

Author's Proof

Before checking your proof, **please read the instructions below**

- Carefully read the entire proof and mark all corrections in the appropriate place, using the Adobe Reader commenting tools ([Adobe Help](#)). Do not use the Edit tool, as direct edits could be missed (the PDF was blocked for editing to prevent this); annotate your corrections instead.
- Provide your corrections in a single PDF file or post your comments in the Production Forum making sure to reference the relevant query/line number. Upload or post all your corrections directly in the Production Forum to avoid any comments being missed.
- We do not accept corrections via email or in the form of edited manuscripts.
- Do not provide scanned or handwritten corrections.
- Before you submit your corrections, please make sure that you have checked your proof carefully as once you approve it, you won't be able to make any further corrections.
- To ensure timely publication of your article, please submit your corrections within 48 hours. We will inform you if we need anything else; do not contact us to confirm receipt.
- Note that the column alignment at the bottom of each page is not ensured during this Author's Proof stage. The columns will be correctly aligned in the final PDF publication. You may therefore notice small differences in the structure of the Author's Proof PDF versus the final publication.

Do you need help? Visit our [Production Help Center](#) for more information. If you can't find an answer to your question, contact your Production team directly by posting in the Production Forum.

NOTE FOR CHINESE-SPEAKING AUTHORS: If you'd like to see a Chinese translation, click on the  symbol next to each query. **Only respond in English** as non-English responses will not be considered. Translated instructions for providing corrections can be found [here](#).

Quick checklist

- Author names** - Complete, accurate and consistent with your previous publications.
- Affiliations** - Complete and accurate. Follow this style when applicable: Department, Institute, University, City, Country.
- Tables** - Make sure the meaning/alignment of your Tables is correct with the applied formatting style.
- Figures** - Make sure we are using the latest versions.
- Funding and Acknowledgments** - List all relevant funders and acknowledgments.
- Conflict of interest** - Ensure any relevant conflicts are declared.
- Supplementary files** - Ensure the latest files are published and that no line numbers and tracked changes are visible.
Also, the supplementary files should be cited in the article body text.
- Queries** - You must reply to **all of the typesetter's queries below** in order for production to proceed.
- Content** - Read all content carefully and ensure any necessary corrections are made, then **upload them** to the Production Forum.

Author queries form

Query no.	Details required	Authors response
Q1	Confirm that the article title is correct and check that it makes sense. 	
Q2	If you are proposing or creating new or updated taxonomy and/or nomenclature, check your taxonomy and/or nomenclature carefully to avoid any errors that may invalidate it. 	
Q3	The citation and surnames of all of the authors have been highlighted. Check that they are correct and consistent with your previous publications, and correct them if needed, noting that the format in the author list should be	

	[First name] [Surname]. Please note that this may affect the indexing of your article in repositories such as PubMed. 	
Q4	Confirm that the email address in your correspondence section is accurate. Any changes to corresponding authors requires individual confirmation from all original and added/removed corresponding authors. Please note Authorship Change Forms are not required for amendments to the correspondence section. 	
Q5	There is a discrepancy between the styling of the author names in the submission system and the manuscript. We have used "Ana F. Duarte" instead of "Ana Filipa Duarte". "João Borges de Sousa" instead of "João Tasso Sousa". Please confirm that it is correct. 	
Q6	Confirm that all author affiliations are correctly listed. Per our style guidelines, affiliations are listed sequentially and follow author order. Requests for non-sequential affiliation listing will not be fulfilled. Note that affiliations should reflect those at the time during which the work was undertaken. If adding new affiliations, specify if these should be listed as a present address instead of a regular affiliation. 	
Q7	Please expand the term "LSTS, LAETA, +ATLANTIC CoLAB, CERENA/DER " if applicable. 	
Q8	The abstract should ideally be structured according to the IMRaD format (Introduction, Methods, Results and Discussion). Provide a structured abstract if possible. If your article has been copyedited by us, please provide the updated abstract based on this version. 	
Q9	Confirm that the keywords are correct, and keep them to a maximum of eight and a minimum of five. (Note: a keyword can be made up of one or more words.) 	
Q10	Check if the section headers (i.e., section leveling) have been correctly captured. 	
Q11	If you decide to use previously published and/or copyrighted figures in your article, please keep in mind that it is your responsibility as the author to obtain the appropriate permissions and licenses to reproduce them, and to follow any citation instructions requested by third-party rights holders. If obtaining the reproduction rights involves the payment of a fee, these charges are to be paid by the authors. 	
Q12	Ensure that all the figures, tables, and captions are correct, and that all figures are of the highest quality/resolution. You may upload improved figures to the Production Forum. If so, please describe in visual terms the exact change(s) made to help us confirm that the updated version has been used in the finalized proof. Please note that figures and tables must be cited sequentially, per the author guidelines 	
Q13	Check that all equations and special characters are displayed correctly. 	
Q14	Provide a URL for the LOOP profile for the following authors if they wish this to be linked to the final published version. If they are not yet registered, ensure that they register with Frontiers at the provided link.  "Lucrezia Bernacchi" If a URL is not provided, the profile link will not be added to the article. Non-registered authors and authors with profiles set to "Private" will have the default	

	profile image displayed. Note that we will not be able to add profile links after publication.	
Q15	We have replaced the section heading "Acknowledgments" with "Funding" as the text appears to discuss funders of the work. Please confirm that this is correct. 	
Q16	Ensure all grant numbers and funding information are included and accurate (after publication it is not possible to change this information). All funders should be credited, and all grant numbers should be correctly included in this section. Note that if you add any commercial funding, please ensure that the funders involvement/non-involvement in the manuscript is declared. If you provided a positive funding statement but don't provide funding details, then the statement will be updated to say no funding was received. 	
Q17	Confirm that the Data Availability statement is accurate. Note that this statement may have been amended to adhere to our Publication Ethics guidelines. 	
Q18	Confirm that the details in the "Author Contributions" section are correct. If any contributions need to be added/edited, choose the appropriate CRediT roles from the list available here and indicate which one(s) apply. Please be aware that writing roles ("Writing – original draft" and/or "Writing – review & editing") are a requirement for authorship. 	
Q19	Confirm if the text included in the Conflict of Interest statement is correct. Please do not suggest edits to the wording of the final sentence, as this is standard for Frontiers' journal style, per our guidelines : The authors/remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest 	
Q20	Provide the journal title, volume number and page range for the following references, if applicable. "Bernacchi et al., 2025", "Toth and Vigo, 2002".	
Q21	Provide the publisher location and accessed date for the url (accessed [Month D, YYYY]) for "European Union-Copernicus Marine Service, 2017", "Gould et al., 2013".	
Q22	Provide the accessed date for the url (accessed [Month D, YYYY]) for "Paull et al., 2012", "Petillo et al., 2010", "Pinto et al., 2018", "Teixeira et al., 2021".	
Q23	The image used in Figure 8 has part labels a-e; however, the part labels e, f are missing in the caption. Please either provide a revised figure or modify the caption, as necessary. 	
Q24	<p>Please confirm that the below Frontiers AI generated Alt-Text is an accurate visual description of your Figure(s). These Figure Alt-text proposals won't replace your figure captions and will not be visible on your article. If you wish to make any changes, kindly provide the exact revised Alt-Text you would like to use, ensuring that the word-count remains at approximately 100 words for best accessibility results. Further information on Alt-Text can be found here.</p> <p>Figure 1 Alt-Text – A globe highlights the location of Portugal, with a zoomed-in map pinpointing Nazaré. Panel a) shows a temperature map off the coast, indicating values from 17°C to 29°C. Panel b) provides a closer view detailing temperature ranges between 17°C to 17.8°C. Panel c) is a bathymetric map of the area, showing ocean depth contours ranging from -5000 to 0 meters.</p>	

Figure 2 Alt-Text – Flowchart illustrating an autonomous underwater vehicle (AUV) operation model. The top section, "First day modelling," shows predicted temperature and uncertainty leading to an AUV mission. The bottom section, "Next day prediction," involves a mean model feeding into an auxiliary AUV assimilation layer, which updates predicted temperature and uncertainty. A backup AUV mission is also indicated. Timeframes reference 14 days prior and plus one day increments.

Figure 3 Alt-Text – Vertical and horizontal 3D temperature graphs. Panel a) shows a vertical color-coded temperature variation over 14 days, ranging from 14 to 20 degrees Celsius, with a color gradient from blue to red. Panel b) is a zoomed-in horizontal view of the top section, emphasizing detailed temperature distribution. Scales are provided for both axes.

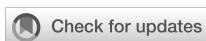
Figure 4 Alt-Text – Map series showing a region with temperature and concentration gradients. Panel (a) displays temperature data in a color gradient from yellow to purple, ranging from 16°C to 20°C. Panel (b) illustrates concentration levels, with values from 0 to 0.2 on a blue scale. A red rectangle highlights a specific area on both maps. Panel (c) presents a zoomed-in view of the highlighted region, detailing specific concentration values with a marked path, dot, and star.

