

General observations:

The manuscript evaluates the impacts of different observation types on the Corsica Channel transport in a simulation of the NW Mediterranean Sea using the ROMS model with 4DVar data assimilation scheme. It follows a well-established methodology that have been applied many times to different regions around the globe. It represents an interesting contribution, that can be helpful in setting up a forecast system and to define future investments in the regional ocean observing system.

I feel the text would benefit from a major revision to strengthen its conclusions and improve the readability and figures. In particular, the discussion is fragmented and the figures are inconsistent in use of axes and colorbar ranges, making them hard to compare.

The conclusions summarize well the main results, but lack a closing statement that clearly defines the contribution of this paper and the expected importance in the larger picture.

Detailed comments:

Introduction:

Highlight the differences of your simulation to the Mediterranean Sea Physics Reanalysis, apart from resolution. What difference does your higher resolution makes? Why is this necessary?

Introduction lacks flow. In special the transition between the examples of 4DVar systems and the oceanography of the north-western Mediterranean Sea at lines 28-29.

Lines 29-35 – Will become more readable if broken up into smaller ones.

Lines 58-60 - This sentence is somewhat misleading, since modelling studies without DA can't by definition assess the impacts of DA. I recommend reviewing and potentially removing the reference to the non-DA simulations here.

Lines 64-65 - I don't understand how the observation impacts can be used to evaluate the dynamical scales. Could you, please, explain.

Lines 66-69 – It gives the impression that there are new observations assimilated that represent the “improvement” in relation to Bendoni et al. (2023). However, the first paragraph of session 2.1 gives a different idea. Therefore, this paragraph could use some rephrasing for more clarity.

Session 2.1:

Line 76 – Please specify the version of ROMS you used.

Line 111 – Can you justify your choice for the number of outer/inner loops and assimilation window length? Where there any sensitivity experiments?

Lines 112-114 – This interchangeable use of the terms “forecast” and “background” is confusing. Background is the accepted term used in data assimilation literature, while “forecast” is reserved to the “free run” initialized from an analysis. Please, review this and correct it through the text.

Moreover, I don’t understand how this FR was run. Did you do a forecast after each analysis? Why not use your previous 4DVar analysis as initial conditions for the next assimilation window? Did you consider overlapping analyses? Please, clarify how this was run and the rationale behind it.

Session 2.2:

You give no explanation for how your observation errors were defined. The values look like what I would expect for instrument error, where representativeness errors would be expected. In addition, there is no spatial structure to the errors. Could you, please, justify your choices?

Lines 127-135 – The 3-hour low pass filter of the HFR velocities would still contain a tidal signal. However, you did not mention if and/or how you included tides in your simulation. Although the tidal signal is small in this area, this inconsistency can impact a “strong-constraint” DA scheme where all main physical processes are in principle included. Could you, please, elaborate how you deal with this?

Lines 145-168 – Please, break into separate paragraphs by observation type. This will help make it clear the different procedures applied to each observation.

Line 146 – Please, add the reference for CORA.

Lines 154-155 – I am curious why you chose to assimilate a SST product with a coarser resolution in relation to our model grid. Could you, please, justify your choice and the impact of errors that are carried from the SST product?

Lines 160-161 – Is the bias correction for the ADT compatible with the Reanalysis used for boundary conditions?

Lines 164-169 – Is the background QC applied to all the observations, or only the ADT? If only the ADT, why?

Section 2.3:

Lines 189-191 – It could be helpful to show the transect position in a map. Figure 1 is a good place to add this. Also, could you provide the rationale for choosing this particular section?

Section 3.1:

Lines 199-204 – Since you use the ratio between the eigenvalues as a guiding metric to identify overfitting, it would be better to show it in Figure 2.

Line 205 – This is the first occurrence of “FR”. Please, add the definition.

Line 217 – These should be positive values (a negative reduction would mean the simulations are getting worst).

Line 225 – Is the difference in the fit of u and v related to the variability of the velocity components? It would be helpful to show the observation standard deviations next to the RMSEs.

