



Research paper

Efficient data-driven flow modeling for accurate passive scalar advection in submesoscale domains

Karlo Jakac^a ,* , Luka Lanča^a , Ante Sikirica^b , Stefan Ivić^a 

^a Faculty of Engineering, University of Rijeka, Vukovarska 58, Rijeka, 51000, Croatia

^b Center for Advanced Computing and Modelling, University of Rijeka, Radmile Matejčić 2, Rijeka, 51000, Croatia



ARTICLE INFO

Dataset link: <https://osf.io/wjsb2/>

Keywords:

Velocity field reconstruction
Optimization
Global positioning system drifters
Smart sensors
Scattered measurements

ABSTRACT

Knowing the sea surface velocity field is important for various applications, such as search and rescue operations, where predicting the movement of objects or substances is critical. However, achieving an accurate estimation of these advection processes is challenging, even with modern measuring equipment, such as high-frequency radar or advanced simulations based on oceanic flow models. Therefore, this paper presents a data-driven framework to approximate sea surface velocity from spatially distributed observations, thus enabling efficient probability advection modeling across submesoscale domains. The system uses quasi-steady flow assumptions to approximate transient flows. To overcome the limitations of point measurements in capturing domain-wide circulation, the method employs a fusion of two simplified 2D flow models to approximate submesoscale dynamics, enabling complete velocity field reconstruction from scattered data. To ensure reliable flow dynamics, the approach iteratively adjusts boundary conditions in numerical simulations to align the simulated flow with observations. Experimental validation was conducted in Kvarner Bay using Global Positioning System (GPS) drifters. The results confirmed the system's ability to replace computationally intensive transient simulations by approximating flow fields using model simplifications. The results demonstrate its efficiency in various cases, making it a practical tool for real-life submesoscale applications requiring swift passive scalar advection.

1. Introduction

The continuous growth of the marine economy and increased traffic bring about some negative aspects, including a higher frequency of accidents and the potential for environmental disasters. Hazardous material pollution is one of the most frequent examples, leading to substantial economic and ecological damage annually. The time required for recovery varies depending on factors such as the severity, composition, and the affected ecosystem, with some cases taking decades for full restoration (Elliott and Jones, 2000). The financial impact of incidents like oil spills can reach billions of dollars, posing a significant threat to marine habitats by contaminating the food web and polluting vast coastal areas (Carson et al., 2003). For instance, the Deepwater Horizon explosion in the Gulf of Mexico caused devastating environmental damage and multiple casualties (Mariano et al., 2010). Similarly, the Sanchi oil spill, caused by the collision of the Sanchi oil tanker with a cargo vessel in January 2018, remains one of the most serious and polluting tanker accidents of the 21st century (Chen et al., 2020). Given the serious impact of such incidents, effective emergency preparedness and response for pollution spreading has become a global

priority to mitigate damage, safeguard lives, and minimize economic losses (García-Garrido et al., 2016).

In recent decades, the lack of reliable predictions for the spread of pollution has made it challenging to protect ecosystems and the economy from the consequences of environmental disasters. The increasing frequency of such incidents has raised concerns within the scientific community about the urgent need for accurate and reliable models to forecast the progression of maritime disasters, providing crucial decision-making support during emergencies (Reed et al., 1999; Kirby and Law, 2010; Wang and Shen, 2010; Olascoaga and Haller, 2012). Various numerical models are commonly used to predict the movement of pollutants or their concentration in the water. The results from these transport models are crucial, as they are often the only available approach. However, there are various prediction models with different complexities in calculating pollution movement. The effectiveness of each model depends on its structure, applied methods, and, most importantly, the accuracy of input data (e.g., sea currents, wind, and source location), followed by result interpretation. Given potential input errors, these parameter uncertainties must be carefully considered.

* Corresponding author.

E-mail addresses: karlo.jakac@uniri.hr (K. Jakac), luka.lanca@uniri.hr (L. Lanča), ante.sikirica@uniri.hr (A. Sikirica), stefan.ivic@uniri.hr (S. Ivić).

Generally, the primary sources of uncertainty in pollution predictions and Search and Rescue (SAR) operations are the inherent complexities of ocean dynamics (Sayol et al., 2014). Furthermore, the ocean exhibits chaotic behavior. Its motion is often approximated using the nonlinear Navier–Stokes equations. As a result, even small variations in initial conditions can lead to significant deviations in forecast fields over time. Additional uncertainties arise from the drifting dynamics model, as an object's motion at sea is influenced by the ratio of its surface area exposed to wind and water, with the wind drag coefficient typically determined empirically through trial and error. The accuracy of oceanic environmental data also poses a challenge, as operational ocean and atmospheric forecasting models typically do not explicitly resolve fine-scale motions; instead, these are represented through sub-grid-scale closure schemes or require grid refinement, both of which increase model complexity and computational demand. These factors, combined with incorrect initial conditions during the early stages of an incident, contribute to the inherent uncertainty in ocean drift predictions and further affect forecasting accuracy. As a result, numerous studies have proposed various prediction schemes and developed drift prediction models to address these uncertainties. One of the most widely used approaches is the probabilistic model based on the leeway dynamics model (Allan and Plourde, 1999; Breivik and Allen, 2008), which helps describe uncertainties in the drifting trajectory of objects. A notable advancement was made by Li et al. (2019), who developed numerical simulation scenarios that incorporated different current and wind forcing data into the forecasting system and varied oil release times to identify key error sources. Additionally, several studies (Qazi et al., 2014; Lorente et al., 2016; Paduan and Washburn, 2013) have integrated coastal radar current data for computing ocean surface currents and for the possible spread of pollutants.

Still, all prediction methods rely on sea surface velocity measurements from various sources. Due to its cost-effectiveness and reliability, satellite tracking of deployed drifters is one of the most widely used methods. In the past two decades, the use of drifters for monitoring purposes has grown considerably, with numerous deployments in various oceanic regions (Haza et al., 2018). These floating sensors, which are designed to collect surface current data, have been extensively studied, and the resulting data have been analyzed in numerous publications (Chaturvedi et al., 2020; Gulakaram et al., 2018). Similarly, satellite-tracked drifters have been used in semi-enclosed seas, like the Adriatic Sea, to examine circulation patterns (Poulain, 2001; Ursella et al., 2006). Equipped with Global Positioning System (GPS), General Packet Radio Service (GPRS), and Very High Frequency (VHF) signals, these drifters offer valuable insights into ocean circulation. Unfortunately, they can drift outside the target areas, resulting in scattered measurements across vast regions. Additionally, even though drifter data can be useful, it does not provide a complete picture of the circulation across the entire domain, as it only offers point-based measurements.

In order to enhance model accuracy and incorporate Lagrangian data more effectively, two primary methods have been explored for reconstructing and integrating Lagrangian data into models. The first method estimates velocities by calculating the ratio of observed position changes over time, and then uses these velocities to adjust model predictions (Hernandez et al., 1995). The second approach employs an observational operator based on the particle advection equation, optimizing the Eulerian velocity field by minimizing discrepancies between observed trajectories and model results (Molcard et al., 2003). Earlier studies (Bennett and Clites, 1987) indicated that Euler methods could introduce trajectory errors in non-uniform flow fields. Therefore, modern pollutant prediction models rely on more physically informed formulations. This approach maintains a consistent Eulerian framework, calculating slick thickness using layer-averaged Navier–Stokes equations and simulating pollution dynamics with the advection-diffusion equation (Tkalic, 2006).

A key challenge with Lagrangian drifter data is that the drifters move with ocean currents, leading to an uneven distribution, with drifters often clustering or drifting out of areas of interest. While Lagrangian data provides trajectories over space and time, research often shifts from trajectory analysis to reconstructing Eulerian velocity fields, with many studies utilizing drifter data for this purpose (Rao and Schwab, 1981; Eremeev et al., 1992; Cho et al., 1998). While satellite measurements are valuable, they face challenges such as limited availability and occasional atmospheric interference. As a result, high-frequency (HF) radar has been used for near real-time surface velocity measurements, often serving to validate ocean current models (Solano et al., 2018; Marmain et al., 2014). Still, HF radar has its limitations, such as shallow depth penetration, vulnerability to interference, difficulties in achieving high spatial resolution, and the need for substantial financial investment.

To overcome most of these challenges, machine learning (ML) has emerged as a promising alternative to measurement interpolation and optimization methods for flow field predictions (Ghalambaz et al., 2024). Studies such as Grossi et al. (2020) and Song et al. (2024) have utilized artificial neural networks (ANNs) to analyze temporal patterns in drifter trajectories and predict long-term movement, reducing errors in drifter models. Additionally, deep learning (DL) has been applied to surrogate modeling of fluid flows (Sun et al., 2020; Tang et al., 2021), enabling fast and efficient estimations without the reliance on extensive computational fluid dynamics (CFD) simulations.

Given the complexity of reconstructing surface flow and modeling pollution spread, we adopt an approach that combines sea surface velocity approximation with a passive scalar formulation. This general framework can be extended to incorporate additional processes, such as decay, reaction, or source terms, if more complex physical behavior is needed, offering a significant advantage for rapid response scenarios. In this work, we propose a methodology for estimating the sea surface velocity field over an entire domain using only sparse and scattered drifter data, while preserving the spatial complexity of the surface flow. The method integrates two simplified two-dimensional models representing bounded and open domains to form a surrogate capable of replicating submesoscale flow patterns and scalar transport. The bounded domain simulates the realistic region of interest, including coastlines and inlet/outlet boundaries, while the open domain – modeled as a circular zone – captures external flow influences. These surrogates are solved independently and then fused, enabling the simplified 2D model to represent broader hydrodynamic behavior. This surrogate-based approach significantly improves computational efficiency by reducing the need for dense measurement coverage and enabling short simulation times. To maintain real-time applicability, the initial flow model omits external drivers such as wind, waves, tides, and temperature, which are subsequently integrated through a secondary surrogate. This hybridization compensates for initial simplifications and provides sufficiently accurate flow dynamics to support passive scalar transport simulations.

