Reply to the reviewer's comments: TransEBM v 1.0: Description, tuning and validation of a transient model of the Earth's energy balance in two dimensions (gmd-2020-237)

Elisa Ziegler, Kira Rehfeld

February 23, 2021

Overall description of actions taken

We have

- implemented the suggestions from all referees, and in addition
- made changes to Figures 5, B1, and B2 (before the insolation plotted was effective insolation including the albedo, now the albedo has been removed, the overall descriptions and conclusions hold),
- and fixed some additional typos.

RC 1: Review from Kelin Zhuang

Reply to the referee

We thank Kelin Zhuang for his highly favorable review and for his comments.

A detailed response to the valuable comments as well as a description of the actions taken is given below.

1 Detailed response

(Original report cited in italics)

1) L12, Abstract, Buhler et. al., better followed by ",in prep" or a footnote to show it is unpublished

We will modify the Buhler et al. citation so that its status is apparent.
Action: The reference has been updated.

2) L13 I would prefer to combine the two simple sentences together, such as "We find that the EBM lacks internal climatic variability mostly due to its reduced descriptions of heat transport and the hydrological cycle."

We will combine the two sentences in L13 as suggested.
Action: We have combined the sentences.

3) L18, "fill gaps" should be changed into "fill in gaps"
4) L23, "state" into "states"

We will correct the typos in L18 and L23.
Action: We have corrected the typos.

5) L60, with regard to hysteresis explored by the EBMs, I would suggest the authors add another publication. Zhuang et al., 2014, Hysteresis of Glaciations in the Permo-Carboniferous, Journal of Geophysical Research ATMOSPHERES

We will add the suggested reference to the discussion of hysteresis in EBMs, which indeed fits this section very well.
Action: We have added the suggested reference.

RC 2: Review from Shaun Lovejoy

Summary of changes

We thank Shaun Lovejoy for his detailed feedback on our manuscript, which raises many highly interesting aspects. In response, we now

- provide an estimate of the relaxation times and discuss them in the revised manuscript and
- extend the manuscript to discuss in particular the heat transport, restarts, linearity, and tuning further, as well as the other many interesting points that the referee raised.

A detailed response to the valuable comments as well as a description of the actions taken is given below.

2 Detailed response

(Original report cited in italics)

2.1 General comments

A). The authors are commended for pursuing the development of energy balance models, making them more realistic. This is a welcome counter to the increasing trend of using GCMs to answer all climate questions. The authors justify their approach by invoking the flexibility of such "low complexity models" with respect to GCMs (they allow for "fast and repeated" simulations). But there are other advantages to their approach and there is no reason that high model complexity is a sine qua non for realism. In actual fact, the development of GCM alternatives is very timely. This is because it is increasingly clear that each (increasingly complex) GCM has its own climate - presumably none have the same climates as the real world.

By exploiting large amounts of historical data, EBMs have the potential of reproducing the real world climate, thus providing results with both lighter computations but that are also more reliable. The authors might mention that it has been proposed that models of EBM type can be thought of as high level models that attempt to account statistically for huge numbers of interactions, of details (e.g. [Lovejoy, 2019]). EBMs and kindred approaches are therefore not just "poor man's" GCMs.

While I have a number of technical questions that I would like the authors to address, overall the paper is well written and the public availability of the code makes it especially appealing. If the authors can answer the questions below, I recommend it for publication, it will be welcome addition to the literature.

We thank Shaun Lovejoy again for this favorable assessment of our work. We agree that EBMs can provide insights different to those provided from GCMs. We will update the introduction with the mentioned reference to introduce the framing of EBMs it proposes and emphasize the value of EBM-type models further.

Action: The values of EBMs was emphasized more. Unfortunately, the provided reference was not available to us and we could not find any other suitable reference.

