We are grateful for the reviewer's careful reading and comments which have served to sharpen the manuscript. Below we respond to the comments point by point.

Reviewer 1

1. There are a few issues with the order in which things are mentioned.

- Lines 29-34 do not mention the sections in the right order and one of the section numbers isn't rendered.

We have fixed the ordering and broken cross reference.

- It is customary to put the tables and figures in the same order as you refer to them in the paper, and this isn't how they are ordered at the moment.

We have reordered the figures.

- Lines 125-130 are mostly a repeat of lines 29-34.

Moved lines to 29-34 from 125-130

- Figure 4 does not seem to be mentioned in the text.

All figures are now mentioned in the text.

2. I suggest citing a couple of other packages for comparing models with observations, e.g.

- Castruccio F. S., 2021: NCAR/metric: metric v0.1. doi/10.5281/zenodo.4708277

- Roberts, C.D., 2017: cdr30/RapidMoc: RapidMoc v1.0.1. doi:10.5281/zenodo.1036387

We have integrated these citations, along with some other packages in a revised Introduction section. Please see lines 24 - 37.

3. Line 85: "By pairing it with dask" What does "it" refer to? Please rephrase.

Sentence rewritten:

" By integrating Dask and Xarray into COAsT, the user has access to a powerful system that provides lazy loading, chunking and parallel code. "

4. Line 86: "the user also has access to data analysis concepts". I'd argue that these are tools or applications. I have access to concepts without the package because concepts are entirely cerebral.

Replaced ‘data analysis concepts’ with ‘data analysis tools’

5. I found the description of CRPS to be better, but it would still be helpful to clarify: is the integral over y an integral over time, space, or both?

Added an extra sentence:

“The difference is integrated spatially over data from the radial neighbourhood.”

6. I’d love one more sentence about what the “mask_maker” class is.

Added an extra sentence (line 338):

This class contains predefined region definitions that help standardize spatially aggregated metrics.

7. Could the caption of Figure 3 please describe the meaning of the legend in panel a?

Added clarification that legend is in days. (Now Figure 1a)

Typos
1. Lines 106: “re used” You do not need a space
2. Line 113: eliminate “However”

Fixed both typos.

Reviewer 2

First, I would like to thank the authors for considering most of our comments and suggestions and for their frank answers. I regret, however, that nothing was added in the introduction to highlight how the package fits among the many python packages dedicated to ocean analyses.

We have added a paragraph to the introduction, and rewritten some other parts of the introduction to provide this context. Please see lines 24 - 37.

Nevertheless, given the useful philosophy developed throughout the article and the valuable validation methods, the paper is worth publishing.

A few (~12) corrections and improvements are still requested and are listed below:

L31 Remove (see Section-??) and join the two sentences. "The Gridded class is used to read, represent and manipulate output from two NEMO model runs, and its use interactively with the Profile and Tidegauge classes allows a comparison between the model and observed data.

Missing cross reference has been fixed.

Sentence replaced as suggested.

L34. Rewording: Information on the observation and modeled data used are given in Section 2.

L34-36 Please delete the paragraph, it is a conclusion, not an introduction. Also, "previous section" is non correct.

Deleted and reworded.

The sentence L34 "The package is open source and all of the code is freely available via Github (www.github.com)" could be replaced after L352.

We remove this line from the introduction and better describe the package availability in the Code Availability Statement.

The sentence "Using COAsT provides a level of transparency which aids knowledge sharing and discussion" could be moved to L366 after slight rewording.

This sentence has been removed.

L48 Even though the package aims to be extended to other ocean numerical models, it is not yet the case (see the documentation, the answer to reviewer 1, and the fact that analyses shown are based on NEMO only). Experiences always prove that despite the best prerequisites, extensions are not that straightforward. So please remove the reference to ROMS or at least add "in the future".

Reworded these sentences to read:

Typically, this kind of data would come from the output of a numerical model. At the time of writing, the package has been tested and used with output from the NEMO model, although this could be extended to other models in the future. The data should be stored in any xarray compatible file (e.g. netCDF, zarr).

L67-L68 The sentence "It labels these dimensions with 1-dimensional time coordinates, 2-dimensional latitude and longitude coordinates and 3-dimensional depth coordinates." is unnecessary (repetition of L47). Please remove it.

Removed.

Moreover, the dimensions do not fully comply to the C-Arakawa grids.

This is true within a single instance of the class. When using related grids with offsets, an instance of the class can be made for each grid. This is not mentioned in the paper as we do not consider variables on these other grided.

L71. Could you please add a few words about the analysis classes? I am not sure whether the tidal and profile analyses are directly available from the COAsT package. It seems they are saved apart under https://github.com/JMMP-Group/NEMO_validation.

or

Move paragraph L65-L71 ("At the time... Table-2) to L62, just before "Analysis classes...

Additional clarity on the COAsT analysis classes is added (L72): “For example, analysis classes include *GriddedStratification*, *ProfileAnalysis*, *TidegaugeAnalysis*, *Transect* and *Contour*”. Furthermore improved tracability of the code used, including the COAsT and NEMO_validation repositories, is provided in Code Availability Statement section.

L82-L87 add no more information. Please consider just removing them ("All three of ... chunking and parallel computation")

Removed.

L116 The three previous sentences are general and do not give any indication of what has been done in COAsT. After "module dependencies", consider adding a sentence stating that in COAsT, unit tests cover 67% of the package.

We have added a sentence:

For COAsT, unit tests cover approximately 67\% of the package at the time of writing.

Section 3. The numbering is unbalanced. Why not numbering L180 as follows?:

3.1 Harmonic Analysis
3.1.1 MHA
3.1.2 Harmonic uncertainty estimation where MHA is not possible (L203)

We have changed the section titles as recommended.

L229-230 Please clarify.
- Was MHA applied at all locations where it was possible and basic HA everywhere else?
- Was harmonic uncertainty estimated at every location (even where MHA was possible) and then used to "mask" correct assessments?

For the analysis presented in Figures 2 and 3, a normal harmonic analysis was applied everywhere, and uncertainty applied to ‘mask’ small differences between the models. We have added to the text to clarify this:

“Here we perform a normal harmonic analysis at each point using all available model data. Then, where differences between the model and the observations are smaller than the uncertainty, they are deemed insignificant and coloured grey in the figures. “

L237 In the figureS

Fixed.

Figure 1. Again the numbering is confusing because Figure 1 is detailed after Figures 2 and 3.

Figure ordering has been fixed.

Figure1/4 description. Please specify which method (MHA or/and just HA) has been applied. (see comment L229-230)

We have edited the captions of Figures 2,3 and 6 to add more information on harmonic method used.