The paper is organized as follows: Section 1 introduces the motivation for tracking and predicting passive scalar fields, along with a review of related work. Section 2 explains the two-dimensional fusion flow model. Section 3 presents the formulation of the model fitting problem. In Section 4, the optimization procedure is described, including initialization and the implementation of an adaptive diffusion coefficient. Section 5 covers the acquisition of measurements for flow reconstruction and the prediction of passive scalar field movement. Section 6 assesses the accuracy of passive scalar field advection. Section 7 discusses the results from synthetic cases and the provided experiment. Section 8 highlights the limitations and provides a discussion. Finally, Section 9 concludes the paper.

2. Simplified two-dimensional flow model

Simulating transient ocean behavior requires substantial computational resources and extended simulation times. To address this, we

adopt a fused flow model based on steady-state, incompressible flow. This approach deliberately omits dynamic factors like wind, waves, tidal fluctuations, and temperature changes, focusing on a fast yet sufficiently accurate approximation of sea surface flow. However, these environmental influences are accounted for through a fusion approach.

This simplified methodology offers an efficient alternative to approaches dependent on large-scale data, as it avoids the need for extensive datasets, complex interpolation procedures, and detailed domain models, thereby substantially reducing computational demands and processing time. By leveraging CFD simulations, the model captures essential fluid flow physics, enabling the calculation of accurate velocity fields and supporting advection/diffusion processes for passive scalar distribution.

The proposed approach is particularly useful in scenarios that require rapid computation of passive scalar field distributions, such as pollutant dispersion or object tracking. The model's design leverages the quasi-steady nature of oceanic flows – where changes are gradual rather than abrupt – making it an effective tool for real-time or near-real-time applications. It strikes a balance between computational efficiency and the ability to capture transient-like flow dynamics, offering a practical solution for rapid response scenarios without the complexity of fully transient simulations.

2.1. Steady-state 2D flow model

The proposed steady-state flow model governs the movement of fluids at low to medium velocities within a connected computational domain $\Omega \subset \mathbb{R}^2$, utilizing the incompressible Navier–Stokes equation for steady flow (Gunzburger, 2012; Lions, 1996; Kronbichler, 2009):

$$\rho(\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla p + \mu \nabla^2 \mathbf{u} + \rho \mathbf{f} \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (2)$$

The vector \mathbf{u} represents the velocity of the fluid, while p denotes its dynamic pressure. Given that the approach models two-dimensional surface flow, it assumes hydrostatic balance in the vertical direction. As a result, the surface pressure is primarily represented by the dynamic pressure component. This assumption is common in surface flow approximations but may limit the model's applicability in scenarios where vertical stratification or pressure gradients significantly influence horizontal dynamics. In this context, ρ stands for the fluid density, and μ is the fluid's dynamic viscosity. The assumption of incompressibility implies that density remains constant. The term \mathbf{f} accounts for external forces acting on the fluid.

To accurately represent the interactions between the observed section of the sea and the surrounding marine environment, a specific set of boundary conditions i.e. tangential velocities and pressure values, are prescribed. The proposed combination must not be randomly assigned; their range should align with realistic conditions to replicate the physical characteristics of actual flow.

In submesoscale oceanic regions such as the Adriatic Sea, which is in the focus of this study, surface currents exhibit significant variability due to interactions between mesoscale and smaller-scale processes. Surface velocities in the Adriatic Sea, derived from radar data, infrared satellite imagery, and numerical simulations, range from below 0.1 m/s to above 0.5 m/s (Cosoli et al., 2013; Notarstefano et al., 2008; Bolaños et al., 2014). The average surface velocities in most areas are typically between 0.1 and 0.2 m/s. Although stronger currents exceeding 0.6–1.0 m/s are known to occur in regions such as the Strait of Otranto (Gucel and Sakalli, 2024), this study focuses primarily on the more coastal and semi-enclosed basins that characterize much of the Adriatic Sea.

We formulated the problem as a pressure-driven flow, which is suitable when exact inlet and outlet regions of the boundary are not known. Specifically, a zero gradient condition for the velocity is assigned

when fluid is flowing out of the domain, while inlet velocity is calculated from the flux in the patch-normal direction. Combination of these boundary conditions require setting velocity component tangential to the boundary \mathbf{u}_t , as well as the total pressure p_0 along the boundary where fluid interacts with the rest of the sea. The connection between pressure, total pressure, and velocity does depend on the direction of the flow, and it is adjusted according to:

$$p = \begin{cases} p_0 & \text{for outlet} \\ p_0 - 0.5\rho|\mathbf{u}|^2 & \text{for inlet.} \end{cases} \quad (3)$$

The combination of these boundary conditions ensures a robust simulation and flexibility to achieve a variation of possible inlet/outlet regions and intensities.

After setting the tangential velocities and total pressure values, the next step is to establish the velocity and pressure profile functions. This process involves interpolating the tangential velocity and pressure values along the boundary while ensuring that the values at the coastline edges are set to zero. The procedure is updated during the optimization process until the boundary condition generates a velocity field that aligns with the velocity measurements from the drifter.

2.2. Turbulence

Turbulence is modeled using the k - ω shear stress transport model (Ferziger and Peric, 2012), which combines the benefits of both k - ω and k - ϵ models to improve the precision and robustness in simulating complex turbulent flows. This hybrid approach splits the flow domain into two regions: the near-wall and outer regions. For the near-wall region, a wall function is used to accurately capture turbulence near the wall. In the outer region, the model behaves like a free-stream model, ensuring reliable turbulence predictions away from the wall (Menter, 1992). The turbulence variables (k , ω) are determined using:

$$k = \frac{3}{2}(|\mathbf{u}| I)^2 \quad (4)$$

$$\omega = \frac{k^{0.5}}{C_\mu^{0.25} L} \quad (5)$$

where k is turbulence kinetic energy, I is turbulence intensity, ω is the specific dissipation rate, C_μ is a turbulence model constant and equals 0.09 while L is the turbulent length scale.

Given the importance of real-time capabilities in the system, a balance between computational speed and accuracy must be achieved. The use of an appropriate turbulence model facilitates improved convergence, thereby enhancing both the numerical stability and responsiveness of the simulation which are key factors for achieving real-time performance.

2.3. Fusion model approach

Simulating the realistic movement of a passive scalar within a specific domain, where environmental factors play a significant role, is often computationally demanding and challenging to implement using conventional CFD approaches. To overcome these limitations, a fusion based approach that combines two two-dimensional flow surrogates is adopted. The resulting fusion model serves as a surrogate for submesoscale dynamics and incorporates a correction offset to partially account for flow simplifications.

As noted, the fusion approach combines two distinct domains. The bounded domain is defined for the realistic region of interest and incorporates all relevant domain elements i.e. coastline and inlets/outlets. The open domain, which encompasses the bounded domain, is defined as a fully open circular zone in which external factors act. Both domain flows are solved separately and subsequently combined, allowing the simplified two-dimensional flow surrogate to partially account for broader flow dynamics (Fig. 1).

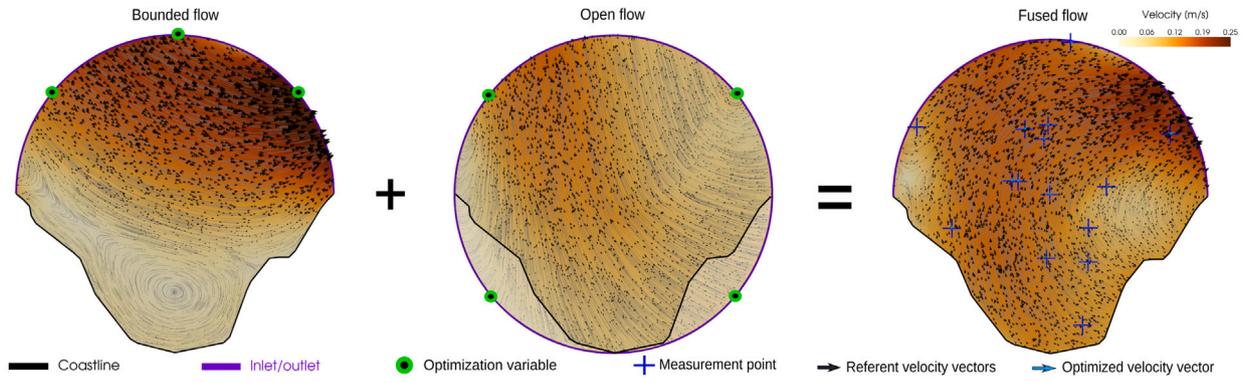


Fig. 1. The figure presents an overview of the fusion model methodology. The left side of the figure shows the bounded domain, which includes the coastline and defines the primary area of interest. The center illustrates the open domain that surrounds the bounded area but does not explicitly model the coastline. This domain accounts for broader influences, enabling the simulation to incorporate realistic conditions that affect surface flow. The right side of the figure presents the fused velocity field, obtained by summing the velocity fields from both simulations for overlapping region. In the overlapping region (seen on center of figure), velocity values from the two domains are interpolated and summed at each grid point within the bounded domain to ensure continuity and a smooth transition between the simulations.

The open domain uses four control points (eight additional variables for the optimization vector) to capture environmental variations, which are incorporated into the flow approximation to enhance the overall accuracy for the entire domain. By integrating velocity fields from both the bounded domain and open domain simulations, the open domain adjusts the flow field to better reflect real-world conditions. To achieve this superposition, velocity values at the node locations within the bounded domain are first extracted. These corresponding velocity values are then retrieved at the same locations in the open domain, ensuring alignment between the two fields. Finally, the velocity fields are summed, effectively incorporating corrective adjustments into the overall flow representation:

$$\mathbf{u}_{fused} = \mathbf{u}_{bounded} + \mathbf{u}_{open} \quad (6)$$

where $\mathbf{u}_{bounded}$ and \mathbf{u}_{open} are velocity fields obtained for bounded and open flow models, respectively.

