Reviewer #1

answers in ***bold italic***

This is the review of the manuscript “Generation of super-resolution gap-free ocean colour satellite products using DINEOF” by Aida Alvera-Azcárate and co-authors.

The manuscript is written with clear English, addresses a very interesting topic and I surely recommend it to be published in EGUsphere. However, I suggest the authors to go through my comments below which are sincerely meant to improve the quality and readability of the work of which I recognize its scientific value.

At the end of section 3 – after introduction, data and methodology – the reader expects to have a clear idea of what has been done, why and how. Unfortunately this is not the case here and I strongly suggest the authors to reshape the information in a more coherent way to make the reading smoother. For example, lines 198-204 in the results are purely methodological and a reader like myself would likely expect them to appear somewhere in section 3. This surely facilitate going through the paper more easily without the need of reading the entire paper any given time. There are other such examples that prevent the reader to fully get what has been done by the authors: please, carefully address this; the quality of the work will surely improve.

We thank the reviewer for their comments. We have tried to streamline the contents to make the reading easier. However, on the example given on lines 198-204, we think that this information, which pertains to this specific implementation and not the overall implementation of the method, should be included when presenting the results, as these choices influence the outcome.

The second important comment that I would like the authors to take into consideration is given by the fact that results are presented as a series of case studies which on one side do provide the reader with effective and impactful examples but on the other lack of giving a quantitative and statistically robust evaluation of the outcomes of the analyses. I am personally aware that this is not the case but still the authors need to prove it in a more robust and convincing way.

We have followed the suggestion by the two reviewers to include percentage change maps in the examples given. The statistically robust evaluation is performed for the validation with in situ data.

As clearly stated in the acknowledgements, this work was developed in the context of the Copernicus Marine Service MultiRes project, and the way results are presented looks typical of a project final report. Personally, I do not have anything against the large number of figures within a paper; still, twenty figures, for the amount of information they convey, appear definitely too many. Perhaps, authors may consider some of them to be condensed.

We have removed figures 8 and 9 from the main paper and put them as supplementary material. Also figure 12, which contained information already present in figure 11, has been completely removed. Figure 20, which showed a Hövmuller diagram along a transect has been also removed.

Here below the detailed list of comments.

Section 1 – Introduction

The problem is well-posed and the line of reasoning clear. I would have expected the authors to choose the area of study in such a way to promote a comparison with the neural network approaches they mentioned in the introduction or at least to use it in one example of the many they show. This would provide robustness to the approach without the need of implementing the other techniques which may require a considerable effort.

Comparison with other techniques needs to be done with equal datasets, validation settings and validation data, so a direct comparison with any of these approaches would require a new implementation. While not shown on this paper for brevity, we have applied the approach to several regions, obtaining accurate results in all settings.

Line 24 – the sentence should read (?) "Super-resolution approaches aimed at increasing the spatial resolution of geophysical datasets and have been developed …"

We have changed the sentence to:

"Super-resolution approaches THAT ARE aimed at increasing the spatial resolution of geophysical datasets have been developed"

to make it clearer, as it might indeed confuse the reader (the phrasing proposed by the reviewer was not correct).

Section 2.1 – Study area

The area of study is well characterized and provides a useful background for all those unfamiliar with the complex dynamic system of the Belgian waters.

Thank you

Section 2.2 – Satellite data

This section is a synthetic overview of the not at all trivial preprocessing approach, very useful.

One important information that is missing or too scattered is the space-time resolution of the various datasets that are used in the work along with their temporal coverage. The spatial domain is well depicted by Figure 1. For example, it is not clear why the validation of the Rrs product against in situ data is performed from 2019 to 2022 (not including 2023 and 2024 data) and the generation of the super-resolution data only covers the 2020.

The application of the DINEOF approach can be done in any period of time. We wanted to use a 1-year time series, and for 2020 we had the largest unbroken timeseries of Hypernet data at the Oostende RT1 station allowing us to validate the super-resolution DINEOF product and its potential in capturing the coastal Turbidity dynamics. We have included this information in section 2.2 to make clear which year we are working with.

For the validation of the Remote Sensing product we extended the used Hypernet data to 2019 and 2022 to increase the number of matchups so solidify the assessment. For the years 2019, 2021 and 2022 less Hypernet data was available making it less useful to evaluate the super-resolution product.

Figure 2 caption – please, substitute "atmospheric correct algorithms" with "atmospheric correction algorithms"

Done

Section 2.2.1 – Remote sensing reflectance and pixel classification

Since this is a crucial element of the S2 & S3 preprocessing, it would be very useful for the reader if the authors could provide a synthetic overview of the C2RCC to ACOLITE/DSF pixel-based switching which is fully described in Van der Zande et al. (2023).

