compute the PBAs of the modified DD16 model, M-DD16 hereafter, to investigate the dynamical stability of its glacial cycles over the past 2.6 Myr.’*

Finally, in the revised version of the manuscript, we have include a table that comprises conceptual glacial cycle models of low dimensionality that consistently reproduce an MPT-like behavior. Doing so, we aim to stress that many plausible mechanisms for the MPT have been proposed, and that other models merit investigation along the lines of the present approach.

1. The term “interaction” is used in this paper as often as we encounter it in general. However, i don’t really understand what is meant by this so often, including in this paper. Please clarify, or, if it is not possible, avoid using this language. As i understand, interaction is about two-way causality, which only makes sense in terms of couplings in governing equations.

We thank the referee for highlighting this linguistic inaccuracy. We agree that the term ‘interaction’ should be reserved for situations in which two dynamical variables influence each other.

There were several occasions in the original manuscript, where we aimed to say that only the combination of external forcing and internal variability of a system can explain the observed variability of the forced system. Wherever we used formulations like '*interaction between external forcing and internal variability*' we replaced this wording by a different formulation, like '*the combination of external forcing and internal variability*' or the '*the external forcing modifies the system's variability*'. The individual changes we implemented with respect to this comment are documented in our answers to the list of the referee's minor comments further below.

2. I also don't understand the paper's distinction between stochastic and deterministically chaotic sources of (internal) variability.

3. In contrast, i think we should distinguish between external forcings versus influences under internal variability. Stochastic terms in equations are not meant to represent external forcing.

We believe this comment mostly refers to the statement in line (194 – original manuscript)

'*In returning to the "fundamental question #2" in Box 1, one must recall that — apart from deterministic chaos à la Lorenz (1963), as obtained by H. Le Treut and colleagues (Le Treut and Ghil, 1983; Le Treut et al., 1988) and shown here in Fig. 4(a) — stochastic contributions à la Hasselmann (1976) to the continuous part of the paleoclimatic spectrum must also play an important role.*'

which was also commented upon by the referee as follows:

I thought Hasselmann uses the stochastic framework only to model time scale separation. Otherwise, what appears as noise is clearly deterministic irregularity due to chaos. Perhaps this distinction should be avoided, or, made precise if possible. The distinction should have to do something objective -- not just subjective, whether we bother to resolve time or not.

We beg to differ: it is well known that (i) "one person's signal is another person's noise"; and that (ii) there is no real spectral gap in atmospheric, oceanic and climatic variability; see, for instance, Nastrom & Gage (*J. Atmos. Sci*., 1985). The use of deterministic vs. stochastic description of certain processes depends on the availability of data, the need for certain levels of detail and accuracy, and other modeling considerations; see, for instance, Palmer & Williams (eds.), *Stochastic Physics and Climate Modelling*, Cambridge U. P., 2009.

On the slow time scale of glacial–interglacial cycles, influences coming from fast processes such as weather, volcanic eruptions, and decadal fluctuations of solar irradiation can be treated as stochastic forcing. Many of these, as the referee correctly noted, are in fact deterministically chaotic processes and need to be resolved as such in numerical weather prediction, for instance. However, on the large time scales of Pleistocene climate, fast processes can be treated as stochastic to the degree of detail describable by the available data. On the other hand, as we show in the manuscript, deterministic chaos can govern the dynamics on the large time scales relevant for Pleistocene climate; see Sections 3 and 4 of the paper. Both the fast forcing, modeled as stochastic, and the relevant chaotic dynamics can contribute to the continuous spectrum of the record.

In response to this comment we slightly modified the above text which now reads (line 225 – revised manuscript):

‘In returning to the “fundamental question #2” in Box 1, one must recall that, on the paleoclimatic time scales of interest — *apart from deterministic chaos à la Lorenz (1963), as obtained by H. Le Treut and colleagues (Le Treut and Ghil, 1983; Le Treut et al., 1988) and shown here in Fig. 5(a) — stochastic contributions à la Hasselmann (1976) to the continuous part of the spectrum must also play an important role.*’

4. Point (ii) in the definition of a pullback attractor might be redundant. This might be the correct conclusion to be drawn from Fig. 11. The situation might be parallel with a time-dependent version of the system that is the normal form for a pitchfork bifurcation. I suppose that the PBA of that system — that goes back to negative infinity in time — is a pullback fixed point that is associated with the single stable solution and unstable solution before and after the bifurcation point, respectively. I suppose that after the bifurcation, we have PBAs that do not go back to negative infinity. This should have serious implications for climate projections.

The point (ii) – or Equation (15) – is in fact required in the definition of a pullback attractor. The first equation guarantees the invariance of the family A_t with respect to the dynamics of the system, while the second equation guarantees pullback attraction – note that X_0 in (ii) is not required to be an element of A_s, as it was the case in (i).

This definition might differ in terms of notation from other definitions; it is, however, the standard definition used by many textbooks, e.g., Caraballo and Han (2017).

In fact, in (ii) the $\rightarrow$ (line 270 – original manuscript) needed to be replaced by an equal sign (line 302 – revised manuscript). Thank you for inadvertently helping us detect this typo.

## Minor comments by Referee 1 and according changes in response

l.6 ‘*We introduce herein a unified framework for the understanding of the interplay between internal mechanisms and orbital forcing on time scales from thousands to millions of years.*’

The forcing is not influenced by the climate by definition. I thought interplay imples that causal influence goes both ways.

See our reply to Comment #1. We have changed the sentence to:

‘*We introduce herein a unified framework for the understanding of the orbital forcing’s effects on the climate system’s internal variability on time scales from thousands to millions of years.*’

l.23 ‘*Specifically, Hays et al. (1976) were able to create a composite record — back to over 400 kyr b2k, i.e., over 400 000 yr before the year 2000 A.D. — from two relatively long marine-sediment records of the best quality available in the early 1970s.*’

I only found the boyband b2k on the web.

We were not aware of the boy band, however, b2k is the currently standard way to denote ages, especially in the ice core community. For example, see the papers by Sune Rasmussen or Anders Svensson from the University of Copenhagen.

l.23 We replaced 2000 A.D. by 2000 CE.

l.25 *‘The authors demonstrated therewith that precessional and obliquity peaks near 20 kyr and 40 kyr were present in this record’s spectral analysis; see Fig. 1.’*

Power spectra are quite notorious for their noisiness. Did Hays et al. or anyone later perform some stat test for these freq components?