By summing the velocity fields from both simulations, a more dynamic and accurate flow model is obtained, reflecting the complex environmental interactions within the domain. In real-world scenarios, certain unrealistic flow patterns can appear near the coastline. Such observations have been reported in several HF radar surface current studies (e.g., Bellomo et al., 2015; Kovačević et al., 2004; Berta et al., 2014), where surface currents occasionally seem to originate from land. Such behavior cannot be reproduced using standard two-dimensional CFD models, which lack the ability to account for intricate three-dimensional coastal effects. Since our method focuses on reconstructing surface flows from observed data, the fusion model is specifically designed to accommodate both realistic and anomalous patterns seen in measurements. This allows for adaptive, real-time updates to the flow field, enhancing the accuracy of passive scalar reconstruction. By incorporating these environmental influences, the simulation becomes more responsive to dynamic conditions, offering a closer approximation to real-world marine scenarios.

2.4. Passive scalar transport

Along with the flow modeling, a scalar transport object was incorporated into the simulation to govern the advection and diffusion of the passive scalar field at each iteration over discrete time incremented by period T_s . The passive scalar field is advected on the velocity field of fused flow (6) as described by the following scalar transport equation:

$$\frac{\partial s}{\partial t} + \mathbf{u}_{fused} \cdot \nabla s - D \nabla^2 s = 0 \quad (7)$$

where s is the scalar field, \mathbf{u}_{fused} is the velocity field, t represents time and D is the diffusion coefficient. The diffusion coefficient is determined using:

$$D = \alpha_1 v + \alpha_2 v_t \quad (8)$$

where α_1 and α_2 are constants and are equal to 10^{-9} and 0.714 respectively. v_t represents turbulent viscosity. Kinematic viscosity ν is set to $10^{-6} \text{ m}^2/\text{s}$.

This scalar transport equation is crucial for replicating the realistic movement of the passive scalar field throughout the domain, ensuring that the behavior of the scalar field is appropriately modeled in relation to the evolving flow dynamics. This combined approach allows for the accurate representation of both the fluid flow and the transport of the passive scalar field, ensuring the robustness of the transient flow simulation.

2.5. Numerical implementation

The computational domains in the considered test cases are designed to be either synthetic, for controlled experiments, or to represent a realistic geographical region for practical applications. For realistic regions, data acquisition is performed using Sentinel Hub's polygon extraction tool, accessible via [Sentinel Hub](#). The Sentinel Hub L2A NDWI (Normalized Difference Water Index) visualization is utilized, focusing on NDWI TIFF images of the specified region. These images enable precise identification of water bodies and coastlines.

To define the domain boundaries, polygons corresponding to the region of interest are created. The process involves extracting coastline points based on NDWI indices, which are then refined and smoothed to ensure accuracy. Following the polygon extraction, a stereolithographic (STL) model is created, which serves as a geometric baseline to create a computational mesh.

Subsequently, a two-dimensional numerical mesh is generated using the cfMesh library (Juretic, 2015). This mesh is utilized to define a numerical problem using the open-source CFD toolkit OpenFOAM (Foundation, 2016). This workflow ensures precise representation of geographical features and enables the simulation of realistic fluid dynamics within the specified domain.

Velocity at the coastline is defined using a no-slip (Dirichlet) boundary condition. When the fluid exits the domain at a boundary face, a Neumann boundary condition is applied to the velocity, i.e. the fluid velocity at the boundary is extrapolated from the velocity inside the domain. For the fluid entering the domain, the open sea boundary is set to a Dirichlet condition, with the velocity computed based on the flux in the patch-normal direction. Additionally, tangential velocities are specified, as the incoming flow may not be perfectly aligned with the inlet boundaries. Defining the tangential velocity allows for more realistic inlet flow conditions, accounting for the potential swirl or tangential motion of the fluid.

The pressure at the open sea boundary is modeled using a Dirichlet boundary condition with a specified range of values, while at the

Table 1
An overview of the boundary conditions employed.

Field	Inlet/Outlet	Coastline
u	pressureInletOutletVelocity	noSlip
p	totalPressure	zeroGradient
k	fixedValue	kqRWallFunction
ω	fixedValue	omegaWallFunction

coastline, a Neumann boundary condition is applied. To maintain a consistent pressure field within the domain, a reference cell is selected and assigned a pressure value of zero. This reference point is then used to compute the pressure gradient across the entire domain.

Initial values for turbulent kinetic energy and specific rate of dissipation are estimated according to the defined equations. For walls, appropriate wall functions are used. A summary of the boundary conditions for all test cases is provided in Table 1.

To model the quasi-steady flow in this study, we opted for the simpleFoam steady-flow solver within OpenFOAM, performing multiple short steady-flow simulations to approximate transient behavior. A variant of the semi-implicit method for pressure-linked equations (SIMPLE) algorithm (Patankar and Spalding, 1983), SIMPLEC (SIMPLE-consistent), was used. For turbulence modeling, the k - ω shear stress transport model is employed to accurately capture the flow characteristics.

In the OpenFOAM simulations conducted for steady-state flows, second-order accurate gradient and Laplacian schemes were used. To enhance stability in areas with steep gradients, a first-order accurate scheme (upwind) was applied. This is also true for divergence terms associated with convective transport. The meshWave method was used to compute distances to the nearest wall. The boundary conditions and simulation setup remained consistent across all test cases.

Further details on the numerical implementation for both bounded and open domain cases, including information on the numerical grid, cell distribution, discretization, and modeling schemes, are available in The Open Science Framework repository: <https://osf.io/wjsb2/>

3. Model fitting problem formulation

The proposed methodology employs an optimization algorithm that iteratively refines the values in the optimization vector to reduce the error, i.e., to minimize the discrepancy between the simulated flow field and the reference flow. Consequently, the accuracy and reliability of the flow field reconstruction are highly dependent on the optimization vector \mathbf{b} , which includes the tangential velocity and pressure values at the boundary control points:

$$\mathbf{b} = (\mathbf{u}_{t,1}, p_1, \dots, \mathbf{u}_{t,n_{CP}}, p_{n_{CP}})^T \quad (9)$$

where n_{CP} represents the number of boundary control points.

To capture realistic surface flows and address the high variability of surface currents, particularly in submesoscale domains, the optimization variable bounds are set to -0.5 to 0.5 m/s for tangential velocity and -0.05 to 0.05 m²/s² for pressure at the boundary control points, as discussed in 2.1. These bounds are defined for numerical stability, with initial optimization candidates randomly distributed within this range, as they do not significantly affect the qualitative outcomes. It is worth noting that the final values for total pressure and tangential velocity at the boundary may deviate slightly from the assigned values due to adjustments based on the results of the Navier–Stokes equations within the internal domain. After defining boundary control point values, cubic spline interpolation is used to compute velocity and pressure values for each cell along the boundary. During optimization, the velocity profile is iteratively updated as the optimization vector evolves until the output error aligns with the target fitness.

3.1. Objectives

For each evaluation, a complete OpenFOAM case is created to compute the velocity field across the entire domain. The velocity vectors at the measurement point coordinates, representing drifter positions, are then extracted and treated as reference values to match at those specific locations. To quantify the deviation, the cost function calculates the drifter error, E_d , which aggregates the discrepancies at these discrete points (drifters) in meters per second, as defined by the following equation:

$$E_d(\mathbf{b}) = \frac{1}{n_{MP}} \sum_{i=1}^{n_{MP}} (\mathbf{u}_{r_i} - \mathbf{u}_{s_i}(\mathbf{b}))^2 \quad (10)$$

where n_{MP} represents the number of measurement points, \mathbf{u}_{r_i} denotes the referent velocity vector, while \mathbf{u}_{s_i} is the simulation velocity vector at the measurement point location s_i .

3.2. Constraints

To obtain feasible solutions within this simulation-based optimization framework, it is crucial to establish suitable constraints. Constraints are tied to the simulation residuals, which guide the optimization process toward achieving the desired fitness and ensuring numerically stable results. When the residuals are below the specified thresholds, the constraints are deemed satisfied. If the residuals exceed the limits, a penalty is imposed for failing to meet the criteria.

The pressure residual constraint, given by:

$$r_p(\mathbf{b}) \leq 10^{-3} \quad (11)$$

helps ensure consistent pressure values during the optimization process, preventing pressure imbalances that could result in unrealistic behavior. The velocity residuals ensure controlled and physically realistic fluid motion and are defined for both velocity components:

$$r_{u_x}(\mathbf{b}) \leq 10^{-4} \quad (12)$$

$$r_{u_y}(\mathbf{b}) \leq 10^{-4}. \quad (13)$$

The k residual constraint, defined as:

$$r_k(\mathbf{b}) \leq 10^{-4} \quad (14)$$

ensures that the turbulent kinetic energy remains within acceptable bounds. The ω residual constraint, set as:

$$r_\omega(\mathbf{b}) \leq 10^{-4} \quad (15)$$

limits the specific dissipation rate of turbulence, maintaining alignment with the fundamental physics. These constraints collectively guide the optimization process, ensuring physical realism, stability, and relevant fluid dynamics representation. All five constraints are assessed across all test cases.

4. Optimization procedure

The selection of the most suitable optimization algorithm for a similar flow field reconstruction method was previously addressed in Jakac et al. (2024), where the performance, efficiency, robustness, and scalability of various methods were evaluated, leading to the adoption of PSO as the preferred approach for this type of modeling. Although the authors noted that optimization outcomes are case-dependent, they observed similar results across different test cases due to the synthetic nature of the simulations, despite variations in data sources.

As outlined in Jakac et al. (2024), the mean square difference of the velocities at measurement points, denoted as E_d , is used as the fitness

function in all optimization tests. The optimization is thus defined as follows:

$$\underset{\mathbf{b}}{\text{minimize}} \quad E_d(\mathbf{b}) = \frac{1}{n_{MP}} \sum_{i=1}^{n_{MP}} (\mathbf{u}_{r_i} - \mathbf{u}_{s_i}(\mathbf{b}))^2 \quad (16)$$

$$\text{subject to} \quad \mathbf{b}_l \leq \mathbf{b} \leq \mathbf{b}_u.$$

Convergence of the optimization process is considered to be achieved when the drifter error threshold, $E_d = 10^{-4}$, is reached, corresponding to a drifter velocity error in m/s. All reconstruction results for which this threshold is met are deemed satisfactory. The (maximum) number of iterations of the fitting optimization is not set.

