Response on "Bedymo: a combined quasi-geostrophic and primitive equation" by Clemens Spensberger, Trond Thorsteinsson, and Thomas Spengler

The study introduces a new modeling framework, aimed at solving the QG and PE equations as closely as possible. The QG model does not capture the dynamics outside the midlatitudes accurately and the primitive equation solvers are mostly formulated in the context of simulating the global circulation and are thus constructed differently making a direct comparison with a QG model elusive. The framework, designed in cartesian coordinates, is tested using 4 tests including both the atmosphere and the coupled atmosphere-ocean. The tests are carefully designed and succeed in demonstrating the strong similarities and differences between QG and PE by analyzing responses in baroclinic growth, coupling and stationary multi-wave forcings. I detail some fairly minor and some major comments below that can help improve the manuscript. I suggest accepting the manuscript, pending these revisions.

1 Minor Comments

1. Using sigma-coordinates in the vertical is common practice among models. Please either mention why is using sigma-cordinate in the vertical unique? or remove this sentence altogether.
2. L15 : "1950s"
3. L42 : as well $\rightarrow$ separately
4. L244 : should be "in line with"
5. In Figures 1,3,6,7 and 8, the x and y labels are interchanged. The x axis is the zonal distance and the y axis is the meridional distance. This should be corrected.
6. L246-7 : ... "results of Simmons and Hosking (1978), for instance, which" ...
7. Section 3.1.2 : maybe I am missing something but how many levels are you using for the storm track test case? Were the original 26 vertical levels used for the PE case?
8. : I am confused by the writing here. There has to be zonal symmetries because you are computing the eddy covariances next. Please reconcile this. The stork tracks are not symmetric - so please revise the sentence.
9. L276 : and the ones actually observed.
10. L296 : , however, do not

2 Major Comments

1. I suggest a re-writing of the abstract because it is misleading with regards to the contents of the manuscript. There is no mention of the tracers in the manuscript, only the abstract. Thus, the sentence should be removed. There is no discussion about the graphical interface in the manuscript either, just a brief mention in the introduction section.
2. The Introduction is inadequately written and I suggest revising it. No historical or existing literature has been discussed. It's severely lacking any conviction on the importance of the study and how it fits within recent modeling initiatives, apart from the mention that the models are being solved as closely as possible.

(a) Please discusses key differences between QG and PE and the traditional methods used to numerically solve them.

(b) Since this is the key foundation the manuscript stands on, please also elaborate the key differences between a QG model and the PE model, when used for atmospheric analysis (apart from the well-known fact that the QG framework works well only in the midlatitudes).

(c) Horizontal and vertical coupling in models is challenging. Moreover, diffusion, which affects momentum balance strongly on long timescales, can also affect the model performance and is strongly dependent on the employed numerical schemes. In a QG model using sigma coordinates, it can introduce strong ageostrophic fluxes. Acknowledging these issues is important in the introduction and so is connecting to the other well established studies aimed at testing dynamical cores, to put the study into a proper perspective. I recommend adding some discussion connecting this study to the more recent studies on evaluation of dynamical cores in much more detail :

Lin, S.-J., Harris, L., Chen, X., Yao, W. & Chai, J. (2017). Colliding modons: A nonlinear test for the evaluation of global dynamical cores. Journal of Advances in Modeling Earth Systems, 9, 2483 2492. https://doi.org/10.1002/2017MS000965

Gupta, A, Gerber, EP, Lauritzen, PH. Numerical impacts on tracer transport: A proposed intercomparison test of Atmospheric General Circulation Models. Q J R Meteorol Soc. 2020; 146: 3937 3964. https://doi.org/10.1002/qj.3881

Ma, J., Xu, S., & Wang, B. (2020). Reducing numerical diffusion in dynamical coupling between atmosphere and ocean in Community Earth System Model version 1.2.1. Journal of Advances in Modeling Earth Systems, 12, e2020MS002052. https://doi.org/10.1029/2020MS002052

Held, I. M., & Suarez, M. J. (1994). A Proposal for the Intercomparison of the Dynamical Cores of Atmospheric General Circulation Models, Bulletin of the American Meteorological Society, 75(10), 1825-1830.

3. Since the model is developed ground up and prepared for educational and research purposes, it is important to provide a performance analysis with a traditional dynamical core. Thus, I suggest the authors to add another a discussion/figure and/or qualitatively compare the runtime of Bedymo with a Held-Suarez like model, for instance, with similar numerical schemes.

4. The study would make for a much more valuable contribution to the existing literature if there was a functionality to use the PE framework in spherical coordinates as well, considering the overall curvature of the planet. Is that a possibility with the framework so far? If not, then the use of the model might be quite limited.

5. I strongly encourage the authors to elaborate more on the future possibilities using the model in the conclusion/discussion section. Is there a plan to introduce more features in the model (this might be a good point to talk briefly about the tracer module)? How does having a "live" graphical interface in the model give it an edge over the traditional frameworks (even modern dynamical cores from prominent modeling centered allow visualisation using certain software packages)?