Figure 5 Alt-Text – Scatter plot comparing AUV and CMEMS temperature data in degrees Celsius. Points represent different depths using colored shapes, with hour intervals shown as various symbols. A diagonal red dashed line serves as a reference.

Figure 6 Alt-Text – Three heat maps labeled a, b, and c display temperature distributions over a grid. Panel a uses a gradient from yellow to purple, showing a temperature range from sixteen to eighteen degrees Celsius. Panel b features shades of green, indicating subtle variations. Panel c combines colors from the first two panels, suggesting a more complex distribution with values ranging up to zero point two. Axes denote distances in kilometers.

Figure 7 Alt-Text – Three heat maps labeled a, b, and c, depicting temperature variations across a grid. Each map shows coordinates in kilometers along the x and y axes from four thousand three hundred seventy-six to four thousand three hundred ninety-six kilometers. Panel a uses a yellow to purple gradient, panel b uses a green gradient, and panel c uses a blue to yellow gradient. Color bars indicate temperature in degrees Celsius, ranging from sixteen to eighteen degrees in maps a and c, and values from zero to zero point two in map b.

Figure 8 Alt-Text – Panel of images depicting geostatistical predictions and temperature data. (a) and (d) show colored maps with black path lines over a gradient background indicating kilometers and temperature in degrees Celsius. (b), (c), (e), and (f) are scatter plots comparing AUV data and geostatistical predictions of temperature, with points color-coded by depth and symbols representing different hour intervals. A key on the right explains the color and symbol codes.



OPEN ACCESS

EDITED BY

Chengbo Wang,
University of Science and Technology of China,
China

REVIEWED BY

Christoph Waldmann,
Technical University of Applied Sciences
Luebeck, Germany
Shaoxiong Qiu,
Northeastern University, China

Q4

*CORRESPONDENCE

Ana F. Duarte
filiptamasoduarte@tecnico.ulisboa.pt

RECEIVED 28 July 2025

ACCEPTED 29 September 2025

PUBLISHED xx xx 2025

CITATION

Duarte AF, Bernacchi L, Mendes R,
de Sousa JB and Azevedo L (2025)
Geostatistical uncertainty maps for
real-world efficient AUV data collection.
Front. Mar. Sci. 12:1674989.
doi: 10.3389/fmars.2025.1674989

COPYRIGHT

© 2025 Duarte, Bernacchi, Mendes, de Sousa and Azevedo. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

KEYWORDS

AUV path planning, uncertainty mapping, geostatistical modeling, autonomous underwater vehicles (AUVs), ocean, predictive models, data assimilation

Geostatistical uncertainty maps for real-world efficient AUV data collection

Ana F. Duarte^{1*}, Lucrezia Bernacchi², Renato Mendes^{2,3,4},
João Borges de Sousa^{2,3} and Leonardo Azevedo¹

¹CERENA/DER, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal, ²Underwater Systems and Technology Laboratory (LSTS), Faculdade de Engenharia, Universidade do Porto (FEUP), Porto, Portugal, ³Associate Laboratory for Energy, Transports and Aerospace (LAETA), Faculdade de Engenharia, Universidade do Porto (FEUP), Porto, Portugal, ⁴ATLANTIC CoLAB, Museu das Comunicações, Rua do Instituto Industrial, Lisbon, Portugal

Autonomous Underwater Vehicle (AUV) trajectory planning for oceanographic surveys is challenging and requires comprehensive and efficient data collection for enhanced mission success. By strategically navigating and targeting high-value data points, the AUV can operate longer and gather more essential information for numerical ocean model calibration. Here, we propose a geostatistical modelling workflow with two complementary objectives. First, to jointly predict ocean temperature and spatial uncertainty maps, representing regions with limited knowledge about the ocean properties of interest, from where optimized navigation paths can be devised and updated. Second, to efficiently assimilate the collected data and update an ocean model with the new data. An autonomous oceanographic survey performed off W. Portugal illustrates the proposed modelling workflow. We use the CMEMS product of Atlantic-Iberian-Biscay-Irish-Ocean Physics Analysis and Forecast as *a priori* and conditioning data of the spatial predictions. During the survey, the data acquired by the AUV are assimilated and used in new geostatistical predictions for the day after the data acquisition. The results show that the proposed methodology efficiently predicts daily ocean temperature and its spatial uncertainty, allowing data assimilation from different sources (i.e., numerical models of ocean dynamics and AUV sampling). This approach enables the assimilation of AUV measurements and the model prediction to have higher value and greater reliability.

111 Q10 1 Introduction

Numerical models of ocean dynamics are essential tools for predicting the spatiotemporal distribution of ocean properties such as temperature and salinity. Accurate ocean predictions are essential for different areas related to climate modelling and forecasting, pollution monitoring, the blue economy, including resource management, and predicting ocean behavior. However, the quality and forecasting capabilities of numerical models of ocean dynamics are highly dependent on the number, quality and spatial distribution of *in situ* observations of ocean properties, as these oceanographic measurements are assimilated during the calibration of numerical models of ocean dynamics (Gould et al., 2013; Sloyan et al., 2019). Knowing how often and where to sample the ocean is non-trivial due to the vastness and harsh oceanographic environment, costs and complex logistics involved during data acquisition.

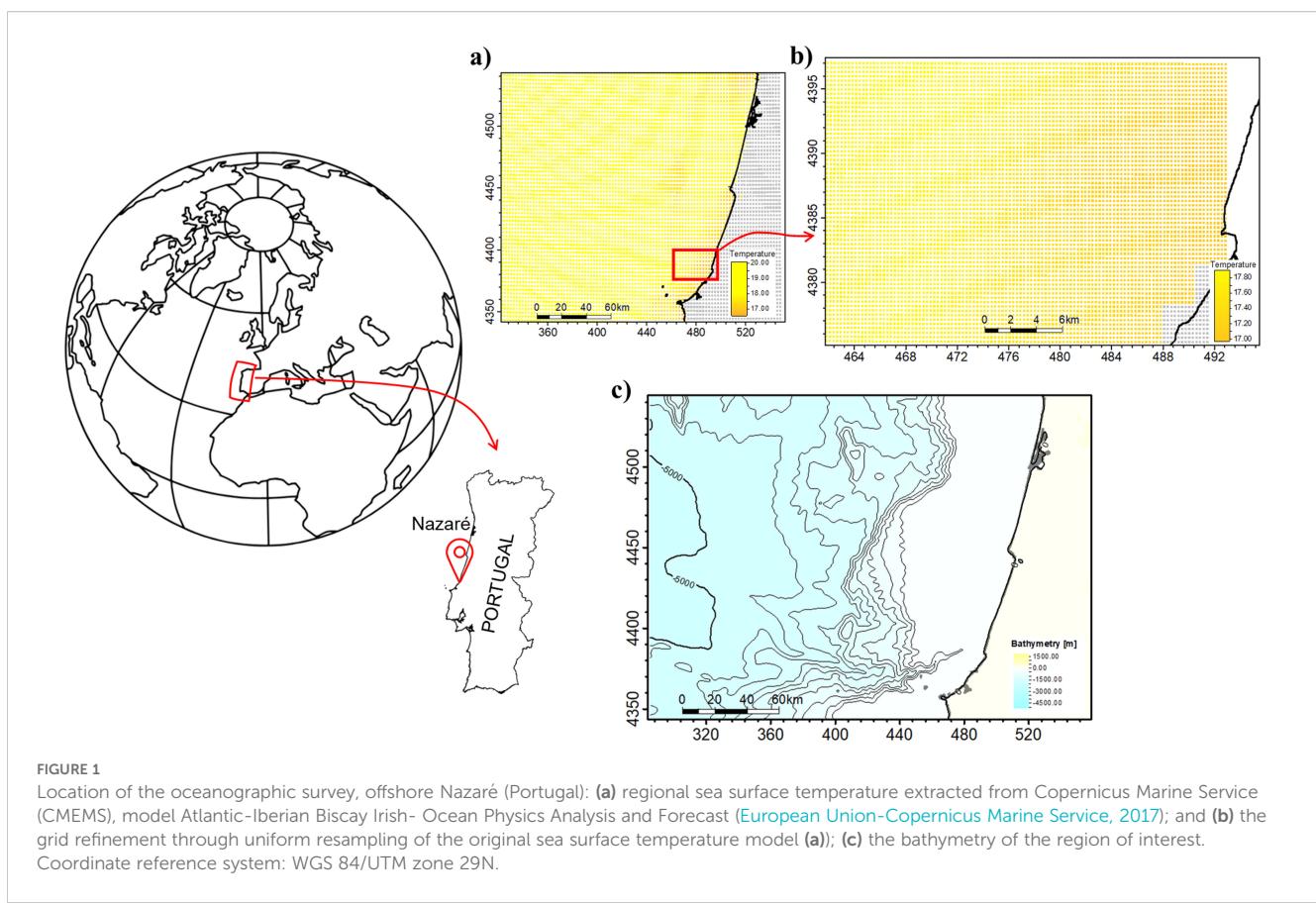
Diverse technologies exist to sample the ocean. Traditionally, *in situ* measurements are obtained using casts deployed from ships or moored at buoys, and provide accurate high-resolution measurements along the vertical direction for ocean model calibration. However, they are logistically complex and costly to operate and, while providing high-vertical resolution, they are spatially sparsely distributed (Kite-Powell et al., 2008). Complementary, Earth observation data are a great source of information as they cover large extents of the ocean near-surface but are subject to the first tenths of meters below the sea surface, have low spatial resolution and depend on atmospheric conditions (Minnett et al., 2019; O'Carroll et al., 2019; Mahdavi et al., 2021).

