Funded by
the European Union

PRIN 2022 PNRR Call for Proposals (D.D.1409 of 14/09/2022)

AIMS

Artificial Intelligence to Monitor our Seas

Project number: P2022587FM

Starting date: 30th November 2023 – Duration: 24 months

Deliverable D3.2
Report on training and validation with spatial-varying data

DOCUMENT INFORMATION

Deliverable number	D3.2
Deliverable title	Report on training and validation with spatial-varying data
Work Package	WP3
Deliverable type[1]	Report
Dissemination level[2]	Public
Due date	28.11.2025 (Month 24)
Pages	19
Document version[3]	3.0
Lead author(s)	Edoardo Pasta, (POLITO)
Contributors	Leonardo Gambarelli (POLITO), Giuseppe Giorgi (POLITO)

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

1 Type: R: Report; D: Dataset
2 Dissemination level: C: Confidential; P: Public
3 First digit: 0: draft; 1: peer review; 2: peer review 3: coordinator approval; 4: final version

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

DOCUMENT CHANGE HISTORY

Version	Date	Author	Description
DRAFT			
0.1	15/10/2025	Edoardo Pasta (POLITO)	Creation
0.2	7/11/2025	Leonardo Gambarelli (POLITO)	Consolidation of input from contributors
FIRST PEER REVIEW			
1.1	14/11/2025	Edoardo Pasta (POLITO)	Consolidation of input from reviewers
SECOND PEER REVIEW			
2.1	26/11/2025	Giuseppe Giorgi (POLITO)	Final version

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

SHORT ABSTRACT FOR DISSEMINATION PURPOSES

Abstract

This deliverable finalises the comparison of spatial interpolation methods for reconstructing Significant Wave Height (Hs) within the AIMS framework. We assess the three retained algorithms—Thin Plate Splines (TPS), Random Forests (RF) and Gaussian Process Regression (GPR)—using gridded moored buoys as independent test points and satellite altimetry plus a synthetic glider as input/training data. The algorithms are evaluated over January 2023, quantifying their performance through Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca).

TABLE OF CONTENTS

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca).

LIST OF PARTNERS

N°	Logo	Name	Short Name	City
1	Politecnico di Torino 1859	Politecnico di Torino	POLITO	Torino
2	ROMA TRE UNIVERSITÀ DEGLI STUDI	Universitá degli studi di Roma Tre	ROMA3	Roma
3	Italian National Research Council	Consiglio Nationale delle Ricerche	CNR	Firenze

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

ABBREVIATIONS

Acronym	Description
ANN	Artificial Neural Network
GPR	Gaussian Process Regression
MAE	Mean Absolute Error
RBF	Radial Basis Function
RMSE	Root Mean Square Error
RF	Random Forest
TPS	Thin Plate Spline

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

LIST OF FIGURES

LIST OF TABLES

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

EXECUTIVE SUMMARY

This deliverable completes the gate-clearance process initiated in D2.2, D2.5 and consolidated in D3.1, by testing, on spatially varying observations, the three algorithms shortlisted for operational deployment: Thin Plate Splines (TPS), Random Forests (RF) and Gaussian Process Regression (GPR).

Building on the December 2022 hindcast-based benchmark, where TPS and RF emerged as the most accurate methods and GPR as a robust backup option, we now assess their performance in reconstructing Significant Wave Height (Hs) from realistic observing networks.

The experimental setup uses the gridded moored buoys as independent test points, while along-track glider data and multi-mission satellite altimetry provide the input/training information. Ten time-instants are randomly selected within January 2023, and for each snapshot the algorithms interpolate Hs from the combined glider–satellite observations to the buoy locations. Model skill is quantified by the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) at the buoys.

The results confirm the good performance of GPR and TPS in handling heterogeneous, spatially sparse observations, while the RF achieves lower performance this time, probably due to the reduced amount of data in this setup, where no numerical model fields are used.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca).

1. Introduction

This deliverable extends the benchmarking work carried out in D3.1 "Report on training and validation on spatial-static data" [1], where a common testbed based on numerical hindcasts and fixed moored buoys was used to compare five candidate algorithms for spatial reconstruction of sea-state parameters. In that study, Thin Plate Splines (TPS), Random Forests (RF), Gaussian Process Regression (GPR), Radial Basis Functions (RBF) and Artificial Neural Networks (ANN) were evaluated on their ability to reconstruct Significant Wave Height (Hs) and Peak Period (Tp) from subsampled model grid points and a subset of buoys, with performance measured at independent hold-out buoys over December 2022.