We have added the following paragraph to provide an overview of the approach:

"To combine the two approaches, a comprehensive, region-independent, and pixel-based automatic switching scheme is required, along with a technique for achieving a seamless transition between the two algorithms. The C2RCC to ACOLITE/DSF pixel-based switching is performed by means of band comparison of the RRS560 and RRS865 products (defined as green-nir ratio) as provided by the C2RCC processor. The green-nir ratio can be modelled using a logarithmic regression curve which starts as linear for the smaller reflectance values, but bends at the point where the saturation of the most sensitive band (i.e. RRS560) occurs. C2RCC pixels which deviate from the logarithmic model are considered erroneous outputs. The ACOLITE/DSF processor has the ability to provide higher RRS ranges compared to C2RCC while being noisier for lower RRS values thus highlighting the complementary between the two approaches. The green-nir ratio value of 45 is selected as the transition point between C2RCC and ACOLITE/DSF products. To ensure a smooth transition between the different ACs, a weighted transition is applied between the green-nir ratio boundaries of 50 and 40 based on the method described by Novoa et al. (2017). The C2RCC to ACOLITE/DSF pixel-based switching is described in detail in Van der Zande et al. (2023). Compatibility between the Sentinel-2/MSI products..."

Line 106 – from the text it appears that the IDEPIX software is applied soon after the implementation of the two atmospheric correction schemes but from Figure 2 they are at the same level as if the two steps were run simultaneously: it is confusing.

We changed Line 106 to be less confusing:

The IDEPIX software (v2.2.10, algorithm update 8.0.3), available as a The Sentinel Application Platform (SNAP) processor, is used for pixel classification, including cloud masking, cloud shadow identification, sea ice, floating vegetation, sub-pixel objects (ships, small islands and rocks), and the land-water distinction taking temporary water bodies (e.g. intertidal areas, lagoons) into account. SNAP is a software developed by the European Space Agency (ESA) designed for processing and analysing Earth observation data, particularly from the Sentinel satellites. It provides a common architecture for all Sentinel Toolboxes and enables the application of the C2RCC and IDEPIX processor on both Sentinel-2 and Sentinel-3 images.

In this or in the next section I would have expected to find relevant info about the resolution that is being used for the two sensors.

We have added this information in section 2.2.2.

Section 2.2.2 – Turbidity and Suspended Particulate Matter

If I understand it correctly, the algorithms used to generate SPM and TUR are an updated version of Nechad et al. (2010) that account for the "switching single band algorithms" developed by Novoa et al (2017). Still the way this paragraph is written is not very clear and I suggest rephrasing it for a better readability.

We have rephrased this paragraph for better readability

The SPM and the TUR products were generated using the generic multi-sensor algorithm described by Nechad et al. (2010). This algorithm provides the theoretical basis for SPM and TUR as a function of reflectance (RRS) at a single band, and provided calibration coefficients for all wavelengths, between 520 nm and 885 nm. It defines a relationship where RRS increases monotonically with SPM/TUR, at first linearly and then tends towards an asymptotic or "saturation" reflectance. This means that RRS becomes insensitive to changes in SPM/TUR which has led to the development of "switching single band algorithms" (Novoa et al., 2017) using using different wavelengths at different SPM concentrations to avoid the saturation effect and typically a smooth weighting between two adjacent spectral bands to avoid image artefacts. The Novoa et al. (2017) approach is applied to both the SPMand TUR products providing a multi-band SPM and TUR product using two bands (red: 665 nm and near-infrared: 865 nm). An example of the TUR products for the Belgian Coastal Zone region is provided in figure 3 showing a good correspondence between both the Sentinel-2/MSI and Sentinel-3/OLCI products providing information at different spatial and temporal resolutions.

Figure 3 – How do the two maps quantitatively compare? A scatterplot between the two would provide the reader with a better mean to interpret and compare the two products. This figure could then be cited also at lines 103-104 when talking about compatibility between the two sensor products. Please make the numbers on the colorbar larger, they are almost unreadable.

We have improved the readability of Figure 3 and added a scatter plot to illustrate how both images compare to each other.

Line 123 – what is the different temporal scale represented by the panels of Figure 3? From the caption it seems that both refer to the 5^{th} April, 2020.

Both panels are on 5 April 2020 as indicated on the caption. This date was chosen to show the same field with the two sensors, to ease the comparison.