B). My main disappointment is that the authors didn't provide much theoretical guidance to understanding their results (nor indeed for justifying the numerical constraints such as choice of time step and "spin-up" time). Indeed (ignoring the "restarts") their EBM is linear so that standard linear analysis could be made. This is facilitated by the excellent North and Kim monograph that develops the theory for the (admittedly simpler) 1-D case with constant coefficients.

Specifically, equation 1 could be Fourier (or – depending on the application – Laplace) transformed in time to reveal the key time scales. For example for deviations from the *mean, we take A = 0 and obtain:*

$$\tilde{T}(\omega, \underline{r}) = \frac{\tilde{F}(\omega, \underline{r})}{(i\omega\tau) + 1 - (\nabla \cdot D\nabla\tilde{T})/B}; \tau(\underline{r}) = C(\underline{r})/B(\underline{r}); \tilde{F}(\omega, \underline{r}) = S_0\tilde{S}_F(\omega, \underline{r})a(\underline{r})/B(\underline{r})$$

Where the tilde indicates Fourier transform in time and F is the effective forcing (incidentally, using the notation S 0 and S F for quantities with different units is not good practice). North and Kim develop essentially this equation for the case where B, C, D are constants and then expand the temperature in Legendre polynomials. This leads to:

$$\tilde{T}_n(\omega) = \frac{\tilde{F}_n(\omega)\tau_n^{-1}}{(i\omega\tau_n) + 1}; \tau_n = \frac{\tau}{1 + (D/B)n(n+1)}$$

Where the subscript "n" is for the n th polynomial. In the 2-D case discussed here, the case with homogeneous coefficients can instead be dealt with either full spherical harmonics (or make the flat earth approximation and perform spatial Fourier transforms). The result is a typical relaxation time scale τ_n than depends on the spatial scale ($\approx 1/n$).

Although the above simple analytic result is strictly valid for constant coefficients, in fact the model constants are homogeneous (or slowly varying in the case of D) over wide swathes so that - except for region boundaries - we should be able to use the above estimates to obtain the basic time scales for adjustment. The basic relaxation times (= C/B) are for ocean and land where (using values from table 2) we obtain $\tau \approx 10$ days (land), ≈ 6 years (ocean). These are fundamental model time scales should help justify the time step (≈ 1 week) and (should) help explain the "spin-up" time (this is really a nonlinear model concept, in this linear model it is more accurately, the time for the classical, exponentially decaying transients).

Another useful consequence of the above is that it explains rather naturally the annual cycle phase shifts: at critical frequencies $\omega = \omega_n = 1/\tau_n$, the phase shift is $-\pi/4$ indicating that the temperature lags the forcing by $365/8 \approx 46$ days. At high frequencies, the lag is $\approx -\pi/2$ whereas at low frequencies it is ≈ 0. Putting in numbers, for annual forcing, we therefore anticipate for ocean regions, the temperature lags by $\approx 60 - 80$ days, whereas over land, $\approx 10 - 20$ days, numbers that compare reasonably with the simulations (fig. 11).

The same equations should (presumably) explain the amplitudes of the annual cycles.

The referee raises many interesting points here that we will be happy to take into account in the discussion of the revised manuscript. With respect to an analytical solution of the model at hand, we do not see how this would be possible in light of the non-homogeneous, spatially-resolved boundary conditions that represent the surface types.
We agree that an estimation of the relaxation time scales (as provided previously for the

model in Zhuang et al. (2017)) would strengthen the manuscript and provide valuable insight into the time scales at work and how they relate to the numerical constraints and the simulated seasonal cycle. We will incorporate this in the manuscript upon revision.

Action: The relaxation times were computed and are provided and described in Sec. 2.2. Furthermore, we have included the computation of theoretical values for the amplitude and phase lags as outlined by the referee in our discussion of the seasonal cycle in Sec. 3.1 and discussed them there.