The results highlighted TPS and RF as the best-performing methods, with GPR providing intermediate but robust performance and native uncertainty estimates, while RBF and ANN consistently underperformed. On this basis, TPS, RF and GPR were advanced to a final comparison stage, with RBF and ANN dropped from further consideration.

The present report focuses on the final comparison, moving from spatial-static, model-based fields to spatially varying observational data. The objective is to verify whether the ranking and qualitative conclusions obtained in the hindcast-driven benchmark remain valid when the algorithms are tasked with fusing heterogeneous measurements acquired by different observing platforms. In particular, we consider:

- the gridded moored wave buoys dataset [2] as the reference system and primary test set;
- the along-track glider dataset [3] that emulates an autonomous platform sampling the sea state along a moving transect as a first training set;
- the satellite altimetry dataset [4] providing additional, spatially sparse estimates of Hs as second training set.

All datasets are considered on the same geographical domain used in D3.1, ensuring continuity with the previous analysis and facilitating a like-for-like comparison of algorithmic performance. Nonetheless, it is important to highlight a key difference: in this setup, no numerical wave data are used for additional training, unlike in the previous D3.1, where data from numerical

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca).

wave model [5] was used for training the algorithms. This implied that in this setup the amount of available data is lower.

The three considered datasets (gridded moored buoys, glider and satellite altimetry) share January 2023 as their only common temporal overlap. For this reason, the analysis is restricted to this month, by selecting ten random time instants within January 2023. Each "time instant" is in practice defined as a 6-hour time-lapse, and all measurements falling within the corresponding 6-hour window are treated as if they were acquired at a single snapshot in time. The three algorithms—TPS, RF and GPR—are trained independently at each instant using latitude and longitude as predictors and Hs as the target variable, and are then used to reconstruct Hs at the buoy locations. Only Hs was considered as output parameter due to the fact that it is the only variable measured by all 3 datasets.

In order to assess the robustness of the methods with respect to the sampling strategy of the autonomous platform, two scenarios were considered for the glider. In the first scenario, the glider follows a deterministic, pre-defined trajectory, emulating an optimally planned survey pattern. In the second scenario, the glider trajectory is randomly perturbed, mimicking a less controlled, more opportunistic sampling strategy. In both cases, the moored buoys and satellite altimetry datasets are kept unchanged, and the full interpolation–validation procedure is repeated, thereby enabling a direct comparison between deterministic- and random-path glider configurations.

Model performance is quantified through the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) computed at the test buoys for each time.

This design mirrors the evaluation framework adopted in D3.1 but replaces the numerical hindcast with realistic, multi-platform observations and a glider-like sampling pattern. In doing so, the deliverable aims to answer two key questions: (i) whether TPS, RF and GPR retain their relative ranking when confronted with real-world observational geometry and measurement noise, and (ii) how sensitive each method is to the specific spatial configuration and sparsity of the input data.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

The conclusions of this report will inform the final selection of interpolation/regression tools to be integrated in the AIMS operational workflow for monitoring Significant Wave Height from heterogeneous observing systems.

2. Experiment Setups

From all the available datasets, ten time instants were randomly selected within January 2023, each corresponding to a 6-hour window. Denoting each instant as (dd-hh), with *dd* the day of the month and *hh* the starting hour of the window, the selected snapshots are:

- 03-12
- 06-06
- 14-00
- 16-12
- 16-18
- 21-12
- 22-12
- 24-12
- 26-00
- 29-12

Figures 1 and 2 show the spatial distribution of measurements for the time window starting on 22 January at 12:00, for the deterministic glider path and the random glider path, respectively.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