4.1. Initial flow fitting

The idea behind recreating passive scalar field advection and diffusion based on a stationary reconstructed flow is to advect and diffuse the passive scalar field according to the times at which the measurements are collected. As the movement of the passive scalar field is heavily influenced by the initial reconstructed flow, and since errors accumulate throughout the process, it is crucial to extend the optimization duration for the first flow reconstruction. Since we are dealing with real-time measurements the optimization does not rely on a predefined maximum number of iterations; instead, it is constrained by a fixed time limit which corresponds to the expected period of real-time measurement acquisition T_s to ensure real-time evaluation of the velocity field and drifter trajectories enabling the system to assimilate new data and treat the flow as quasi-steady within each interval.

To minimize the error accumulation, the initial flow reconstruction is given a duration of $5T_s$. This ensures the flow is accurately reconstructed before being used in subsequent steps. After the first reconstruction, the optimization time is shortened to match the acquisition period of new measurements, which are used to update the flow for further advection and diffusion calculations. Since the flow field is not expected to change significantly between periods, this approach, starting with a longer reconstruction and followed by shorter ones based on updated measurements ensures accurate reconstruction of the transient flow dynamics. As a result, the movement of the passive scalar field is reliably replicated by using consecutive stationary flow reconstructions.

4.2. Optimization initialization

The optimization process seeks to find the best solution by iteratively adjusting the optimization vector. This means that each simulation starts with internal field values initialized to zero, while the boundary conditions differ. However, certain configurations of optimization vector values may sometimes result in simulations that struggle to converge or take considerably longer to reach convergence, increasing the overall time required for optimization.

Given that the flow field is not expected to change significantly between periods, the initial bounds for optimization variables are applied only during the first flow reconstruction. For subsequent reconstructions, the optimization bounds are updated and limit the search range within 30% of the current best optimization result i.e:

$$\begin{aligned} \mathbf{b}_l^* &= \mathbf{b}_{opt} - 0.3(\mathbf{b}_u - \mathbf{b}_l) \\ \mathbf{b}_u^* &= \mathbf{b}_{opt} + 0.3(\mathbf{b}_u - \mathbf{b}_l) \end{aligned} \quad (17)$$

where, \mathbf{b}_l^* and \mathbf{b}_u^* represent the updated optimization bounds for the optimization variables. This refinement reduces the search space for subsequent iterations, enhancing the optimization process by improving optimization/fitting convergence and concentrating on the most favorable parameter ranges for each period.

Additionally, a flow-based initialization method, inspired by Jakac et al. (2024), is applied. In this approach, the optimization process starts with the best solution from the previous step, using its internal

field values and optimization variables to initialize the new simulation. By combining different initialization approaches and the reduced search space concept, the overall number of iterations is reduced, leading to quicker simulations and overall faster convergence.

4.3. Adaptive diffusion coefficient based on reconstruction flow error

To reduce errors in flow reconstruction and their impact on passive scalar transport, we use the mean square displacement, denoted as S^2 , of a Brownian particle to estimate a compensating diffusion coefficient. S^2 represents the average squared displacement of a diffusing particle over time and is defined as:

$$S^2(t) = 2 \cdot n_{dim} \cdot D \cdot t, \quad (18)$$

where n_{dim} is the number of spatial dimensions ($n_{dim} = 2$) in this case, D is the diffusion coefficient, and t is the time interval.

Reconstruction errors introduce deviations between the reconstructed flow field and the referent flow field, resulting in a net displacement over a period T_s . This displacement is calculated using the reconstruction error E_d incorporated into the MSD relationship for two-dimensional diffusion to determine the compensating diffusion coefficient:

$$D_c = \frac{(\sqrt{E_d} \cdot T_s)^2}{4 \cdot T_s}. \quad (19)$$

Here, $\sqrt{E_d} \cdot T_s$ represents the mean displacement associated with the drifter error over the period T_s . This compensating diffusion coefficient quantifies the uncertainty introduced by inaccuracies in flow reconstruction.

Therefore, the adaptive diffusion coefficient for the reconstructed flow, D_{adp} , can be expressed as:

$$D_{adp} = D_{base} + D_c. \quad (20)$$

In this formulation, D_{base} denotes the diffusion coefficient of the referent flow, while the adaptive diffusion coefficient D_{adp} is recalculated at each period using newly acquired measurements. This adjustment compensates for reconstruction errors, ensuring that the advection-diffusion process more closely aligns with the behavior of the passive scalar field under the referent flow conditions.

5. Acquiring measurements

5.1. Measurement point advection

As drifters move through the domain under the influence of currents, we opted to model their movement as a quasi-steady advection problem. This process involves advecting the positions of the measurement points, or drifters, within the domain. The positions are updated based on the reconstructed velocity field, with the movement of each drifter determined by the velocity field from the preceding simulation step. The position is then adjusted using Euler's method for solving the particle advection equation:

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) + \mathbf{u}(\mathbf{x}(t))\Delta t \quad (21)$$

where $\mathbf{x}(t)$ represents the position of the measurement point at time t , $\mathbf{u}(\mathbf{x}(t))$ is the velocity at that position, and Δt is the integration time step. This approach ensures that positions of measurement points are simulated according to the flow dynamics at each time step, enabling the tracking of their positions over time. The advection of the measurement points is critical for obtaining measurements at new locations, thereby allowing for the reconstruction of the velocity field and capturing the evolution of the flow. By iterating this process across different flow states, the movement of the drifters is effectively simulated, providing a reliable representation of transient flow dynamics in the domain.

5.2. Absent measurements

In real-world scenarios, absent measurements can arise for several reasons. One common cause is that drifters may drift outside the monitored area, leading to a loss of data. Additionally, signal interruptions can occur, particularly when drifters travel over large distances or encounter environmental factors such as rough sea conditions or physical obstructions that disrupt communication, compounded by the limitations of the tracking technology. Despite these potential data deficiencies, the proposed flow reconstruction method is designed to function independently of the number of available measurements at any given time. By leveraging the inheritance approach discussed in earlier sections, the method compensates for missing data, as the flow field is not expected to change drastically between periods. This approach allows for effective reconstruction of the flow field, even in the presence of scattered or incomplete measurements, ensuring that the overall accuracy of the flow approximation remains reliable.

6. Assessment of passive scalar field advection accuracy

To evaluate the accuracy of passive scalar advection, we compare two advection/diffusion processes on a synthetic benchmark case: one driven by referent flow which represents the referent scalar field movement we want to replicate, and the other simulated with quasi-steady flow assumptions. The comparison is based on the areas of advected/diffused probability by comparing the resulting scalar fields. Both scalar fields, which represent the passive scalar field, are initialized at the same location with identical values to ensure consistency.

To generate the referent advection process, we synthetically create referent flow cases with time-varying boundary conditions, solving the scalar transport equation alongside the referent flow to track passive scalar movement. These boundary conditions produce transient flow dynamics, which we aim to approximate using quasi-steady flow assumptions.

The quasi-steady methodology is an inherently iterative process that approximates unsteady flow by solving a sequence of steady-state problems. By reconstructing the velocity field for each discrete time step incremented by period T_s , the system effectively captures flow variations and provides a reliable basis for modeling passive scalar advection.

To assess the accuracy of passive scalar advection, we use two metrics. The first metric, called the intersection metric, quantifies the portion of the referent scalar field that is captured by the simulated scalar field. It is defined as:

$$I = \frac{\int_{\Omega_{intrsc}} s_{sim} d\Omega}{\int_{\Omega} s_{ref} d\Omega} \quad (22)$$

where Ω_{intrsc} represents the intersection area between the simulated and referent passive scalar fields, and s denotes the passive scalar values. The intersection metric is scaled between 0 and 1, where a value of 1 indicates perfect alignment between the simulated and referent scalar fields, meaning the reconstructed flow completely replicates the transient dynamics. A value closer to 0 indicates a greater discrepancy between the fields, highlighting the limitations of the reconstructed flow in capturing transient advection effects.

The second metric, called the coverage metric, measures the extent to which the simulated scalar field is covered by the referent scalar field. It evaluates how much of the referent probability field is represented by the simulated probability field. The coverage metric is calculated as:

$$C = \frac{\int_{\Omega_{intrsc}} s_{sim} d\Omega}{\int_{\Omega} s_{sim} d\Omega} \quad (23)$$

A value of 1 indicates that the referent passive scalar field fully covers the simulated probability field, while a value closer to 0 indicates less coverage.

Table 2

Characteristics of the validation test cases.

Case	Simple bay	Lošinj case	Cres case
Test case type	Synthetic	Realistic	Realistic
Domain area [km ²]	24.6	96.49	55.62
Number of boundaries	1	4	1
Total boundary length [km]	9.4	9.63	7.22
Coastline length [km]	9.1	57.12	31.74
Number of boundary control points	5	6	3
Max velocity in the domain [m/s]	0.25	0.4	0.5
Number of cells	4625	7530	15 833
Average cell size [m]	73.02	113.21	59.28

Both the intersection and coverage metrics provide a quantitative evaluation of how well the reconstructed passive scalar field (based on stationary flow) corresponds to the reference passive scalar field (based on transient flow). These metrics provide insight into the effectiveness of the flow reconstruction and the precision of the advection process.

7. Results

To capture evolving flow dynamics in a computationally efficient manner, transient flow is approximated through a sequence of steady-state reconstructions, each performed at regular measurement intervals $[t, t + T_s]$. For each interval, boundary conditions are recomputed based on reference flow measurements, enabling the framework to track gradual system changes without resorting to full transient simulations. Such an approach addresses time complexity, while the usage of a fusion model helps mitigate spatial complexity. This iterative reconstruction process allows the simulated passive scalar field to more accurately reflect real-world transport dynamics, particularly in scenarios where transient flow characteristics significantly influence advection patterns. In order to illustrate the sequential structure of the proposed method, the complete optimization-simulation workflow for each measurement interval is outlined in the flowchart shown in Fig. 2.

To validate our methodology, we designed three distinct test cases, each within a different domain: a synthetic domain, a realistic domain, and a domain where we conducted experimental field measurements. These cases were selected to systematically evaluate our approach under varying conditions, ranging from idealized to real-world scenarios. The specific characteristics of each test case are summarized in Table 2. This diverse set of cases enables a comprehensive assessment of our method's performance across different environments and scales.

