July 30, 2026

Dear Referee:

Thank you very much for your careful evaluation of our manuscript **egusphere-2025-6389** entitled "A Continuous Implicit Neural Representation Framework with Gradient Regularization for Sea Surface Height Reconstruction From Satellite Altimetry".

The following document reproduces the second-round comments in a point-by-point format. Point-by-point responses to all comments are provided below. The corresponding revised manuscript text is marked blue.

With kind regards,

Liming Pan
(on behalf of co-authors: Dongshuang Li, Liming Pan, Zhaoyuan Yu, and Linwang Yuan)

Point-by-point response to the second-round comments

Report of the Referee

A few minor questions and suggestions:

Comment 1 (Abstract): "*TV and SIREN should be defined in the abstract when they are introduced.*

"provides improved effective-scale diagnostics" is confusing until you read that part of the paper. I suggest "provides improved effective feature resolution"."

Response: We thank the reviewer for these helpful suggestions. We have revised the abstract to define the two abbreviations at their first appearance: SIREN is now written as "sinusoidal representation network (SIREN)", and TV is written as "total-variation (TV)". We also replaced "provides improved effective-scale diagnostics" with "provides improved effective feature resolution", as suggested, to state the intended result more directly. The revised text in the abstract is as follows:

- The framework uses a sinusoidal representation network (SIREN)-based coordinate network to learn a continuous space–time representation of SSH from sparse along-track observations, and includes total-variation (TV)-based spatial-gradient regularization to discourage unsupported local variations.
- The method maintains a level of global accuracy comparable to existing interpolation and data-assimilation approaches and provides improved effective feature resolution under the adopted benchmark protocol.

Comment 2 (Introduction): "*Line 55: "functions of coordinates" is a better description than "continuous coordinate-to-SSH mapping" on lines 70, 114, and later.*"

Response: We thank the reviewer for this suggestion. We agree that describing the INR as a function of its spatial and temporal coordinates is clearer than repeatedly using the phrase "coordinate-to-SSH mapping". We have therefore revised the Introduction and subsequent methodological and discussion passages for consistent wording. For example, the revised Introduction now states:

- Motivated by these advantages, this study leverages INRs to represent SSH as a continuous function of spatial and temporal coordinates from sparse satellite altimetry observations.
- In this study, we adopt SIREN as the INR backbone for representing SSH as a continuous function of spatial and temporal coordinates from sparse satellite altimetry observations.

The same wording is now used in the framework description, results discussion, and limitations section.

Comment 3: "*Line 180: This description of the gradients overlaps with the discussion and equations starting on line 380. Could they be made more consistent, please?*"

Response: We thank the reviewer for noting this possible ambiguity. We have revised the two passages to distinguish explicitly between (i) normalized-coordinate derivatives used only in the dimensionless TV regularization and (ii) physical SSH-gradient magnitudes used only as visualization diagnostics in the OSSE. The former are not converted to physical distance, whereas the latter are rescaled to longitude, latitude, and physical distance before plotting. The revised text now states:

- The second component of the proposed framework uses normalized-coordinate gradients of the reconstructed SSH field to reduce unsupported rapid local variations. Since the SIREN-based INR is differentiable with respect to its normalized input coordinates, these derivatives can be computed with respect to normalized longitude and latitude. They are used only to define the dimensionless TV regularization term below and are not interpreted as physical SSH gradients per unit distance.
- The physical SSH-gradient magnitude shown in the bottom row is a visualization diagnostic distinct from the normalized-coordinate gradients used in the TV regularization. It is derived from automatic differentiation of the SIREN representation and explicitly rescaled to physical distance.