Hays et al. (1976) did actually show that the spectral peaks corresponding to the orbital frequencies are significant. Such a significance test must be done avoiding the effect of orbital tuning. Huybers and Wunsch (*Nature*, 2005) as well as Huybers (*Nature*, 2011) did show evidence for orbital forcing in records that were not orbitally tuned; the latter papers were not based on power spectra but on the timing of terminations.
In response to this comment we have change the sentence

‘*The authors demonstrated therewith that precessional and obliquity peaks near 20 kyr and 40 kyr were present in this record’s spectral analysis; see Fig. 1.*’

to (l.35 – revised manuscript)

*'The authors demonstrated therewith that significant precessional and obliquity peaks near 20 kyr and 40 kyr were present in this record's spectral analysis; see Fig. 1.'*

l.50 *'On the other hand, it also became clear that a model whose only stable solutions were stationary, could not reproduce very well the wealth of variability that the proxy records were describing.'*

A stationary solution in the sense of a fixed point attractor means that no variability is present at all, let alone DO events.

We thank the referee for pointing out this inaccuracy. We have changed the sentence accordingly:

(line 66 – revised manuscript)
'On the other hand, it also became clear that a model whose only stable solutions were stationary could not reproduce very well the wealth of variability that the proxy records were describing, not even in the presence of stochastic forcing'

l.54 '*For instance, the models of Ghil and associates (Källén et al., 1979; Ghil and Le Treut, 1981) captured the phase differences between peak ice sheet extent and minimum temperatures suggested by Ruddiman and McIntyre (1981) in the North Atlantic, while the work of Saltzman and associates (e.g., Saltzman and Maasch, 1988) captured the asymmetry of the glaciation cycles with their more rapid "terminations" (Broecker and Van Donk, 1970).'*

Is it not uncommon to refer to the self in third person?

Since we are a team of four authors, neither 'we' nor 'I' would correctly refer to Michael Ghil exclusively, so we do not see any alternative to the current formulation.

l.60 '*Hence, they could not capture the wealth of spectral features, with their orbital and other peaks, of the paleorecords available by the 1980s.'*

What would be the difficulty with adding in these simple models the variability in solar forcing? If none, then would they still not capture some of the associated features of the paleo record?

Adding the orbital insolation forcing to simple climate models was the next evolutionary step in the history of paleoclimate modeling. To clarify further this point, we slightly modified the sentence (line 64):

*'In this paper, we try to show a path toward resolving the four fundamental questions listed in the box below. In the next section, we summarize existing results on how the climate system's intrinsic variability arises at Quaternary time scales and on how this variability interacts with the time-dependent orbital forcing.'*

and instead now write (line 84 – revised manuscript):

‘*In this paper, we try to show a path toward resolving the four fundamental questions listed in the box below. In the next section, we summarize existing results on how the climate system’s intrinsic variability arises on Quaternary time scales and on how this variability is modified by the time-dependent orbital forcing, which was added to the previously autonomous climate models as the next step in paleoclimate modeling evolution; see, for instance, Le Treut and Ghil (1983) and Le Treut et al. (1988) vs. Ghil and Le Treut (1981).*’

Note that this modification also takes into account the referee’s comment on the term ‘interaction’.

Fig.3 ‘*Courtesy of N. Boers*’

He is an author of this paper. Or, has this same diagram appeared in a specific past publication? In that case i thought the precise reference should be given.

The phrase was added at an early stage in the development of the ms. and has now been deleted. Thank you for noticing.

Box1. ‘*How does the dominant peak of the observed variability near 100 kyr arise, given the rather diffuse orbital forcing at this periodicity?*’

In what sense is it dominant? Does this peak exceed a hypothetical background level more than the other peaks. At the higher freq peaks, this background is at a lower level. Exceedance is meant not in a log but lin scale.

In fact, the 100-kyr periodicity can be seen by the unaided eye to dominate the late Pleistocene’s benthic d18O records (e.g., LR04). Also, it is widely agreed upon in the literature that the 100-kyr peak is the dominant one after the MPT, as already stated by Hayes et al. (Science, 1976).

We agree with the referee that this point is not sufficiently explained until the point where Box 1 is shown in the manuscript. We have added, therefore, to the sentence in line 30

‘*The work of Hays et al. (1976) and of the subsequent CLIMAP and SPECMAP projects resulted in a much more detailed spatio-temporal mapping of the Quaternary and extended the belief in the pacemaking role of orbital variations into the more remote past.*’

the following statement (line 41 – revised manuscript) :

‘*The spectral peaks near 20 kyr and 40 kyr have been widely interpreted within the geological community as evidence for a linear response of the climate system to the orbital forcing (Imbrie and Imbrie, 1986). A third spectral peak at 100 kyr was, however, the most pronounced, but much more difficult to reconcile with the orbital theory of Quarternary glaciations. Since no* *sufficiently pronounced counterpart can be found in the spectra of the seasonal insolation forcing, Hays et al. (1976) hypothesized a nonlinear response of the climate system in order to explain this dominant periodicity of the late-Pleistocene glacial–interglacial cycles.'*

Box 1 *'What are the contributions of the orbital forcing and of the climate system's intrinsic variability to items (1)–(3) and how exactly do the two interact?'*

What two? Can we say that variance at different frequencies, or different Fourier modes, intercat? Isn't it variables that can interact?

The following reformulation implemented in the revised manuscript should resolve the issues rightfully raised by the referee:

*'What are the contributions of the orbital forcing and of the climate system's intrinsic variability to items (1)–(3) and how does the former one modify the latter?'*

l.80 Eq. (1b)

Why would a positive temperature anomaly build up ice volume?

The positive influence of $T$ on $\dot{V}$ is due to the precipitation-temperature feedback, which states that the higher the temperature, the more moisture can transported by the atmosphere and as a consequence, the higher will be the amount of precipitation over land masses leading to net growth of the ice sheets.

This is explained in detail in the following, so we refrained from changing the manuscript with respect to this comment.

l. 84 *'The EBM represents the well-known ice-albedo feedback used by both Budyko (1969) and Sellers (1969), while the ISM relies on the precipitation-temperature feedback postulated by KCG and used also by Ghil and Le Treut (1981), who coined the term.'*

If we have only +ve feedback, the system would be unstable. In the EBM we indeed have -ve feedback too.