In recent years, the use of autonomous underwater vehicles (AUVs) for ocean sampling has brought significant benefits and garnered attention. These vehicles can carry various sensors that take quasi-synoptic high-resolution measurements in continuous time of several properties in large areas at a relatively low-cost. However, AUVs have endurance limitations, depending on the size of the vehicle and navigation and ocean conditions. On average, the endurance of the AUVs used in the field application shown herein is up to 8 hours sailing at 3 knots. Due to these operational limitation the AUV's paths should be concentrated within areas of greater relevance for numerical ocean model calibration and validation (e.g., areas of high spatial variability and/or uncertainty) (Gafurov and Klochkov, 2015). Besides, the high variability of ocean dynamics and the continuous influx of measured data by AUVs require the vehicle to perform near-real-time intelligent trajectory planning (i.e., adaptive sampling) (Petillo et al., 2010; Yu et al., 2022). In adaptive sampling, the objective is to predict the types and spatiotemporal locations of new sampling locations that are expected to be the most useful (i.e., informative) given previously acquired data and the vehicles' characteristics in terms of navigation and autonomy capabilities. Adaptive sampling strategies depend on the objectives of the surveys, such as mine countermeasures (Hwang et al., 2019), sampling for water mass classification (Paull et al., 2012), numerical model solutions (Lermusiaux, 2007), and tracking coastal upwelling and river plume fronts (Zhang et al., 2012; Pinto et al., 2018; Mendes et al., 2021; Teixeira et al., 2021; Ge

et al., 2023; Mo-Bjørkelund et al., 2025), and the number and type of vehicles used. Various types of multi-variable optimization algorithms can be used to plan intelligent AUV trajectories (Yu et al., 2022).

Here, we do not focus on the optimization algorithm (Bernacchi et al., 2025) applied for adaptive sampling, but on building near-real-time relevant information regarding the spatial uncertainty of ocean modelling when performing AUVs' path planning. As for the adaptive ocean sampling strategy, we use a lightweight multi-vehicle path planning algorithm based on the Prize-Collecting Vehicle Routing Problem (PCVRP) (Toth and Vigo, 2002; Toth and Vigo, 2014) capable of considering practical constraints typical of real-world oceanographic surveys, such as endurance limitations for each vehicle, distinct deployment and recovery locations and known obstacles within the operational area. As adaptive sampling requires near-real-time information about the spatial distribution of ocean properties, the use of conventional numerical models of ocean dynamics is challenging, as these simulators are computationally expensive and their execution is slow, hampering their usability in real-time planning tools. Alternatively, we propose a computationally efficient proxy to generate spatial uncertainty models and assimilate new data through geostatistical simulation (Deutsch and Journel, 1998). The multi-vehicle path planning is applied over three-dimensional models of spatial uncertainty of ocean temperature generated by computing the pointwise standard deviation distance of a set of ocean temperature models predicted with geostatistical simulation (i.e., geostatistical realizations). The pointwise standard deviation model allows the identification of areas with high uncertainty when forecasting the model's behavior (i.e., grid locations are considered uncertain if they exhibit within the entire ensemble of geostatistical simulations high variability). The geostatistical realizations are generated based on calibrated models of ocean dynamics up to a given day and existing direct ocean observations. The use of geostatistical simulation alleviates the computational burden of conventional numerical ocean models (i.e., they act as a proxy of the full numerical model of ocean dynamics), allowing for the easy assimilation of new direct observations obtained during oceanographic surveys. This approach provides good approximations for short-term ocean temperature predictions and the corresponding spatial uncertainty.

The proposed methodology was subjected to real-world conditions, where it was thoroughly tested and validated by a fleet of Light Autonomous Underwater Vehicles (LAUVs) in an oceanographic survey located offshore Nazaré (Portugal) (Figure 1), in the context of the FRESNEL (Field expeRiments for modEling, aSsimilatioN, and adaptive sampLing, <https://lsts.fe.up.pt/project/fresnel>) field campaign conducted in late October 2024. In this oceanography survey, the AUV's path is daily planned based on ocean temperature predictions constrained by the data acquired on that day. The experimental results obtained using the proposed methodology demonstrate its effectiveness in operational contexts, the benefits of integrating it with adaptive ocean sampling strategies and potential future implementation on board the AUVs.



In the next section, we describe in detail the proposed methodology to assimilate direct observations of ocean temperature acquired by AUVs and how to update spatial uncertainty maps. Then, in Section 3, we show the results obtained in the FRESNEL survey. Section 4 presents the results obtained and explores potential pathways for the proposed methodology. The last section summarizes the main conclusions of this work.

2 Methodology

The proposed methodology uses geostatistical simulation (i.e., direction sequential simulation ([Soares, 2001](#)) to predict three-dimensional models of ocean temperature and corresponding spatial uncertainty. The predictions are short-term, for a relatively small number of consecutive days, but they use limited computational resources and avoid running full numerical models of ocean dynamics. The temperature predictions are based on pre-existing, calibrated deterministic numerical models of ocean dynamics, which extend up to a certain day before the survey starts (or the day of interest). [Figure 2](#) summarizes the proposed methodology, described in the subsequent sub-sections.

2.1 Geostatistical prediction of ocean temperature

In stochastic sequential simulation ([Deutsch and Journel, 1998](#)) all cells of a numerical model represented by a grid are sequentially visited following a random path. At each location along the random path, the kriging estimate and variance computed jointly from observed data (i.e., *in situ* measurements) and previously simulated grid cell locations within a neighborhood are used to define a probability distribution function of the variable of interest from which a value is drawn. As each run considers a different random path, and therefore the conditioning data is modified along the simulation path, it results in diverse predictions (i.e., geostatistical realizations). Different stochastic sequential simulation methods exist and differ in their *a priori* assumptions ([Deutsch and Journel, 1998](#)). In the proposed methodology, we use direct sequential simulation ([Soares, 2001](#)) (detailed in Appendix I) due to its flexibility and the ability to use the observed data directly without any Gaussian transform. In direct sequential simulation, the kriging estimate and kriging variance are computed following a spatiotemporal covariance matrix represented by a variogram model fitted to the experimental variogram, often computed from observations. In the field application shown below we use a variogram model fitted to the CMEMS data due to their spatially exhaustive nature. The imposed spatiotemporal continuity model (i.e., covariance matrix) is an *a priori* assumption about the natural

