

Author Response

Note: Reviewer comments in italics:

This manuscript introduces the Earth Science Box Modeling Toolkit (ES-BMTK), a Python library designed to streamline the build and operation of box models. The toolkit incorporates many commonly used processes such as air-sea exchange, marine carbonate chemistry, and isotope calculations. Its feasibility and robustness are validated through replicating the work by Boudreau et al., 2010.

Given the significant utility of box models in the Earth Science, ESBMTK has considerable potential for applications in both teaching and research settings. Below are several suggestions that, hopefully, could help further enhance the quality and impact of this project:

We would like to thank Shihan Li for their time and thoughtful comments. Below is our response to the specific points raised by this review.

- 1. The toolkit's ability to separate model geometry from its underlying numerical implementation is highly convenient, particularly for users without extensive programming experience. However, this abstraction also hides the core functions, making debugging more challenging. One potential enhancement could be enabling the model to output the codes for governing equation it generates during runtime. Such a feature would not only improve the model transparency but also serve as a valuable resource for educational purposes, allowing users to better understand the mechanics of the model.*
 - The ability to save the equations for debugging purposes already exists in the model, it is however not well documented. We are thankful for the suggestion and added a section to the manual explaining how to do this.
- 2. Several processes and their parametrizations are described in Section 2. However, I found it somewhat challenging to clearly link these processes to the toolkit framework as depicted in Figure 1. It would be helpful if the authors could integrate these processes explicitly within the framework illustration or provide some example codes.*
 - This is a good point. We revised Fig. 1, and now provide a code example in the MS as well.

3. *I believe perturbation experiments are crucial for exploring system behavior under various conditions. I was happy to learn from the documentation that the model accepts external forcings. I suggest highlighting this feature more prominently and giving brief introduction on its implementation in the main text*

- We are thankful for the suggestion and added an additional paragraph to the MS.

4. *Several things might need further clarifications.*

a & b) *Line 99, the authors state that “the above parameters at the beginning of each run and assumes that they are constant over the integration interval.” However, it is unclear what “the above parameters” specifically refer to. My interpretation is that this might relate to thermodynamic and kinetic constants. However, the carbonate system is typically time-dependent and calculated in each integration. ... Again, from the same paragraph, the authors state that modeling temperatures remain constant throughout the runtime. However, this seems inconsistent with examples where transient and long-term carbon cycles are simulated. Since temperature can affect thermodynamic constants in the carbon system, this assumption may not be realistic. If my understanding is incorrect, additional explanation would be appreciated to resolve this confusion.*

- Yes, currently the thermodynamic and kinetic constants are calculated only once at the beginning of the run, and then held constant. We have done this since in many cases (e.g., the glacial-interglacial ocean) the changes to the thermodynamic constants of the carbonate system are for all practical purposes fully compensated by the alkalinity changes that result from ocean volume changes (Zeebe and Wolf-Gladrow, 2001). It is thus a common practice for many of the published box-models, since the calculation of the thermodynamic constants is computationally expensive. We are aware however that in models where ocean temperature changes faster than ocean volume (e.g., as a result of anthropogenic carbon release) the above can be a limitation. We therefore are actively working to implement a feature that will recalculate the thermodynamic constants if pressure/temperature change during the model run. However, this will have to wait for a future version of the library. To address the current concern,

we have revised the working in the MS and now clearly state the current approach, its rationale and its limitations.

5. *Line 105, It is noted that $[H^+]$ is initially calculated using the `pyco2sys` library, and subsequently, the iterative approach of Follows et al. (2006) is applied. I'm confused by the rationale behind using two different methods for this calculation.*
 - The carbonate system is fully defined if we know at least two of its components. If we, e.g., know the concentration of total alkalinity (TA) and dissolved organic carbon (DIC), we can calculate all other carbon species. This is however computationally expensive. Follows et al. 2006 showed that if we have a reasonable guess for the initial $[H^+]$ at time $i - 1$, one can directly compute the concentration of all carbon species with sufficient precision based on the concentrations of TA and DIC at time i . This is computationally highly efficient, but relies on a suitable initial estimate for $[H^+]$ which we obtain from `pyco2sys`.
6. *Carbon isotopes are crucial tracers in carbon cycle modeling. In Example 4 (https://github.com/uliu/ESBMTK-Examples/blob/main/Examples_from_the_manual/po4_4.png), the steady-state surface box ^{13}C reaches approximately 8‰, which is not realistic. Since the Boudreau et al. (2010) model does not incorporate isotope modeling, I recommend that the authors conduct additional validation work for this component.*
 - The carbon-isotope example in the manual was never meant to be realistic (for that matter neither is the P-cycle model), it was merely a demonstration on how to do it. Achieving realistic DIC values would require that we also add carbonate precipitation and dissolution which would result in a rather complex model. To address this concern, we revised the manual based on the suggestions of reviewer 1 and now show simple examples where the outcome of the isotope fractionation process is predictable (i.e., mixing of two reservoirs, equilibrium fractionation etc.) We would like to note that the isotope calculations are covered by unit-tests and that internal testing also involves an unreleased ESBMTK implementation of the LOSCAR isotope model (Zeebe, 2012) which allows us to compare the to the carbon isotope calculations against published results.

References

- Zeebe R.E. (2012) LOSCAR: Long-term Ocean-atmosphere-Sediment Carbon cycle Reservoir Model v2.0.4. *Geoscientific Model Development* **5(1)**, 149–166. doi:10.5194/gmd-5-149-2012.
- Zeebe R.E. and Wolf-Gladrow D.A. (2001) *CO₂ in Seawater: Equilibrium, Kinetics, Isotopes*. Number 65 in Oceanography Book Series, Elsevier.