We are not sure, what exactly is meant by the referee, here. Note that the term EBM only refers to Equation (1a), while the term ISM refers to (1b). Together the two equations constitute an oscillatory climate model.

l.104 Why does this i not have a dot on top?

This is a pretty standard character for the imaginary unit; its LaTeX representation is $\imath$.

l.115 Perhaps you don't want to start a new para here because that would leave you with a one sentence para.

Yes, we agree and have removed the paragraph in the revised manuscript.

l.170 ‘*We start this section by describing some fairly simple ways in which the orbital forcing might have interacted with intrinsic climate variability, thus helping to solve the mismatch between Figs. 3(a) and 3(b) in Section 1.*’

interacted

In line with our answer to Comment #1, we have replaced the above sentence by:

‘*We start this section by describing some fairly simple ways in which the orbital forcing might have modified intrinsic climate variability, thus helping to solve the mismatch between Figs. 3(a) and 3(b) in Section 1.*’

l.174 ‘*These authors found that, as expected for a nonlinear oscillator, its internal frequency f 0 interacts with the forcing ones, {f 1 , . . . , f 5 }, to produce both nonlinear resonance and combination tones (Landau and Lifshitz, 1960).*’

See comment 1.

Here, we have used the term ‘affected by’ instead of ‘interact.’ (line 204 – revised manuscript)

‘*These authors found that, as expected for a nonlinear oscillator, its internal frequency f 0 is affected strongly by the forcing ones, {f 1 , . . . , f 5 }, resulting in both nonlinear resonance and combination tones (Landau and Lifshitz, 1960).*’

l.195 ‘*In returning to the “fundamental question #2” in Box 1, one must recall that — apart from deterministic chaos à la Lorenz (1963), as obtained by H. Le Treut and colleagues (Le Treut and Ghil, 1983; Le Treut et al., 1988) and shown here in Fig. 4(a) — stochastic contributions à la Hasselmann (1976) to the continuous part of the paleoclimatic spectrum must also play an important role. In fact, the theory of random dynamical systems touched upon in the next subsection provides an excellent framework for a “grand unification” of these two complementary points of view (Ghil, 2014, 2019).*’

I thought Hasselmann uses the stochastic framework only to model time scale separation. Otherwise, what appears as noise is clearly deterministic irregularity due to chaos. Perhaps this distinction should be avoided, or, made precise if possible. The distinction should have to do something objective -- not just subjective, whether we bother to resolve time or not.

Please see our response to the Comments #2 and #3.

l.201 ‘*The highly preliminary results on interaction between external forcing and internal variability summarized in Sec. 3.1 encourage us to pursue in a more systematic way the interaction between orbital forcing and intrinsic climatic variability that may have contributed to generate the rich paleoclimate spectrum on Quaternary and longer time scales (e.g., Westerhold et al., 2020).*’

somewhat redundant language, beside the problem with the concept of interaction.

In the revised manuscript we have reformulated as follows in order to address both issues (line 235):

*'The highly preliminary results summarized in Sect. 3.1 encourage us to pursue in a more systematic way the effects of the orbital forcing on intrinsic climatic variability, effects that may have contributed to generate the rich paleoclimate spectrum on Quaternary and longer time scales (e.g., Westerhold et al., 2020).'*

l.209 *'On the road to including deterministically time-dependent, as well as random effects, one needs to realize first that the climate system — as well as any of its subsystems, and on any time scale — is not closed: it exchanges energy, mass and momentum with its surroundings, whether other subsystems or the interplanetary space and the solid earth.'*

Does Earth climate affect the sun? I wouldn't have thought. Surely, on certain time scales, we can treat some effects as external forcing with a very good approximation. Volcanic eruptions are surely external to the climate system.

We are not sure whether we understand this comment correctly. Of course, the Earth does exchange energy and momentum with the sun, however, the size of the sun makes the influence of the earth on the sun negligible, while the contrary is true for the moon.

However, the aim of the above sentence is to emphasize, that neither energy nor momentum in the climate system are preserved quantities. We believe that the original formulation is suited to convey this message and stuck to it in the revised version of the manuscript.

l. 215 *'Alternatively, the external forcing or the parameters were assumed to change either much more slowly than a model's internal variability, so that the changes could be assumed to be quasi-adiabatic, or much faster, so that they could be approximated by stochastic processes. Some of these issues are covered in much greater detail by Ghil and Lucarini (2020, Sec. III.G).'*

I think it's important to make a distinction between the very slow and fast processes other than their time scale. A slowly evolving ice sheet should be possible to model as an external forcing; but a fast atmosphere regarding the upper ocean "of interest" is considered part of the internal variability. In fact, I can't easily think of a fast process that is not considered as part of the internal variability but rather as an external forcing.

We have a relevant criticism of the OCCIPUT project in Sec. 4.4 of

https://link.springer.com/article/10.1007/s10955-019-02445-7

Thank you for reminding us to cite the Tél et al. (*JSP*, 2020) paper in the revision, which we definitely plan to do. In the broader perspective hinted at in this comment, it might be true that, in the case of paleoclimate, the fast processes which can be modeled as noise are mostly internal to the climate system itself – though for volcanic eruptions at least this is rather debatable. However, such an assertion is far from true in general and there are a number of several cases in which fast external forces act on internally slow systems. Thus, we will refrain from classifying fast processes that can be modeled as noise as being ‘internal’ processes opposed to forcings that are necessarily ‘external and slow,’ as proposed by the referee, to the extent that we understand this comment.

Also, the waxing and waning of ice sheets constitutes the dynamics of interest here, and thus it would not make much sense to consider global ice volume as a slow changing parameter. However, it might be, that we did not interpret this comment correctly.

l.217 ‘*The key concepts and tools of NDSs and RDSs go beyond such approaches that rely in an essential way on a scale separation between the characteristic times of the forcing and the internal variability of a given system; such a separation is rarely, if ever, actually present in the climate sciences.’*

I think the Hasselmann view hinges on this.