Comment 4: "*Line 195: Optimizing for α breaks the independence of the test data.*"

Response: We thank the reviewer for identifying that the previous wording could suggest that the regularization coefficient a was optimized using the final test data. We have clarified the experimental protocol. The OSE and OSSE final test sets follow the predefined Ocean Data Challenges benchmark protocols and are common evaluation datasets for all participating methods. Hyperparameters, including a, network, and optimization settings, were determined using separate validation datasets. For the OSE, the validation dataset comprises the non-CryoSat-2 input missions during 16–31 March 2021 and the corresponding CryoSat-2 observations. For the OSSE, pseudo-altimetric observations and the corresponding NATL60 SSH fields from 2 January to 30 September 2013 form the validation dataset for hyperparameter selection. The final OSE and OSSE results are then reported for the prescribed 2021 CryoSat-2 and 2012 NATL60 benchmark evaluation datasets, respectively.

- The coefficient a, together with the network and optimization settings, was determined using the validation datasets described in Sect. 5.
- The revised text in Sect. 5.4.1 (OSE Data Preparation) is as follows: The OSE configuration is conducted over the Western Mediterranean Sea within the domain $[1°\mathrm{E}, 20°\mathrm{E}] \times [30°\mathrm{N}, 45°\mathrm{N}]$. It assesses the reconstruction method under realistic satellite sampling and measurement conditions. This domain defines a common reconstruction target. Following the SSH MapMed OSE benchmark, the reconstruction model is fitted using real along-track SSH observations from

SARAL/Altika (alg), Haiyang-2B (h2b), Jason-3 (j3), Sentinel-3A (s3a), and Sentinel-3B (s3b) from 1 January to 15 March 2021. The benchmark evaluation uses CryoSat-2 new-orbit observations (c2n) from 15 January to 15 March 2021. The same non-CryoSat-2 input missions from 16 to 31 March 2021 and the corresponding CryoSat-2 observations form the validation dataset for hyperparameter selection.

- The revised text in Sect. 5.5.1 (OSSE Data Preparation) is as follows: The OSSE configuration is conducted over a dynamically active portion of the Gulf Stream region, spanning 65°–55° W and 33°–43° N. It uses the NATL60 SSH simulation, based on the Nucleus for European Modelling of the Ocean (NEMO), as the known reference field. This domain defines a common reconstruction target. Following the NATL60 OSSE benchmark, the reconstruction model is fitted using pseudo-altimetric observations from the four-nadir-plus-SWOT configuration during the assessment period from 22 October to 2 December 2012. The complete NATL60 SSH simulation over the same domain and period provides the reference for full-domain evaluation. Pseudo-altimetric observations from the four-nadir-plus-SWOT configuration and the corresponding NATL60 SSH fields from 2 January to 30 September 2013 are used as the validation dataset for hyperparameter selection.

Comment 5: "*The section from line 217 to 229 might be clearer if put in the experiment descriptions in Sections 5.4 and 5.5.*"

Response: We thank the reviewer for this helpful suggestion. We have moved the relevant text to the OSE and OSSE experiment descriptions in Sects. 5.4 and 5.5.

Comment 6: "*Starting line 220: why start CryoSat-2 new-orbit observations from Jan 15 2021, instead of Jan 1?*"

Response: We thank the reviewer for requesting this clarification. The 15 January start date is prescribed by the SSH MapMed OSE benchmark for the CryoSat-2 new-orbit evaluation observations, which run until 15 March 2021. We have made this role explicit in the OSE data-preparation description:

- The benchmark evaluation uses CryoSat-2 new-orbit observations (c2n) from 15 January to 15 March 2021.

Comment 7: "*Equation (12) and Fig. 5 discussion starting line 332: How is* $\mathrm{PSD}_{\mathrm{err}}$ *determined for the OSE, since CryoSat-2 is only on sparse tracks?*"

Response: We thank the reviewer for identifying that the previous wording could suggest that Eq. (12) is also evaluated for the OSE. Equation (12), which was revised in response to the previous review round, is retained unchanged and defines the wavenumber–frequency score only for the OSSE, where the complete NATL60 reference field is available. We have instead clarified the OSE calculation directly in the Figure 5 discussion and caption. We follow the publicly available SSH MapMed OSE benchmark evaluation procedure: the reconstructed SSH is sampled at the locations and times of the withheld CryoSat-2 observations, and $\mathrm{PSD}_{\mathrm{err}}$ and $\mathrm{PSD}_{\mathrm{ref}}$ are then calculated as one-dimensional spatial PSDs of the along-track reconstruction error and CryoSat-2 reference SSH, respectively. Continuous CryoSat-2 tracks are separated at time gaps greater than 4 s and partitioned into nominal 300 km segments with 50 % overlap; Welch spectra are computed from the pooled valid segments. The 0.5 crossing provides only the OSE spatial effective scale λ_x.

