Report on 'the IAS2024 coastal sea level data set and first evaluations' by Fukai Peng et al.

This study proposes a new reprocessing of altimetry data from Jason-1, 2 and 3 in the world coastal zones over 2002-2022. It provides a new useful data set of altimetry-based sea level time series at a large number of coastal sites, complementing the limited tide gauge network. The coastal data are provided along satellite tracks at 20hz resolution, from the coast up to 20 km offshore. They are based on retracking of the radar waveforms using 3 different retrackers (ALES, WLS3, Adaptative Retracker), in addition to the standard retracker MLE4 used for the open ocean geophysical data records (GDRs). A total of 1548 'virtual' coastal stations are built along the world coastlines. The data set is validated by using available tide gauge data. It is also compared to a previously published, similar coastal product (called here ESA CCI product).

In principle, this proposed new dataset merits to be published. However, important clarification concerning the data processing is needed. The comparison with the ESA CCI product also requires significant clarification. Thus, the manuscript requires major revision before being suitable for publication in ESSD. Overall, the manuscript lacks a lot of crucial information about the data processing. Besides, the recurrent claim, all along the paper, that the IAS2024 is superior than the ESA CCI product (which is possibly true but not fully convincing) suggests competing interests rather than the desire of objectively presenting a new useful product.

Major and minor concerns are listed below (with major concerns in bold).

- Page 6, Calculation of altimeter SSH: To my knowledge the **GPD+ wet tropospheric correction** is not available beyond mid-2021. Which correction is applied beyond that date?
- Page 6, Seamless combination of SSH estimates from multiple retrackers: it is not clear **how the different retrackers are used**. What is the bias (wrt the MLE4-based GDRs) you are talking about? Is it estimated for each 20Hz measurement? What is the Dijkstra algorithm? In fine, **how are the ranges estimated? Are they based on different retrackers depending on the shape of the radar waveforms?** Please clarify. I think it would be useful to provide some discussion on the respective performances of the retrackers. The retracking procedure is crucial for building coastal altimetry data sets.
- Page 7, Assessment of the IAS2024 dataset : After your 3-sigma editing, how are the **temporal distribution of the data per mission**? For example, the Jason-1 mission suffers more for missing data than Jason-2 and Jason-3. **These missing data affect trend estimates**. Please discuss.
- Page 8, Assessment of the IAS2024 dataset: Little is said about the **intermission bias estimation**. Since you are focusing on coastal altimetry data, how do you compute it? Locally? regionally? This point is also crucial because any error in the intermission bias directly affects SLA trend estimates.
- Page 8 and beyond, Assessment of the IAS2024 dataset : **How is the coast defined**? Which data set is used and **how is the distance to the coast computed**?

  Do you apply a smoothing to your SLA time series?

  Could you show the distribution of closest distance to the coast for the 0-10km segments?

- Page 9 and beyond , Application of the IAS2024 data set: You define three sets of along-track segments (0-10km, 5-15 km, 10k-20 km) and assign virtual stations to each segment**. From Table 4 (page 16), I note about the same number of virtual stations are found for each segment, which is surprising.** I would expect more virtual stations for the offshore segment (10-20 km) than for the onshore segment (0-10km) because more noisy data should have been eliminated in the latter case. Please discuss.

- There is no indication on how the 10-day data are projected onto fixed points along a nominal ground track. How do you proceed?
- Page 10, Figure 4: indicate on the panels the name of the satellite. The figure caption should be more detailed.