Yes, Hasselmann’s approach does hinge on such a separation, but he is wrong, Nobel prize notwithstanding. In particular, as far as modeling the Quaternary’s glacial–interglacial cycles, the time scales of climatic oscillations, of about 40–100 kyr, cannot be separated from the time scale of the external forcing, of about 20–40 kyr. We agree that the Hasselmann Brownian-motion model of climate that relies on such a time scale separation produces the widely observed continuous red-noise spectrum. This point, though, does not necessarily prove that there exists a clear time scale separation in the weather and climate phenomenology; see, for instance, Huybers & Curry (*Nature*, 2006) or Lovejoy & Schertzer (CUP, 2013).

l.224 ‘*Readers who are less interested in this mathematical framework — which allows a truly thorough understanding of the way that orbital forcing interacts with intrinsic climate variability on Quaternary time scales — may skip at a first reading the remainder of this section and continue with Section 4.’*

perhaps some editing issue here?

We must admit that we do not understand this comment. However, the term ‘interacts with’ has been replaced here too by ‘acts on’ (line 255 – revised manuscript).

l.257: We have realized that there is a typo in the ‘family of snapshots’ in the subscript $t \in R$.
The subscript $R$ will be replaced by a $\mathbb{R}$.

l.320 *‘The noise processes may include “weather” and volcanic eruptions when X(t) is “climate,” thus generalizing the linear model of Hasselmann (1976), or cloud processes when we are dealing with the weather itself: one person’s signal is another person’s noise, as the saying goes.’*

The "weather is not independent of the climate" so there is no point in considering the pullback attractor for the SAME random noise realization, as mentioned above.

We really do not know how to interpret this comment. It is not stated anywhere that weather was independent from climate, in contrast, by introducing $G(X)$ in Equation (22), we explicitly allow for multiplicative noise. Certainly, there are still more complex forms of noise, which are not covered by Equation (22), but we don't think that is the referee's point here?

The pullback attractor $A(\omega)$ in fact assumes a fixed noise realization. However, in this picture $A(\omega)$ itself is random and its distribution is determined by the distribution of the noise and the internal system dynamics.

We have not done any changes in response to this comment.

l. 407 '*Figure 9 shows a trajectory of the FHN model for which the sinusoidal forcing used in Fig. 8 was replaced by a rescaled time series of atmospheric CO 2 concentrations retrieved from Antarctic ice cores (Bereiter et al., 2015):*'

Why that? There has been no physical meaning given to (24) yet so that we could guess why CO2 is the driver.

The aim here is to provide an illustrative application of the NDS theory to a concrete paleoclimate modeling example, where we have prioritized concise math over physical consistency. However, we agree, that the introduction of the CO2 forcing is a little bit ad hoc and therefore have added explanation at the beginning of Section 3.3. We have also made an effort here to better connect the section with Sections 2.1 and 2.2, by emphasizing that the external forcing truly acts as a bifurcation parameter switching on and of internal oscillations. Since the entire section 3.3. has been subject to quite some changes, we refrain from presenting all of these here. Most important with respect to the referees comment might be the paragraph we inserted in line 393 of the revised manuscript:

'*We discuss the example of the FHN model at some length in order to illustrate how external forcing can act on a system's internal variability and thereby give rise to more complex dynamics. This model's concise mathematical formulation and its widespread application in paleoclimate modeling and other fields make it ideally suited for this goal. We start with a description of the autonomous model, with no time-dependent forcing. Subsequently, we introduce a simple sinusoidal forcing and numerically compute the corresponding PBA. We then extend these consideration into the realm of random dynamical systems by adding stochastic forcing and discuss the resulting random attractor. Finally, we replace the synthetic forcings by one that corresponds to a paleoclimate proxy record of past CO 2 concentrations retrieved from Antarctic ice cores (Bereiter et al., 2015) and show that this setup brings the model's trajectories into good qualitative agreement with the D-O patterns observed in δ 18 O records from Greenland ice cores. In doing so, we pay less attention to the physical interpretation of the model's variables, while focusing on the detailed explanation of model behavior and on the role of the forcing in the resulting dynamics.*'

l.448 ‘*For this purpose, we apply the Daruka and Ditlevsen (2016) model of glacial-interglacial cycles with slight modifications.*’

We hope that this point has been sufficiently addressed in our answer above.

l.451 ‘*The model’s variables are a global temperature anomaly y that is proportional to minus the global ice volume and an effective climatic memory term x that represents the internal degrees of freedom.*’

This sounds a bit like eq. (1). However there the proportionality is between the tendency of temperature and minus the ice volume. Why?

The proportionality is not causality as in Eq. (1) but a simple (anti-)correlation such that the glacial has low temperature and large ice volume.

I don't really know what to think of by this. Some more details would be helpful.

As mentioned above and as was the case for the FHN model: We have set out to provide the reader with mathematically easy to understand non-autonomous dynamical systems with some relevance to paleoclimate modeling. In fact, we provide little physical interpretation. Nevertheless, we hope that the introductory text on the DD16 model (see beginning of this answer) we have incorporated in the revised manuscript provides a better guidance for the reader.

l.476 ‘*The PBA in this case is simply a moving fixed point, as plotted in Fig. 11(a), since the model dynamics is predominantly stable in the long time interval prior to the MPT that is situated around 1.2–0.8 Myr b2k.*’

This means that the PBA has no useful application here. Also, as we pointed it out in https://link.springer.com/article/10.1007/s10955-019-02445-7 a "pullback fixed point" is known in classical ODE theory as a "particular solution".

Please note, though, that in classical ODE theory, a "particular solution" does not have the stability properties that a "pullback fixed point" has, nor does it lag the forcing necessarily in the same way. We agree that the result up to this point is not very impressive; this being so, however, could not be known in advance. More importantly, our investigation shows that the situation changes across the MPT, a point that is central to our paper. We believe that the modified DD16 model’s change in PBA behavior across the MPT regime — given an external forcing whose characteristics do not change at or near the MPT —is a nice illustration of the point that the combination of internal dynamics and external forcing is crucial for the resulting variability.

fig. 11(a) Is this really the very same integration whose result we see in Fig. 10c? In there, after the MPT, the wiggles seem to indicate irregularity, i.e., instability of the trajectory, i.e., chaos.

