Review of the revised GMDD manuscript 'The Matsuno baroclinic wave test case'
Authors: Ofer Shamir, Itamar Yavoby, and Nathan Paldor

General comments:

The manuscript introduces a test case for shallow water models that is built upon analytic solutions of the linearized shallow water equations on the equatorial beta-plane, originally published by Matsuno (1966). The research is very interesting and will add a very valuable contribution to the literature.

The authors thoroughly revised the manuscript based on the first round of reviews. The revised version has addressed many of my questions and concerns, but a few inconsistencies, a potential initial data code bug, and omissions still remain in the manuscript. It needs to be revised further. For example, a reader of this manuscript is still not able to initialize a shallow water model without additional information from the initialization codes and there is a mismatch between the initial data codes and Eq. (6a) in the manuscript. These issues are all listed in the specific comments 1-7 below. Even after further corrections of the manuscript, it will still be paramount to also publish the codes. I strongly argue against dropping the codes as the authors suggested.

Unfortunately, the manuscript creates a big new question mark. I had requested that the authors also demonstrate that the test case can be used in global shallow water models on the sphere, not just in tropical channel models (which are very rare). My expectation was that the results are more or less identical to the tropical channel model, and wanted to see the confirmation. I am now stunned to see completely different results for the global shallow water model, and the discussion of the new global results is extremely sparse and insufficient. Why are the results so different? The authors do not even comment on the fact that e.g. all colors are flipped when comparing the Rossby wave in the tropical channel model (Fig. 1) to the Rossby wave of the global shallow water model (Fig. 4). I do not see a reason why even the 'initial states' (comparing hour 0 in Fig. 1 to hour 4 in Fig. 4) show opposite colors. The 4-hour difference in the time snapshots, which I find unnecessary and very irritating, cannot be the reason due to the long wave period of 18.5 days. This looks like a sign error in the initialization to me. I therefore suspect that one of the model setups (either the channel or the global) were initialized incorrectly and need to be reassessed. More details are provided in the specific comments 12-14 below.

Specific comments:

1) Page 3, Eq. (1): Add the definition of the imaginary unit *i*. In addition, the manuscript (Eq. (1)) fails to state that only the real parts of Eq. (1) and also Eq. (3) are used for the initialization of u, v, Φ, $\omega_{n,k,j}$ as coded in the initialization codes.

2) Page 3, line 19: The manuscript omits the definitions of the physical constants Ω, *a* and *g*. Please add these to promote completeness and reproducibility, since there are many possibilities to set these constants. I assume that it is paramount that the tested shallow water model also needs to use the same constants. Make the user aware of this.

3) Page 4, Eq. (6) and line 20: The expressions for the symbols H_n and $\hat{H}_n$ are missing, and the connection between H_n (in Eq. (6)) and the normalized $\hat{H}_n$ (line 20) is unclear. Do you imply H_n = $\hat{H}_n$ as suggested by the code? In case they are the same why are two different symbols used (maybe typo)? The manuscript points to the Press et al. (2007) 'Numerical Recipes' book for the

explanation of the Hermite three-term recurrence relation for H_n. I think this is a major barrier for the adoption of this test case by others, especially if the publication of the initialization codes is dropped. Taking a look at the initialization codes, the exact definition and normalizations for H_n used here are short enough so that they should be provided in an Appendix to this manuscript. Please add this information for completeness.

4) Page 4, Eq. (6a): There is a mismatch between the definition of $\hat{v}_n$ (formerly ψ_n (Eq. (4)) in the first version of the manuscript) and the Fortran/Matlab/Python initialization codes. In the original Fortran/Matlab/Python initialization codes (e.g. line 158 in the Fortran program) ψ_n is initialized as

$$(\hat{v}_n =)\ \psi_n = AH_n \exp\left[-\frac{1}{2}\varepsilon^{1/2}\left(\frac{y}{a}\right)^2\right]$$

but the manuscript defines this quantity as (see the definition of $\hat{v}_n$ in Eq. (6a)):

$$\hat{v}_n = \psi_n = AH_n\left[\varepsilon^{1/4}\left(\frac{y}{a}\right)\right]\exp\left[-\frac{1}{2}\varepsilon^{1/2}\left(\frac{y}{a}\right)^2\right]$$

The factor $\varepsilon^{1/4}\left(\frac{y}{a}\right)$ is missing in the codes. Is this a typo in the manuscript or an error in the initialization codes? If it is an error in the initialization codes then all simulations and analyses need to be repeated.

