**Response to RC2 by Hans Cederwall**

Please see our response to each comment below in blue.

This is a good paper trying to merge a benthic model with the older pelagic model by Savchuk and Wulff. The result is interesting, furthering modeling work is encouraged.

Remarks

Line 3, 22 and other places: Bioturbation in strictu means the mixing and turning over of the sediment. In the Baltic Proper that is dominantly carried out by the two amphipod species, who lies borrowed in the sediment during daytime, but swims about searching for food on the sediment surface during nighttime (they normally bioturbate the uppermost cm, down to ca 5 cm). To some part also Saduria takes place in the bioturbation but only on the suface layer. Species like Macoma, Halicryptus and Marenzelleria who are more or less permanently burrowed in the sediment (often deeper than 5 cm). They are more of bioirrigators. It should be mentioned early in the paper that the authhours in their term bioturbation includes bioirrigation.

It is true that the definition of bioturbation often excludes bio-irrigation. We have chosen to use the definition of Kristensen et al. (2012), who proposed that for consistency, bioturbation (by fauna) should be defined as an umbrella term that covers "all transport processes carried out by animals that directly or indirectly affect sediment matrices. These processes include both particle reworking and burrow ventilation." We added this definition to the introduction:

"Here, we define bioturbation as all biological processes that affect the sediment matrix, including burrow ventilation (bio-irrigation) and reworking of particles (Kristensen et al., 2012)."

Line 65: Here is stated that the penetration depth of oxygen in Baltic Sea sediment is usually measured in mm rather than cm. This is only true for deeper bottoms from ca 60m or deeper. It could of course be found at shallower depth i archipelago areas, but those areas has been excluded in this work. I have measured redoxpotential at zoobenthos stations for more than 30 years and only on deep stations found oxygenation only in the top mm. The model here includes a depth interval of 0-120 m, so in the uppermost 60 m oxygenation is definitely better than what is said here.

We agree that the refences given (Almroth-Rosell et al., 2015; Bonaglia et al., 2019; Hermans et al., 2019) are biased towards deeper parts of the Baltic Sea, and may not give a representative picture. We therefore removed the statement.

Line 108: Limecola balthica has changed back to Macoma balthica (Caroline Raymond, pers. comm.)

We have happily changed the species' name to the newly (re)accepted *Macoma balthica* throughout the manuscript, following the update in the World Register of Marine Species: https://www.marinespecies.org/aphia.php?p=taxdetails&id=141579

Line 110: Pontoporeia is misspelled

Spelling corrected.

Line 124: Misspellings: basalmaintenance and biomas s

These errors were due to a problem with the pdf conversion, and will be corrected in the next version of the manuscript.

Line 183: Why hindcast simulations for benthic variables.? There was a lot of data collected in the 1970s. For example in the Joint sampling programme (Elmgren 1978) and by revisiting the Hessle-stations (Cederwall & Elmgren 1980). There should also be a lot of finnish data collected in the 1970s by Lassig and Andersin from the Finnish Institute of Marine Research. As far as I know these data were transferred to SYKE when the FIMR was closed down. Data from the Joint Sampling Programme and the Hessle project is stored in the benthic database BEDA. Contanct Caroline Raymond or Mats Blomqvist.

The model needs 1349 initial conditions for each benthic variable, one for each depth meter in each basin. While it would potentially be possible to use data collected around the 1970s to initialize the model, it would include a lot of extrapolation and guesswork and would probably not cause any significant changes to the main results of the current study focusing on the period 2000-2020, as the variables have time to turn over several times from the start in 1970. However, we consider this a very interesting proposition and will look into the availability of historical data and possibility to compile and use them for future studies.

Line 213: I strongly suggest you leave at least the Arkona Basin out, beause the fauna here contains several other species than the ones you have mentioned in this preprint. Stick to the Gotland Basins, Gulf of Finland and Gulf of Riga who has a similar set of species.

We agree that the model is best applicable from the Bornholm Basin and north, where the composition of fauna is similar. We moved the biomass comparison for the Arkona Basin from the main manuscript (Fig 3) to Appendix C and changed the sentence on line 213 to: "Simulated mean biomasses of the individual functional groups and the groups combined were mostly within one standard deviation of observed means from the Bornholm Basin (basin 8) in the south to the Bothnian Bay (basin 11) in the north, although it should be noted that the spread of observed data is large."

