

Egusphere-2025-657 – Answer to reviewers

REPORT#1

Review of "Enhancing coastal winds and surface ocean currents with deep learning for short-term wave forecasting", manuscript egusphere-2025-657

This manuscript presents a practical approach to improving the performance of numerical wave models by correcting their forcing fields—namely wind and surface currents—using Artificial Neural Networks trained on remote sensing data such as SAR and HFR. The methodology is applied and validated at multiple pilot sites, demonstrating consistent and significant improvements across several key metrics. The corrected forcings lead to better wave height and period predictions, both under normal and extreme conditions. Overall, this work is methodologically sound, relevant to the field of operational ocean forecasting, and contributes meaningful advancements in the integration of remote sensing with data-driven modeling techniques. Therefore, after making some appropriate revisions (mainly formatting issues), I believe this manuscript is suitable for publication. Here are some of my comments about the manuscript.

We thank Reviewer #1 for the positive feedback and for taking the time to review our manuscript. We have tried to address all your suggestions in this revised version, and we believe they have significantly improved the manuscript. We hope this revision meets with your approval.

Major comments:

1. I noticed that a GAN-based architecture was used for wind field correction, while an autoencoder-like structure was adopted for surface current correction. Could the authors comment on the rationale behind selecting these different architectures for the two tasks? Also, were other model types explored or compared during the development process, e.g., if a CNN-based super-resolution network was used directly instead of a GAN model (i.e., SAR is used directly as a target, with the network output calculating an RMSE-like loss), would this be any less effective? (Note that there is no need for the authors to add additional experiments here, just a brief discussion)

We thank the reviewer for this comment regarding our architectural choices. We agree that the rationale behind selecting different network types for the ANN wind and current corrections requires a clear explanation. We have incorporated the following text into the Discussion section:

Several architectures were tested for the Wind and Surface current ANNs, but the proposed GAN and AE networks provided the best performance. For the Wind ANN, better results were achieved using a GAN than a CNN-based super-resolution network, such as a U-NET (Ronneberger et al., 2015). Although GANs are more complex, they partially alleviate the problem of defining a suitable loss function. Furthermore, the Wind ANN requires a Generative architecture because it not only increases the image resolution but also attempts to reduce persistent biases. Conversely, for the Surface Current ANN, the problem involves corrections in both space and time, making a GAN with Convolutional-LSTM layers difficult to train; therefore, an Autoencoder was selected.

2. Figures 2 and 3 clearly illustrate the model architectures, and they are generally well-presented. However, the diagrams could be further improved by including more detailed information on the data dimensionality. For instance, adding the input and output shapes at the beginning and end of each model—either directly in the figures or in the accompanying text—would help readers better understand how the data is transformed through the network. This additional context would make the architecture more transparent and informative, especially for those interested in replicating or adapting the models.

We thank the reviewer for this constructive suggestion. We agree that providing clear information on data dimensionality is standard practice and greatly enhances the reproducibility and clarity of the figures.

We have addressed this by adding further explanations regarding the input and output shapes in the manuscript (specifically in Subsections 2.2 and 2.3).

Regarding the direct inclusion of tensor sizes in Figures 2 and 3, we opted to keep the diagrams schematic for the following reasons:

- Surface Currents ANN: Each pilot site model architecture uses different tensor sizes for input and output because these dimensions depend on the unique spatial extent of each High-Frequency Radar (HFR) domain. Since the tensor size varies across the three pilot areas (e.g., the HFR at Galicia has a larger extent than the one at Gran

Canaria), including these variable values directly in a single figure could easily lead to confusion for the reader. Therefore, for the sake of clarity, we chose to detail these site-specific dimensions within the accompanying text instead.

- ANN Wind: This model involves a super-resolution task where the input images are a fixed 10×10 pixels, and the output is a fixed 100×100 pixels.
3. I suggest the authors include a brief subsection (such as **2.6 Error Metrics**) in Section 2 that summarizes all the error metrics used throughout the manuscript. This summary should provide the definitions and explicit formulas for each metric (e.g., RMSD, bias, correlation, etc.). Doing so would enhance clarity and help readers better understand the evaluation criteria, especially those who may not be familiar with all the statistical indicators applied.

