Response to reviewer 3

I. **Methodology**

The reviewer begins from the formulation three "necessary steps required to assessing the future glacial inception under different levels of carbon dioxide emissions":

1. *Using a low-rder model to capture the very long term (millennial) of climate, under physical constraints. For instance the the three-dimensional stochastic system of Saltzman and Maasch (1991).*

Apart from the fact that the reviewer recommends us to use the Saltzman and Maasch (1991) model instead of our own (which is lacking "*any strong argumentation/reason/justification*"), the meaning of this sentence with several typos and missing words is not clear to us. Not mentioning the fact that the model is designed not for millennial but for orbital and longer time scales. As far as the choice of the modelling approach is concerned, we must state that while we have a great respect to the works Burry Saltzman and his colleagues made during the 90's, at present this and similar models are only of historical interest. Saltzman's model is based on the assumption that glacial cycles represent self-sustained oscillations in the Earth system, and to produce such oscillations a cubic power term has been introduced in the equation for carbon dioxide without any justification. Since the 90's our understanding of Earth system dynamics advanced significantly but no one was able to simulate self-sustained oscialltions in the Earth system with realistic models and no one ever discovered this cubic power term. Needless to say that Saltzman and Maasch (1991) and similar models cannot be used for simulation of the impact of anthropogenic CO_2 emission on climate. To the contrary, our simple model is to a large extend based on the results of our own comprehensive Earth system model CLIMBER-2 which is able to simulate successfully not only the latest glacial cycles (Ganopolski and Brovkin, 2017) but also all Quaternary glacial cycles (Willeitet al., 2019).

2. *Treating the estimate of the model parameters and the forecast probabilistically.*

As we show in our paper, the problem is not the treatment of model parameters "probabilistically" but the fact that past climate data provides no sufficient constrains on model parameters suitable for future simulations. We found that the majority of model versions which successfully simulate past glacial cycles have unrealistic relationship between the critical insolation threshold and CO_2 concentration and thus cannot be applied for the future simulations. In such a situation, any attempts to attach "objective" probabilities do different model realisations is nothing more than quackery.

3. *Verify the accuracy and validating the model statistically and by checking the reproduction of physical phenomena.*

Which *physical phenomena* reviewer means here we do not know, but we fully agree with the next reviewer's sentence: "*Different physical assumptions may lead to dynamical systems with dynamical properties that are similar enough to produce a convincing visual fit on palaeoclimate Data*". Indeed, during the recent decades different workers proposed a number of completely different mathematical manipulations which transform some combinations of the Earth's orbital parameters into curves with variability patterns more or less similar to the glacial cycles of the late Quaternary. Moreover even the use of a "*Bayesian framework*" did not help M. Crucific (Crucifix, 2012) to distinguish between the right models simulating glacial cycles as nonlinear response to orbital forcing, and the wrong models where glacial cycles originate from self-sustained oscillations. This is, of course, not surprising – the correct model cannot be derived solely from paleodata.

The approach which we employ in this study, and which represents a further development of the method used in Archer and Ganopolski (2005), is based on using a combination of paleoclimate data and the results of physically-based Earth system models. We believe, this is the only feasible alternative to the use of complex Earth system models, which are by far too computationally expensive for this task. In the "*critical comment which should be absolutely addressed*" #8 the reviewer wrote "*validating a work using results from another simulation [Ganopolski et al., 2016] does not seem accurate to me*". What "*accurate*" means in this context we do not understand. And, of course, we did not validate one model by another one. Instead, we use the results of the physically-based and well-tested Earth system model CLIMBER-2 to constrain parameters of a simple semi-empirical model which cannot be constrained by paleodata. We do not believe that there is an alternative to our approach.

In the rest of the review, the reviewer repeats time and time again that the right approach is the approach described in Crucifix and Rougier 2009 (hereafter CR09, the reviewer cited this paper nine times) and that our approach is absolutely unjustified (the reviewer used expressing containing "(un)justified" and "justify" more than 20 times!). We were glad to learn from Crucifix's comment that the reviwer#3 is not Michel Crucifix. We have great respect to Michel Crucifix whom one of the authors (AG) knows for 20 years since the time when Michel Crucifix was PhD student and AG was a member of his PhD committee. However, the methodology described in CR09 is absolutely inappropriate for our purposes. The main reason is that, although CR09 manuscript is entitled " On the use of simple dynamical systems for climate predictions", the authors of this manuscript used the term "climate predictions" with a meaning different from the one usually used. CR09 is about modelling of future glacial cycles without any anthropogenic influence. Of cousrse, "climate prediction" usually means modelling of climate response to the anthropogenic perturbation. This is obviously the central goal of our study. The model which has been used in CR09 is not suitable for this task and the methodology described in CR09 is of no use for development and testing of such a model.

