In the following response, the original review is shown in black and our response in orange.

Comment on gmd-2021-250
Victor Stepanenko (Referee)

Referee comment on "Effects of dimensionality on the performance of hydrodynamic models" by Mayra Ishikawa et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-250-RC2, 2021

The paper brings valuable information for the limnological community on the relative performance of models of different spatial dimensionality applied to an artificial reservoir with significant longitudinal variation of water properties. 1D, 2D and 3D models are compared to measurements carried out at Brazilian drinking reservoir, in terms of water level, temperature, velocity, turbulence and passive tracer. Results elucidate capabilities and limitations of models used, which allowed authors to formulate recommendations on their further applications. I suggest to publish the paper after minor revisions.

We thank Victor Stepanenko for his valuable comments to improve the study.

My two major concerns on the paper are:

What was the model calibration procedure? Please clearly describe for each model. Usually, in intercomparison studies some parameters are set the same for all models, the others are allowed to be calibrated individually. Please provide reasoning on the choice of such parameter groups in your case.

A better description of the calibration procedure will be provided as a subsection of section 4. We indeed predefined some parameters to be equal in all models, but later we noticed that the model performance can be improved by adjusting their values as part of a model calibration. By now we can say that the main parameters were the bulk exchange coefficients for heat and momentum (wind drag coefficient), which are the typical calibration parameters of models.

1D models by construction simulate the horizontally averaged fields, e.g. temperature; this means it is strictly speaking incorrect to state that 1D model simulates the temperature of point observations better or worse, than 2D/3D model, because there is a possibility in a latter case to take temperature from a cell nearest to observation location; this is what should be mentioned in discussion on results of the model intercomparison for intake region; my suggestion is also to add 1D model

comparison to horizontally averaged data from reservoir-wide surveys which you use in Section 6.2.2

We agree with the comment and will follow your suggestion. We will add a comparison of horizontally averaged temperature from the 2D and 3D models, and estimate the error with the measured temperature profile.

Specific comments:

The title: hydrodynamic models of what? (reservoir?)
We agree that the type of system needs be better specified, as the scope of the journal is very broad. We will change it to: Effects of dimensionality on the performance of hydrodynamic models for stratified lakes and reservoirs.

Line 29: please change “identical”
It will be changed to: “While the mechanistic description of underlying physical processes are similar in all models …”

Line 125: I suggest to change: lateral -> transversal
The wording will be changed accordingly.

Line 130: explain the choice of turbulent scheme in this study
Information will be added. We simply chose the most common model (k-ε), which was also available for the 3D model.

Line 135: shallow water equations are 2D in space; this is not the same as 3D dynamics with hydrostatic approximation
The shallow water equations used in Delf3D are 3D in space. It is assumed that the horizontal length scales are much larger than the vertical ones. But the third dimension (z) is considered and the calculated parameters change over all 3 dimensions (i.e. 3D-application).
The momentum equation in z-direction is simplified by considering negligible vertical acceleration, which leads to the hydrostatic equation for pressure, i.e. hydrostatic pressure assumption. Then the vertical velocity is calculated via the continuity equation for the case of 3D models.
We will rewrite to: “3D Reynolds-Averaged Navier-Stokes equations with the hydrostatic approximation for the vertical direction”.

Line 138: I suggest to change: resolved -> parameterized
The wording will be changed accordingly.

Line 169: As one can judge from this section, downwelling longwave radiation was not measured and used to force models, rather empirical formulae applied; this might be one of error sources, please indicate in discussion

We will follow the suggestion and mention the lack of longwave radiation measurements as one source of uncertainty in the discussion. However, detailed analysis of the different formulations used by the models to estimate the radiation balance is certainly beyond the scope of the present manuscript.

Table 1: I suggest to change: wind coefficient -> drag coefficient

The wording will be changed accordingly.

Table 1: the light extinction coefficient was put different in models; have you had any measured transparency properties like Secchi disk?

This was one of the parameters that initially was going to be the same in all models and later it was decided to be used as a calibration coefficient. Secchi disk depths were measured over the campaigns, and its average along the longitudinal and over time was 2 m, which by coincidence is the default in Delft3D, and leads to a light extinction coefficient of 0.85 m^{-1}. This information will be added to the calibration description.

Table 1: “branches width” or “segment width” in 2-d raw?

A branch is a collection of segments, in our case we had 2 branches. The first is the main axis of the reservoir, and the second is the left side arm. So it is the segment width. An indication of the branches will be added in Figure 3 to make it clear.

Table 1: what is 0.85 m in raw 3? grid spacing?

The thickness of the vertical cell.

We will change the wording in the first column to make it clear.

Table 1: specify compiler in computational time section

We will add the information that all source codes were written in FORTRAN.

Table 1: longwave radiation schemes are different, what are the implications?

See our response to your former comment. We will follow the suggestion and mention the lack of longwave radiation measurements as one source of uncertainty in the discussion. However, detailed analysis of the different formulations used by the models to estimate the radiation balance is certainly beyond the scope of the present manuscript.

Table 1: “Kinematic viscosity of water” -> did you mean molecular viscosity?
Yes. We will make this clearer by rephrasing to “molecular kinematic viscosity”

Table 1: Lines “Vertical eddy viscosity” and “Vertical eddy diffusivity” should contain coefficients not simulated by “Turbulence closure model”, namely, background values; this is not clear by formulation “Computed”, etc.
We will remove the word computed and provide only the background values of viscosity and diffusivity.

Table 1: please replace “Computed …” by concrete information on computation scheme
We believe that a description of the computational scheme (especially for diffusivity) would become too complex for Table 1. We will modify the table to list the background values of eddy viscosity and diffusivity, respectively.

Heat exchange with sediments neglected, what are implications, esp. for shallow zones?
We briefly mention this in the discussion at Line 517, where it says: “Regarding the heat exchange with sediment, Stepanenko et al. (2013) showed that it did not have significant influence on simulations of bottom water temperature of a shallow lake for a comparable temperature range as observed in Passaúna Reservoir.”

Line 223: water level is not a boundary condition (understanding boundary conditions in mathematical sense as additional constrains at boundaries for partial differential equations)
The water level at the spillway was used as a boundary condition as it represents an open boundary in the 3D model. This information will be added to the manuscript.

Line 325: do evaporation differences explain level discrepancy between models?
This can be one of the explanations, but since we do not have good measurements of evaporation it is not possible to affirm which model was the best. This was discussed at Line 500.

Line 364: p-value is given for which hypothesis? Please clearly explain so that the reader understands the hypothesis being tested every time you mention p-value
The information will be added accordingly.

Fig. 6 b,c,d: better to add regression line
Regression lines will be added.

Lines 466-468: I can't follow this sentence

We will rephrase the sentence: "The GLM model was set up with a maximum water depth of 17 m, while at the point of analysis water depth was ~12 m. For this reason, we present model outputs up to a maximum depth of 12.5 m. If the maximum of the tracer concentrations was below this depth, the inflow regime was categorized as underflow. For CE-QUAL-W2 and Delft3D the closest cell to the station was selected, which represents the actual water depth at the monitoring site."

Line 469: what is interflow, underflow, overflow?

A small description of them will be added at the beginning of the section. They are a classification of the flow path of the inflowing waters within the reservoir according to their location over the depth. Overflows have a path along the surface, underflows along the bottom and interflow in intermediate depths.

Lines 483-485: better to put in beginning of section as definitions of terms used

Agreed, as mention in the previous answer.