2.2 Detailed comments

1. In the paper, the transport equation (eq.1) has a diffusion term $\nabla \cdot D\nabla T$ where the 2-D gradient operator is used, D is the thermal conductivity. In the North and Kim book (and earlier papers going back to Sellers 1969), the diffusive term is $\frac{\partial}{\partial \mu} D(1-\mu^2)\frac{\partial T}{\partial \mu}$ where μ is the (dimensionless) cosine of the colatitude. A consequence is that in the North and Kim treatment, D has the same dimensions as the sensitivity: W/(m^2 C), it is a conduction coefficient per radian, the value is the same as that given in the 1981 review: D = 0.67 W/(m^2 C). In the paper (table 3), values of the order of 1 are given for D and the units are in W/C. The obvious explanation for the discrepancy is that the model is 2D so that the model uses values multiplied by the average grid area which is 6.2x10 10 m^2. While this would have the correct units, it is very far from the values given. Therefore, please express D in standard units of thermal conduction W/(m C). (I checked the Zhuang paper and it doesn't give the values either, I suppose they were somewhere in the Fortran code?).

Indeed, as the referee assumed, both the values and computation of D (different from that used in other references from North et al. we consulted, including North and Kim (2017)) reflect those in the Fortran code accompanying Zhuang et al. (2017). We will investigate the possibility of expressing the thermal conductivity parameters in terms of the suggested units.

Action: Since the diffusion term D appears as $\nabla(D\nabla T)$ resulting in mixed terms with respect to the derivative converting it to the suggested units is not possible. For comparison with other literature, it should be easier to convert the values there.

2. P11. Explain restart a little better. It seems to essentially be a way of introducing nonlinearity, but this is not clearly expressed. Can the mathematics not be indicated (explicitly with an equation) with forcing that depends on the temperature?

Since they are a central part of our modifications to the model, we will gladly improve the explanations of the restarts. Restarts allow the changing of the boundary conditions, relating these adaptations of the boundary conditions (i.e. changing ice sheets and sea level) to the GMT at the end of the previous run would be possible, but has not been implemented at this point.

Action: The explanation of the restarts in Sec. 2.3 was extended.

3. P17, phases, amplitudes: It would be useful to estimate these for homogeneous regions (see the above).

There are latitudinal dependencies which break the homogeneity leading to differing amplitudes and phases in nominally homogeneous regions. Therefore, we find it hard to determine regions, for which these calculations can be done.

4. Tuning: this paper follows the tradition of guessing, then tuning model parameters to attempt to fit the data. There are so many parameters that this is hazardous. Is there no way to try to estimate the parameters directly? (Using standard units for thermal conductivity would make this more transparent!). Also I'm surprised (table 8) at how little tuning affects the initial guess parameters. I'm not sure a lack of sensitivity is a good thing? Please comment.

The tuning process did indeed not use a regression minimizing the tuning metrics across the whole multi-dimensional parameter space. We agree that automating and improving this procedure is desirable and have therefore since submitting this manuscript worked on a procedure to do just that. As it stands, the tuning did sample the parameters spaces of the individual parameters. In conjunction with the, by comparison, limited amount of parameters, the tuning result is not pure guess work, and, in our opinion sufficient for the applications presented here. The sensitivity of the model to changes to the parameters varies as shown in Figures 8 and 9: Most parameters can produce large responses in the simulated temperature profile, in a number of cases even small changes to the parameters have a drastic effect, e.g. any parameter related to the ocean or outgoing radiation. Table 8, on the other hand, reflects the fact that the initial parameterization provided by Zhuang et al. (2017) was already quite quite good with respect to our tuning goals. Overall, the model does not, in our assessment, seem to suffer from a lack of sensitivity.

Action: We extended the Discussion (Sec. 5) to discuss the tuning.

5. P18: You mention nonlinearity, the equation is linear so that presumably, the origin is in temperature dependence of parameters (e.g. albedo), and this occurs via restarts but this is not mentioned. Intuitively, this means that the model has a slow nonlinearity, but it would be important to state this mathematically more clearly. Please explain (see point 2 above).

As suggested, we will expand on the (limited) non-linearity in the model.