Besides, the Bayesian approach is not the only one possible or correct. Parameter estimation can be approached either through the frequentist or Bayesian point of view. In the frequentist framework point-estimates of unknown parameters are obtained and it is not possible to assign probabilities to the parameter values. It is assumed that there are enough measurements to derive useful information on the parameters. In the Bayesian approach the unknown parameters are treated as random variables and the measurements are complemented with information about a prior belief about the parameter values. Results may vary depending on which prior is selected. We opted for the frequentist approach and, therefore, there is no probability associated with the parameter estimation.

2. **No analogue problem**

The main reason why CR09 cannot be used for the design of the models suitable for "climate predictions" is the "no-analogue problem" or, in other words, the past is not the future. (See also discussion in the reply to Reviewer 1). The fact is that during the last 800 kyr for which reliable reconstructions of CO_2 concentration exist, CO_2 concentration was below 300 ppm, and most of time it was even below 250 ppm. At the same time, at present CO_2 concentration is already above 420 ppm and it is expected that at the end of the century it will be somewhere inbetween 500 and 1000 ppm. Assuming no negative net CO_2 emission in the future, CO_2 will stay for the next 100 kyr above 300 ppm even for optimistic 1000 PgC cumulative emission, which is higher than over the past 800 kyr. In the case of 5000 PgC emission, CO_2 will stay above 300 ppm for nearly 1 million years! Thus during the period of time in the future considered in our study, CO_2 will stay above the range its natural variability observed during the past 800 kyr. This is why it is not surprising that paleoclimate data are unable to constrain the most critical for future prediction parameter K (slope of critical CO_2-insolation relationship). After all, statistics is a not magic - it cannot extract from the data information which the data do not contain. Of course, we fully agree with the first reviewer that accurate paleoclimate reconstructions from a warmer climate state, for example late Pliocene and earlier Pleistocene would be very useful. Unfortunately, all CO_2 reconstructions prior to 800 kyr are very uncertainty and cannot be used to constrain model parameters. (To get an idea what "uncertain" means, one can make a look on Fig. 5 in Berends et al., 2021).

This is why we do not see any alternative to our approach, which, of course, is fundamentally different from CR09. The essential elements of our approach are:

1) we constructed a set of model equations based on general understanding of climate dynamics and the results of simulations with CLIMBER-2. This ensures that our simple model has stability properties and dynamical behaviour similar to CLIMBER-2. In particular, similar to CLIMBER-2, the simple model has two stable equilibrium states (glacial and interglacial), under orbital forcing simulates strongly asymmetric glacial cycles which are phase-locked to eccentricity and depend only weakly on the initial conditions, etc.
2) the antrhopogenic CO_2 perturbation has been calculated using results of another EMIC (cGENIE)
3) we calibrated the model against paleoclimate data for the last 800 kyr and generated a large set of model realizations which simulate past glacial climates with the required accuracy
4) we rejected all model realisations which simulate glacial state at present
5) we applied a strict constraint on critically important parameters (slope of critical CO_2-insolation relationship) derived from CLIMBER-2 and thus arrived to a much narrower ensemble suitable both for past and future simulations.

Needles to say is that such approach represents a significant step forward compared to CR09 because CR09 described methodology for the calibration of the model suitable only for modelling of the past while we developed a model suitable for modelling past and future.

3. Response to reviewer's comments

Below is our response to the reviewer's "*Critical comments and questions to be absolutely addressed*" and some more specific comments.

Comment #1. "*Neither the 100ky duration of ice ages, nor their saw-tooth shape were predicted by Milankovitch**. Please check literature and update the knowledge***"

Why the reviewer decided that we are not aware about these limitations of the classical Milankovith theory – we cannot even guess. Obviously our paper is not a review of the astronomical theory of glacial cycles. There are numerous publications, including those were AG was co-author, which present useful reviews of the current status of the understanding of glacial cycles such as Berger (2012), Past Interglacials Working Group of PAGES (2016), Berends et al. (2021). In our manuscript we devoted only one sentence to the Milankovich theory: *"The astronomical theory of glacial cycles (Milankovitch, 1941) postulates that growth and shrinkage of ice sheets is primarily controlled by changes in Earth's orbital parameters (eccentricity, obliquity and precession)..."*. This stement is obviously correct. However, in others of our publications, we not only discussed these facts: "*One of the major challenges to the classical Milankovitch theory is the presence of 100 kyr cycles that dominate global ice volume and climate variability over the past million years* " (Ganopolski and Calov, 2011); "*Of particular interest is the transition between 1.25 and ~0.7 Ma ago, ..., from mostly symmetric cycles with a period of about 41 thousand years (ka) to strongly asymmetric 100-ka cycles"* (Willeit, et al. 2019), but also provided possible explanations for these facts.