- **Page 11-14: Comparison with the ESA CCI data set: I have a number of comments on this comparison.**
  1. First of all, it is not clear at all which version of the ESA CCI product has been used here. Page 8, it is written that the ESA CCI v2.3 data set is used (with data provided over January 2002 to June 2021). If this is correct, why only ~800 sites are correlated in Figure 6.a (I exclude the left hand part of the histogram centered at zero which seems to correlate data from IAS2024 at sites where ESA CCI has no data; i.e., in high latitude regions, Red Sea, tropical islands –see my comment below-). The ESA CCI v2.3 version contains 1189 virtual stations. Thus the histogram should show 1189 counts. The authors should indicate the date at which they downloaded the ESA CCI data to be clear about which product they considered.
  2. It is not fully clear how the ESA CCI data are used for the comparison. From what is written page 12, lines 261-264, it looks like that the ESA CCI along-track SLAs are averaged over the same 3 segments as for the IAS2024 data and then compared. But according to Fig.6a 'correlation coefficient plots' this seems not be the case. How many common points are considered in each segment? Does this comparison refer to the common time span of the two data sets?
  3. Fig.6a is totally unclear and likely misleading. To which virtual stations correspond the two histograms? As the processing of the raw coastal altimetry data are rather similar (except that IAS2024 uses several retrackers compared to ESA CCI which only uses ALES, but all other corrections being the same), **I do not understand the large number of cases with zero correlation. This seems totally unlikely**. My understanding is that the left hand side histogram of Fig.6a centered at zero corresponds to the purples dots of Figure 7 (i.e., where the ESA CCI product has no data). If this is the case, then the histogram centered at zero is not only misleading but definitely wrong! Please remove it!
  4. Please show examples to SLA time series at the same location from the two products. If, for example, your SLA time series are smoothed and if you compared to unsmoothed ESA CCI data, the correlation will indeed decrease. Please explain exactly how you have done.
  5. **Fig.7a is misleading**. As the ESA CCI product has no data in high latitude regions, red sea and tropical islands, **the map should show the correlation ONLY at the common sites**. All purple points (where ESA CCI has no data) need to be removed to not give the false impression that there are numerous sites where the correlation is zero.
  6. Fig.7b: **same comment as for Fig.7a. It is misleading to show high rms (red points) where the ESA CCI product has no data!**
  7. Page 13: Comparison with ESA CCI; **trend differences**. As for the SLA time series, **it is totally unclear how the comparison is done**. To trust this comparison exercise, you must be clear on what you do.

     I recommend you **provide a comparison of your trends computed over the 10-20km segments with trends estimated from classical gridded products** (e.g., Copernicus datasets, CMEMS or C3S). This would be another highly useful way of validating your product (in terms of trends). This has been done for the ESA CCI product for the offshore data (>10km from the coast) with trend differences < 1-2 mm/yr in general with the C3S gridded product.

- Page 14-16, Altmeter-based virtual stations: comparison with tide gauges. **How many tide gauge are used**? Justify the **200 km distance threshold** considered for the virtual stations-tide gauge comparison. Do the tide gauge records have continuous data over the overlapping time span or not? If not, how do you proceed? Please explain.
- Table 4 as well as Fig.14d and page 21: Concerning the trend comparison, what means 'trend differences'? Do the authors compare VLMs estimated from the differences altimetry-based SLAs and tide gauges, and GNSS-based VLMS? If this is the case, how many tide gauges have collocated GNSS stations? In many cases, GNSS-based VLM solutions often show significant discrepancies between computing centers. You chose the ULR7a solution here, which is fine. But do you have an idea of the associated VLM uncertainties?
- Pages 20-21, Vertical land motions: This section estimates VLMs from the differences altimetry-based SLAs and tide gauge data. This is ok. But over the rather short time span of analysis (20 years), **many coastal SLAs display strong interannual variability, masking any long term trend**. Only a few coastal regions show dominant linear sea level trends over this time span. Thus what are the errors associated with the trend estimates by this approach?
- The plots in Figures 15d and 16b are unclear. What are supposed to show the boxes?
- The color bar of the maps should be modified in order to see differences of the parameters from one region to another. Most points are green…
- The f**igure captions lack precision**. They do provide enough information and details on what the figure shows.

- **Summary:** in proposing a new coastal sea level product, this study would be of interest for the oceanographic and climate communities. However, in its current version, it lacks important basic information related to the data processing and to the comparisons with the ESA CCI product and the tide gauges. As a consequence, the results are not always convincing. I strongly recommend **major revision** and re-review of this manuscript and ask the authors to explain in details what they exactly did at the different steps of their analysis.