Section 2.3 – Multi-sensor chlorophyll data

Line 126 – please spell out CMEMS as Copernicus Marine Environment Monitoring Service (even in braces is fine).

Reading of the text let understand that cmems_obs-oc_atl_bgc-plankton_my_l3-multi-1km_P1D only covers the period February-October 2022. Please, rephrase this sentence.

Thank you for these two comments, we have changed the text in the manuscript.

Section 2.4 – In situ data

How many in situ-satellite rrs matchups were extracted in the period 2019-2022? How come the 2023 and six months of 2024 were not included in the analysis?

More information on the number of match ups and total processed satellite images is available in section 4.2 Validation (Line 258-264), e.g. 59 matchups were available for S2/MSI and 179 for S3/OLCI. At the moment of analysis for this manuscript the 2023 and 2024 Hypernet data were

not yet available to us. The WaterHypernet systems are still in their development stage with the team working on its quality control and optimization resulting in the fact that the data is not released in NRT mode but rather in batches once the quality can be assured.

Section 3.2 – Generation of super-resolution data

It is not entirely clear why the authors only use data from 2020?

This is just a choice of period as we could have chosen another one. With DINEOF we do not need long time series to obtain accurate results, the analyses can be done in short-term periods of a few months. As mentioned in a comment above (and included in the manuscript), we have specified in the Data section that 2020 provided the most in situ data for validation, so we chose that period.

Figure 4 caption – please, substitute spare with square or box.

Thank you, it has been changed in figures 4 and 5.

Section 4.1 – Super-resolution data

The analysis presented in this section is effective but very qualitative; some hint on how to make it more quantitative at a reasonable cost is provided below.

Lines 202-204 – The only two optional … – what is the range of variability in the results associated with the settings of these two parameters?

We have added the following text in the paragraph: "As shown in Alvera-Azcárate et al (2009), the use of this filter can result in more EOFs being retained as optimal, which in turn results in a higher variability in the final results. Several tests were performed for values α = 0.01 to α = 0.1 and n = 1 to n = 10 and the combination that maximized the number of EOFs was retained."

Line 210 – isn't there also a peak in January?

Light during mid-December to mid-January is very low in the North Sea, which does not allow to accurately derive ocean colour variables. The time series starts therefore on mid-January, and this does not allow to properly resolve the peak that appears to occur in January. We have however mentioned the January peak in the text.

Line 210 – please substitute "apaprent" with "apparent"

Done

Figure 6, 7, 8 – even if these figures do provide a mean to interpret the overall results and the added value of the super-resolution data, this entire analysis lacks of robustness as it only refer to single case studies from which inferring a general rule might be difficult. An important missing information is the data density around the specific examples, that is, how many observations, both high and low resolution, are present in the previous and following days? This, along with the temporal distance between observations and interpolated data, would help explaining where the smaller scale present in the super-resolution data comes from. Probably, a more effective way to evaluate the outcome of the DINEOF interpolation would have been to randomly (or regularly along the time series) remove some day data (both high and low resolution) from the initial time series and use them for a more robust and statistically significant comparison: involving a larger number of observations.

It is indeed not easy to assess the 3D results with a handful of figures. However, the examples here are chosen to illustrate what the results are on a few days with different initial settings (Sentinel-2 or Sentinel-3 data, total absence of data, presence of noise…). As DINEOF uses the full 3D matrix to infer the final reconstruction, the presence or not of data just before and after, although very impactful, do not represent the full situation. As for the evaluation of the outcome retaining some information, we tried to do what the reviewer suggests in section "4.3 Scale Assessment", in which initially high resolution data are downgraded and the high resolution is not shown to DINEOF in a setting similar to the Sentinel-3/Sentinel-2 combination done in the present section. This allows to compare the high resolution original data with the DINEOF reconstruction, and the results (especially the zoom shown in figure 14) show that indeed the results have small-scale variability not present in the initial image.

Section 4.2 – Validation

Figures 10, 11 and 12 all have to do with the validation of satellite Rrs (both from S2 and S3) agaist in situ measurements using hypernet data. There is however some inconsistency between the statistics in figures 10 and 11 and those reported in figure 12. Please verify that the numbers in the figures are correct.

Figures 10 and 11 are consistent and we have omitted figure 12 to reduce the total number of figures in this manuscript.

Caption of figure 11 – please substitute "Ostend" with "Oostende" or the other way around, consistently with the rest of the manuscript.

Done

Figure 12 – I would expect some more degree of spectral consistency between rmse and mape plots, the lack of which could depend on the uneven distribution of the relative error characterized by long tails: perhaps, the median rather than the mean relative error would provide values more in line with respective rmse.