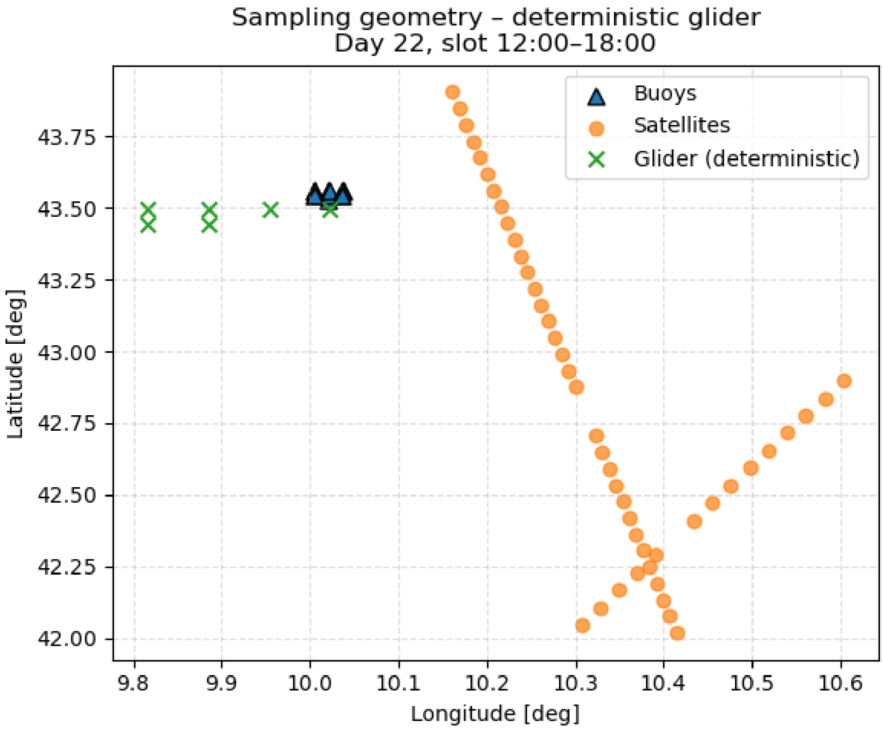

Figure 1 - Example of the spatial distribution of the measurements for the deterministic glider path

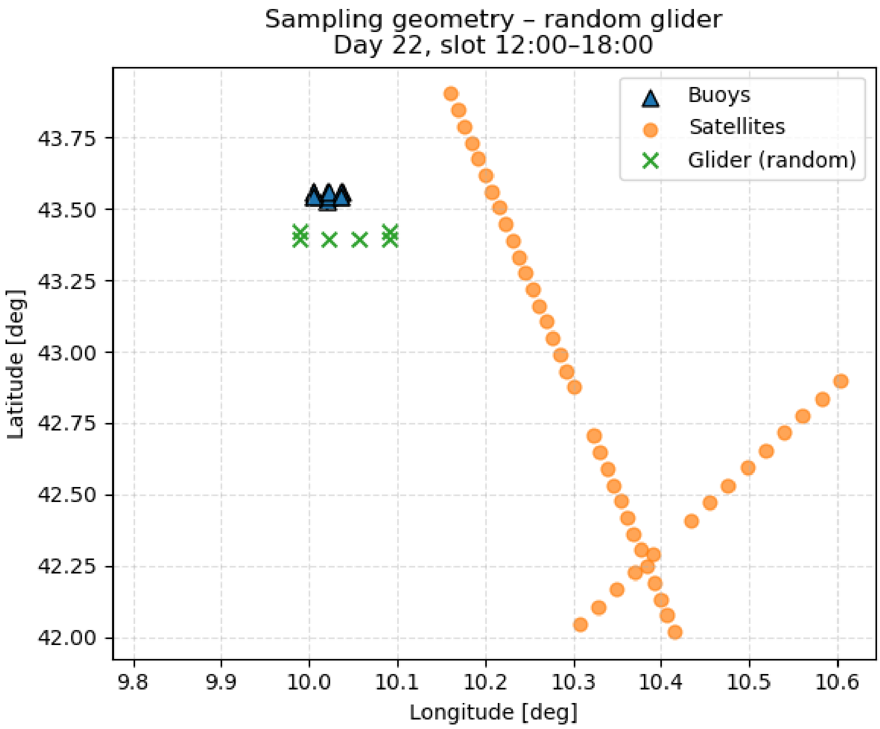

Figure 2 - Example of the spatial distribution of the measurements for the random glider path

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

From Figures 1 and 2 it is evident that, under the deterministic trajectory, the glider is able to sample a wider area compared to the random-path configuration.

Model performance was quantified, for every experimental setup, by the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE), computed on the buoys used for testing.

$$MAE = \frac{1}{T}\sum_{i=1}^{T}|z_i - \hat{z}_i|$$

$$RMSE = \frac{1}{T}\sum_{i=1}^{T}(z_i - \hat{z}_i)^2$$

3. Obtained Results

After running all three algorithms on each of the selected time instants, the reconstruction error at the test buoys was evaluated, allowing us to compute a MAE and a RMSE for every algorithm over the whole month of January 2023.