As far as the "recommendation" to "*check literature and update the knowledge*" given by the reviewer#3 to the scientist (AG) who in 2011 received the EGU Milankovitch medal for "for his pioneering contributions ... to the understanding of the role of climate system feedbacks and the link between Milankovich forcing and global glaciation", published more than 30 papers directly related to mechanisms of glacial cycles and Milankovitch theory – such "recommendation" cannot be considered anything by rudeness. While the authors do not know who reviewer is, he/she knows the names of the authors even though is unable to spell them properly (the second author never published papers under the name "*Ganopolsky*").

In the "Crucial comment #7" the reviewer demonstrates the "knowledge" of the theory of glacial cycles by telling us to use the term "*pacemaker*" instead of "*control*" or "*driver*" " when referring to the astronomical forcing and went further explaining that "*The theory of ice ages has already evolved*".

While we fully agree that the theory did evolve, it evolved in the opposite direction to what the reviewer thinks. The term "pacemaker" in application to glacial cycles first appeared already in a paper by Hays et al. (1976) entitled "Variations in the Earth's Orbit: Pacemaker of the Ice Ages". Since then results of numerous simulations with physically-based models clearly demonstrated that orbital forcing is not just a pacemaker (this can mean essentially everything) but the real driver of glacial cycles. This was formulated in one of our paper as: "*Here... we demonstrate that both strong 100 kyr periodicity in the ice volume variations and the timing of glacial terminations during past 800 kyr can be successfully simulated as direct, strongly nonlinear responses of the climate-cryosphere system to orbital forcing alone...*" (Ganopolski and Calov, 2011). This result has been confirmed by numerous works done by Andre Berger, Ayako Abe-Ochi, Axel Timmermann and others.

This follows (comment #9) the amazing recommendation "*Please, refer to the most up to date theory of the astronomical forcing instead of Milankovich. check the paper CR09 for a detailed explanation and the theory and its history*".

First, we do not understand what the reviewer has against citing Milutin Milankovitch. We have a great respect to Milankovitch for his extraordinary achievements. Second, as far as the "detailed explanation and the theory and its history" of "astronomical forcing", does the reviewer really believe that this 12 - years old paper contains up-to-date review of the theory of glacial cycles? Does the reviewer believe that nothing substantial has been achieved during the past decade? What about works by Berger and Loutre (2010), Ganopolski and Calov (2011), Abe-Outchi et al. (2013), Ganopolski and Brovkin (2017), Willeit et al. (2019)? We strongly suspect that it is the reviewer who should "*update the knowledge*".

Comment #3 contains numerous repetitions of the reviewer's believe that CR09 approach is the right one and ours is not. We believe, we presented already strong evidence that the reviewer is fundamentally wrong.

Furthermore, the reviewer heavily criticises the use of correlation in sentences like: "correlation is not an adequate criteria to assess goodness of fit in time series the way you did it" "calibration using maximization of the correlation coefficient?! this really need to be explained and justified and proved working." "Correlation should not be used as a validation criteria!" "[correlation] it is not a good way to assess relation or association between time series" "correlation is insufficient by itself, and it assumes linear relations only" "Why did not you considered any Least-Squares (Model Fitting) Algorithms?"

The reviewer states that maximising correlation is not a proper fitting technique. This statement is incorrect. Please see Livadiotis and McComas (2013) who present the maximization of the correlation fitting method. Those authors show that the method is mathematically well defined under certain conditions and that it should be preferred over the classical least squares fitting in situations in which the data sets exhibit variations that need to be described, such as the variations that concern us here: glacial cycles.

Comment #4. "*Carbon dioxide curves: your choice of the evolution need to be justified. why should it be decreasing exponentially*?"

