Reply to Reviewer 1

1) on page 3, line 80, the coefficients s are denoted as functions, which is misleading since it is then stated just below that these are constants. The word 'functions' should be replaced with 'coefficients'.

Done

2) Even though the paper's results do not depend on this, for completeness, on page 7, around line 160 an equation of state should be introduced, since no relationship between density and temperature is given anywhere else.

We have elucidated the relationship between density and salinity using the UNESCO equation of state at line 98.

3) on page 7, line 159, a nabla operator is apparently missing in the definition of the b gradient at intermediate levels

Thank you. We have corrected the typo.

4) the statement 'The $z-$layers are a particular case where the interfaces do not depend on time and space' on line 206 of page 9 is unclear. From the previous definitions of the Z terms, these do not seem to be dependent on time, so it is unclear why the z-layer case should be a particular time independent case. This point should be clarified.

We refer to the actual interfaces, not to the reference one. Contrary to the models introduced earlier, the actual interfaces $z_{\alpha\pm1/2}$ do not depend on time. In the text, we have been more precise about this point, without alluding to "particular cases".

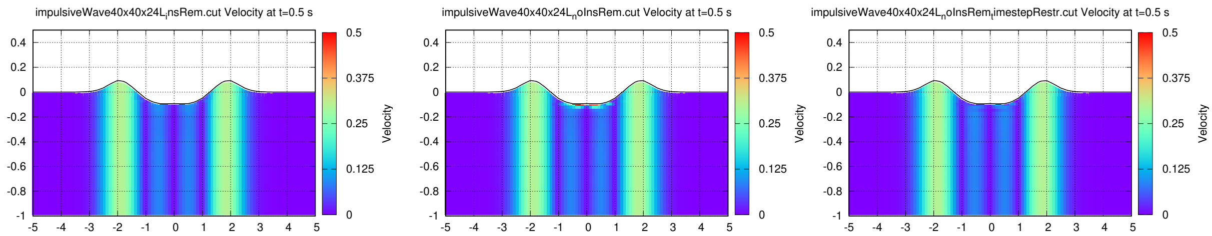
5) Starting on line 268 of page 11, the authors should introduce some changes in the discussion of the semi-implicit time discretisation. Firstly, stating as done by the authors that semi-implicit time discretizations are standard for ocean models is not correct. Several important ocean models use split explicit time discretizations. In my opinion, semi-implicit methods are superior for a number of easily provable mathematical reasons, but still this is not really the standard in this literature. Furthermore, the advantage of using two different implicitness parameters is debatable. The use of two different thetas makes the presentation more complicated, while the parameter values seem to have always been taken equal to each other. I would suggest to use a single theta. In order to help the reader following the derivation, it would also be appropriate to introduce in this context the fully discrete mass equation, which is instead presented only later. Finally, the authors should state that the time discretisation they use is a well known one, referring explicitly to the widely cited paper "Casulli, Vincenzo, and E. Cattani. "Stability, accuracy and efficiency of a semi-implicit method for three-dimensional shallow water flow." Computers & Mathematics with Applications 27.4 (1994): 99-112." in which it has essentially been first introduced in the community of coastal and ocean modelling.

Following your suggestions we have made small changes. We have removed the expressions "as it is standard for ocean models" and "popular choice for many coastal ocean models". Since the introduction, we have added the reference about the semi-implicit method for the shallow water flows together with the reference for the staggered finite element, on which SHYFEM relies. Moreover, the text has been rewritten with one single θ.

6) On line 364 of page 15, the authors state that the problem with thin vertical layers is that the vertical diffusion matrix becomes ill-conditioned. This seems to be the least important issue, compared to the stability restrictions of the explicit advection discretisation. Removing thin layer is important for a number of reasons, but the authors are suggested to provide different motivations for doing so, also because really extreme ratios between the thickest and thinnest layer should be reached for this conditioning issue to be a serious one.