We thank the reviewer#1 for this suggestion. We have included a new subsection 2.6 Error Metrics.

Detailed comments:

1. The authors should pay close attention to citation formatting throughout the manuscript. For instance, in Line 70, the citation "(Gurgel et al., 1999)" is correctly formatted, but in other places (e.g., Line 52: "Hauser et al. 2023"), the comma after "et al." is missing. Such inconsistencies should be carefully checked and corrected. Additionally, figure references should follow the format of the journal—"Figure X" is appropriate at the beginning of a sentence, while "Fig. X" should be used elsewhere. Some citations are also inconsistently bolded, which should be standardized to maintain uniform formatting. Issues like this hopefully the authors can address them in a revised manuscript

We thank the reviewer#1 for this remark. We have revised the whole manuscript and we have standardized citation formatting and Figure references.

2. In Line 188, the authors refer to "training/validation datasets" in the context of evaluating model performance. However, if the dataset mentioned here is used solely for post-training evaluation rather than during model training for purposes like early stopping or hyperparameter tuning, it would be more accurate to refer to it as a "test dataset" rather than a "validation dataset". Similarly, the term "validation period" used later in the manuscript should be revised to "test period" or "evaluation period" which may help avoid confusion.

We thank the reviewer#1 for this remark. We have revised the entire manuscript and changed all references to "validation dataset" to "testing dataset."

3. In Lines 301–314, multiple date formats are used inconsistently, such as "January 2021 – January 2023", "Sep 2021 – Jan 2022", "25th – 27th November", "November 26–27, 2021", and "20th–23rd January 2022". I recommend standardizing the date format throughout the manuscript for consistency and improved readability.

We thank the reviewer#1 for this remark. We have standardized the date format throughout the manuscript with this convention: 20th January 2022.

4. Many formatting inconsistencies can be noted in Fig. 5 and Fig.9. For example, for the scatterplot, while the scale intervals are numerically the same, the x-axis has a sparser scale density than the y-axis. Also, the gridlines are either present or absent. The unit notation is also different between the two: one uses '[m/s]' while the other uses '(m/s)'. In addition, the 1:1 reference line is drawn in different colors - red on the left and green on the right - which may cause unnecessary distractions. Standardizing these visual elements will enhance the overall coherence and presentation quality of the charts.

We agree with Reviewer#1 that Figs.5 and 9 could have been improved. Thus, we have redone them, following Reviewer #1 recommendations.

5. In Fig. 7, which displays both positive and negative deviations, I suggest adjusting the color bar so that the central (white) point is explicitly labeled as 0. Additionally, using symmetric tick values for positive and negative ranges—ideally with a limited number of decimal places (e.g., [..., -7.6, -3.8, 0.0, 3.8, 7.6, ...])—would improve both the readability and the aesthetic quality of the figure. Also, the word spacing in the subheading of this image is odd.

We agree with Reviewer#1 that Fig. 7 required improvement. Thus, we have redone it, following Reviewer #1 recommendations. We have replaced the colour scale by a symmetric one.

6. In 13–16, there is a noticeable mismatch in the color tone between the plot lines and their corresponding legend entries—for example, while both may be shades of blue, one appears significantly lighter or darker than the other. If the legends were added during figure post-processing, using a color picker tool to precisely match the tones would improve the visual coherence. Although this does not affect the scientific interpretation, ensuring consistency in color tones would enhance the professionalism and clarity of the figures.

We thank Reviewer #1 for this remark. We have revised Figures 13 through 16 to correct the mismatch between the plot lines and their corresponding legend entries.

7. Throughout the manuscript, there are noticeable inconsistencies in figure formatting that should be addressed. For example, multiple styles are used for subfigure labels, including (1), [1], and (i), which creates confusion and detracts from the overall professionalism. Additionally, figure and table titles vary in formatting—some are in italics while others are in regular font, which should be standardized. Moreover, the resolution of several figures appears to be quite low, with visibly pixelated text and labels.

We thank Reviewer #1 for this remark. We have tried to address the inconsistencies in figure formatting. We have also tried to increase the quality of those figures with pixelated text and labels.