RC1

Melcher and colleagues present an extended dynamical systems model designed to simulate Dangaard-Oeschger events. The aim of the model extension is to capture the irregular periodicity of observed DO-events which traditional modelling approaches have been unable to capture so far. The extended model is based on a Stommel-type box model formulated by Vettoretti et al (2020) and includes a novel control parameter that feeds back on AABW production. By varying the control parameters in time, the model is able to realistically capture the highly variable nature of stadial and interstadial duration. The authors thoroughly assess the behaviour of the new model and make a convincing case for the inclusion of the new control parameter.

Overall, the manuscript is well written, the figures are well prepared and the results are very relevant for the wider DO- and AMOC community. In the current form of the manuscript it difficult to get through the technical details in some places, and some points need clarification or additional context. Below I list my detailed comments (which are mostly minor).

> *The authors would like to thank the reviewer for their constructive and helpful review, using their precious time to review our work. The new manuscript have been elevated significantly by their inputs.*

Derivation of the model:
I appreciate the effort the authors put into introducing the model in small steps and thoroughly illustrating the baseline model from Vettoretti et al (2022). At the same time, the order of things did not always work for me and confused me in parts. Especially the presentation of the equations was confusing. In three different places the text referred to a different set of equations as the model's equations (eq1&2, eq3&4, eq8-10). I see how the three sets are related but while reading I stumbled a lot and things only became clear at the very end. I am not sure about the best sequence of things. Perhaps it could work to keep the first part of 2.2 free of equations and discuss only in terms of processes. Regarding the presentation of the equations, for me the logical sequence would have been to start with the general form of equations (currently eq3&4), introduce the Vettoretti 2022 model (eq8&9) and then introduce the updated model including alpha (eq8&10). I am sure there are multiple valid ways of restructuring Section 2.2 and its subsections. I would however ask the authors to revisit the section and improve its clarity.

> *We have reorganised and trimmed the model description in response to these comments. Although we appreciate the idea of keeping section 2.2 free of equations, we find that explaining the physics of the Vettoretti et al. (2022) model and understanding Figs. 3-4 is best supported with some use of symbols and equations. However, we have revised the presentation to (1) make it clearer when we talk about the Vettoretti et al. (2022) model and when we talk about our improved model (with slope parameter), (2) have reduced repetition and small variations in notation, and (3) have integrated equations 3 and 4 into the text as they do not play a role in our model derivation but mainly serve to make the link to the FitzHugh-Nagumo model and the Mitsui and Crucifix (2017) paper.*

Control parameters:
For non-experts in dynamical systems theory it would be helpful to always state which exactly are the control parameters in each context (instead of just referring to "the control parameters"). And also to always clarify whether a control parameter is constant throughout a simulation or being varied. In case of Fig.3: Would the system remain constant in time (except for noise) in the cases of a and c? What makes the system move along the delta_b manifold in case of constant control parameters?

> *Every instance of control parameters now only refers to all three, else we have specified which and removed 'control parameters'.*
>
> *The system would indeed be constant in time (except for noise) in the cases a and c. We try to convey this in line 150, along with the figure text of figure 3, writing 'the system is in a stable interstadial- or stadial-state'. What makes the system move along the manifold is addressed later in this document.*

Precursor Events:
How many of these precursor events are present in the actual ice-core record? Can you elaborate a bit more on why it is important to be able to replicate them? I am also somewhat confused by the discussion of the different precursor types in Section 4. Can the Type I event be distinguished at all from a regular event? Is the existence of the Type 1 events purely theoretical? Do you see all three types of precursor events in different realisations? Fig9 only shows the one example.

> *We have extended the discussion around lines 65-70 where we more thoroughly introduce precursors events following Capron et al. (2014). There are definitely two GI-23.2 and GI-21.2 as found by Capron et al (2014), and we follow Capron's suggestion to also include the MIS 3 events (numbered GI-17.2,16.2,and 15.1 by Rasmussen et al., 2014). We have highlighted this in figure 1. The reason why we care about precursor events is explained in line 70-72.*
>
> *As Type I events requires a specific configuration of the nullclines, which has an impact on the shape of the produced trajectory, it should technically be distinguishable from a regular event. However, an in-depth analysis of the actual differences is a study on its own, and not done in this manuscript.*
>
> *These points have been added in the introduction, and the discussion. Along with a discussion on why both types 1 and 2 are included.*
>
> *There are not always precursor events in the realisations. We cannot be sure whether there will be precursor events in a specific simulation, however, we can at least discern between types I and II and we do see both in different realisations.*

Alpha:
In some places it is unclear, whether the reproduced non-periodicity comes from introducing alpha or from *varying* alpha? From my understanding, the non-periodicity comes from varying alpha. This is especially important in the comparison with more complex GCMs (e.g., l419-423). Most GCMs that have produced self-sustained DO oscillations have done so with constant forcing. I would therefore think, that the self-sustained oscillations are more comparable with a constant alpha. The complex models are able to simulate the bipolar sea saw, therefore I think that some feedbacks comparable to alpha are already present in the models (whether they are realistic is another question of course). Some clarification and more detailed discussion of what the processes would represent alpha in the real ocean would be helpful here.