7.1. Simple bay

This synthetic domain of 25 km² provides a controlled environment with predefined flow conditions, allowing for a precise assessment of the reconstruction method's performance without external uncertainties. The domain consists of a coastline and an inlet/outlet, which are designed to simulate realistic boundary interactions.

Since the goal of the proposed method is near-real-time simulations, where extensive numerical computations must be avoided, a coarse mesh must be used. To assess the reliability of the simulation results under these conditions, a grid convergence study was performed using the conventional mesh generation approach. Three meshes were generated for the synthetic Simple Bay test case: coarse (4625 cells), medium (10 564 cells), and fine (23 296 cells), with a uniform refinement ratio of 1.5. The Grid Convergence Index (GCI) was evaluated at all grid points using vorticity values, following the approach outlined in Roache (1998), which employs Richardson extrapolation to estimate the convergence order. The mesh sensitivity results, summarized in Table 3, confirm that the solution lies within the asymptotic range of convergence. This is supported by the estimated order of convergence of $p = 1.79$ and the calculated mean correction factor $\bar{c}_c = 0.976$,

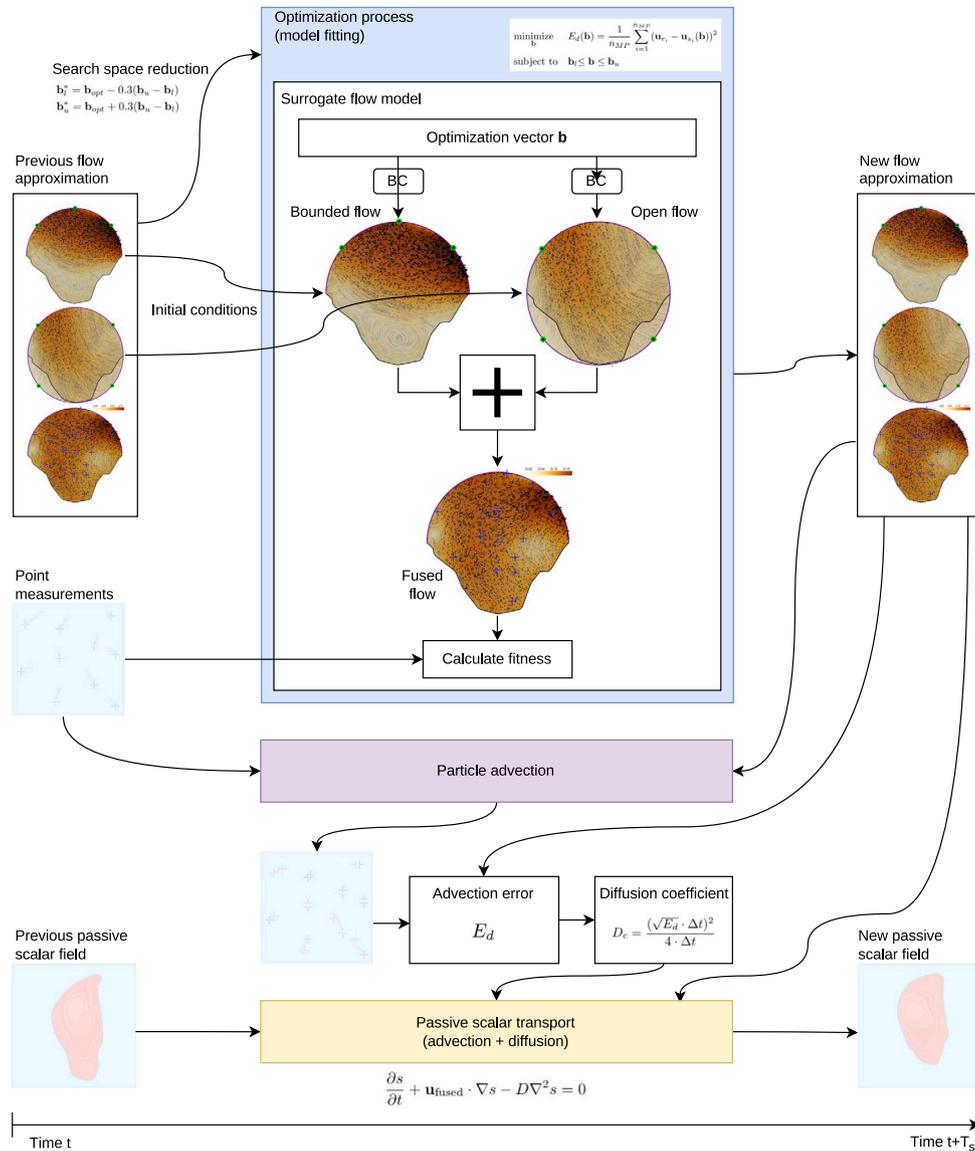


Fig. 2. Flowchart of the optimization-simulation procedure performed started at time t to make scalar transport assessment for $t + T_s$. The workflow includes input of velocity field measurements, dual-domain simulation using OpenFOAM, boundary condition optimization, fusion of velocity fields, error evaluation, and iterative minimization of the drifter-based error E_d . This process is repeated for each subsequent discrete time step, incremented by T_s , to reconstruct transient flow fields.

Table 3

Grid details and GCI for the synthetic Simple bay case. Assessment is made at all grid points for the vorticity values.

Mesh	Number of cells	Refinement ratio (r)	Relative error (%)	GCI (%)
Coarse	4 625	—	—	—
Medium	10 564	1.5	1.53	1.09
Fine	23 296	1.5	0.76	1.34

which exceeds 0.95, further indicating that the numerical error behaves predictably with mesh refinement.

Fig. 3 illustrates the evolution of the referent flow over 32 400 s, highlighting variations in both direction and magnitude that we aim to reconstruct. By using this synthetic domain, we can isolate and assess the specific impact of boundary condition updates on the accuracy of passive scalar field reconstruction, focusing on the method’s effectiveness in handling temporal flow changes. The case provides a clear validation of the methodology under controlled conditions, establishing a baseline for comparison with more complex real-world scenarios. While drastic flow changes would not typically occur over a short

period in real-world cases, the synthetic nature of this domain allows us to test our methodology under extreme conditions and adaptability in reconstructing flow fields subjected to evolving boundary.

Since this is a quasi-steady approach, selecting an appropriate T_s is crucial. It must be long enough to allow the optimization to reconstruct the flow field accurately, but not so long that significant flow changes occur, which would lead to large errors in passive scalar advection. Passive scalar advection is highly sensitive to the accuracy of the reconstructed flow, and its error accumulates over time. To assess the impact of T_s , we evaluated the reconstruction accuracy by computing the root-mean-square error of the entire velocity field ($RMSE_f$) over a range of T_s values spanning 300 to 1500 s. For each value of T_s , we performed 10 independent optimization runs to account for the stochastic nature of the algorithm and calculated the median $RMSE_f$. The results of this analysis are presented in Fig. 4.

Fig. 5 compares four different approaches for flow reconstruction and the resulting passive scalar advection. Due to the dynamic nature of the referent flow, the stationary-fit approach (B), which reconstructs the flow using only the initial 15 drifter measurements from the referent flow and then advects the passive scalar field over 32 400 s without

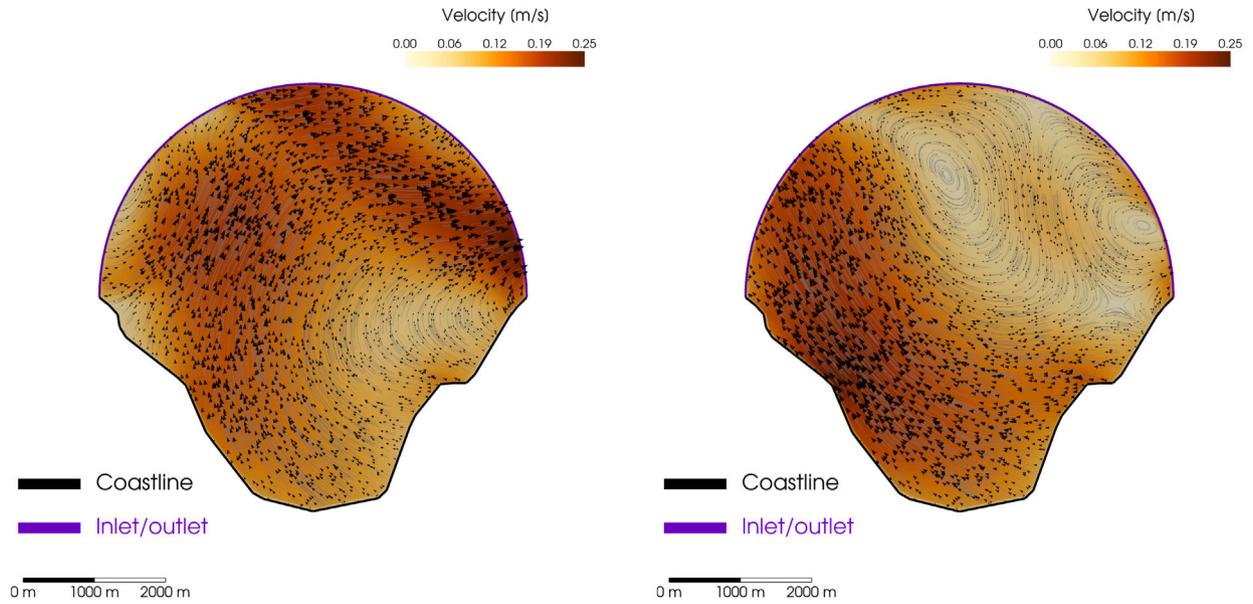


Fig. 3. Figure represents one type of the evolving referent flow in the synthetic Simple bay case after 32 400 s, highlighting the difference between the initial flow (left side) and the final flow (right side). As observed, the flow undergoes significant changes, demonstrating how the dynamics of the system can shift over time while making the flow reconstruction more challenging.