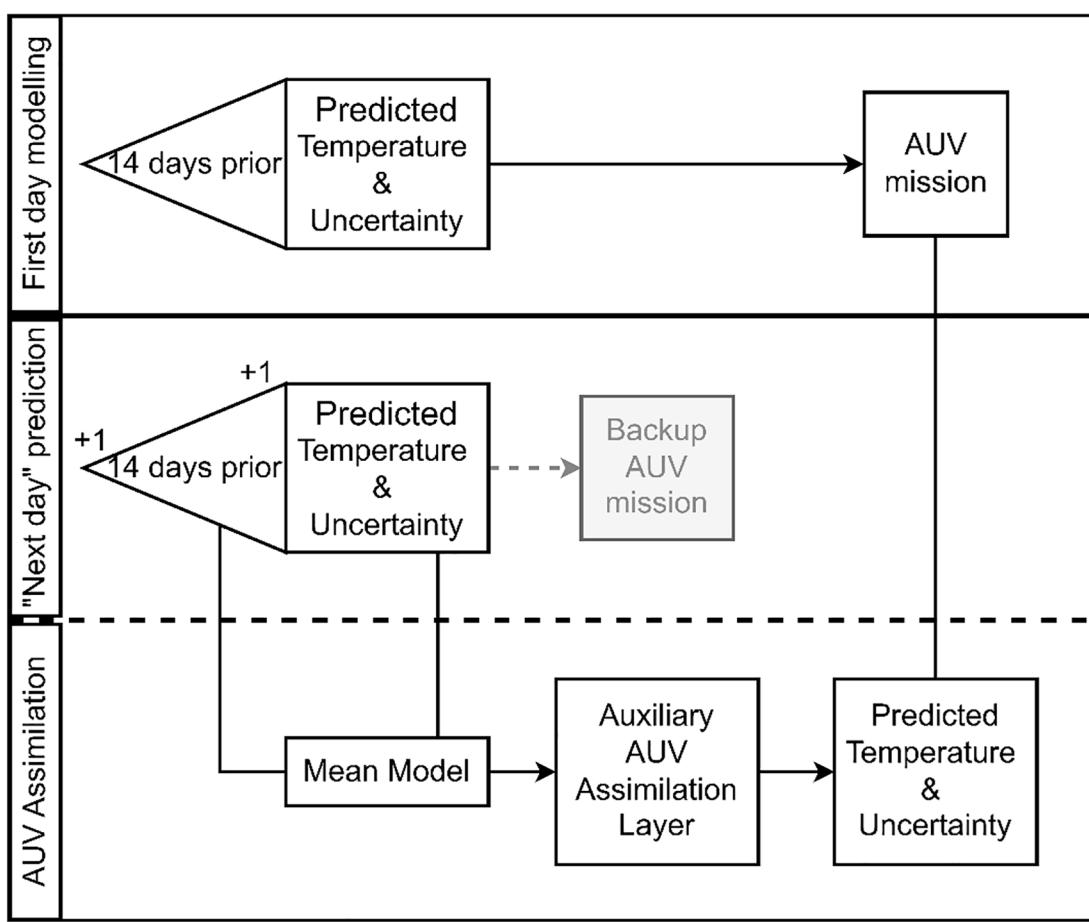


FIGURE 2
Schematic representation of the proposed methodology.

phenomena being modelled and exactly reproduced in all geostatistical realizations within a given ensemble of models.

The proposed methodology starts by retrieving the expected spatiotemporal continuity pattern for ocean temperature based on the *a priori* information provided by a deterministic calibrated ocean model for the study area over a long period (Figure 3a). These data are used to explore and model the spatiotemporal continuity pattern (i.e., directions of maximum spatiotemporal continuity). In the proposed methodology, the horizontal dimension represents the geographical coordinates of the numerical model of ocean dynamics (i.e., a deterministic model), while the vertical direction represents the temporal evolution of the model at a given depth. Specifically, the long-term calibrated ocean temperature is collated at each depth of the ocean numerical model (Figure 3a). Then, the three-dimensional experimental variogram is computed and fitted with a model. The range of the fitted variogram model along the vertical direction represents the maximum temporal continuity expected at that specific depth, and the horizontal range its spatial extent. The spatial analysis continues by repeating this process for the next depth of the numerical model of ocean dynamics (or down to the depth of interest). The independent variogram model per depth sample simultaneously capture of the spatiotemporal distribution of ocean temperature.

To forecast ocean temperature at a given depth, we use a set of fourteen previous days from the deterministic calibrated ocean model as observations while predicting the subsequent day (Figure 3b). Hence, the simulation grid is composed of fourteen layers as represented by the *a priori* data (i.e., the calibrated ocean model), followed by the layer corresponding to the forecasting day. Alternatively, additional *a priori* days might be considered depending on the oceanographic complexity of the study area. The selection of the grid size will mainly affect the computational cost of the simulation. After trial and error, and considering the study area specification, we opted for fourteen days, as this number allows for an efficient prediction of the subsequent day within a reasonable computational cost that could eventually be supported onboard by an AUV. Due to computational limitations, each depth of the numerical model is independently simulated. The described procedure is then applied to each depth of interest.

Multiple runs of the geostatistical simulation generate multiple geostatistical realizations of ocean temperature. The ensemble of geostatistical realizations approximates the posterior distribution of the predictions. It can be analyzed to assess the spatial uncertainty and variability of ocean temperature predictions. The spatial uncertainty maps can be obtained by computing the pointwise inter-quartile distance, or standard deviation, within the ensemble

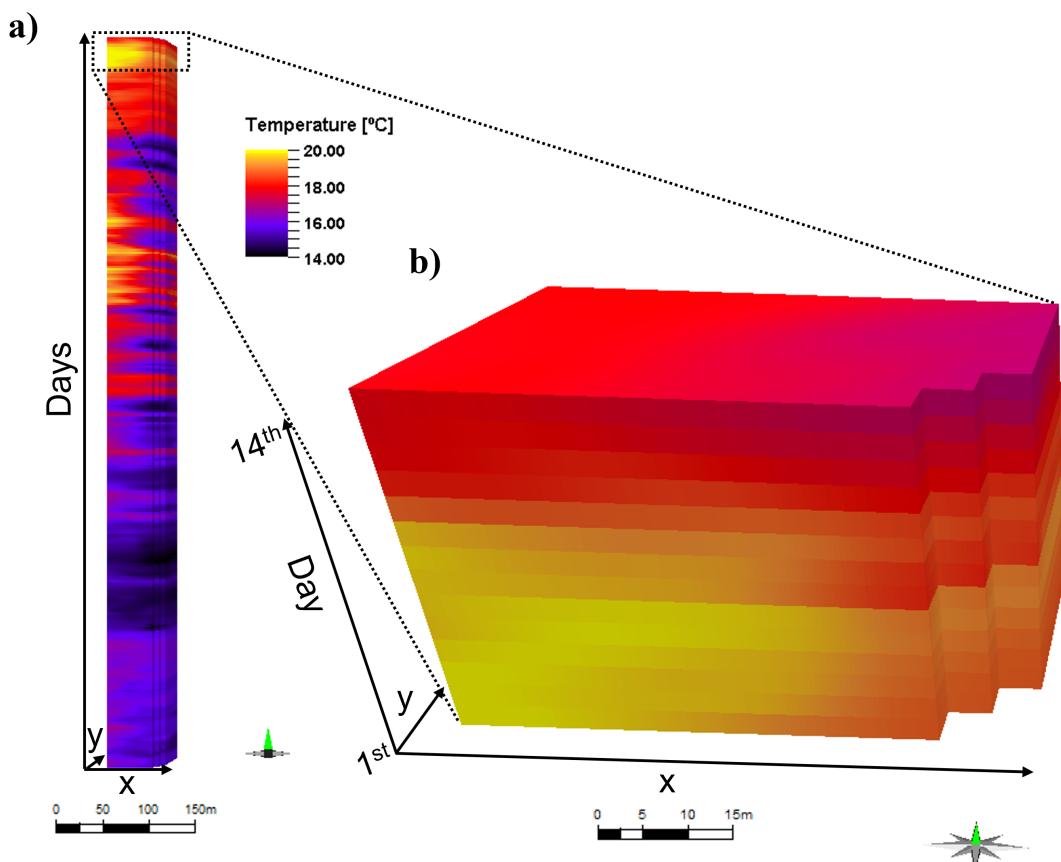


FIGURE 3
(a) Ocean temperature predicted from a calibrated deterministic temperature model for a given depth considering a long temporal window. Each layer represents a given day in the long-term prediction; **(b)** for the same depth, zoom-in of the most recent fourteen consecutive days predicted by the deterministic model. This information is used to predict the subsequent one in the application example shown herein.

of geostatistical realizations. The predicted uncertainty maps enable the identification of areas that are more complex and/or less characterized, and can be used for intelligent path planning algorithms where high-uncertainty regions have higher rewards.

2.2 Data assimilation and stochastic model update

Besides predicting spatial uncertainty models for AUV path planning, the proposed geostatistical framework allows the seamless integration of direct measurements acquired by an AUV during an oceanography campaign. The data acquired on a given day of the oceanographic survey updates the ocean temperature forecast for the subsequent day. At this stage, we aim to predict the ocean temperature for the next day by combining previous predictions with newly measured data. In this case, the simulation grid is defined with only two layers (i.e., time steps). The first layer corresponds to the direct measurements of the AUV at its corresponding spatial location, and the second layer comprises the time step at which the new prediction will occur. As the spatial sampling of the AUV is frequently larger than the spatial resolution of the simulation grid, the data values that fall into the

same cell are averaged to accommodate for resolution differences. In other words, we use a simple arithmetic upscaling between the AUV measurements and the grid cell value. Alternative methods could be applied depending on the objectives of the oceanographic survey (e.g., select the most extreme measurement for the upscaling). In highly heterogeneous areas (e.g., front regions) gradient-aware upscaling methods might produce more robust results. To incorporate the spatiotemporal patterns of the predictions for the day when the direct measurements were obtained, we use stochastic sequential simulation with local means (Soares, 2001) at the time step corresponding to the data acquisition. In direct sequential simulation with local means, the simulated value is based on its conditional distribution, considering previously simulated values, sample data and the local expected value as represented by the local mean model. This approach enables the update of the entire ocean model while simultaneously incorporating direct AUV measurements and previous *a priori* knowledge.