In order to reduce the number of figures in this manuscript we have omitted figure 12.

Line 282 – "very well" should be backed up by some statistics.

In this sentence we use the word "capture" to indicate that we are talking about the qualitative assessment of the results. The statistics are shown in figure 14. We have removed "very well" in the text to avoid confusion.

Figure 13 – the figure and associated discussion has some potential which should however be supported by some statistics, otherwise the entire paragraph is too qualitative and the reader might see it as only speculative.

To my understanding Figure 13 and Figure 14 contain the same data; figure 13 and associated discussion is too qualitative. On the other hand, Figure 14, if supported by an associated statistics, is more robust. Perhaps, dots in Figure 14 could be coloured according to time (months or seasons) condensing the information and reducing the number of figures.

Thank you for this comment. We do think that showing the time series of figure 13 is important, since the improvement in temporal resolution is clearly illustrated. The information on figure 14 is also important, as pointed out by the reviewer, in order to give quantitative metrics about the performance of the results.

Section 4.3 – Scale assessment

Figure 15 – showing the comparison on a selected transect provides a good idea of the outcome of the analysis, which unfortunately falls short of statistical robustness. Perhaps a relative error map instead of the very similar maps (for which it is almost impossible to find differences) would better complement the transect view. This comment also applies to figures 6, 7, 8 (perhaps not crucial because of the cloudiness), 9 and 16.

The statistical robustness of the results is mostly shown in the validation section. But the reviewer is right that showing change maps can help understand what the technique does. To avoid the large number of figures, we have moved figures 8 and 9 to the supplementary material, and we have added the difference field and the relative percentage change to the two examples of figures 6 and 7.

On the day with initially Sentinel-3 data (9 May, figure 6), there are changes observed along the whole domain, with a dominance of along-coast structures following the sandbanks present in the region and that cause changes in turbidity. Most of the changes between the initial and final product affect therefore the structure of these sandbanks. In the day with initially Sentinel-2 data (20 May, figure 7) we observe a similar behaviour, with long and thin structures shown in an along-shore direction, showing that the turbidity around the sandbanks is being modified.

Lines 318-319 – this sentence is perfectly in line with my previous comment about figure 15. I believe it mostly has to do with the way these results are presented. Another aspect that I would suggest to take into consideration is to try to condense the information as much as possible trying to avoid specific case studies which provide the reader with useful insights but are difficult to be used to derive general rules.

In that section, 4.3. we show the results in the whole domain to show how they look like overall and then proceed to show a zoom so that the details of the super-resolution are shown. With 3D matrices of more than a hundred time steps, it is difficult to show the details of the results without actually showing a few examples of the final product. The overall quality of the reconstruction is shown in the comparison with in situ data, but to actually show the final details of the super-resolution dataset we think showing the actual data is the best way. The addition of the relative change percentage images does contribute, we hope, to provide a quantification of the results resolution.

Line 325 – is there any reference or figure to back up this sentence? And perhaps with the drawback of reducing spatial variability by smoothing the field (as mentioned at line 300).

We have performed a few tests with linearly interpolated data, and they solve the issue of the squares, although we need to assess how this interpolation affects the final variability. We have therefore modified the sentence which now reads:

“A linear interpolation could avoid such pattern in the final results, although this would need to be tested.”

Section 5 – Submesoscale variability in the Belgian Coastal Zone

This section is a bit controversial: from one side it is presented as the right and expected application for the super-resolution data but on the other it is soon discovered that the data are not suitable to answer the question because of the low temporal resolution. I am not fully sure that this section actually adds value to the work, at least the way it is presented.

Lines 358-359 – even if it is somehow intuitive to assume that during boreal summer river outflows are at their minimum in Europe, it would also be preferable to have a reference to back up this sentence.

We agree that the lack of high temporal data hinders some of the analyses shown in this section. We have decided to remove figure 20 (the Hövmuller diagram) and related discussion which is the one that indeed would need this higher temporal resolution to be more conclusive. The two examples of specific dates and how the transect changes and is influenced by the presence of the sandbanks are retained, since they are a useful example of how these data can be used to better assess the influence of turbidity at small spatial scales.

Section 6 – Conclusions

Lines 362-366 – as they just mentioned in the previous paragraph, authors should also mention here the importance of the high temporal variability which, unfortunately, is still not covered by the ocean colour sensors currently on orbit.

We have added this sentence to the end of the third paragraph in the conclusions:

“Satellite data lack however the high temporal resolution that would be needed to study the variability of these small-scale features at adequate temporal scales.”

Line 385 – please substitute “omre” with “more”

Done