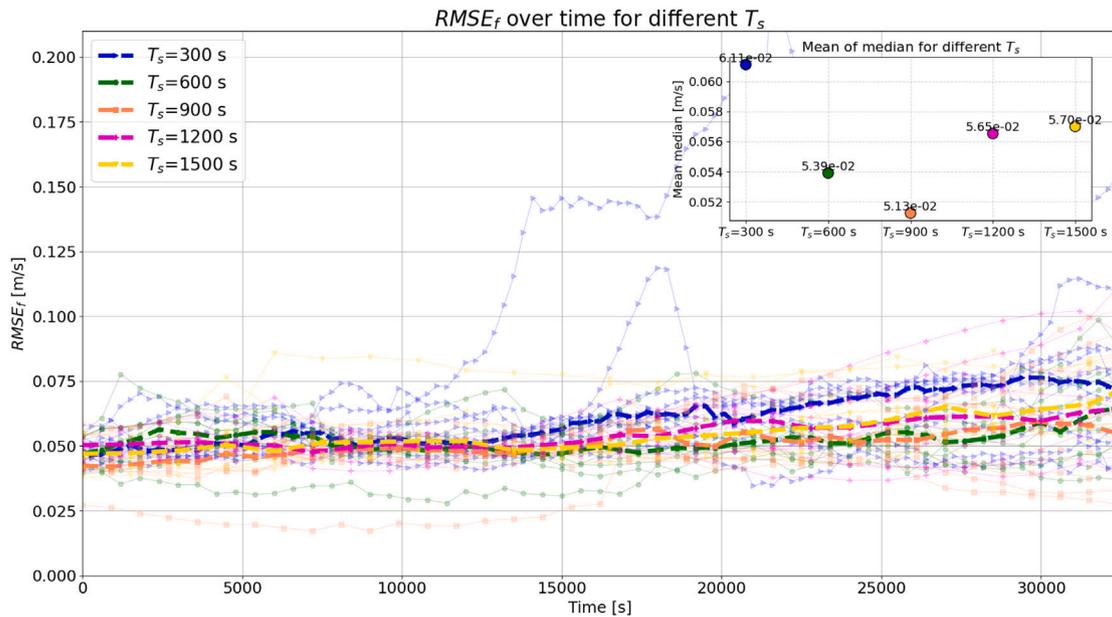


Fig. 4. Figure represents $RMSE_f$ of the reconstructed velocity field for five separate analyses using different values of T_s . Transparent lines represent individual optimization runs, while the thicker dotted line indicates the median $RMSE_f$ for each analysis. The inset plot in the top right shows the mean of median $RMSE_f$ values across all five analyses, highlighting the trade-off between allowing sufficient time for flow reconstruction and minimizing errors caused by temporal flow variability.

updates, led to significant errors, ultimately resulting in 0% intersection with the referent field. The transient-fit approach (C), which reconstructs the referent flow at each T_s period across multiple steady-state flow optimizations, was also unable to capture the complex referent flow due to the limitations of the bounded CFD domain, again resulting in 0% intersection. Notably, after 32 400 s, only four drifters remained in the domain, meaning measurements were available from a limited number, further reducing the accuracy of the reconstructed flow.

To improve accuracy, the fusion model was introduced into the fitting process (E), enhancing the reconstructed flow accuracy and achieving an 86% intersection. To further improve accuracy an adaptive diffusion coefficient (F) was introduced, with the fusion model producing even better results, reaching 90% intersection after 32 400 s.

The analysis of success (D) highlights the effectiveness of each approach in capturing the passive scalar field, providing insights into their respective strengths and limitations. Furthermore, the best-performing approach, which combines the fusion model and adaptive diffusion, not only maximized the intersection percentage but also optimized the balance between covered and excessive area, ensuring the most accurate passive scalar field reconstruction.

7.2. Loinj case

In order to assess the proposed methodology for a realistic domain, we selected Unije Bay near the island of Loinj, covering an area of 96.5 km². This domain features four distinct inlet/outlet areas, which can

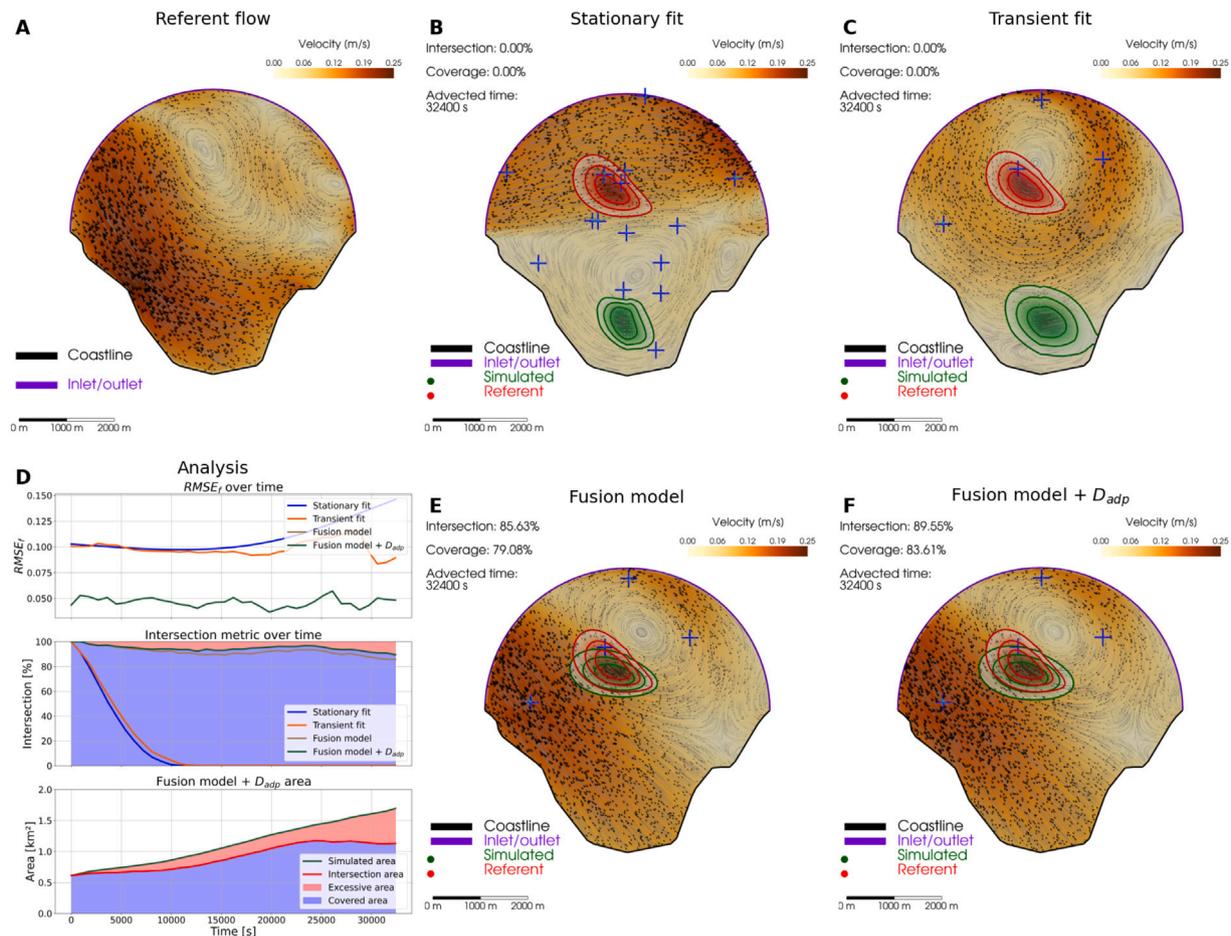


Fig. 5. (A) Evolution of referent flow after 32400 s, which we aim to reconstruct. (B) Stationary reconstructed flow based on 15 drifter measurements from the initial flow where the flow was reconstructed using initial state and conditions. (C) Flow reconstruction with periodic measurement corrections (updates) where the fusion approach is not considered. (D) Intersection analysis comparing different approaches for passive scalar field advection simulation. (E) Flow reconstruction with updated measurements and fusion model. (F) Flow reconstruction with updated measurements, fusion model, and adaptive diffusion compensation to account for errors in flow reconstruction.

induce complex and dynamic flow patterns. Unlike synthetic, the realistic domain introduces naturally occurring flow variations influenced by external environmental factors, providing a more challenging test for the reconstruction methodology (see Fig. 6).

To demonstrate the effectiveness of our methodology in a realistic domain, Fig. 7 compares four different approaches for flow reconstruction and the resulting passive scalar advection. As observed in the synthetic case, both the stationary fit (B) and transient fit (C) failed to replicate the referent passive scalar movement. Surprisingly, in this instance, the transient fit performed even worse than the stationary fit, indicating that optimization can sometimes converge to incorrect solutions.

Incorporating the fusion model (E) into the transient fitting significantly improved accuracy, achieving nearly 93% intersection with the referent passive scalar field after 36 000 s. Further improvement was achieved by introducing an adaptive diffusion coefficient (F), which yielded an intersection metric of 94% and a coverage metric of 100%.

The success analysis (D) evaluates the effectiveness of each approach in capturing the passive scalar field, revealing a slight improvement with the inclusion of the adaptive diffusion coefficient and unexpectedly better performance of the stationary fit compared to the transient fit. Moreover, the best-performing approach – combining the fusion model with adaptive diffusion – covered some excessive area. However, the coverage metric indicates that this excess region contains a minimal trace of passive scalar, which can partially be attributed to numerical errors, and can thus be disregarded when predicting the movement of the bulk of the passive scalar.