By incorporating these data sources, the simulation captures spatial trends and patterns that might otherwise be overlooked. It allows us to integrate new measurements without losing the overall spatial features obtained during the first step. The uncertainty of these predictions can be assessed as previously described (e.g., by

computing the pointwise inter-quartile distance, or standard deviation, from a set of geostatistical realizations).

The temperature and spatial uncertainty predictions, obtained with the proposed methodology, are then used for intelligent path planning of AUVs and are sequentially updated based on the direct measurements acquired by AUVs during the oceanographic survey. The adaptive sampling algorithm (Bernacchi et al., 2025) concentrates measurements on locations with higher uncertainty as predicted by the ensemble of geostatistical realizations (i.e., spatial uncertainty maps). The use of this geostatistical simulation method allows to assimilate the direct observations and stochastically update the *a priori* models.

3 Field demonstration

The proposed methodology was tested and validated in a field oceanographic survey off W. Portugal (Figure 1) during the FRESNEL campaign for the last two weeks in October of 2024. The campaign aimed to test the autonomous maritime robots' capacity to sample the ocean water, adapting their path based on the predictive models' outputs, and to assimilate the acquired data for the subsequent predictions. The campaign involved the simultaneous deployment of multiple assets, including three Light Autonomous Underwater Vehicles (LAUVs) developed and operated by the Underwater Systems and Technology Laboratory (LSTS), University of Porto. The AUVs were equipped with conductivity, temperature, and depth (CTD) sensors that sampled the water from the surface until a depth of approximately 40 meters (to a maximum of 100 meters in specific cases where the bathymetry allows) following a “yo-yo” descending-ascending movement.

3.1 Data

The deterministic numerical model of ocean dynamics used to conditioning the geostatistical simulation was downloaded from the Copernicus Marine Service (CMEMS), model Atlantic-Iberian Biscay Irish- Ocean Physics Analysis and Forecast (European Union-Copernicus Marine Service, 2017), covering an area from 11.75°W to 8.43°W and from 39.00°N to 41.16°N with horizontal resolution of 0.028° in both directions, approximately 2.37km and 3.09km in x and y, respectively (Figure 1a), at Nazaré latitude. Vertically, the model has 17 depth samples, with decreasing vertical resolution with depth, covering from the sea surface to a depth of 40 meters. Temperature prediction is performed for each of these depth samples individually and independently. Since the spatial resolution of the AUV measurements is much higher than the deterministic CMEMS model, we increased the model grid spatial resolution eight times, by uniformly resampling the original grid, to better discretize the data to be assimilated throughout the grid model (Figure 1b). We applied the proposed methodology considering October 29th, 2024 as acquisition day, to predict the next day (i.e., October 30th, 2024).

3.2 Results

On the first survey data (October 29th, 2024), ocean temperature predictions and spatial uncertainty models were generated using information from the 14 previous days of the CMEMS product for the study area. This step yields seventeen 3-D datasets (one per depth sample considered), where the horizontal dimensions represent the x and y coordinates of the study area, and the vertical dimension represents the time step (i.e., a day). The CMEMS information was used as experimental data to model the three-dimensional variogram models at each depth sample and to constrain the geostatistical simulation of 100 realizations per depth. Figure 4 shows the resulting temperature and corresponding spatial uncertainty represented by the pointwise median (Figure 4a) and standard deviation (Figure 4b) computed from the ensemble of 100 geostatistical realizations, respectively. This information was used to select the deployment location and to define the AUVs' path planning (Bernacchi et al., 2025) on the first day of the survey (Figure 4c).

After deployment and AUV recovery, the measured temperature values at their corresponding coordinates and depths were upscaled to fit into the refined model grid. In general, the CMEMS model overpredicts the ocean temperature measurements acquired by the AUV, with its performance decreasing with depth (Figure 5). Also, the range of AUV temperature measurements per grid cell increases with depth. This effect is expected as the vertical size of the CMEMS model grid cells increases with depth, allowing for more AUV samples per grid cell.

The upscaled temperature values were then assimilated into the models as conditioning data for predicting ocean temperature and spatial uncertainty for the next day.

3.3 Assimilation of temperature values and prediction

To highlight the importance of data assimilation and model update in obtaining more reliable short-term predictions, we begin by illustrating the temperature and spatial uncertainty predictions without AUV data assimilation for the next day (October 30th, 2024). Figure 6 illustrates temperature predictions based exclusively on the 14 days prior to the prediction day, CMEMS data, and does not consider any *in-situ* data. The pointwise median of 100 geostatistical realizations of temperature (Figure 6a) shows higher temperature values to the west of the model than in the opposite direction. The same region is characterized by high spatial uncertainty values (Figure 6b).

The next day's predictions after assimilation of the temperature values measured by the AUV are shown in Figure 7. The main spatial patterns of temperature are similar to those obtained without the assimilation (Figure 6), as the geostatistical simulation includes the *a priori* information provided by the calibrated oceanographic model for the survey area and the *in situ* measurements. However, the spatial uncertainty map (Figure 7b) changes considerably, illustrating the update of the *a priori* information with the data

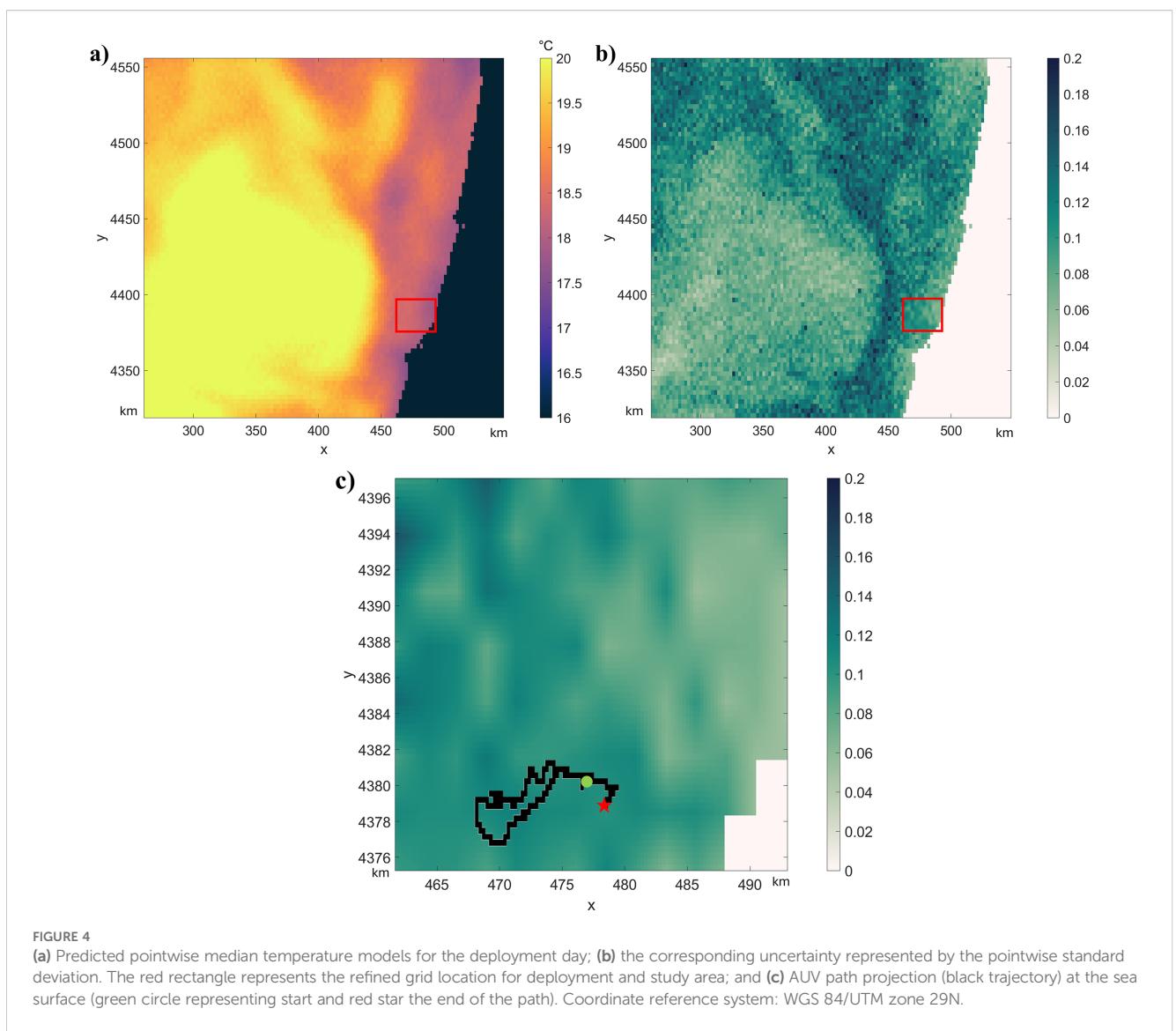


FIGURE 4

(a) Predicted pointwise median temperature models for the deployment day; (b) the corresponding uncertainty represented by the pointwise standard deviation. The red rectangle represents the refined grid location for deployment and study area; and (c) AUV path projection (black trajectory) at the sea surface (green circle representing start and red star the end of the path). Coordinate reference system: WGS 84/UTM zone 29N.

acquired during the survey. The highest temperatures (i.e., ~18°C) are located in the western part of the model. Still, the uncertainty values are low for the area surveyed by the AUV and larger in the opposite direction. Additionally, the predictions obtained from the CMEMS model are corrected to include colder water for the region sampled by the AUV (Figure 5).