That's true. The dynamics becomes unstable after the MPT. This instability, however, is not very strong. As a result, we can observe a fairly stable solution that is consistent with the proxy record.

l.480 ‘*However, when keeping the parameters α and β fixed at their post-MPT values α = 0.7 and β = 3.9 throughout the simulation interval and repeating the computation of the PBA, a more complex picture arises. In the latter case, Fig. 11(b) shows a bunching of trajectories into separate clusters, subject to the quasi-periodic orbital forcing of Fig. 10(a).’*

Are these really separate clusters? One question is if there is one attractor, or, there are coexisting attractors.

Thank you very much for this comment! Given this comment, we looked at the chaotic PBA closely and observed that there were no separate clusters, as in the work of Pierini & Ghil on a simple model of the wind-driven ocean circulation or in the coupled ocean–atmosphere VDDG model (Vannitsem et al., *Physica D*, 2015). In the revised manuscript, we will rephrase the term “separate clusters” as “separate fuzzy clusters”; see Ghil & Robertson (*PNAS*, 2002) (line 623 – revised manuscript).

If the latter, only one of these can be accessed with time-dependent parameters given that the PBA is a moving fixed point. It is a kind of "intransitivity".

Prior to the MPT the PBA is a moving fixed point and yes, this means that also after the MPT the PBA as defined in equations (14) and (15) certainly remains a single moving fixed points, since the pre-MPT dynamics would cause trajectories starting from different initial conditions to converge if the the starting time is taken to minus infinity.

‘in case of the latter’ means that there are coexisting attractors – this contradicts ‘given that the PBA is a moving fixed point’.

This raises the issue that, out of the two, the chaotic PBA originates not in the infinite past but at a particular point in time. That is, your point (ii) given by eq. (15) to define a PBA seems redundant.

The first statement is true, but the second isn’t (see our reply to the comment #4 ).

The chaotic PBA bursts into existence from a single point in phase space, i suppose.

Regardless whether this is the case here or not, clearly, it is a possibility, and so point (ii) needs to be revised. One only needs to consider a system for which for some fixed values of a parameter there is only one attractor, and for other fixed values of the parameter there are more. Then the autonomous system can be defined in a way that the parameter is made time-dependent, taking values in the single attractor regime earlier and progressing in time to values where there are co-existing attractors of the autonomous system.

Perhaps the simplest example can be given by x' = c + mu*x-delta*x^3, mu = atan(t)*2/pi

With fixed mu and c = 0, this is the normal form for a pitchfork bifurcation. With c>0, we have an imperfect pitchfork.

Alternatively, you have in (29) more like a perfect pitchfork scenario, c = 0, with a continuity of the nonchaotic and chaotic attractors across the regimes, the bifurcation point being the point of continuity. You can simulate this in Matlab by

```
[t,x] = ode45(@(t,x)(atan(t)*2/pi * x - x.^3), [-50 30], linspace(-2,0,11));
```

In Fig. 11a, you continue to have the 40 trajectories clumped "completely" together perhaps because the trajectory data is represented by the SAME number of finite machine precision, as a consequence of starting your simulation sufficiently far back in time. Perhaps, continuing the simulation longer, you would start to have chaotic trajectories and so a broadening of the ensemble. However, that might just be because of imprecise number representation on the computer, as i suppose your PBA that is initiated in the infinite past, is really just a moving fixed point. (That would be the unstable middle pin of the perfect pitchfork.)

The chaotic "part of the" (?) PBA does not go back to the infinite past.

We like to try to answer to this comment, with the proviso that we might not have understood it in all it’s details.

The PBA of the M-DD16 with the parameter shift is certainly just a single moving fixed point. As the referee correctly mentioned – other possible trajectories, that burst into existence across the MPT, cannot be accessed, because the pre MPT dynamics force trajectories that start from different initial conditions to converge arbitrarily strong.

The M-DD16 model with fixed post-MPT parameters has a more complex PBA comprised of fuzzy clusters. This is also what we explain in the manuscript and this does not conflict with our definition of the of the PBA which is correct as is.

In fact, the presence of the more complex yet in the deterministic setting inaccessible structure of potential trajectories after the MPT in the M-DD16 model with shifted parameters could be investigated in future work. The ‘inaccessible’ trajectories could become highly relevant, once noise is added to the system.

l.483 ‘*First, post-MPT dynamics is much more irregular and unstable than the stable, quasi-periodic dynamics prior to the MPT.*’

You said earlier it was a moving fixed point; that's not the same as quasi-periodic.

We removed “quasi-periodic" because, strictly speaking, the motion is not quasiperiodic, as you point out, i.e., it is not a sum of periodic motions with irrationally related frequencies. We use instead “more stable.” (line 625 – revised manuscript).

l.498 *‘In Sec. 3, we presented first results on the interaction between the orbital insolation forcing of Sec. 1 and the intrinsic variability of Sec. 2, and proceeded to introduce the novel concepts and tools of the theory of nonautonomous and random dynamical systems (NDSs and RDSs) that can help to better model and understand this interaction.’*

In line with our response to comment 1 we replaced the sentence by

*‘Sect. 3, we presented first results on the effects of the orbital insolation forcing of Sect. 1 on the intrinsic variability of Sect. 2, and proceeded to introduce the novel concepts and tools of the theory of nonautonomous and random dynamical systems (NDSs and RDSs) that can help to better model and understand these effects.’*

l.499 novel – novel in what sense?

In the sense that this theory was only introduced into the climate sciences by Ghil et al. (*Physica D*, 2008) and by Bódai et al. (*NPG*, 2011), as stated in the first paragraph of Sec. 3.2 herein.

## Comments by Referee 2 and according changes in response

1. It is not clear to me what level of mathematical knowledge this paper is targeting. The paper begins assuming very little knowledge, by examining Hopf Bifurcations in equations (3--5) which are a topic covered in any dynamical systems course. Yet later in the paper readers are assumed to know what the `Hausdorff semi-distance' is. I think it would be better to assume less mathematical knowledge than more, perhaps the Hausdorff semi-distance could be replaced by a more informal comment about the system approaching $\mathcal{A}_t$. Another place the analysis could be streamlined without loss of understanding is by setting $\beta = \mu/2$ in equation (17) thereby reducing the number of parameters. The paragraph starting on line 304 provides conclusions without justification, which are only obvious to people familiar with dynamical systems. Perhaps a figure would help here?

We thank the referee for this constructive comment. Given that Referee #3 made a closely related comment, we would like to give a combined answer to both comments.