Action: We have adapted the descriptions of the non-linearity in the Abstract and in Sec. 3.2.

6. P19: You mention "spin-up". This is a term normally used for nonlinear models such as GCMs. Your model is linear so that you are discussing transients, and these can be analyzed by classical methods (see above). I think it would clarify how the model really works.

We thank the referee for pointing out this imprecision in our wording and will improve this upon revision.

Action: The word spin-up was removed from the manuscript and instead we describe this model behavior as decay of the transients (c.f. Sec. 2.4, 3.3).

7. Section 3.3: Here and as earlier with spin-up the key is the effective global scale relaxation time: the parameters are no so different from North and Kim so that the relaxation times should not be very different. I expect the Last Millenium simulation to thus be a low pass filter of the forcings at time scales longer than the longest (ocean) relaxation time, hence that their long term statistics for example will be the same. This is of course not true if the restarts introduce enough nonlinearity. The only complication is due to the diffusion term that is variable in space. But it's magnitude is not in fact so variable.

We will provide an estimate of the relaxation times and discuss them in the revised manuscript.

Action: As described above, we have estimated the relaxation times.

8. Table 1. – The caption of the table says the heat capacity values are taken from Zhuang. et al. (2017b) while in the table it suggests they are from Zhuang. et al. (2017a), but they do not correspond to the values given in Zhuang. et al. (2017a).

We thank the referee for spotting this error, indeed the values are taken from the model code (Zhuang et al., 2017b in the manuscript) and we will change the references accordingly.

Action: We have changed the reference.

9. Line 231 – Why was minimizing the RMSE prioritized over agreement with GMT?

In our assessment, for a two-dimensional model to provide added value to zero- or one-dimensional EBMs, it is important that the spatially resolved temperature field produces realistic features, which is why the latitudinal profile was prioritized over the absolute value of the GMT. The latitudinal profile furthermore showed considerable deviations from observation in regions of interest (e.g. the polar regions), whereas the GMT was already very close to that found in the ERA climatology. Similar arguments hold for the seasonal profile. There are, of course, applications where the GMT, or more so the development of the GMT anomaly would take precedence. Depending on the applications the relative importance of the tuning metrics would change.

Action: We have extended Sec. 2.4 and 5 to discuss this prioritization in the tuning process.

10. Table 7. – Clarification on why what the bracketed (x) represents as opposed to just x?

We are grateful that the referee noticed this omission in the manuscript. The bracketed x is meant to represent a notably lesser, but not negligible, influence than the other parameters. We will update the description accordingly.

Action: We now provide the explanation in the Table caption.

11. Line 246, Figure 10a, could the choice of 1960-1989 climatology – a period with strong aerosols which mask the CO2 warming – be the reason for underestimate of warming? What are the consequences of changing other climatologies, say per-industrial or closer to present (less aerosols)?

We agree that the choice of reference period for the tuning will have an influence on our results. The underestimation is also definitely related to the forcing considering that a similar underestimation is not an issue in the past millennium simulations. So far, we had not considered the aerosol forcing as a possible source, but it is certainly worth looking into and examining further, we thank the referee for raising this interesting point.

Action: We discuss the issue of tuning period and choices with respect to the tuning in Sec. 2.4 and 5. With respect to the aerosol forcing, we cannot, at this point, conclude that it is the reason for the underestimation of the warming.

Figure 10. – It appears the tuning gives little improvement (GMT anomaly and average latitudinal temperature) or actually hinders emulation of the reanalysis (seasonalamplitude)? What is the reason for the much lower amplitude of the seasonal cycle in both hemispheres?

Due to the prioritization of the latitudinal profile (and reducing its deviation from observations) in the tuning, this is the metric with the most notable improvement, whereas the GMT was already in quite good agreement beforehand and therefore only improved in decimals. The amplitudes of the seasonal cycle are mostly affected by the changes to the heat capacities of ocean and atmospheric column, which also promoted the slight shift and the better agreement in the southern hemisphere, but came at the expense of the agreement in amplitudes.