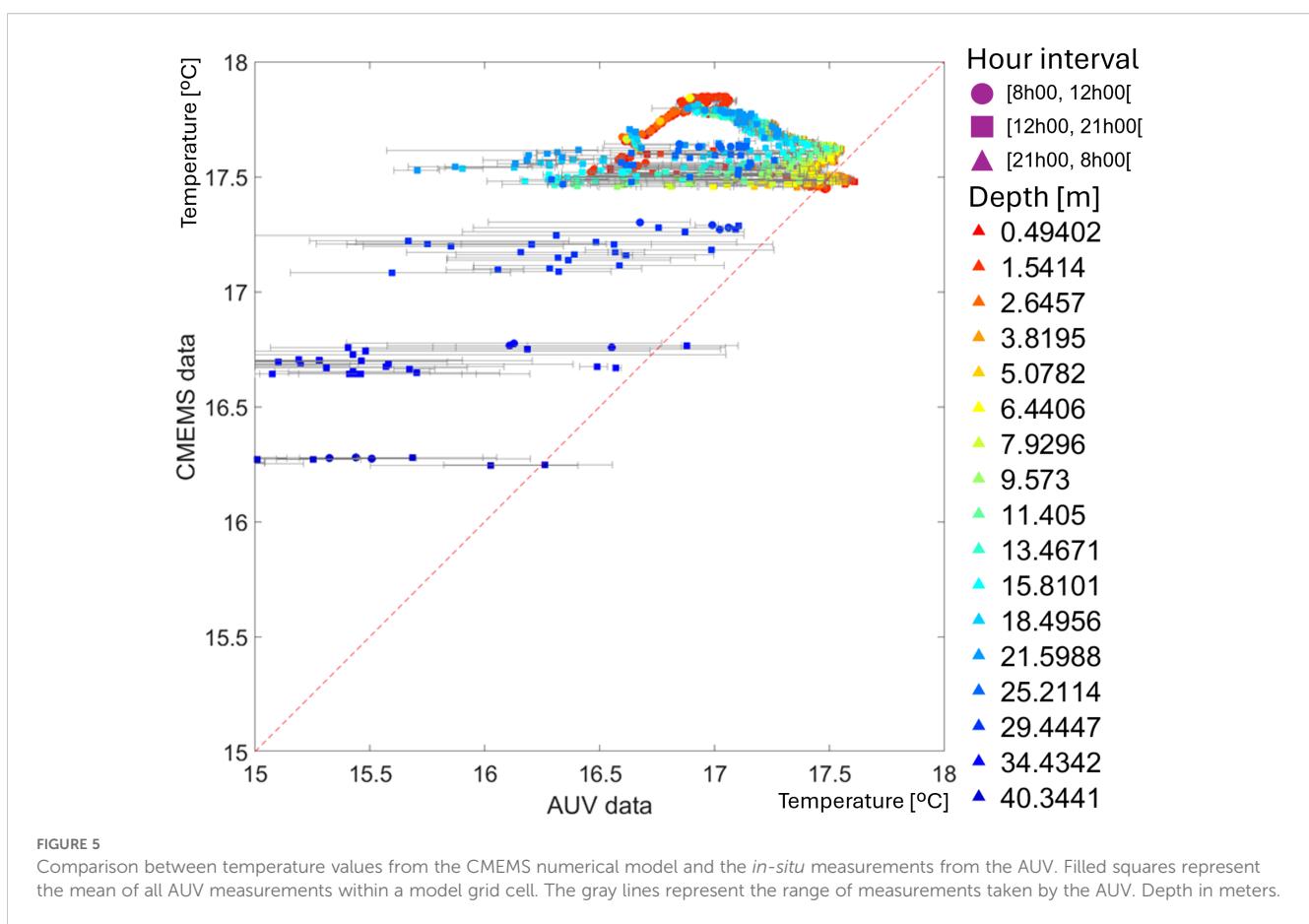
On October 30th, 2024, two different AUVs measured ocean temperature within the study area. One AUV followed a hand-planned navigation path, while the other used an automatically defined path planning over the three-dimensional spatial uncertainty model (Figures 8a, d). As we are not focused on the benefits and limitations of manual versus automatic AUV path planning, we assess the robustness of the data assimilation capabilities of the proposed methodology. Figure 8 shows the match between the direct measurements of both AUV versus the original CMEMS model and the updated one with the information collected on the previous day. The model update allows better predictions with higher match rates between the model and actual

observations, allowing for compensation for the cold waters missing in the original CMEMS model.

4 Discussion

We proposed herein a geostatistical framework with a threefold objective. The first approach is to incorporate uncertainty into short-term predictions by utilizing full numerical models of ocean dynamics as conditioning data. The second objective is to assimilate direct measurements of ocean properties and update the initial ocean model. The third is to be computationally efficient, requiring low computational power aiming at the deployment of the proposed methodology in an autonomous vehicle for real-near-time data assimilation and model update.

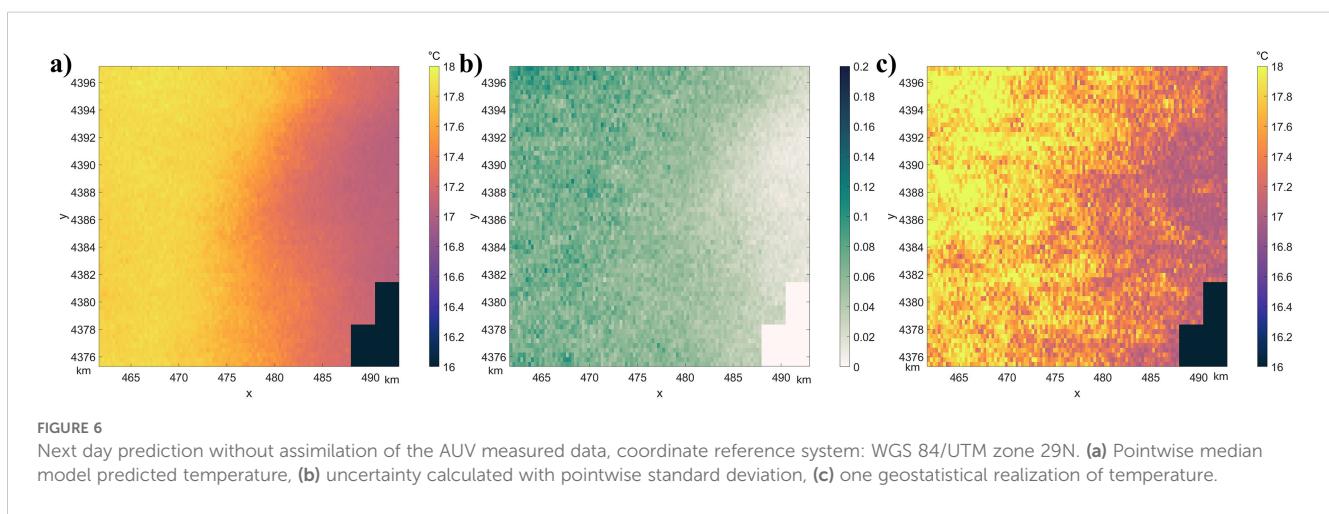
The field application example demonstrates the capability of the predicted spatial uncertainty model to serve as input for an automatic AUV path planner. Areas associated with high spatial uncertainty are those where the calibrated deterministic numerical



model of ocean dynamics exhibits more temporal variability, and therefore, mismatches between actual observations and predictions are expected. The application example illustrates an AUV navigation path through uncertain areas (Figure 4) and an overestimation of ocean temperature (Figure 5). The assimilation of the acquired data results in a change in the spatial uncertainty pattern (Figure 6, Figure 7), illustrating the impact of data assimilation on the prediction. Additionally, the data assimilation produces predictions that are closer to the actual observations

(Figure 8), even for navigation paths that do not depend on the predicted spatial uncertainty.