Thank you for the remark and for elucidating this point. We have tried to run some of the tests without *vertical viscosity (in order to exclude the ill-conditioning issue). Without the removal operations, we had to reduce strongly the time step otherwise oscillations in the velocity appear near thin layers. In figure we share the results for the impulsive wave test. We have changed the discussion about this point.*

Figure title: horizontal velocity profiles at $t = 0.5\,s$



with layer removal, $\Delta t = 0.005$ no layer removal, $\Delta t = 0.005$ no layer removal, $\Delta t = 0.0001$

7) On line 522 of page 23 the statement ' Flooding is thus performed with a 1-layer shallow water model with the classical wetting/drying algorithms that may be deployed in dry or nearly dry areas (e.g. positivity limitation, discharge regularization, etc...).' is unclear and should be clarified with respect to the previous and the following statements. Do the authors refer only to the 1d model that is used as a reference? From the way the statement is formulated, it seems that this 1d approach is also used in the z-adaptive method, but it is unclear how, since as discussed before this statement the z-adaptive method inserts more layers as the simulation goes on. Before more layers are introduced (e.g. in the wetting phase) one could imagine that the same 1d approach is used, but it is unclear how this can be done in the drying phase, when multiple layers are present.

As you said the sentence is referred only to the flooding phase. For the drying phase, under a CFL condition based on the vertical velocity, we believe that also drying should occur with one layer only. However, due to presence of finite thresholds for the identification of the dry elements, it may happen that, for very high vertical resolutions, more layers become dry in one time step. We have thus removed the ambiguous sentence. Moreover, the dry region in SHYFEM is treated in a simplified manner: a mass-conserving barotropic momentum is computed from a "flattened" free-surface and it is then distributed across the layers. Within this simplified algorithm, having more layers does not present any specific issue.

8) The results of the realistic simulation of the Po delta are interesting and promising. However, it is a but awkward to refer to a mesh with 24 layers as 'coarse', as opposed to a 'fine' mesh with 27 layers. While it is clear that 27 layers are enough to trigger the adaptation algorithm, it would be appropriate to remove these 'coarse' and 'fine' labels and, possibly, to perform further simulations with an even larger number of layers, thus using a mesh that can honestly be called 'fine' with respect to the 24 layer one.

We have removed the "coarse" and "fine" labels. We were not able to embark on a complicated convergence study.

9) On line 566, a possible larger truncation error of the vertical advection scheme is mentioned. It is unclear to what the 'larger' refers, since all the advection schemes employed are first order upwind.

Yes, the order of accuracy is the same for both the runs. However the truncation error depends also on the constant that multiplies the grid spacing, which in turn depends on the advection velocity. The mass-transfer function may have different values between the two runs: with the $z-$layer it coincides with the vertical velocity and with $z-$star it depends also the derivatives of the grid interfaces. We are speculating that (it's merely a conjecture) for the z-star run, the mass-transfer function is larger then the vertical velocity, triggering a larger error. We have tried to enhance the clarity of this possible explanation.

Suggested minor corrections

We have implemented all these minor corrections. We have verified the English using an online tool. We have corrected another typo in formula (18).

Reply to Reviewer 2

The authors have massively revised the article and I am satisfied with the changes now. Thank you for addressing my concerns. I only have one minor question:

In figure 1 and 2, if alpha is one, then it looks like there are some surfaces with the same name. For instance in Figure 1, the free surface is $s_{1/2}$. If below you let alpha be one, then third surface from the top is also $s_{1/2}$. Can you address this?

Thank you for the comment. We have added vertical ellipsis to stress the fact that α *indicates a generic layer index between* 1 *(first layer) and* N *(last layer). The remaining notation in both figures seems correct to us. If the first layer is* $\alpha = 1$*, then the third surface from the top is* $s_{3-1/2} = s_{5/2}$ *and not* $s_{1/2}$.

Also, there is a typo on line 30, "the" should be "they".

Done