7.3. Cres case

To properly validate the predictive algorithm, a sea experiment was conducted in the larger area of Valun Bay on the island of Cres, located in the Kvarner Bay region of the northern Adriatic Sea, Croatia. This location covers an area of 5 km by 10 km in longitudinal and latitudinal directions and was selected for its balanced environmental conditions, making it well-suited to represent submesoscale flow. While more exposed than smaller enclosed bays, Valun Bay provides sufficient natural barriers to maintain manageable conditions while allowing the influence of moderate currents. Its moderate size also facilitates efficient tracking and retrieval of drifters. Data collection was carried out using 12 TB-560 tracking beacons from Alltek Marine Electronics Corp (AMEC) (Marine, 2024), with all available drifters deployed to maximize measurement coverage. These beacons are simplified versions of Automatic Identification System (AIS) Class B units, offering a cost-effective solution for vessel tracking. AIS is widely used in maritime navigation to automatically share important information, including vessel identity, position, course, and velocity. Data exchange between vessels and shore stations improves maritime safety, aids in collision avoidance, and traffic management. Despite their basic design, these beacons efficiently transmit AIS data, making them suitable for small vessels, recreational boats, fishing boats, and buoy-based applications. AIS messages contain various categories of information essential for maritime communication and navigation, with standardized formats ensuring compatibility across vessels and shore-based stations. Table 4 outlines the main types of AIS messages.

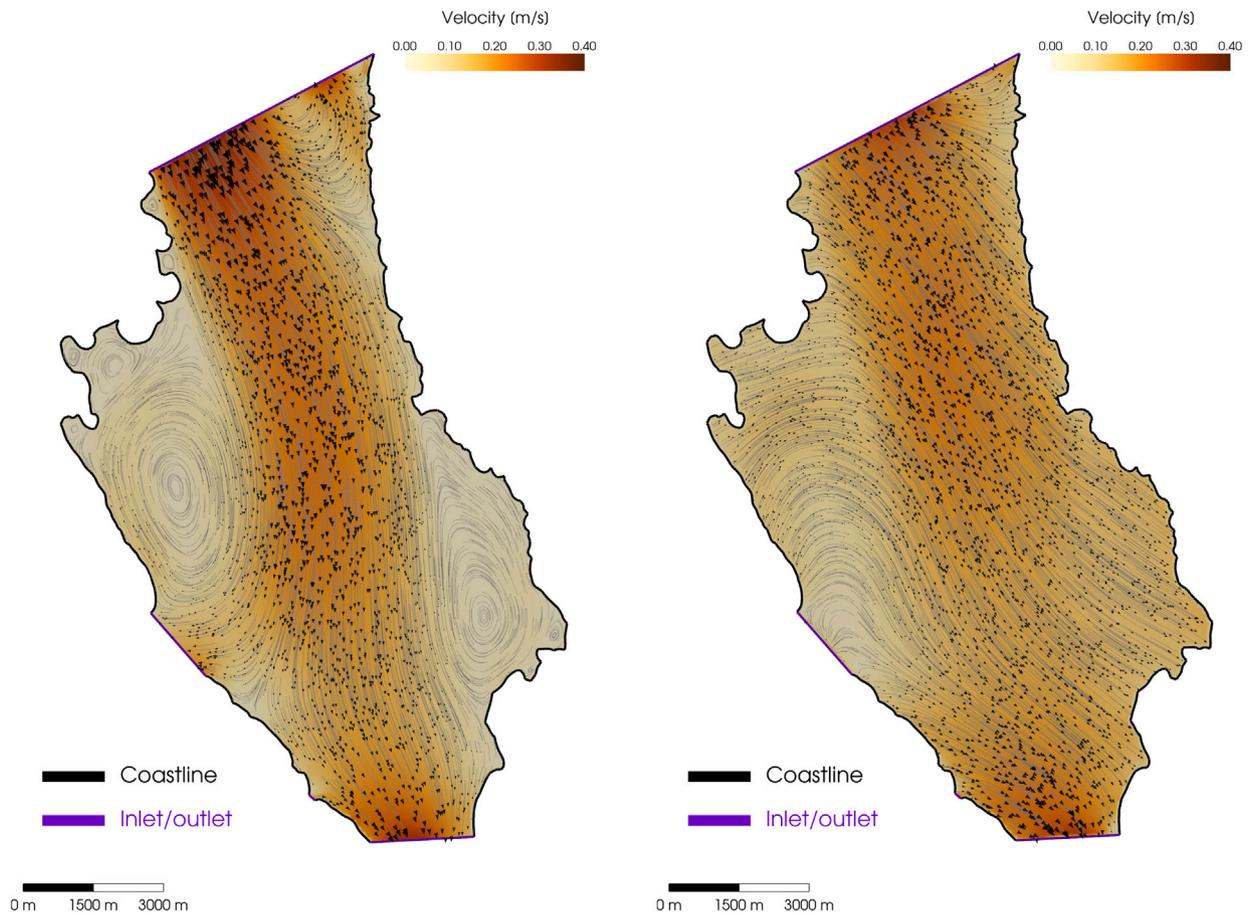


Fig. 6. Evolution of the referent flow in the realistic Loinj case after 36 000 s, illustrating the differences between the initial flow (left) and the final flow (right). This case features four inlet/outlet areas, allowing for the formation of multiple vortices within the domain. Initially, the flow is directed toward the western and southern outlets, whereas after 36 000 s, it shifts toward the eastern coastline and the southern outlet.

Table 4
Summary of AIS message types.

Message	Description	Example use cases
Type 1	Position Report Class A	Real-time vessel position (Class A)
Type 5	Static and Voyage-Related Data	Vessel identification and voyage details
Type 18	Standard Class B Position Report	Real-time vessel position (Class B)
Type 24	Static Data	Essential vessel identification information
Type 27	Position Report for Long-Range AIS	Extended-range position updates
Type 8	Binary Broadcast	Custom data transmission for specific applications
Type 22	Channel Management	Channel management for AIS transmissions

Outfitting floating buoys with AIS devices allows real-time tracking of sea surface velocity. Buoys transmit collected data at set intervals, offering authorities and rescue teams valuable insights into ocean currents and surface conditions. Once activated, the TB-560 usually acquires a GPS position within a minute and sends position reports using AIS message types 18 and 24.

7.3.1. Drifter preparation and deployment

To measure sea surface velocity using AIS devices, a custom buoy platform with a mount for tracking beacons was designed and built. This involved developing a draught system to reduce the impact of waves and acquiring all necessary parts for constructing the drifters. A dedicated workstation, equipped with an AMEC Cypho-150 receiver

and a 550 VHF antenna, was deployed along with the necessary software to decode and receive AIS messages from the tracking beacons. By combining these components, operational drifters were successfully built to capture sea surface velocity data. The entire process, from initial design to final construction, is shown on the left side of Fig. 8. The right side of the figure depicts the loading of the equipment onto the boat and the subsequent deployment of all 12 available drifters during the experiment, which took place on September 21st and 22nd, 2024, to maximize our data collection.

7.3.2. Experimental data

Each drifter is assigned a distinct Maritime Mobile Service Identity (MMSI) number, which enables the retrieval of GPS coordinates linked to that MMSI from satellite systems. The drifter uses these GPS readings to calculate its speed and heading, transmitting the data every 10 s to a workstation through AIS messages 18 (position reports) and 24 (static data). During the data processing phase, these messages are decoded to extract the drifters location and velocity.

Since velocity measurements are taken every 10 s and rely on satellite GPS, some inaccuracies are unavoidable. To improve the accuracy of the data, we use a moving average filter on the raw measurements. Specifically, we compute the average velocity over 1-min intervals, which involves averaging all the readings for each drifter within that period. This approach helps to minimize short-term variations, filtering out noise and highlighting more consistent trends. Averaging the data over a minute results in more accurate velocity measurements, reducing the impact of transient errors from individual data points. This method provides a more dependable depiction of surface velocity dynamics in our analysis.