The resulting performance metrics are reported in the following figures and in the accompanying summary table.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

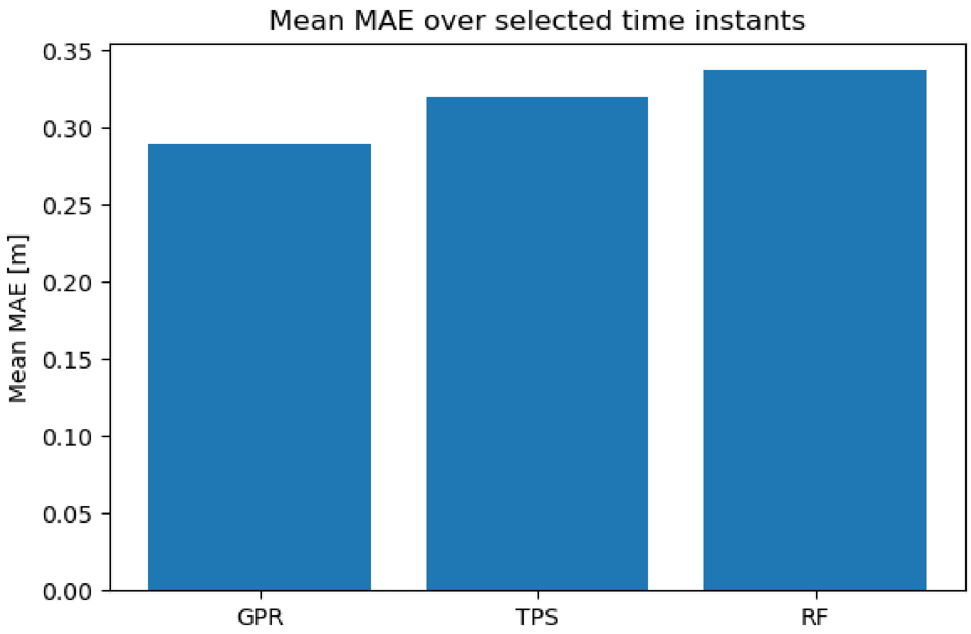

Figure 3 - MAE achieved by the reconstruction algorithms for the deterministic glider path.

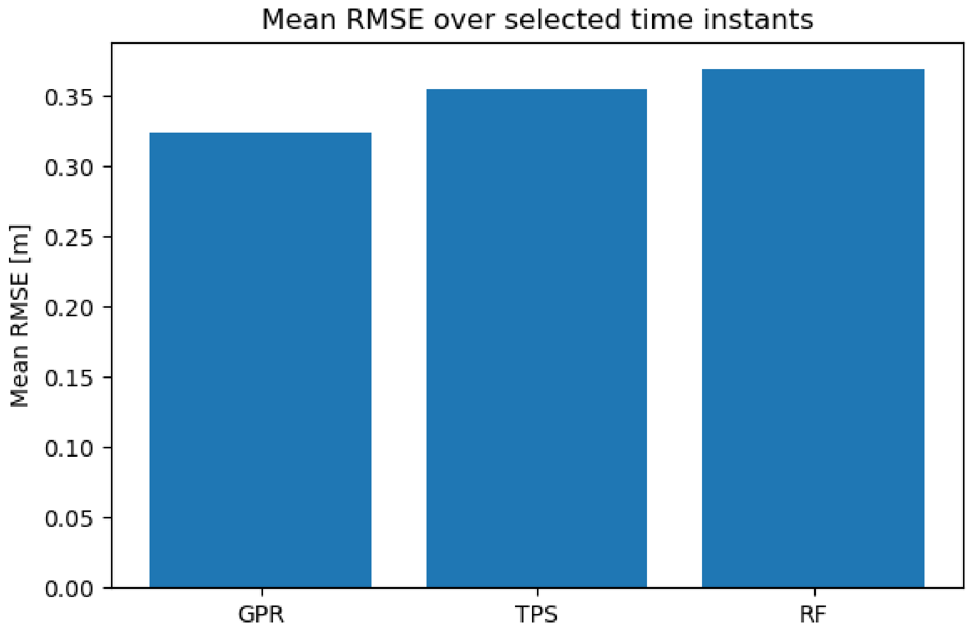

Figure 4 - RMSE achieved by the reconstruction algorithms for the deterministic glider path.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