While the geostatistical approach does not include any physical constraints in terms of oceanographic behavior, its computational costs are negligible when compared with full numerical models of ocean dynamics. The computational cost of a geostatistical realization depends exclusively on the number of grid cells that comprise the model covering area of interest. This aspect opens the door for the deployment of these methods in the AUV's onboard



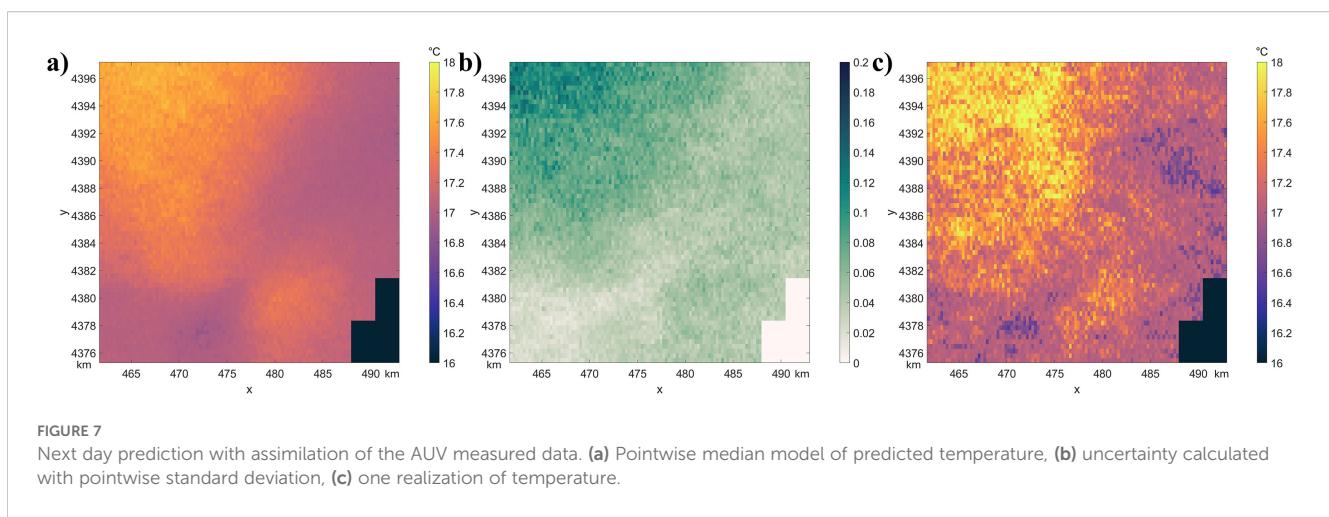


FIGURE 7

Next day prediction with assimilation of the AUV measured data. (a) Pointwise median model of predicted temperature, (b) uncertainty calculated with pointwise standard deviation, (c) one realization of temperature.

processing unit, as well as for automatic data assimilation, model update, and spatial uncertainty prediction in near-real-time and data cycles shorter than those illustrated herein.

On the other hand, the quality of the prediction depends on the quality of the existing *a priori* ocean model and its resolution, as this information will be used as constraining data, as well as the ability to retrieve reliable variogram models from the existing historical data. Using an incorrect variogram model will decrease the quality

of predictions and the usefulness of the proposed methodology, and a lower resolution may cause problems when identifying submesoscale structures. Additionally, the independent prediction at each depth sample is a disadvantage, as it may overlook existing vertical dependencies. This aspect might be addressed with dimensionality reduction techniques as proposed in prediction problems related to geophysics (Azevedo, 2022) or with the development of a full spatiotemporal model. With the proposed

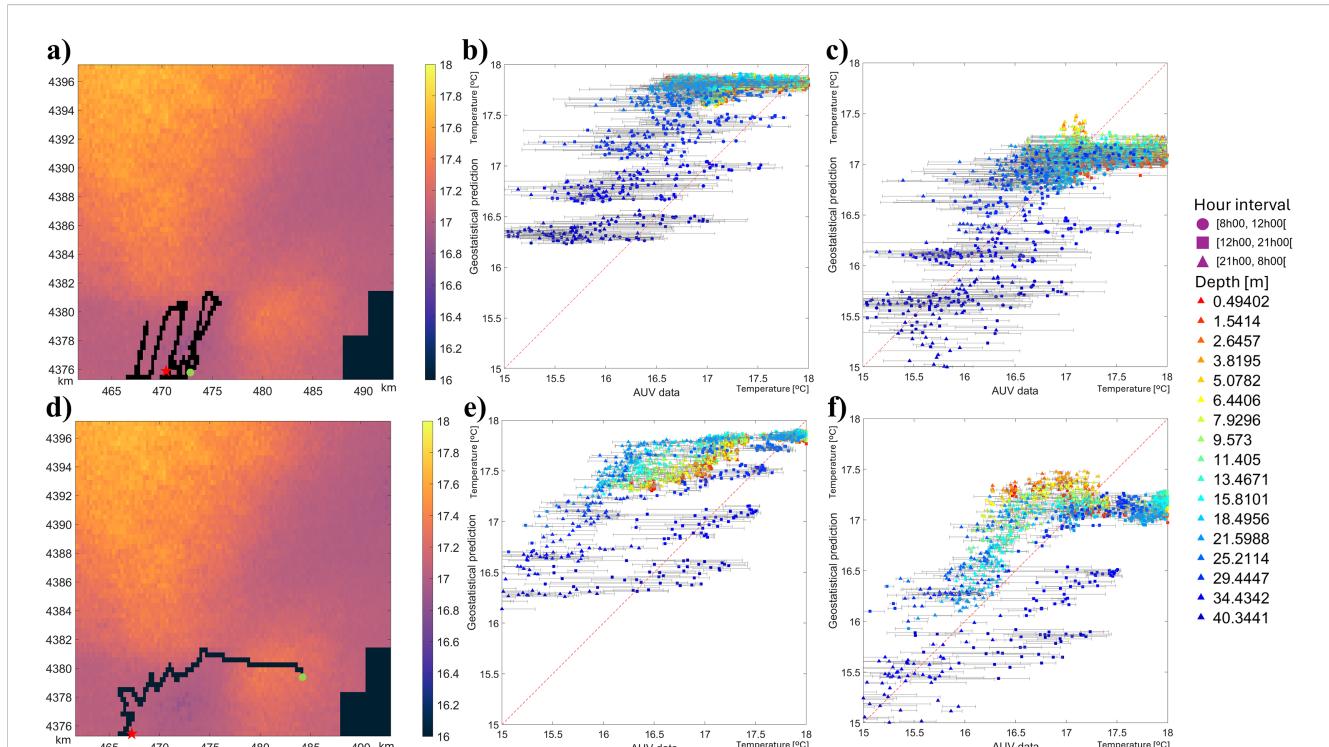


FIGURE 8

(a) LAUV-Xplore 5 path from the 30th of Oct., and geostatistical temperature prediction for the same day after assimilation (green circle representing start and red star the end of the path). (b) Comparison of measurements and geostatistical model prediction for LAUV-Xplore 5 data location. (c) Comparison of measurements and geostatistical model prediction with AUVdata assimilation from LAUV-Xplore 5 data location. (d) LAUV-Xplore 2 path from the 30th of Oct., and geostatistical temperature prediction for the same day after assimilation (green circle representing start and red star the end of the path). (b) comparison of measurements and geostatistical model prediction for LAUV-Xplore 2 data location. (c) comparison of measurements and geostatistical model prediction with AUVdata assimilation from LAUV-Xplore 2 data location.

approach the vertical dependencies are implicitly modelled by the *a priori* information used as data conditioning during the prediction.

Finally, the geostatistical framework enables a flexible approach. While the application example shown herein deals with ocean temperature, the property of interest could be any biogeochemical property if there are historical data for this property in the study area and sensors to measure it in the field. Given its low computational cost, the geostatistical model can also be implemented onboard the AUVs, allowing it to be updated more frequently with *in situ* data and to optimize operations on the fly using as priors to adaptive sampling algorithms.

5 Conclusions

In this work, we introduced a geostatistical modelling approach designed to support smart data collection and efficient trajectory planning for AUVs. By predicting ocean temperature and spatial uncertainty, the method qualifies AUVs to identify and prioritize regions where data collection would be most valuable. This targeted approach increases knowledge while reducing unnecessary exploration, ultimately enhancing the scientific return of oceanographic missions.

Although, this methodology can be applied as is to an oceanographic campaign, it would benefit from some upgrades such as using an *a priori* model with higher resolution, overcoming the necessity of increasing the resolution by interpolation. This will allow to better capture sub-mesoscale structures and interesting features to sample.

Due to the method's ability to work with lightweight computations, making it well suited for real-time decision making in the field, the integration of the algorithm into the AUV during the deployment also brings autonomy to the sampling, and makes use of a more efficient path planning that have previous measurements in consideration and accommodates real-time conditions and unexpected changes in the environment.