Action: The discussion of the tuning was extended as described above.

12. Line 83: The refence to [Rypdal et al., 2015] for scaling and EBMs is not accu-

rate. When [Rypdal et al., 2015] made their fractional modification of [North et al., 2011]'s heat diffusion model they obtained it precisely by removing the critical energy balance term! The [Rypdal et al., 2015] model neither deals in energy, nor does it allow for balance. It is unstable with respect to infinitesimal step forcings, it's ECS is infinite. In the expression "EBM", one thus must eliminate the "E" and the "B". To correct this fundamental fault, one must reinsert the linear balance term that was removed and that corresponds physically to black body emission, and in this case one obtains the Fractional Energy Balance Equation ([Lovejoy, 2019], see the discussion of this point in [Lovejoy et al., 2021] http://www.physics.mcgill.ca/ gang/eprints/eprintLovejoy/esubmissions/QJRMS.FEBE.revised.3.11.20.pdf). However the introduction of the necessary balance terms comes at the cost of having two different high and low frequency scaling regimes. Incidentally, the authors could also mention that energy balance models have been used to make global scale climate projections to 2100 [Procyk et al., 2020]. http://www.physics.mcgill.ca/ gang/eprints/eprintLovejoy/neweprint/esd-Procyk.discussion.2020-48.pdf.

We will modify the discussion of the reference in response to the referee's comment. Adding a discussion on the use of EBMs for projections is indeed valuable and we will make the requested improvement to the manuscript.

Action: We have amended the discussion of the Rypdal et al. 2015 reference. We have also included future projections as a possible application of EBMs in the Introduction. Since the peer review process for Procyk et al. is currently under way, we have opted for an earlies reference from the group: Hébert et al 2019, on which Procyk et al builds.

13. In the abstract it is stated that the EBM lacks internal variability. I find the statement a little odd since a linear model cannot generate internal variability: none would be expected?

We will correct this statement and emphasize that, while the EBM is not fully deterministic, in particular when considering randomized forcing, it does not represent internal variability of the atmosphere-ocean system.

Action: We have changed this part of the abstract.

14. Line 95: The issue of energy storage was not mentioned. Please discuss where the energy is stored in this model.

The model does not have an explicit storage of energy, the effect that storage has on the climate is approximated using the heat capacity parameters.

Action: We point out that energy storage happens implicitly in the discussion of the heat capacity parameters in Sec. 2.2.

15. In eq. 1, a is the co-albedo, not the albedo, although eq. 4 uses it as an albedo.

Please fix this. Also, as indicated above, the notation S_0 and S_F with different units is confusing. Why not use the North and Kim notation?

The notation with respect to the albedo is indeed inconsistent between the equations and we thank the referee for raising this point and will correct this. With respect to insolation, the notation in the manuscript follows the implementation of the model as it is provided in Zhuang et al. so as to ensure that the manuscript and model code match.

Action: Eq. 1 now contains the albedo, not co-albedo.

16. As mentioned above, in table 3 please use correct units for thermal conductivity.
Please refer to the reply to point 1.

17. Eq. 2: where does this form for D come from? It seems a bit weird?

This is in accordance with the model code provided for and documented in Zhuang et al. (2017). The computation of D was not changed in our extension of the model.

18. Given that the longest time scale in the model is the ocean relaxatjon time of about 6 years (see above), presumably, the last Millenium simulation is just a low pass filter of the forcing? (or is nonlinearity somehow important?).

These are very interesting points, which we will investigate further.

RC 3: Review from an anonymous reviewer

Summary of changes

We thank the anonymous reviewer for taking the time to assess our manuscript and for providing many valuable and in-depth comments. In response to the suggestions by the reviewer, we have

- improved the readability and flow of the text in sections 2.2 and 2.3 and
- improved the presentation of figures and tables, in particular with respect to color choice, as well as table layout and descriptions.