We also changed a sentence in the Appendix to reflect the limited applicability of the model in Arkona: "The model is not applicable to the high-salinity areas at the entrance to the Baltic Sea (Kattegat, Öresund and Arkona Basin), as it does not include the high diversity of functional groups present in these areas (Figs. C1–C2)."

Line 221, and Table 1: The benthic database BEDA contains primary data for 7 stations sampled in 1976 for the Joint Sampling Programme. In the mid 1990s a mapping of the macrofauna of the Gulf was done within the NMR financed Gulf of Riga Project. The results are published (Cederwall et al. 1999). Possibly the primay data are still held by Vadims Jermakovs, Latvian Institute of Aquatic Ecology (An institute where your colleague Bärbel Muller-Karulis earlier worked).

We thank Dr. Cederwall for the suggestions for additional resources on benthic fauna biomasses. We also acknowledge that the SMHI Sharkweb database relies on quality checks by the data deliverers and has a very limited quality check within itself. However, we would still argue that our approach using only published and/or open-access data is a reasonable choice. The main purpose of the biomass validation in this study is to confirm that the validation performed by Ehrnsten et al. (2020) still remains valid in the modified model version. Therefore, we believe it would be beyond the scope of the current study to make a new, comprehensive collection of data from unpublished sources. We have made extensive quality checks of the data from both Finnish and Swedish databases to reduce the error introduced by data quality issues, as explained in Ehrnsten et al. (2020) and in Appendix C.

To improve model validation for the Gulf of Riga, we now extended the validation with data from Gogina et al. (2016). In their appendix, data on wet biomasses for benthic species is provided. The data is based on a comprehensive compilation from several sources up to 2013 and is provided as means for 5 km$^2$ squares. Within the Gulf of Riga excluding the shallow coastal areas, data for 95 squares were used. We classified and analysed this data in the same way as the other validation data and now include the Gulf of Riga in the quantitative validation in Figs. 3, C1 and C2. Even though the data is slightly more aggregated than the previously used data sources, we believe this gives a good picture of the general biomass range in the Gulf together with the comparison in Table 1 (extended with two publications, see response under Table 1 below). Conforming with the original manuscript, the new analysis shows that simulated mean total biomass was higher than observed. With the new data, we can see that this is mostly due to higher simulated biomass of the dominating group, *Macoma balthica*. Possible reasons for overestimation are already mentioned in the original manuscript section 4.1 (lines 299-303). The new cost function values vary between 0.11 and 0.75, i.e. simulated biomasses were well within one standard deviation from data. Simulated biomasses are also within the range of the new data in Table 1.

Line 277-279: Is the big difference in sedimentation between the BSAP scenario and the HIGH load scenario reflecting a difference in phytoplankton species composition? Historically the main input to benthic ecosystem has been the sedimentation of diatoms during the spring bloom, not the sedimentation during blue-green blooms. There has however been a shift in species composition in the spring bloom, where diatoms have decreased and other groups (who have lower sinking rate) have increased (Hjerne et al. 2019).

Indeed, the difference in sedimentation in relation to primary production rates reflects differences in timing and composition of primary production. BALTSEM simulates three types of phytoplankton: diatoms, N-fixers (i.e. cyanobacteria) and 'other species', representing mainly summer-blooming flagellates. Diatoms have the fastest and cyanobacteria have the slowest sinking rate. In addition to sinking rate, the proportion of primary production reaching the seafloor depends on zooplankton grazing rates, bacterial remineralisation rates and physical properties of the water column (e.g. stratification). Thus, sedimentation is determined by a suite of interacting processes, and the relative contribution of each is not straightforward to tease out. In general, though, BALTSEM simulates a shift from spring to summer blooms and increase in cyanobacteria in response to eutrophication and warming over the past decades, as briefly discussed in section 4.3.

We have added graphs on primary production and N-fixation in the different scenarios to Fig. 10. These show that the rate of N-fixation increases strongly with nutrient loading, implying a change in phytoplankton composition. N-fixation is about 5 times higher in the HIGH load scenario than in the BSAP scenario in the Baltic Proper, while N-fixation is completely absent in the BSAP scenario in the Gulf of Riga.

Line 397: cf Mäkelin & Villnäs 2022. Could the seasonal variations in benthic stoichiometry have any influence on your modelling work?