5) Page 6, line 8: There is a wrong definition of the planar wavenumber k (with units 1/m). The parameter k is confused with its dimensionless spherical counterpart k_S. The definition k=5 needs to read k_s =5 and, for completeness, I recommend also providing the definition of k= k_s/a (for the equatorial plane) again.

6) Page 6, line 13: When providing the codes, please make sure that the parameters in the codes match the manuscript, e.g. correct the amplitude to amp=10^{-5} m/s in the shallow water model (line 117) to match the information for the amplitude A on page 6, line 13.

7) Page 8, line 7: The manuscripts fails to describe the boundary conditions (in the y-direction) for u and Φ. Please add this information for the channel model.

8) Page 8, lines 31-33, Fig. 2 (g): What does 'appears to be' mean? You seem to suggest that a tiny phase speed error leads to an exact $\pi/4$ (or any integer multiple of $\pi/4$?) shift of the v latitude-time diagram (panel g) after 99 wave periods. I find this speculation questionable, and even if true it would not be a tiny error. If there were a phase error like this, why would this not affect all other fields as well? The authors lost me here. Could there be a plotting problem with panels 2g and also 2j (e.g. wrong output file)?

9) Page 10: add a reference for the GFDL spectral transform shallow water model

10) Page 11, line 7: remove double 'the'

11) Page 11, line 8: remove 'at'

12) Page 11: Almost all users of the test case will use a global shallow water model, so the addition of the new section 4.2 is very much appreciated. However, the current discussion of the global shallow water results is inadequate and insufficient. In my view, the discussion of the global results is way more important than the discussion of the tropical channel results, but this section seems to be added in a rushed fashion. As mentioned above, I am now stunned to see completely different results for the global shallow water model. Why are the results so different in comparison to the channel model?

The authors do not even comment on the fact that e.g. all colors are flipped when comparing the Rossby wave in the tropical channel model (Fig. 1) to the Rossby wave of the global shallow water model (Fig. 4). The same is true for the EIG wave (Figs. 2 and 5). I suspect this is due to an initialization (sign) error in either the channel or the global model. This needs to be investigated, likely reassessed and explained.

13) Figs. 4 and 5 (in comparison to Figs. 1 and 2): I do not see a reason why even the 'initial states' (e.g. comparing hour 0 in Fig. 1 to hour 4 in Fig. 4) show opposite colors. The 4-hour difference in the time snapshots, which I find unnecessary and very irritating, cannot be the reason due to the long wave period of 18.5 days of the Rossby wave. As mentioned in point 12) this looks like a sign error in the initialization to me. Change Figs. 4 and 5 and show time step 0 instead of the simulation after 4 hours. In addition, replot Fig. 4 and label the 18W and 18E points (as in all other figures) instead of the 15W and 15E points.

14) Figs. 4 and 5: Figs. 1 and 2 are normalized with the initial states at t=0. Did you normalize Figs. 4 and 5 with the state after 4 hours or with the state at t=0? If the state after 4 hours was used, correct the normalization (use t=0) to make the normalization procedures identical.

15) Page 14, line 14: correct 'the fact that all …'

16) Page 15: The new spatial convergence study is helpful, but the way it was conducted is questionable. In order to test spatial convergence, one typically selects the smallest time step (to minimize time step errors) and keeps this time step constant. Instead, the authors used the longest time step, kept it constant for decreasing grid spacings and of course observed numerical instabilities at some point. The time step errors could potentially be very large for the shortest grid spacings, and these errors become part of the spatial convergence assessment. I suggest repeating the convergence study and using a very small time step for all simulations to avoid this problem.