A detailed response to the helpful remarks of the reviewer as well as a description of the actions taken is given below.

3 Detailed response

(Original report cited in italics)

3.1 Minor comments, general

I find the flow of the text disrupted in Sect. 2.2, due to the many small tables and long table captions. I ask the authors to make the row/column structure consistent for every parameter and shorten the captions if possible. I suggest rewriting paragraph 2-4 of Sect. 2.3, preferably placing the details related to software, computer type, compilers and processor in the same paragraph.

We appreciate this valuable comment and agree that the tables as they are hinder readability. We will make the table structure consistent and try to shorten the descriptions as much as possible. We will also rewrite Sect. 2.3 to reflect the referee's comment and improve the flow of the text.

Action: To improve the readability, we have combined the small tables for each parameter group into one table (Table 1) for all of them with a much shorter caption. In addition, we have reworked the sections for a better flow of the text. We have also placed the details regarding software, etc. in the same paragraph.

3.2 Minor comments, specific

L. 10, 210,364: when referring to the climatological period, I would prefer to write the complete year 1989 (1960-1989, instead of 1960-89). You switch between both writings throughout the text.

We thank the referee for spotting this, we will write the years out in full consistently upon revision.

Action: All years are now written out in full.

L. 18: "...fill gaps left by proxy and observational records," -¿ isn't a proxy record also an observational record? Consider reformulation.

Indeed, this is misleading, we will change the phrasing.

Action: We have changed the phrasing.

L. 233: "... were only discarded if they produced a change in GMT by several degrees". Please specify to greater detail the cut-off. "Several degrees" is not informative enough.

We agree with the referee that this is too vague and will improve it.

Action: We now state clearly that we mean 2 degrees or more.

L. 127: "...timescales of order 10ˆ2 years and higher", consider replacing "higher" with "longer".

We will implement this replacement.
Action: We have implemented it.

L. 251-252: "TransEBM agrees well with the reanalysis. In particular, it is able to simulate the dip in temperatures around the equator as well as the temperatures in the polar regions."
This appears to be contradicting the sentence on Discussion lines 371-372, or you need to elaborate:
"In the latitudinal temperature distribution, the dip in temperatures near the equator related to the Intertropical Convergence Zone (ITCZ) is not reproduced."

We thank the referee for noticing these inconsistent statements. We will adjust the writing in both cases to refer to the specific simulations: in the first case, the statement refers to the tuning period, for which a better match at low latitudes around the ITCZ was achieved in comparison to the parameterization by Zhuang et al. (2017). The part in the discussion was meant to describe simulations outside the tuning period. In these, the agreement in equatorial regions decreases as shown for example in the simulation of the past millennium. We will improve the statements accordingly.
Action: We have updated the discussion in Sec. 3.1 and 5 accordingly.

L.269-285: consider specifying that the validation in this subsection relies of the implemented restarting extension.

We will improve the phrasing of the subsection to make this clearer.
Action: We have improved the phrasing.

L. 286-306: Similarly, this validation is associated with the extension to transient forcings and transient simulations. Consider highlighting these features.

Many thanks to the referee for this excellent suggestion, we will incorporate it in the text.
Action: We have improved the phrasing.

L. 288: "...follows Neukom et al. 2019, as does CO 2 ." Not clear what you mean by "follows", please reformulate.

We will improve the wording.
Action: We have changed the sentence accordingly.

3.3 Comments on figures, including color choices

The following comments on color choices and contrasts for the figures are given because printed colors appear slightly different than they do on the screen. A high-quality printer was used for printing this manuscript, so the following comments should be generally applicable.

Page 5, Figure 2: the black font on blue background color is difficult to read in printout.

Indeed, we did not test the colors in a printout and are grateful the referee made us aware of this shortcoming of our color palettes. We will adjust and test them during the revision of our manuscript for all figures, in particular those suggested by the referee below.

Action: We have changed the colors to improve them for printouts.