Why should we justify the use of the results of the well-established Earth system model published in a respected scientific journal? Moreover, these results are consistent with the previous findings. The reason for exponential decay of anthropogenic perturbation on very long-time scale is the removal of atmospheric CO_2 by weathering processes.

Comment #5. "*The relationship between critical insolation threshold for glacial inception and CO2 levels is known and must be analyzed using an appropriate sensitivity analysis*".

This relationship is ONLY known from OUR paper (Ganopolski et al., 2016). This paper presents a single equation with a single set of numerical parameters. We do not understand how our formula can be "*analysed using an appropriate sensitivity analysis*".

Comment #6 „Calibration/ Validation need to be done correctly"
Regarding model validation, when producing a model with a predictive aim the gold standard is to evaluate the model predictive ability using previously unseen data (Stone, 1974). As we are dealing with time-series, the common procedure of randomly splitting the data into training and validation sets is inappropriate as it disrupts its temporal structure. We opt then to use an out-of-sample method, essentially holding out the last/earliest part of the time-series for testing (see for example Cerqueria et al., 2020). Given the cyclic structure of the paleo climatic data we are using, the only sensible option is to divide the information into 8 cycles (corresponding roughly to the 8 glacial cycles in the last 800 kyr). In the manuscript we reported the results for the model predictive ability when holding out 50% of the data for validation (i.e. holding out either the last 4 or the earliest 4 cycles). This is a quite stringent test for the model and a more adverse situation than the alternatives of holding out just 3, 2 or 1 of the cycles. We think the model validation methodology we employed is, therefore, adequate and not "weak" as the reviewer claims.

Comment #8. *The glacial inception problem.*
(a) "Validating a work using results from another simulation [Ganopolski et al., 2016] does not seem accurate to me".

As we explained already, we do not "validate" one model by another – we used the results of a complex model to construct and constrain a much simpler one. When measuring complexity in the length of program codes, CLIMBER -2 is thousand times more complex than our semi-empirical simple model.

(b) The glacial inception problem has been treated probabilistically and by using conceptual models. This study must be taken into account: refer to the work by CR09 on On the use of simple dynamical systems for climate predictions: i. How do you position yourself comparing to the work by Crucifix and Rougier 2009? ii. Why not to use the same idea for the modeling part?

CR09 presents a modelling approach which is not suitable for future climate prediction. We presented a model which is suitable for future projections. CR09 is based on fundamentally wrong model, in which glacial cycles represent self-sustained oscillations. Our model simulates glacial cycles the same way as complex models do. How we are supposed to use "*the same ideas*"?

Comment #9. *« This approach, obviously, is not applicable for a possible future Antarctic and Greenland melting under high CO2 concentrations." Why? How do you justify that?*

Not clear what is necessary to justify? Why our model is not applicable to Antarctica? Because the Antarctic ice sheet cannot be described by the same model as the Northern Hemisphere ice sheets, which is obvious. Or does the reviewer ask about potential problems related to neglecting future sea level rise? As we discussed in the response to the first reviewer, for the considered set of emission scenarios, this is not a serious problem.

Juts one more "interesting" statement by the reviewer*: "As stated, this can be useful under the present challenges of climate change requesting **carbon dioxide storage**".*

Our model is designed for supporting development of the "nuclear waste storage" not for "*carbon storage*". This model in principle cannot be used for carbon storage because does not allow negative CO_2 emission.

4. Overall impression about this review

We have a great respect to the work of reviewers. During his 40-years long scientific career AG prepared hundreds of reviews for dozens of relevant journals and received the American Meteorological Society award for his reviewer activity. We always strive to respond to reviewers comments and suggestions in the most constructive manner. But is this a scientific review?

"*This paper really* ***made me sad****"*

*"****Concepts are being mixed*** *and the* ***goal itself is unclear to the authors****"*

"...have been ***inadequately*** *followed and their related approaches* ***incorrectly*** *applied by the authors"*

"work by Talento and Ganopolsky ***does not reflect any aspect of the correct modeling*** *approach towards a probabilistic forecast of climate"*

"This is what the authors tried (wanted?) to accomplish, but ***failed unfortunately****"*

"This work ***has no provided any forecast neither probabilistic forecast*** *of the climate"*

"In addition, the statistical modeling part, ***is applied incorrectly*** *and* ***many chosen assumption are unjustified****"*

"The model selection procedure ***has not been carried correctly*** *either"*

"No future forecast, or prediction (even under scenarios), especially when using observations, ***can be carried out non probabilistically****"*

"This work is ***absolutely not a forecast work*** *and nor a probabilistic forecast"*

"This work embraced a method based on many ***unjustified simplifications*** *and approaches"*

"Honestly, either I did not understand at all what you did, or it is looking more like a ***patchwork using inadequate pieces****!"*

"This is ***inadequate and inaccurate****"*

„… despite the ***low level of the manuscript****…"*

For sure, we are not going to apologize to the reviewer for "making him/her said". In the view that reviewer#3 did not even try to understand the main concept of our work and has a rather limited knowledge about the subject, such rude review is absolutely unacceptable.

