Comment on egusphere-2022-1126
Anonymous Referee #1

The authors investigated the potential role of Data Assimilation in improving the accuracy of barotropic processes induced variant scale/mode sea level anomaly in the Mediterranean Sea. The study is based on the state-of-the-art simulation kernel in SHYFEM. The authors comprehensively investigated the improvement of the astronomical tide, surge and seiches implemented by DA, and promoted the adaptability of SHYFEM with inclusion of EnKF. The manuscript is well written and organized with a sensible logic. However, given I still have these several following major concerns, I cannot recommend an acceptance at its present form.

- Although it is still a nowadays great challenge to DA to treat/improve the hindcast and forecast of sea level anomaly in the region where the SLA oscillation is significant, I'm still wondering why the authors conduct this simulation in a two-dimensional or barotropic configuration? Will the inclusion of, e.g. dynamic height associated with the baroclinic processes be really negligible in the region? If it is not, why the heat fluxes, evaporation and precipitation, as well as riverine discharges are excluded? The larger scale circulation, at least those in the synoptic scale, is another issue related to this concern. Could the authors include some discussion related to the unimportance of these processes? Or, the authors may want to state that they are treating those larger-scaled motions as reference levels already, although I don't think that is a straightforward statement.
- I still have concerns about how did the simulation treat the open boundary condition, although the manuscript did clarify that the authors treated the boundary condition with great effort. If sea level is kind of prescribed at the western boundary, how could the circulation (including their impacts in SLA and currents) be connected with that to the further west of the open boundary, which I think is provided by, for example, the CMEMS reanalyses. I may also suggest the authors include a paragraph to elaborate the way the open boundary condition is implemented or explicitly show the algorithm of the open boundary condition.
- Why the satellite altimetry data is not used as observed data in this research? Are they at least usable for the astronomical tide correction and forecast? If gridded data is problematic, how about the along-track data? There are dataset of harmonic constants extracted from the along-track data by using this operation, and the authors mainly
used much higher resolution records at the surrounding tidal gauge. I mean, there are more observations with much higher spatial coverage may help further improved the DA.

- In the perturbation runs, why the drag coefficient \( C_d \) in the quadratic formulation is not perturbed? Dissipation of energy with the scales smaller than tides through the bottom friction could also be an important process that determines the characteristics of tidal currents, and in this sense, although the authors stated that the current research is focusing on SLA variations, in the current configuration, accuracy in flows will also be an important aspect. Did the authors analyze whether the current design could also improve flows or not?

- In my opinion, it is still important to rely on DA to improve the parameterization in the simulation, since it is not that feasible for operational users to generate a large number of perturbation runs to have that short-term forecast improved.

- It is really hard to intensify the meshes in Figure 1. Could you zoom in to some critically locations to show the spatial variability of resolution?