You use italic fonts for CO 2 for the first time in the caption. Italics are also used later in the main text, but inconsistently. Normal fonts are used e.g. in the abstract. I prefer normal fonts for CO_2, please check throughout the text and make the use consistent.

Again, we thank the referee for spotting this inconsistency. We will eliminate it in the revision process.

Action: $\mathrm{CO_2}$ now only appears in non italics.

Page 12-13, Figures 6-7: the individual colors used to distinguish "changed" and "unchanged" features are too similar for the printout.

Page 14, Figure 8: Yellow-ish colors are difficult to discern for the printout.

Page 20, Figure 13: Please add legends to this figure as well.

We will move the labels into a dedicated legend.

Action: We have changed the colors for "changed" and "unchanged" in Figures 6 & 7. We have replaced one of the yellows with a red. The legend was moved outside of the figure.

Page 21, Figure 14: colors of PAGES2k and CESM time series are too similar to discern for the printout.

Page 29, Figure C1: both panels labeled as (b). Numbers superposed on the maps are difficult to discern for printout.

We thank the referee for discovering the error in the panel labels, we will correct it. Similarly, we will improve the coloring of the superimposed text.

Action: The color for CESM was changed. The panel labels were corrected and the coloring of the superimposed text changed (and the font size increased and labels moved where necessary).

3.4 Comments on tables

Consider shortening captions for tables 1-4 in sect. 2.2.

We agree with the referee that the tables disrupt the readability of the section and will try to shorten them accordingly.

Action: The tables were combined into one so as to not be as disruptive.

Page 19, table 9: consider specifying the context of the zero-sea level of Grant et al. (2012).

We will specify the zero-sea level of the reference.

Action: The zero sea level is now described in the caption.

3.5 Suggested addition and references

Introduction pages 2-3, lines 58-73:
Studies show that EBMs are able to simulate hysteresis and tipping points, but CMIP5 GCMs cannot simulate such strong transitions, exemplified for the Arctic sea ice and the Atlantic meridional overturning circulation (Wagner & Eisenman (2015), Bathinay et al. (2016)). It could be relevant to highlight this capacity of the EBMs compared with more complex models.

We agree that this is a quality of EBMs that it is worthwhile to mention and thank the referee for pointing us to the references. We will adjust the paragraph using the suggestion.

Action: We extended the discussion of hysteresis and tipping points and included the references in the introduction.

3.6 Comments related to the GitHub code repository

Include readme file in repository:
The code is well-documented in the GMDD manuscript and in the GitHub repository. Please ensure that the manuscript and associated documentation can be easily traced from the repository. A readme file visible on the front page of the repository is recommended, referring to the Zhuang et al. (2017) and Ziegler & Rehfeld manuscripts. The readme file could for instance also list the necessary software needed to run the code and repeat the statement of the software license which is included in the manuscript.

We agree that the repository should include a readme (listing references, license information, and necessary software) and will update the repository to that effect.

Action: This readme and the test run description will be included in the next update of the repository.

Suggestion of test code visible on the front page of the repository:
The authors describe the default configuration file on manuscript lines 325-329. This information together with other defaults could be summarized and highlighted in a separate "Test run" file of the repository, instructing the user to an example testable code to help validate their installation.

Another excellent suggestion that we will gratefully implement.

References

North, G. R. and Kim, K.-Y.: Energy Balance Climate Models, Wiley-VCH, Weinheim, 2017.

Zhuang, K., North, G. R., and Stevens, M. J.: Model code: A NetCDF version of two-dimensional energy balance climate model based on the full multigrid method in FORTRAN, URL https://github.com/ElsevierSoftwareX/SOFTX-D-16-00023.

Zhuang, K., North, G. R., and Stevens, M. J.: A NetCDF version of the two-dimensional energy balance model based on the full multigrid algorithm, SoftwareX, 6, 198–202, https://doi.org/10.1016/j.softx.2017.07.003, URL http://dx.doi.org/10.1016/j.softx.2017.07.003, 2017.