The corresponding revised manuscript text is reproduced below:

- *Figure 5 discussion:* Because the withheld CryoSat-2 reference is available only on sparse along-track segments, the Figure 5 diagnostics are one-dimensional along-track spatial spectra rather than the wavenumber–frequency diagnostic of Eq. (12). Following the publicly available SSH MapMed OSE benchmark evaluation procedure, the reconstruction is sampled at the CryoSat-2 locations and times. Continuous tracks are separated at time gaps greater than 4 s and partitioned into nominal 300 km segments with 50 % overlap; Welch spectra are computed from the pooled valid segments. In Figure 5b, $\mathrm{PSD}_{\mathrm{err}}$ is the one-dimensional spatial PSD of the reconstruction error and $\mathrm{PSD}_{\mathrm{ref}}$ is the one-dimensional spatial PSD of the CryoSat-2 reference SSH on these same segments.

- *Figure 5 caption:* Here, $\mathrm{PSD}_{\mathrm{err}}$ and $\mathrm{PSD}_{\mathrm{ref}}$ are the one-dimensional spatial PSDs of the reconstruction error and CryoSat-2 reference SSH, respectively, computed from the valid CryoSat-2 along-track segments following the SSH MapMed OSE benchmark evaluation procedure. Higher scores indicate smaller relative spectral errors.

Comment 8: "*Fig. 8: Why is the temporal period cut off at 2 days, when there seems to be skill at shorter time periods? And why would this skill exist?*"

Response: We thank the reviewer for this important observation. The cutoff is not intended to claim temporal skill below 2 days. In the original figure, the temporal PSD is calculated from 41 daily SSH fields. The positive temporal-frequency grid therefore ends at $20/41\ \mathrm{day}^{-1}$, whose reciprocal is 2.05 days. The black 0.5 contour reaches this lower frequency-grid boundary; because the extraction routine takes the smallest temporal period on that contour, it returns 2.05 days. The calculation contains no PSD values at periods shorter than 2.05 days and does not extrapolate the contour beyond the boundary.

This does not remove the comparative result. All methods use the same daily-sampled PSD grid. INR and its no-TV ablation reach the highest temporal frequency available to this common diagnostic (2.05 days), whereas the non-INR methods cross the 0.5 threshold at longer periods. Thus, the result demonstrates a relative advantage of INR within the benchmark, while not permitting a comparison of performance below 2.05 days or between the two INR configurations at such periods.

We have revised the text, Table 2, and the Figure 8 caption to identify this value as a Nyquist-limited comparative diagnostic rather than an exact temporal resolved scale. In particular, the colors adjacent to the lower boundary are not evidence of verified sub-2.05-day skill. The ratio-based score can also be sensitive where the reference spectral energy is weak. A spectral assessment at shorter periods would require SSH fields sampled more frequently than once per day and a recomputed PSD diagnostic. The revised manuscript text states:

- The temporal axis terminates at 2.05 days because the PSD calculation uses 41 daily SSH fields. Its largest positive temporal frequency is therefore $20/41\ \mathrm{day}^{-1}$, corresponding to a shortest represented temporal period of $41/20 = 2.05$ days. The black 0.5 contour reaches this lower boundary. Thus, within the common benchmark, INR maintains the PSD-score criterion down to the highest temporal frequency accessible to the daily-sampled diagnostic, whereas the non-INR methods cross the criterion at longer periods. This is a valid relative comparison under the shared protocol. However, the boundary-limited 2.05-day value neither establishes skill at periods shorter than 2.05 days nor provides an exact temporal resolution.