**Referee #3**

At lines 102 (section 1) the authors bring in the notion of a Hopf bifurcation with one type of simple system (eqn 5). Then in section 2.2 the description of the subcritical and supercritical Hopf bifurcations are described with another system (eqn 6). I would like to see more diagrams in Section 2.2 (some of us can visualise in our head what is happening when parameters are varied (e.g. through a Hopf bifurcation) but I think it is important to try to improve section 2.1 and 2.2 in a more unified way so at make these sections more accessible to a newer audience that is reading this type of material for the first time. I think a clear illustration with both language and an additional set of figures (possibly using the example systems from equation 5 or 6) would be helpful. For example, one might introduce the sections with language such as, "A Hopf bifurcation occurs when a periodic solution or limit cycle that surrounds an equilibrium point appears or disappears when a (control) parameter is varied. When the stable limit cycle surrounds an unstable equilibrium point, the bifurcation is supercritical. In the case that the limit cycle is unstable and surrounds a stable equilibrium point, the bifurcation is subcritical." And then also illustrated these concepts later on with the simple systems used.

The comments of Referees #2 and #3 convinced us to give a typical illustration of a supercritical Hopf bifurcation in our manuscript to supplement the text (Fig. 4 in the revised manuscript). Referee #2 correctly noted that this is part of any dynamical systems course; we target, however, a readership where not everybody necessarily attended such a course.

In the revised manuscript we avoided the use of the Hausdorff semi-distance, by saying (line 323 – revised manuscript):

'we find that the distance of any trajectory at time t to the set A t tends to zero, as we pullback the initial time t_0 to -\infty'.

We supplemented the paragraph starting at line 304 (original manuscript) with substantial explanation (see revised manuscript line 337) as well as visualization (Fig. 6 (c-e) revised manuscript).

Originally, the paragraph reads (line 304 – original manuscript):

*‘Note that the structure of the system’s trajectories depends on the ratio $\omega/\nu$ and three different cases must be distinguished. If the radius is modulated with the same frequency as the oscillation itself, i.e. $\omega = \nu$, after one period the system practically repeats its orbit. More precisely, the radius of the oscillation does differ from one “roundtrip” to the next, but this difference ($\rho$) tends to zero as $\rho(t)$ asymptotically approaches the PBA $A_t$. If $\omega$ and $\nu$ are rationally related, $m\,\omega = n\,\nu$ with $n, m \in N$, then the same quasi-repetition of the orbit occurs after n periods of the radial modulation and m periods of the system’s oscillation. Such a trajectory will appear as an n-fold quasi-closed loop. Finally, if $\omega/\nu \notin Z$, then the trajectory does not repeat itself but instead covers densely the annular disc $D = \{(\rho, \varphi)\colon \rho \in [\mu - \alpha\beta, \mu + \alpha\beta]$ and $\varphi \in [0, 2\pi)\}$. The trivial evolution of the phase ($\rho$) is depicted in panel (c), while the trajectories of $\rho(t)$ and their convergence to the PBA $A_t$ are shown in panel (d).’*

The revised manuscript now reads:

*‘Panels (c–e) demonstrate a particularity of this system, which is characteristic of dynamics confined to a torus. Namely, the structure of the system’s trajectories depends on the frequency ratio $\omega/\nu$ and three different cases must be distinguished. If the radius is modulated with the same frequency as the oscillation itself, i.e. $\omega = \nu$ as in panel (c), the forcing and the system have a fixed phase relation. That is, for a given phase of the system, its radius is always attracted by the same fixed point.*
*Hence, the system practically repeats its orbit after a short time. More precisely, the radius of the oscillation does differ from ($\rho$) one “roundtrip” around the torus to the next, but this difference tends to zero as $\rho(t)$ approaches the PBA $A\ t$ . If $\omega$ and $\nu$ are rationally related, i.e., $m\,\omega = n\,\nu$ with $n, m \in N$, as in panel (d), then — after n periods of the radial modulation and m periods of the system’s oscillation — the phase relation between the system and its forcing will repeat itself and hence we observe the same quasi-repetition of the orbit after the time $n\,2\pi/\nu = m2\pi/\omega$. That is, such a trajectory will appear as an n-fold quasi-closed loop.*
*Finally, if $\omega/\nu \notin Z$, as in panel (e), then a given phase of the system will never coincide with the same phase of the radius modulation more than once. Hence, the trajectory does not repeat itself but instead covers densely the annular disc $D = \{(\rho, \varphi) : \rho \in [\mu\ 0 - \alpha\beta, \mu\ 0 + \alpha\beta]$ and $\varphi \in [0, 2\pi)\}$.’*

2. The point of this paper is to demonstrate the advantage of the NDS picture over the autonomous picture, yet much of the analysis could be done in the autonomous regime using standard assumptions about timescale separation. In the analysis of the FHM model, $\gamma$ is taken to be a slow oscillation which allows for a discussion (lines approximately 385--400) that would be familiar to people who had only worked with autonomous systems. In the Daruka-Ditlevsen model, $\alpha$ and $\beta$ are again slow parameters, so why shouldn't this be analysed with the classic tools of autonomous dynamical systems? I think it would be useful to emphasise what extra information the nonautonomous picture gives us, and what would go wrong analysing it using the tools from autonomous dynamical systems.

In the case of the M-DD16 model, $\alpha$ and $\beta$ are not meant to represent the external forcing, but instead a slow change of the internal dynamics – discussing the reason for this change is not the aim of this manuscript. Instead, the external forcing is the insolation at 65°N at the summer solstice. While the autonomous system has a single stable fixed point for all $\alpha(t)$, $\beta(t)$, the response of the system to forcing changes significantly over time. The time scales of the resulting dynamics are the same as the forcing time scales, so for this example we believe, there is no other way but treating this nonautonomous system with methods that are well adapted to such system systems.

Regarding the FHN model, we agree that this is a borderline case. The external forcing used in the original manuscript acts like a slow parameter and the time scales are pretty much separated, as denoted in Eq. (27) (original manuscript). We have improved the relevance of this example by a second forcing case with higher frequency to the discussion (line 478 and following – revised manuscript). This increases the entanglement of the external forcing with the internal dynamics and thus helps to underpin the need for an NDS approach.