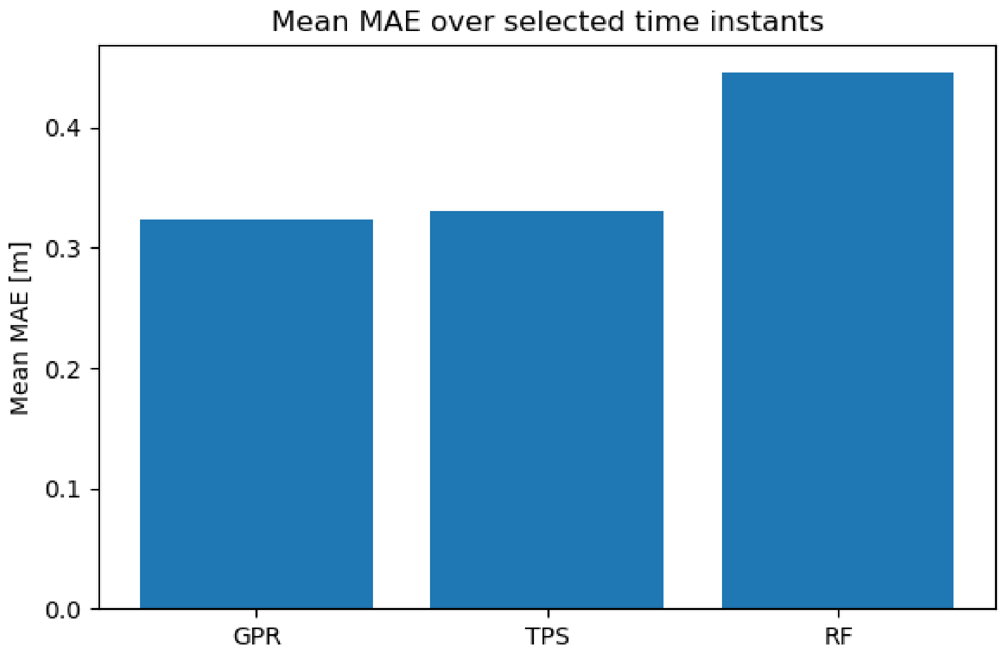

Figure 5 - MAE achieved by the reconstruction algorithms for the random glider path.

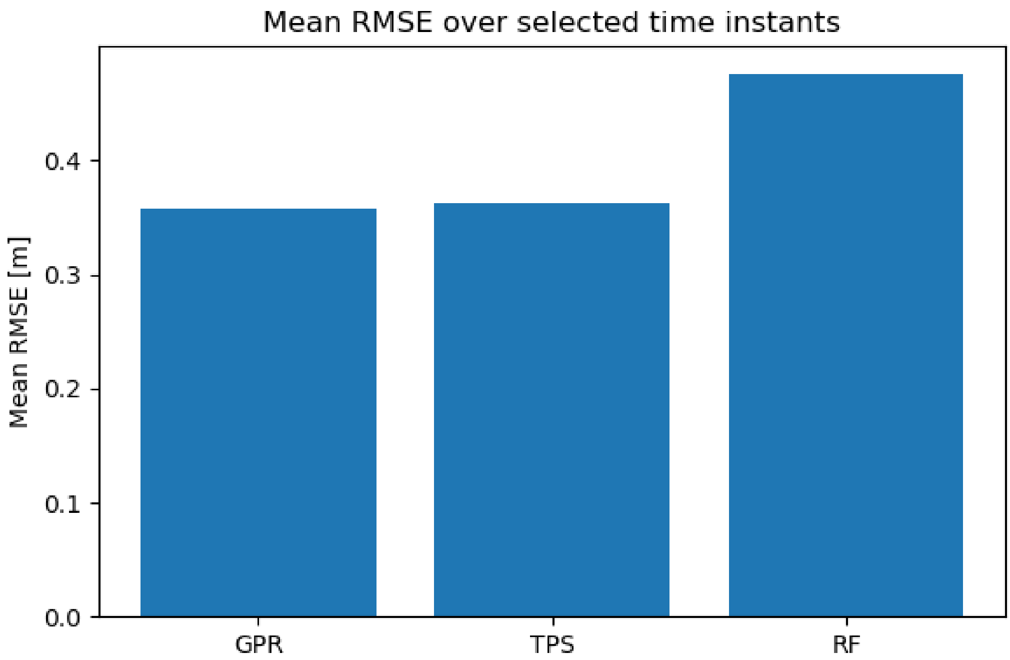

Figure 6 - RMSE achieved by the reconstruction algorithms for the random glider path.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

Table 1 - MAE and RMSE achieved by the reconstruction algorithms for the deterministic glider path.