> *To better explain that constant alpha breaks the original model's periodicity, we added a sentence at line 304. "\alpha allows the model to have two stable fixed points, transforming the system to a purely noise-driven bistable system with aperiod oscillations."*
>
> *In this manuscript, we for the first time propose to add the alpha term to break the periodicity of simple models. A possible interpretation of what this could represent in GCMs is briefly mentioned around line 300, but we think that it is beyond the scope of this paper to explore the connection between alpha and GCM physics.*

Detailed Comments

----

l40-53 How are the two stochastic approaches different? Perhaps it si more useful to group into studies that are based on the Stommel model and those that are not? Which approaches do include the bipolar sea saw and which do not? Perhaps you could mention the FitzHugh-Nagumo model here as well already? Also make it clearer that the baseline for your new model is the model in Vettoretti et al 2022 and explain how the Vettoretti box model relates to previous box models.

> *We have attempted this by completely rewriting lines 36-62. Now we introduce the three distinct approaches, completely ODE (Roberts and Saha, 2017), completely stochastic (Lohmann and Ditlevsen, 2018), and a mixture (Kwasniok, 2013). We have added where FitzHugh-Nagumo and Vettorreti 2022 fits in this partitioning. With regards to distinguishing between the simple model in Vettoretti et al. (2022) and our model, we have made this clearer in Sec. 2.2.*

l54 include also other comprehensive models that oscillate for completeness (e.g. MPI-ESM Klockmann et al (2020,https://doi.org/10.1029/2020GL090361), HadCM3 Armstrong et al (2023,https://doi.org/10.1007/s00382-022-06564-y) or Malmierca-Vallet et al (2023,https://doi.org/10.5194/cp-19-915-2023) for an overview).

> *Added Malmierca-Vallet et al*

l104 if you keep the order of section 2.2. make it explicit that eq 1 and 2 are not the final form, but a more general form of the equations, and that the equations refer to the baseline model without alpha (see general comment on model derivation)

As explained above, we have revised section 2.2 so the difference between the Vettoretti et al. (2022) and new models is clearer.

l109/110 the dependence of q on delta_b is not obvious from eq1

You need equations and arguments in Section 2.3.2 (Model formalism) to derive the new equation (4). Here you see the full nonlinear behaviour. Here in section 2.2, we aim to keep the level of formalism low, but we have added a reference in the text to 2.3.2 to make it clear where the full argument can be found.

Fig3 caption: "as lambda changes" what does "lambda" refer to here? The Eigenvalues have not been introduced yet

Should have been Gamma, changed.

l157-157 is that also what you do or is it in contrast to what you do?

Yes it is sort of the same. Added a line to clarify.

l163 what makes the southern bouyancy flux increase?

Added that the AMOC decreases, which makes the southern buoyancy flux increase

l166 is the salt-advection feedback acting on the northern or the southern box? Not equal to the typical AMOC-North Atlantic salt advection feedback?

Added a paragraph on the freshwater feedback

l162-177 be careful with words. transport/flow/circulation seem to be used for q and AMOC simultaneously?

Added clarification in the text to clarify the difference

What makes the system move along the manifold? I am probably lacking some basic dynamical systems theory here (see general comment on control parameters).

The system is autonomous, meaning that (temporal) changes in B drive (temporal) changes in deltab and changes in deltab drive changes in B. The system is coupled and it produces a relaxation oscillation in dynamical systems parlance. Therefore for each time step the equations describe where the system moves to making it move along the manifold. Rember along the slow manifold there is a zero derivative in the \dot{deltab} direction but not in the \dot{B} direction.

l182 what about the heat loss removing buoyancy in northern box/real ocean?

Added line explaining the effect of heat loss removing buoyancy l193 I find this information confusing here without context.

We hope the new structure solves this problem

l194 link eq3/4 back to eq1/2 (see general comment on model derivation)

We hope the new structure solves this problem

l212 are there haline and thermal modes in this model?