General theory predicts that animals regulate their stoichiometry within a narrow range (Sterner and Elser, 2002), which is the basis for our assumption of a fixed stoichiometry in the benthos. However, as shown for example in this very interesting study, some variation does exist. We are grateful to our colleagues Mäkelin and Villnäs for sharing their preliminary results with us, which were used to parameterise the model. We tested different stoichiometric ratios based on the range found in the study. Indeed, by changing the C:N weight ratio in the benthic fauna from 6 to 7, the growth of fauna shifted from being mostly N-limited to mostly C-limited. However, the quantitative difference in biomass was relatively small. In other words, a variable stoichiometry

could have an impact on growth and excretion estimates, but we would need more information on how, and especially why, the stoichiometry varies over time before implementation in this ecosystem-scale model would be feasible.

Fig 3: You show bars for the depth interval 0-30 m. I strongly suspect you have very few if any stations in the depth interval 0-10 m. This because you have outruled data from archipelagos and open sea research vessels don't like to go into shallow waters. Also these areas are dominated by transport bottoms hard to sample quantitately. Finally in these shallow bottoms you have another set of species than the set you have worked with. You should change to the depthinterval 10-30 m. In the text you mention that you had no shallower (<30 m) data from the Arkona Basin. The benthos database BEDA contains some data from the 1980s (Mats Blomqvist, pers. comm.) There is also a lot of data held by German institutes. But on the other hand I advise you to leave the Arkona Basin out of your paper since the fauna there differs so much from the Baltic Proper.

As suggested, we removed the Arkona Basin from Fig. 3 and modified the text as explained in response to Line 213 above.

While it is true that we have excluded a large part of the shallow areas with a diversity of habitats and communities, we still included data from depth interval 0-10 m, making up ca 11% of the data in the <30 m depth category. We could exclude this data from the comparison, but we believe it is more correct to include it since we cannot exclude this area from the model.

Table 1: Could becompleted with Cederwall et al. 1999.

This excellent publication was surprisingly hard to find, but we eventually got hold of it through the Technical Library of Hamburg. We added the data to Table 1. We also added data from another publication to the table (Gaumiga and Lagzdins, 1995). The data in these publications seem to strengthen the validation of the model: simulated biomasses are well within the range of the reported observations.

Table A4, footnote: Mäkelin & Villnäs is published in Limnol. Oceanogr. 2022. The reference is missing in the reference list.

Reference to the recently published study was added as suggested.

Line 901: You should know that sharkweb/SMHI does not have a benthological quality control of the data delivered to them. They assume that laboratories deliver correct data. This is not always the case (Mats Blomqvist pers. comm.). I suggest you import data from BEDA.

Please see our response to comment to Line 221 above.

Final comment: The macrofauna is not the only part of the benthic ecosystem. The meiofauan for instance can have biomasses of 5-10 g wetweight/m2 (Elmgren 1976). To what extent does the meiofauna influence the sediment chemistry?

It is possible that the meiofauna is important for sediment processes, but since it is an understudied group compared to macrofauna, we do not have the proper means to estimate their effects. In a modeling context, it is also a bit tricky to include the meiofauna as a functional group, as the grouping is based on size rather than function.

**Summing up**: This is basically a good paper well worth publishing after revision.

References in the response

Blackford, J. C.: An analysis of benthic biological dynamics in a North Sea ecosystem model, J. Sea Res., 38(3–4), 213–230, doi:10.1016/S1385-1101(97)00044-0, 1997.

Gaumiga, R. and Lagzdins, G.: Macrozoobenthos, in Ecosystem of the Gulf of Riga between 1920 and 1990, edited by H. Ojaveer, pp. 196–211, Estonian Academy Publishers., 1995.

Gogina, M., Nygård, H., Blomqvist, M., Daunys, D., Josefson, A. B., Kotta, J., Maximov, A., Warzocha, J., Yermakov, V., Gräwe, U. and Zettler, M. L.: The Baltic Sea scale inventory of benthic faunal communities, ICES J. Mar. Sci., 73(4), 1196–1213, doi:10.1093/icesjms/fsv265, 2016.

Kristensen, E., Penha-Lopes, G., Delefosse, M., Valdemarsen, T., Quintana, C. O. and Banta, G. T.: What is bioturbation? the need for a precise definition for fauna in aquatic sciences, Mar. Ecol. Prog. Ser., 446, 285–302, doi:10.3354/meps09506, 2012.