Algorithm	MAE (m)	RMSE (m)
TPS	0.320 m	0.355 m
GPR	0.289 m	0.324 m
RF	0.337 m	0.369 m

Table 2 - MAE and RMSE achieved by the reconstruction algorithms for the random glider path.

Algorithm	MAE (m)	RMSE (m)
TPS	0.330 m	0.362 m
GPR	0.324 m	0.358 m
RF	0.446 m	0.476 m

As can be seen from the figures and tables, GPR systematically emerges as the best-performing algorithm, for both glider configurations and for both performance indicators (MAE and RMSE), while RF consistently yields the largest errors. This behaviour can be largely attributed to the characteristics of the available datasets: the number of training samples is relatively small, and measurements are spatially sparse and unevenly distributed. GPR explicitly models spatial covariance and is known to be robust under data sparsity, remaining effective even when only a limited number of samples are available. In contrast, RF models typically require larger and more homogeneous training sets to adequately explore the input space and properly tune the ensemble of trees.

A second observation concerns the choice of glider trajectory. All three algorithms tend to perform better with the deterministic glider path than with the random one. This can be explained by the fact that the deterministic trajectory provides a more uniform coverage of the study area and avoids collecting many successive measurements in the same location. As a result, the deterministic path offers a more informative and spatially balanced training set, which translates into improved reconstruction skill at the buoy locations.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca).

4. Conclusions

This deliverable has completed the final comparison of three spatial interpolation methods—Thin Plate Splines (TPS), Random Forests (RF) and Gaussian Process Regression (GPR)—for the reconstruction of Significant Wave Height (Hs) within the AIMS framework. Building on the spatial-static benchmark of D3.1, the present study moved to spatially varying observational data, using gridded moored buoys as independent test points and combining satellite altimetry with a synthetic glider as input/training data. The analysis focused on January 2023, selecting ten 6-hour time windows in which the three datasets overlap, and considered two glider configurations: a deterministic, pre-planned trajectory and a randomly perturbed path.

The results show that GPR systematically achieves the lowest errors in terms of both Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), for both glider configurations. TPS performs consistently well and remains close to GPR, while RF exhibits noticeably higher errors, particularly in the random-path scenario. This behaviour is consistent with the limited and spatially sparse nature of the available training data: GPR and TPS are relatively robust to sparse sampling, whereas RF typically benefits from larger, more homogeneous datasets to fully exploit its ensemble structure.

The comparison between deterministic and random glider trajectories further highlights the importance of sampling strategy. All three methods perform better when the glider follows the deterministic path, which provides a more uniform coverage of the study area and reduces redundant measurements in the same location. This suggests that, for an operational AIMS configuration, a carefully designed glider survey pattern can significantly enhance the quality of Hs reconstructions, independently of the specific interpolation algorithm.

Overall, the findings confirm that GPR and TPS are well suited for integrating heterogeneous, spatially sparse observations of Hs, with GPR offering the additional advantage of probabilistic uncertainty estimates. RF appears less competitive under the current data constraints but may remain of interest in settings with denser observational networks or as part of ensemble approaches.

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.

REFERENCES

[1] Deliverable D3.1 Report on Training and Validation on Spatial-static data.
[2] Deliverable D1.3 Report and dataset from moored gridded wave buoys.
[3] Deliverable D1.4 Report and dataset from gliders experimental campaign.
[4] Deliverable D1.2 Report and dataset from satellites.
[5] Deliverable D1.1 Dataset from Wave Models

AIMS: Artificial Intelligence to Monitor our Seas is funded by the European Union - NextGeneration EU within the PRIN 2022 PNRR program (D.D.1409 del 14/09/2022 Ministero dell'Università e della Ricerca). This document reflects only the authors' view, and the Commission and Ministry cannot be considered responsible for any use that may be made of the information it contains.