We aim to follow the nomenclature from classical literature. In Stommel's or Cessi's work, the q>0 is the thermal mode and the q<0 is the haline mode. This is what is being referred to. In Stommel's model the thermal model is driven mainly by meridional thermal gradients dominating, while the haline model is driven mainly by the meridional salinity gradients. This model can not really separate these dominant influences because the dimensionality of the system is collapsed (combining deltaT and deltaS into deltab)
As this is not strictly necessary for the understanding in our manuscript, but adds context for the expert reader, we only lightly touch upon this.

eq7 what are B_c and b_c?

They are characteristic values used to non-dimensionalize the physical equations, and thus also used to redimensionalize them back to physical units. Introduced in line 266 with numerical value and unit

l228-230 mention this earlier?

We like this placement as it gives the reader all the necessary information to understand how q and AMOC is related and what the salt-advection feedback is.

l242 which control parameters do refer to here?

Alpha and gamma, substituted in the text

l274ff state stability criteria up front for convenience, otherwise the reader needs to guess what the criterion is (negative Re(lambda1,2)? ad if yes, what does that mean?)

Done on line 333

Fig5 Is the case with alpha = 0 in c equivalent to the cases presented in Fig3a-c?

This is true as stated in the caption of figure 5

l282 the change from b to d seems quite drastic, not so much from d to f

"Drastically" removed

l296 which control parameter do you refer to here?

All of them, changed to plural

2.2.3 refer to left hand side of Fig.5 throughout the section to illustrate

Added clarification of current references not sure more is needed

l295ff interpretation of b0/value of b0?

Such an interpretation is given on line 136, …,

l314 How do you set the initial conditions? Simply as a set of values for delta b and B?

Yes, (delta b, B) = (1,0), we have added to line 378

l323 how do the 100 members differ? only in the noise realisation?

Yes

l327/l344 the age is impossible to determine from the crosses in Fig6. The caption suggests the crosses should be numbered, but they are not.

They are now numbered. Thanks for catching that

Fig6(especially a/b) How do mu_E=0 and a variable mu_P go together? If there is no event, should mu_p not always be either 0 or 1?

As mu_E and mu_P are ensemble averages, it is cases where the system is highly stable, but for the given set of conditions might have a stochastic jump. So in most cases, E = 0, but sometimes E = 1 and mu_E is close to, but not equal to, 0. Depending on when in the trajectory the jump happens, the P can be higher or lower. We have changed the text to better reflect that mu_E and mu_P are ensemble averages.

l335 "with decreasing gamma, mu_P' moves to lower values" is only true for gamma between 1.5 and 3. for gamma < 1.5 mu_P' increases with decreasing gamma. Or am I reading the plot wrong?

Fixed on line 403 and fixed mistake in figure 6

l338 what exactly do you mean by "extends the arches downward"? Do you mean that each line of constant gamma reaches lower values of mu_E'?

Yes this formulation has been added as a clarification on line 406

l342/343 again not sure what you mean by "increasing sigma moves the entire arch downward"?

Should have said alpha instead. Fixed

l382-385 does this refer to the sequence in time?

Yes, added clarification.

l386 Are 91 and 21ka b2k the right timings? To me 97 and 24ka b2k would seem more appropriate? How relevant is this misfit?

You are right and it has been updated and an additional precursor at 107 ka b2k This is not that important as it changes with every realisation, and not all realisations have precursor.

l389/390 It might be worth including this member also with a separate colour to a and b.

Very good suggestion this has been done.

l425-427 but would you have expected anything else given that the target are summary statistics and not the actual timeseries?

This paragraph reflects our frustration with using RMSE as a measure of performance. Even if our model had generated the actual time series, RMSE would not answer the question of whether our model can truly recreate that time series. RMSE only tells us that, at each time step, the model's output stays within a certain range of the actual data. But this doesn't guarantee that the model can accurately reproduce the entire time series. RMSE assumes at each timestep an individual point distribution and completely disregards the temporal correlation between subsequent points. We have not been able to come up with a better (and still simple) measure, but hope a better solution will be proposed in future work.

Editorial Comments

All corrected with one comment: Origo is the Danish word (from Latin) for the origin, in the specific context of a coordinate system.

----

l33 first mentioning of NGRIP here but the acronym is only introduced in l87

l209 and Eq6 say the same thing

l269 "close to 1" instead of "intermediate"?

l271-273 difficult to read, please reformulate more clearly

l278 It should be "delta_b" in the equation and not "b"

l323/324 move sentence with Runge-Kutta to l316, after "[...]run for 5371 steps."?

l375 Is "linear interpolation" correct? The caption of Fig7 says cubic spline fit

l410 "improve" instead of "determine"?

l412 What is "origo"?

l421 "might be able" instead of "might be possible"

l456 "vertical jumps" instead of "vertical jump"