Q17 Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://resources.marine.copernicus.eu/product-detail/IBI_ANALYSISFORECAST_PHY_005_001/INFORMATION.

Q18 Author contributions

AD: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. LB: Data curation, Methodology, Writing – review & editing. RM: Conceptualization, Data curation, Investigation, Methodology, Resources, Supervision, Validation, Writing – review & editing. JS: Conceptualization, Data curation, Investigation, Methodology,

Resources, Supervision, Validation, Writing – review & editing. LA: Conceptualization, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

Q15

Q16

The author(s) declare that financial support was received for the research and/or publication of this article. LA and AD gratefully acknowledge the support of the CERENA (FCT/04028/2025). AD is funded by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project PRT/BD/154661/2023. JS and RM acknowledge the support of the LAETA within the scope of the project with reference UIDB/50022/2020. The authors acknowledge the FRESNEL project (Field expeRiments for modEling, aSsimilatioN and adaptive sampLing), funded by the Office of Naval Research (ONR) under ONR award number N00014-22-1-2796. The work was also supported by JUNO—Robotic Exploration of Atlantic Waters project—Ref^a 2021/0008—from FLAD, by the DiverSea project, supported by the European Commission under grant agreement 101082004 and the TRAINEE project (<https://doi.org/10.54499/2024.07606.IACDC>) supported by FCT.

Conflict of interest

Q19

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

1101 1102 References

- 1103
1104 Azevedo, L. (2022). Model reduction in geostatistical seismic inversion with
1105 functional data analysis. *GEOPHYSICS* 87, M1–11. doi: 10.1190/geo2021-0096.1
- 1106 **Q20** Bernacchi, L., Duarte, A. F., Mendes, R., Azevedo, L., and Sousa, J. B. D. (2025).
1107 Closing the data loop: real-world AUVs adaptive sampling for improved ocean model
1108 predictions. doi: 10.1109/MARIS64137.2025.11139688
- 1109 Deutsch, C. V., and Journel, A. G. (1998). *GSLIB: geostatistical software library and*
1110 *user's guide. Version 2.0. 2nd ed* (New York: Oxford University Press), 1.
- 1111 **Q21** European Union-Copernicus Marine Service (2017). *Atlantic-Iberian Biscay Irish-*
1112 *Ocean Physics Analysis and Forecast* (Mercator Ocean International). Available online
1113 at: https://resources.marine.copernicus.eu/product-detail/IBI_ANALYSISFORECAST_PHY_005_001/INFORMATION.
- 1114 Gafurov, S. A., and Klochkov, E. V. (2015). Autonomous unmanned underwater
1115 vehicles development tendencies. *Proc. Engineering*. 106, 141–148. doi: 10.1016/j.proeng.2015.06.017
- 1116 Ge, Y., Eidsvik, J., and Mo-Bjørkelund, T. (2023). 3-D adaptive AUV sampling for
1117 classification of water masses. *IEEE J. Oceanic Eng.* 48, 626–639. doi: 10.1109/
1118 JOE.2023.3252641
- 1119 Gould, J., Sloyan, B., and Visbeck, M. (2013). “*In Situ* Ocean Observations,” in
1120 *International Geophysics* (Elsevier), 59–81. Available online at: <https://linkinghub.elsevier.com/retrieve/pii/B9780123918512000039>.
- 1121 Hwang, J., Bose, N., and Fan, S. (2019). AUV adaptive sampling methods: A review.
1122 *Appl. Sci.* 9, 3145. doi: 10.3390/app9153145
- 1123 Kite-Powell, H., Colgan, C., and Weiher, R. (2008). Estimating the economic benefits
1124 of regional ocean observing systems. *Coast. Management*. 36, 125–145. doi: 10.1080/
1125 08920750701868002
- 1126 Lermusiaux, P. F. J. (2007). Adaptive modeling, adaptive data assimilation and
1127 adaptive sampling. *Physica D: Nonlinear Phenomena*. 230, 172–196. doi: 10.1016/
1128 j.physd.2007.02.014
- 1129 Mahdavi, S., Amani, M., Bullock, T., and Beale, S. (2021). A probability-based
1130 daytime algorithm for sea fog detection using GOES-16 imagery. *IEEE J. Sel. Top. Appl.*
1131 *Earth Observations Remote Sensing*. 14, 1363–1373. doi: 10.1109/JSTARS.2020.3036815
- 1132 Mendes, R., Da Silva, J. C. B., Magalhaes, J. M., St-Denis, B., Bourgault, D., Pinto, J.,
1133 et al. (2021). On the generation of internal waves by river plumes in subcritical initial
1134 conditions. *Sci. Rep.* 11, 1963. doi: 10.1038/s41598-021-81464-5
- 1135 Minnett, P. J., Alvera-Azcárate, A., Chin, T. M., Corlett, G. K., Gentemann, C. L.,
1136 Karagali, I., et al. (2019). Half a century of satellite remote sensing of sea-surface
1137 temperature. *Remote Sens. Environment*. 233, 111366. doi: 10.1016/j.rse.2019.111366
- 1138 Mo-Bjørkelund, T., Mendes, R., López-Castejón, F., and Ludvigsen, M. (2025).
1139 Adaptive ocean gradient tracking using an autonomous underwater vehicle with a
1140 boundless model. *IEEE J. Oceanic Eng.* 50, 955–967. doi: 10.1109/JOE.2024.3484577
- 1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
- 1156 O’Carroll, A. G., Armstrong, E. M., Beggs, H. M., Bouali, M., Casey, K. S., Corlett, G.,
1157 K., et al. (2019). Observational needs of sea surface temperature. *Front. Mar. Sci.* 6, 420.
1158 doi: 10.3389/fmars.2019.00420
- 1159 Paull, L., SaeediGharahbolagh, S., Seto, M., and Li, H. (2012). “Sensor driven online
1160 coverage planning for autonomous underwater vehicles,” in *2012 IEEE/RSJ International*
1161 *Conference on Intelligent Robots and Systems* (IEEE, Vilamoura-Algarve, Portugal), 2875–2880. Available online at: <http://ieeexplore.ieee.org/document/6385838/>
- 1162 Petillo, S., Balasuriya, A., and Schmidt, H. (2010). “Autonomous adaptive
1163 environmental assessment and feature tracking via autonomous underwater
1164 vehicles,” in *Oceans’10 Ieee Sydney* (IEEE, Sydney, Australia), 1–9. Available online
1165 at: <http://ieeexplore.ieee.org/document/5603513/>.
- 1166 Pinto, J., Mendes, R., Da Silva, J. C. B., Dias, J. M., and De Sousa, J. B. (2018).
1167 “Multiple Autonomous Vehicles Applied to Plume Detection and Tracking,” in *2018*
1168 *OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)* (IEEE, Kobe), 1–6. Available online
1169 at: <https://ieeexplore.ieee.org/document/8558802/>.
- 1170 Sloyan, B. M., Wilkin, J., Hill, K. L., Chidichimo, M. P., Cronin, M. F., Johannessen, J. A.,
1171 et al. (2019). Evolving the physical global ocean observing system for research and
1172 application services through international coordination. *Front. Mar. Sci.* 6, 449.
1173 doi: 10.3389/fmars.2019.00449
- 1174 Soares, A. (2001). Direct sequential simulation and cosimulation. *Math. Geology*. 33,
1175 911–926. doi: 10.1023/A:1012246006212
- 1176 Teixeira, D., Sousa, J. B. D., Mendes, R., and Fonseca, J. (2021). “3D Tracking of a
1177 River Plume Front with an AUV,” in *OCEANS 2021: San Diego – Porto* (IEEE, San
1178 Diego, CA, USA), 1–9. Available online at: <https://ieeexplore.ieee.org/document/9705995/>.
- 1179 (2002). The vehicle routing problem [Internet]. Society for industrial and applied
1180 mathematics. doi: 10.1137/1.9780987181515
- 1181 P. Toth and D. Vigo (Eds.) (2014). “Front Matter,” in *Vehicle Routing: Problems,
1182 Methods, and Applications, 2nd ed.* (Society for Industrial and Applied Mathematics,
1183 Philadelphia, PA), i–xviii. doi: 10.1137/1.9781611973594.fm
- 1184 Yu, F., He, B., Liu, J., Wang, Q., and Shen, Y. (2022). Towards autonomous
1185 underwater vehicles in the ocean survey: A mission management system (MMS).
1186 *Ocean Engineering*. 263, 111955. doi: 10.1016/j.oceaneng.2022.111955
- 1187 Zhang, Y., Godin, M. A., Bellingham, J. G., and Ryan, J. P. (2012). Using an
1188 autonomous underwater vehicle to track a coastal upwelling front. *IEEE J. Oceanic Eng.*
1189 37, 338–347. doi: 10.1109/JOE.2012.2197272