In the paleoclimate application of the FHN model, we have less freedom in our choice of time scales. However, it should be noted that, here too, the forcing changes on time scales that are relevant for the internal dynamics as well and it directly determines the length of stadials and interstadials, as well as the frequency of the oscillation.

3. I find the section on RDS a bit disconnected from the rest of the paper. There are no concrete paleoclimate applications given and the section introduces concepts such as the Random Attractor which is not defined even informally and are not used in the rest of the paper. I would recommend either cutting this section or adding in a simple paleoclimate example.

The paper aims to give the reader an overview over the mathematical concepts used for the study of non-autonomous dynamical system and thus we would prefer not to omit Random Attractors. An informal 'definition' of a Random Attractor is given by Fig. 7 (revised manuscript).

The referee's comment that the RDS theory being disconnected, motivated us to include a detailed discussion of an FHN model Random Attractor starting at line 489 in the revised manuscript.

4. Figures 8 and 9 are seriously misprinted, e.g. Fig 8e has times labelled as -20000,-000,-0000,-000,0,000,0000.

We are sorry, this must have happened during the uploading process. In the version we have stored locally and which we uploaded as it is, the figures have the correct labeling. There are other details missing in both figures and we will exercise extra care in uploading the modified version, if encouraged to do so by the handling editor.

I realise this is a matter of personal preference, but might the PBA figures e.g. 5a look clearer if projected onto the x-y plane? I always find 3D figures confusing. Figure 5c doesn't seem useful, perhaps it could be replaced by the trajectories of the system with different $\omega/\nu$ values.

Originally, we had a 2D projection of panel (a) in Figure 5, but the different trajectory overlapped so strongly that the plot was rather messy and not helpful. We hence prefer to keep the 3D version of Fig. 5a.

The referee is right that panel (c) of Fig 5. (original manuscript) is kind of trivial. In the revised version of the manuscript, we have, first of all, specified the values for $\mu$ and $\nu$ used for the computations. Second, we have replaced panel (c) by three panels which show projections of trajectories with qualitatively different $\omega/\nu$ ratios in order to visualize the discussion in lines 337–349, as requested by the referee in their first comment.

Minor comments:

1. Lines 31--32 `an the' should be `and'.
Thank you, this was corrected.

2. Around lines 455, what do the parameters mean physically?
We have extended the introductory paragraph on the M-DD16 model (line 580 – revised manuscript), which now reads:

*'Among these glacial-cycle models, the model of Daruka and Ditlevsen (2016, DD16 hereafter) belongs to the more abstract ones, as it is not derived from detailed physical considerations. Still, its concise form, interesting nonlinear dynamics, and ability to simulate glacial cycles, as well as the MPT, make the DD16 model well suited for our illustrative purposes. We first slightly modify this model from its original formulation. We do so mainly in order for the model to better approximate the benthic $\delta^{18}O$ proxy reconstruction of glacial–interglacial cycles due to Lisiecki and Raymo (2005), especially the timing of glacial terminations; compare our Fig. 13 with Fig. 1 in DD16. Thereafter, we compute the PBAs of the modified DD16 model, M-DD16 hereafter, to investigate the dynamical stability of its glacial cycles over the past 2.6 Myr.*
*Our model's variables, following DD16, are a global temperature anomaly y that is proportional to minus the global ice volume and an effective climatic memory term x that represents the internal degrees of freedom. In the deterministic case, the governing equations of the M-DD16 model are given by*

$\tau \dot{x} = \lambda y,$ *(29a)*
$\tau \dot{y} = -\alpha(t) + x - x^3 - \beta(t)F(t)x - \kappa y;$ *(29b)*

*here t is the time in kyr and F (t) is the normalized June 21 insolation at 65 ◦ N, based on the calculations of Laskar et al. (2004), as shown in Fig. 13(a).'*

We hope this clarifies the question.

3. Lines 518--527, why not quote Emiliani and Geiss directly?

The analysis of Emiliani & Geiss in *Geologische Rundschau* (1959) was purely descriptive and is no longer relevant to the level of mathematical discourse in this paper. But the two paragraphs cited from Ghil & Childress (Springer, 1987, Sec. 12) still are very much so.

## Comments by Referee 3 and according changes in response

At lines 102 (section 1) the authors bring in the notion of a Hopf bifurcation with one type of simple system (eqn 5). Then in section 2.2 the description of the subcritical and supercritical Hopf bifurcations are described with another system (eqn 6). I would like to see more diagrams in Section 2.2 (some of us can visualise in our head what is happening when parameters are varied (e.g. through a Hopf bifurcation) but I think it is important to try to improve section 2.1 and 2.2 in a more unified way so at make these sections more accessible to a newer audience that is reading this type of material for the first time. I think a clear illustration with both language and an additional set of figures (possibly using the example systems from equation 5 or 6) would be helpful. For example, one might introduce the sections with language such as, “A Hopf bifurcation occurs when a periodic solution or limit cycle that surrounds an equilibrium point appears or disappears when a (control) parameter is varied. When the stable limit cycle surrounds an unstable equilibrium point, the bifurcation is supercritical. In the case that the limit cycle is unstable and surrounds a stable equilibrium point, the bifurcation is subcritical.” And then also illustrated these concepts later on with the simple systems used.

We thank the referee for this very useful and constructive criticism. We added a visualization of the supercritical Hopf bifurcation to the revised version of our manuscript (Fig. 4 – revised manuscript).

In general the paragraphs are quite short (e.g Line 54). There is no need to start a new paragraph in a lot of places in the manuscript, please try to make the text flow a bit better.

We have carefully revises the manuscript with respect to this comment and combined paragraphs where we thought it would improve the readability.

L243, 249 monotonic, monotonically

Thank you for pointing this out. This was corrected accordingly.

In the section on 3.3 on applications D-O events. Figure 8a is a bit confusing , maybe I missed it , but I don’t see how the abscissa and ordinate are defined , it looks like simply x and y, yet they are both scaled to \alpha?

In fact, there are many labels missing in the uploaded version of the manuscript, but they are present in the local version stored on our computers. We apologize for the inconvenience.