Lines 237-241 – Please, present the standard deviations for all observations to support your affirmation. This could be added to Table 1.

Lines 251-256 – All that the time series show is data was in fact assimilated.

Session 3.2:

There is a lot of jumping back and forth with the figures, which harms the readability of the paper. Please, consider rearranging your discussion.

Lines 260-262 – I disagree that that graphs show that there is no BIAS. A better way to show this is by plotting the bias in Figure 3. Moreover, the graph for in situ S does not look “evenly distributed” for innovations.

Lines 273-276 – Please explain how can the impacts on transport shown in Figure 7 be many times higher than the values and variability shown in Figure 8.

Lines 276-281 – So, the observational studies consider the velocity constant with longitude? And how do you explain the difference to the other model studies? It is not good enough to say additional analysis could be done. Please, explain this better.

Lines 282-287 – What explains this variability and the differences in the assimilating model?

Lines 304-308 – Please, explain why the impact per observation is bigger in situ T and S in relation to HFR.

Lines 309-313 – This calls for a look at the vertical structure (transect) of the increments. You could link to the discussion of Figure 9.

Lines 339-347 – You kind of hint at the potential issue here: The placement of your southern boundary. Given the importance, please expand your discussion.

Line 348 – Here is a jump in your discussion. You should add some explanation and refer to session 2.3.

Lines 352-355 – This again shows the weight given to your southern boundary, on top of potential issues on the way the river flux is added to your simulation. How does the sea surface salinity maps look? Are the river plumes well represented?

Lines 380-395 – This is a very interesting analysis! As the heat flux reflects only what is happening at the surface, it would be good to see what is happening bellow. I believe the water column heat content or vertically integrated density would be helpful here.

Session 4:

I feel there is a missed opportunity here to discuss how effective the different observations are in constraining the simulation and the potential implications for forecasting and the “greater” western Mediterranean Sea – the usual goal of a assimilative system. You have lots of information that can be helpful in driving the evolution of the observing system.

A closing paragraph would do the job.

Table 1 – I am confused by your delta signals. While for RMSE a negative value represents improvement, the opposite seems to be truth for the correlation. I would rather see the actual values for the RMSE, correlation, and BIAS (instead of deltas). This would better make the point if your simulation is actually fitting well to the obs. And, please, be consistent with the +- signals.

Table 2 – Please, improve your caption to clearly indicate what you mean by “Global” and “Datum”.

Figures:

Figure 1 - a) The “HFR” area is not clear. This can be resented by a polygon showing the coverage area. Also, what you present seem to be presenting is the mean stream function - this should be clear in your caption.

Here would be a great place to show the section chosen for the transport calculation used in your observation impact calculation.

b) I feel the total observation density to not be very informative, and would be better plotted as a “pcolor” or “contourf”. Instead, the distribution of each observation type can give a better idea of the coverage and expected impacts.

c) Why do you have an empty depth bin close to 2000m?

Figure 2 – Please, show the ratio between eigenvalues in subplot (b).

Figure 3 – Could you, please, use the same axis ranges for the number of observations in all your subplots? This will make it easier to compare.

Figure 4 – As for figure 3, please use the same axis limits for the number of observations. It would also be interesting to have the std of the (u,v) observations to be able to evaluate how much the variation of the fit by depth is relative to the variability of the currents.

Figure 5 – The reference to the subplots in your label is wrong. Always add the unit after the depth.

Figure 6- Please, add the units for depth.

Figure 7 – Please, have use the same axis and colorbar for all your subplots. This figure is underexplored in the text, which makes me question if it is necessary.

Figure 8 – b) This is unnecessary. You don’t seem to explore this figure enough in your text.

Subplots c) and d) arguably contain the most important results in your manuscript, and should be deeper explored in your text. Try to keep in mind what is the main point you want to make with your paper.

Figure 9 – Please, add your section position to the maps.

Figure 10 – Please, use the same limits for your map colorbars. The colormap you used is not helpful either. More contrast between values would make it easier to spot differences.