References

Abe-Ouchi, A., Saito, F., Kawamura, K., Raymo, M. E., Okuno, J., Takahashi, K., and Blatter, H.: Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume, Nature, 500, 190–194, doi.org/10.1038/nature12374, 2013.

Archer, D., and Ganopolski, A.: A movable trigger: Fossil fuel CO2 and the onset of the next glaciation, Geochemistry Geophysics Geosystems, 6, 7, 10.1029/2004gc000891, 2005

Berends, C. J., Köhler, P., Lourens, L. J., & van de Wal, R. S. W. (2021). On the cause of the mid-Pleistocene transition. Reviews of Geophysics, 59, e2020RG000727.

Berger, A., and Loutre, M. F.: Modeling the 100-kyr glacial-interglacial cycles, Global and Planetary Change, 72, 275-281, 10.1016/j.gloplacha.2010.01.003, 2010.

Berger, A.: A Brief History of the Astronomical Theories of Paleoclimates, 107-129, in Climate Change, A. Berger et al. (eds.), Springer-Verlag, Wien, 2012.

Cerqueira, V., Torgo, L., & Mozetič, I.: Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, *109*(11), 1997-2028, https://doi.org/10.1007/s10994-020-05910-7, 2020.

Ganopolski, A. and Calov, R.: The role of orbital forcing, carbon dioxide and regolith in 100 kyr glacial cycles, Clim. Past., 7, 1415–1425, https://doi.org/10.5194/cp-7-1415-2011, 2011.

Ganopolski, A., Calov, R., and Claussen, M.: Simulation of the last glacial cycle with a coupled climate ice-sheet model of intermediate complexity, Clim. Past, 6, 229–244, https://doi.org/10.5194/cp-6-229-2010, 2010.

Ganopolski, A., Winkelmann, R., and Schellnhuber, H. J.: Critical insolation-CO2 relation for diagnosing past and future glacial inception, Nature, 529, 200–204, https://doi.org/10.1038/nature16494, 2016.

Ganopolski, A., and Brovkin, V.: Simulation of climate, ice sheets and CO2 evolution during the last four glacial cycles with an Earth system model of intermediate complexity, Clim. Past., 13, 1695-1716, 10.5194/cp-13-1695-2017, 2017.

Gregory, J. M., Browne, O. J. H., Payne, A. J., Ridley, J. K., and Rutt, I. C.: Modelling large-scale ice-sheet-climate interactions following glacial inception, Clim. Past., 8, 1565-1580, 10.5194/cp-8-1565-2012, 2012.

Hays, J. D., Imbrie, J., & Shackleton, N. J.: Variations in the Earth's orbit: pacemaker of the ice ages. Washington, DC: American Association for the Advancement of Science, 1976.

Livadiotis, G., & McComas, D. J.: Fitting method based on correlation maximization: Applications in space physics. *Journal of Geophysical Research: Space Physics*, *118*(6), 2863-2875, https://doi.org/10.1002/jgra.50304, 2013.

Paillard, D.: The timing of Pleistocene glaciations from a simple multiple-state climate model, Nature, 391, 378–381, 1998.

Past Interglacials Working Group of PAGES (2016), Interglacials of the last 800,000 years, Rev. Geophys., 54, 162–219, doi:10.1002/2015RG000482.

Stone, M.: Cross-Validatory Choice and Assessment of Statistical Predictions, Journal of the Royal Statistical Society: Series B (Methodological), 36(2), 111–133, https://doi.org/10.1111/j.2517-6161.1974.tb00994.x, 1974.

Willeit, M., Ganopolski, A., Calov, R., and Brovkin, V.: Mid-Pleistocene transition in glacial cycles explained by declining CO2 and regolith removal, Sci. Adv., 5, 8, 10.1126/sciadv.aav7337, 2019.

Wolf, E., et al.: Interglacials of the last 800,000 years, Reviewes of Geophysics. 54, 162–219, 2016