However, the caption of Figure 8, did not sufficiently explain the Figure. We have replaced the caption

*‘FitzHugh-Nagumo (FHN) model with time scales τ f = 2000, τ x = 100, τ y = 60, and α = 2. (a) The cubic term of the fast derivative P 3 (x, y) as a function of y for x = 0 (solid blue line); dashed lines indicate the same function with x = ±2α/ sqrt(27). (b) Trajectories of the nonautonomous model, with γ(t) = sin(t/τ f ), and starting at the times {t 0 = −20 kyr, t 1 = −16 kyr, t 2 = −13 kyr, t 3 = −7 kyr} in the (x, y)-plane, using different colors for t 0 , t 1 , t 2 and t 3 . (c) The time-dependent forcing γ = γ(t). (d, e) The same trajectories as in (b), p but plotted in time, as y = y(t) and x = x(t), respectively. ;*

*in panels (c)–(e), the gray shading indicates intervals during which |γ| > 1/3 and the internal oscillation is suppressed.'*

by

*'FitzHugh-Nagumo (FHN) model with parameters τ x = 100, τ y = 60, and α = 2. (a) The cubic polynomial P 3 (x, y) of the fast derivative as a function of y for x = 0 (solid blue line); the red lines point to the local maximal and minimal values of P 3 (x, y), namely ±2a/ √27, respectively — these are the maximal values by which P 3 can be shifted up or down, while maintaining all of its three roots; the dotted gray lines indicate the shifted function with x = ±2α/ √27. The purple lines labeled y min and y max mark the right and left boundaries for the roots y ` and y r , respectively: y ` and y r can never be located in between the two purple lines. (b) Trajectories of the nonautonomous model with γ(t) = 0.8 sin(t/τ f ) and τ f = 1 000, plotted in the (x, y) phase plane; the trajectories are colored by their starting times {t 0 = −20 kyr, t 1 = −16 kyr, t 2 = −13 kyr, t 3 = −7 kyr} and the initial positions were drawn from a standard Gaussian bivariate distribution. (c) The slow time-dependent forcing γ(t) = sin(t/1 000). (d, e) The same trajectories as in (b), but plotted in time as y = y(t) and x = x(t), respectively; (f–h) Same as panels (c–e), but for the fast time-dependent forcing γ(t) = sin(t/350). The gray shading in panels (c)–(h) indicates intervals during which |γ| > 1/3 and the internal oscillation is hence suppressed.'*

Please note that in response the 2nd comment by referee 2 the Figure was supplemented by a second case for the externally forced FHN model, where the periodical forcing's frequency was increased in order to increase the entanglement of the different time scales involved.

For example at line 371, the description of the gamma and the fixed points that arise. I don't see any description on how the y = \gamma nullcline intersecting the cubic polynomial P_3(x,y) (manifold) is what determines the unstable or stable fixed points of the system. There are a lot of \alpha symbols illustrated on Figure 8a but there is no clear description in my opinion.

Figure 8a shows P_3(x,y), it does not show the nullcline of y with respect to x. However, in the revised version of the manuscript we have added a Figure (Fig. 9 – revised manuscript), that shows the nullclines of x and y in the x-y plane with different values for \gamma. The corresponding paragraph (line 435 – revised manuscript) explains the stable limit cycle vs. stable fixed point behavior of the FHN model on the basis of the nullcline intersection. We thank the referee for inspiring us to implement this complementary explanation.

In FIgure 8c the authors show the non-autonomous forcing for \alpha(t) and then on line 378 they introduce the non-autonomous \gamma and it is not clear what physical implications that \alpha provides although \gamma is related to CO2 eventually, and how the two non-autonomous forcing are related.

We agree that we did not provide sufficient physical explanation on what we are doing with the FHN model. The revised manuscript contains an introductory paragraph that clarifies our intentions (line 393 – revised manuscript).

*'We discuss the example of the FHN model at some length in order to illustrate how external forcing can act on a system's internal variability and thereby give rise to more complex dynamics. This model's*

*concise mathematical formulation and its widespread application in paleoclimate modeling and other fields make it ideally suited for this goal. We start with a description of the autonomous model, with no time-dependent forcing. Subsequently, we introduce a simple sinusoidal forcing and numerically compute the corresponding PBA. We then extend these consideration into the realm of random dynamical systems by adding stochastic forcing and discuss the resulting random attractor. Finally, we replace the synthetic forcings by one that corresponds to a paleoclimate proxy record of past CO 2 concentrations retrieved from Antarctic ice cores (Bereiter et al., 2015) and show that this setup brings the model's trajectories into good qualitative agreement with the D-O patterns observed in δ 18 O records from Greenland ice cores. In doing so, we pay less attention to the physical interpretation of the model's variables, while focusing on the detailed explanation of model behavior and on the role of the forcing in the resulting dynamics.'*

The authors have not referenced or discussed how this work relates to or improves upon the work of Roberts and Saha (2016) which also illustrate non-autonomous dynamics on a FHN type model. In particular Roberts and Saha draw particular attention to how they modulate the slow manifold through time dependent changes and attempt to relate it to physical mechanisms (e.g. insolation forcing). They also introduce the time dependent sinusoidal forcing on the linear nullcline in the slow component of the slow-fast system. I'm not sure which processes are more important in attempting to explain the last glacial cycle millennial scale variability; either through an amplitude modulation of the slow manifold based on obliquity paced variations as Roberts and Saha have done or the time dependant variation of the linear nullcline using CO2 as the authors have done here.

We thank the referee for bringing to our attention the Roberts and Saha (2017) paper which we now cite in the context of our discussion on the FHN model (line 391 – revised manuscript). Since our manuscript intends to showcase the relevance of NDS theory and to provide the reader with a simple to grasp, illustrative example we refrained from presenting a physical discussion on the modeling approach presented in this paper and the one presented by Roberts and Saha.

The legend for t0,t1,t2,t3 in Figure 8 looks incomplete.

Sorry, this must have happened during the update process.

I like figure 8b , I would almost like to see the 4 curves illustrated separately as the red dominates. I'm not sure if there is an easy way to do this.

We have improved the visibility of the different trajectories by changing the linewidths and the line styles.

I like the section on the MPT , but the additional value seems to come from the incremental understanding achieved from relating it to more recent concepts from NDS and RDS. I don't particularly think the title is completely appropriate , but I don't have a good alternate suggestion. The authors mention orbital insolation , but there are also internal mechanisms, plus a lot of discussion on NDS and RDS , but I'm not sure you can formulate this into an adequate short title.

Thank you for the suggestion. After an internal discussion we concluded that we would like to keep the section title as is.