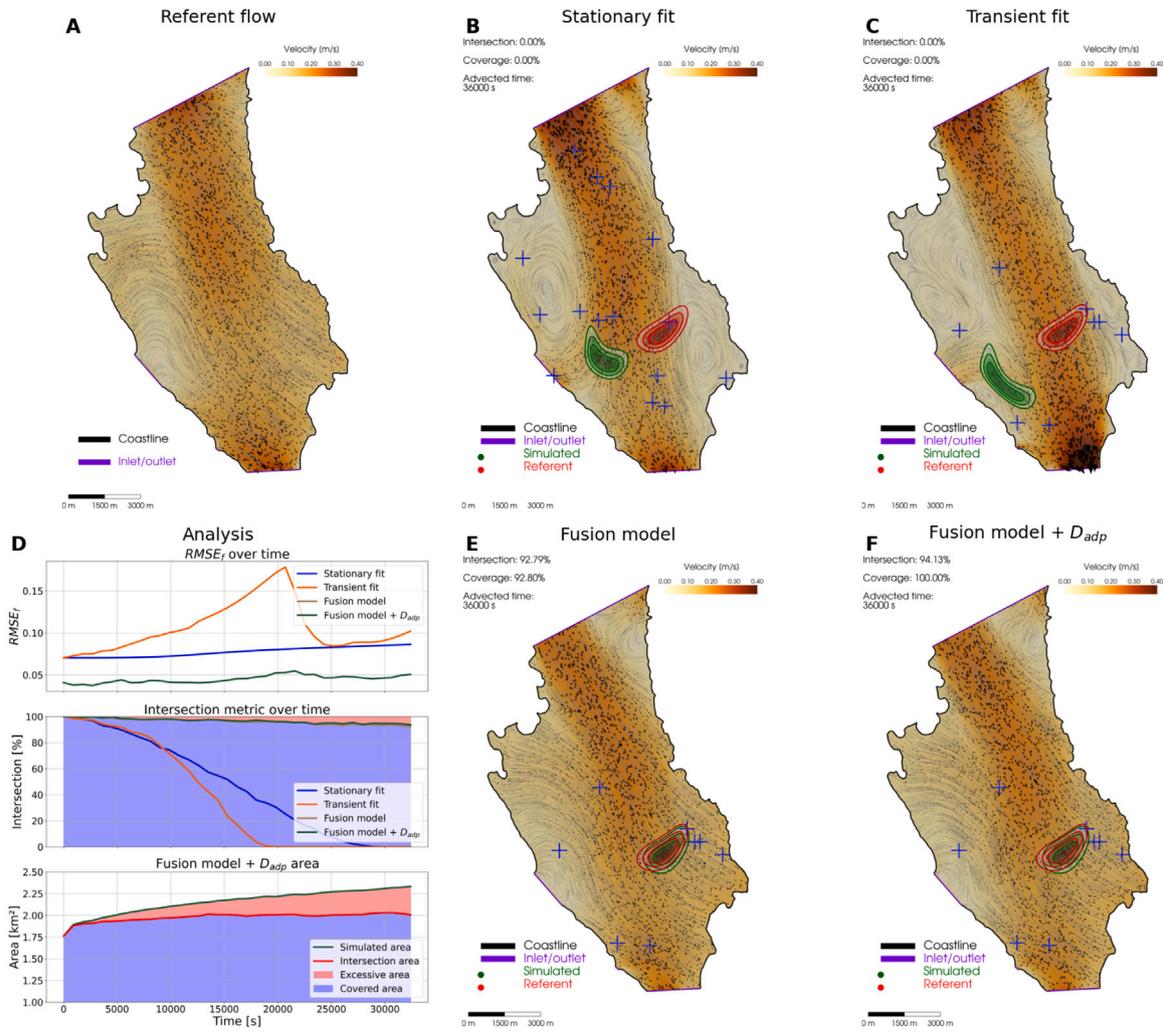


Fig. 7. (A) The reference flow after 36000 s, is predominantly directed toward the eastern coastline, a pattern commonly associated with the Jugo wind in this region. (B) Stationary flow reconstruction based on 15 drifter measurements from the initial flow, resulting in no intersection with the reference passive scalar field. (C) Transient flow reconstruction without the fusion model. An even greater discrepancy between the simulated and reference passive scalar field is observed. (D) Intersection and adaptive diffusion analysis, comparing different approaches for passive scalar advection. (E) Transient flow reconstruction incorporating the fusion model yields excellent results. (F) Flow reconstruction integrating fusion approach and adaptive diffusion compensation. The captured passive scalar field is now more in line with the referent state.

The drifter deployment across the 55-square-kilometer bay took approximately 1 h 30 min, with each drifter carefully positioned at its assigned location. The outcomes of the deployment on September 22, 2024, are shown in Fig. 9, with the movement of each drifter represented in different colors.

While some measurement gaps are present, the overall movement pattern remains clear. These gaps are attributed to signal loss in specific areas of the bay, which may result from obstacles such as landforms, vegetation, or the distance between the drifters and the antenna. Despite these limitations, the deployment successfully produced valuable data on the drifters movement and the dynamic conditions in the bay.

To replicate the trajectories of each drifter based on measurements, we focused on a 2 h time period within the overall 4 h 30 min experiment. To capture drifter movement, we obtained new measurements at 300 s intervals, assuming that no drastic flow changes would occur within this time frame.

Initially, the algorithm with a fixed wind, as proposed in Jakac et al. (2024), where a uniform wind field was added to the CFD simulation for flow reconstruction, was used. However, a key issue arose in the southern part of the domain, where actual velocities are smaller due to the coastline that surrounds the area, creating a closed system. The

added wind velocity adversely affected the sea surface velocity in this region, causing the reconstructed trajectories to be significantly longer than those observed in the experiment.

To address this issue and accurately predict the movement of drifters within the closed system, a fusion model was employed with an updating measurement period of $T_s = 300$ s, based on data from the deployed drifters. At each new time step the measurements from the deployed drifters were acquired. These measurements were then used to reconstruct the surface velocity field using a steady-state solver, with optimization running for T_s s before new measurements arrived. To replicate the drifter movement, the positions of the drifters were advected over the T_s period using the reconstructed flow, ensuring that the flow dynamics remained consistent with the updated measurements at each step. Visual representation of changes at the first and last time step can be seen in Fig. 10.

Based on the results, the conventional bounded domain simulation approach struggles to generate velocity fields that align with realistic measurements, particularly in regions far from the inlet and outlet, such as in closed domains like the Cres case. In contrast, the open domain approximation introduces non-uniform velocity fields and adjusts the overall flow to better match the measurements. The time-dependent

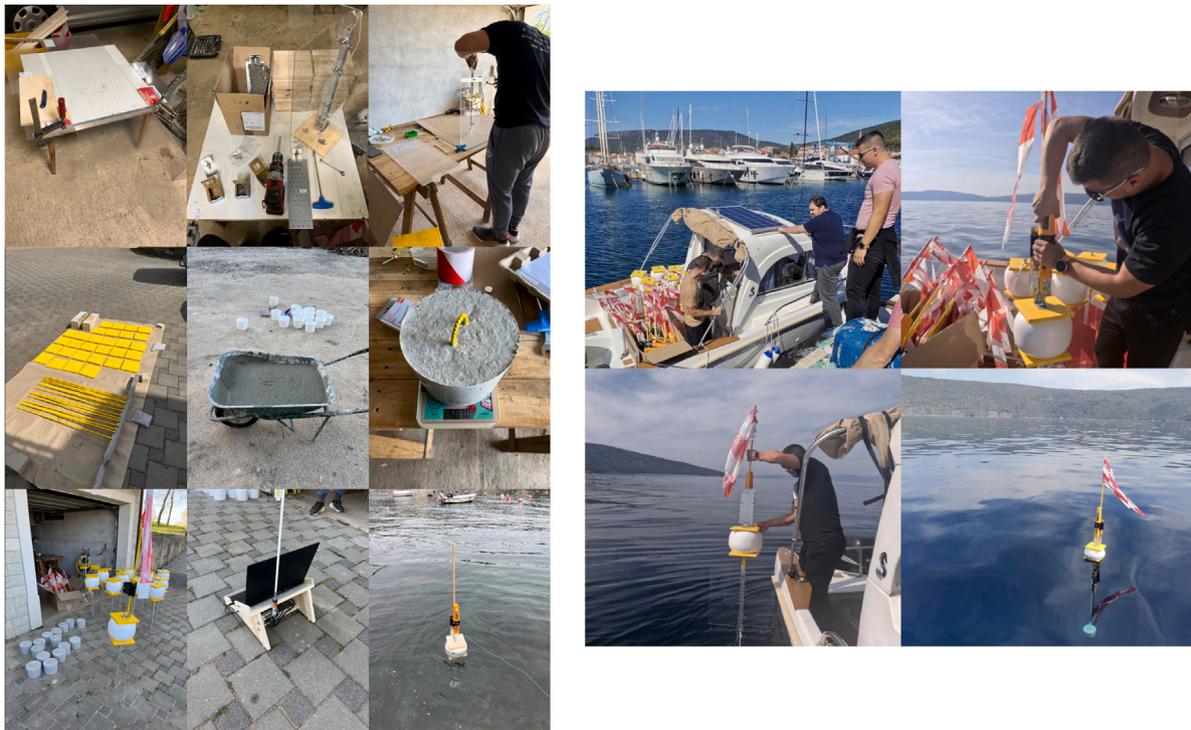


Fig. 8. Figure depicts the process of drifter preparation and deployment. On the left, the preparation process is shown, including the design and final assembly of the drifters. On the right, the loading of the drifters and their transit are illustrated. After reaching the deployment area, the drifters were activated and systematically deployed at pre-determined locations within Cres Bay to ensure optimal distribution for the experiment.

evolution further emphasizes the model's adaptability, improving its ability to represent realistic surface dynamics.

To validate the proposed methodology, we attempted to reconstruct experimental drifter trajectories using both fixed wind from [Jakac et al. \(2024\)](#) and the fusion model. The difference is clearly illustrated in [Fig. 11](#), where experimental drifter trajectories are shown in darker colors, while the simulated trajectories for both correction approximations are displayed in the same color but with increased transparency.

The fusion model approximation provides notably improved results, more accurately capturing the realistic curvature of the trajectories, particularly in the southern part of the domain. In this region, the fixed wind approximation tends to exaggerate surface flow velocities, while the fusion model approximation effectively compensates for this issue. Although several of the reconstructed trajectories show a high degree of similarity to the experimental ones, some discrepancies are still present. To better assess the accuracy of the trajectory reconstruction, the trajectory reconstruction error, defined as the distance in meters between the reconstructed and experimental trajectories, is shown in [Fig. 12](#).

This comparison shows that both the stationary fit and fixed wind approaches have large variations in trajectory areas, with the minimum and maximum error trajectories spanning a broad range over time. In contrast, the proposed methodology using the fusion approach (green-shaded area) demonstrates significantly smaller error trajectory areas, indicating a more consistent and replicable trajectory pattern.

8. Limitations and discussion

The proposed fused methodology combines two simplified two-dimensional flow models to act as a surrogate for submesoscale dynamics, aiming to predict the movement of a scalar through a stationary simulation with continuously updated measurements. The surrogate's accuracy is inherently limited by the spatial resolution of the

underlying computational grid. Although a mesh sensitivity analysis demonstrated satisfactory convergence and well-defined error bounds, a coarse grid was ultimately adopted to prioritize computational efficiency over fine scale flow resolution. This affects the model's ability to capture small-scale and highly localized flow features, which can be critical in coastal applications or when precise trajectory prediction is required. While the proposed concept does not discourage the use of higher-quality meshes, it is important to be aware of the numerical consequences and the increased computational complexity that would result from their application. Additionally, the model intentionally excludes factors like wind, tides, and temperature fluctuations to ensure computational efficiency, with these exclusions being addressed by the fusion model through an open domain flow approximation, thus balancing computational speed with predictive accuracy. The two-dimensional flow assumption limits the applicability in scenarios where vertical flow structures significantly influence horizontal transport. In addition, while innovative, the approach can produce physically unrealistic flow patterns near the coastline. These artifacts, though documented in similar studies, may affect the reliability of predictions in near-coastline regions. However, it is important to recognize that, as a surrogate model, the proposed approach is not intended to deliver absolute accuracy, but rather a level of accuracy that is acceptable within the defined computational and operational constraints.

The proposed methodology can be prone to error accumulation over time. Since trajectory deviations caused by inaccuracies in the velocity field are not corrected between optimization windows, deviations compound progressively, resulting in increasingly divergent trajectory predictions. Furthermore, the methodology has only been assessed for a limited velocity range (up to 1.5 m/s), hence a comprehensive assessment across a broader spectrum of flow conditions remains to be conducted. It is important to note that in smaller domains with large velocity ranges, passive drifters (i.e., measurement points) may leave the domain quickly, reducing the number of available reference

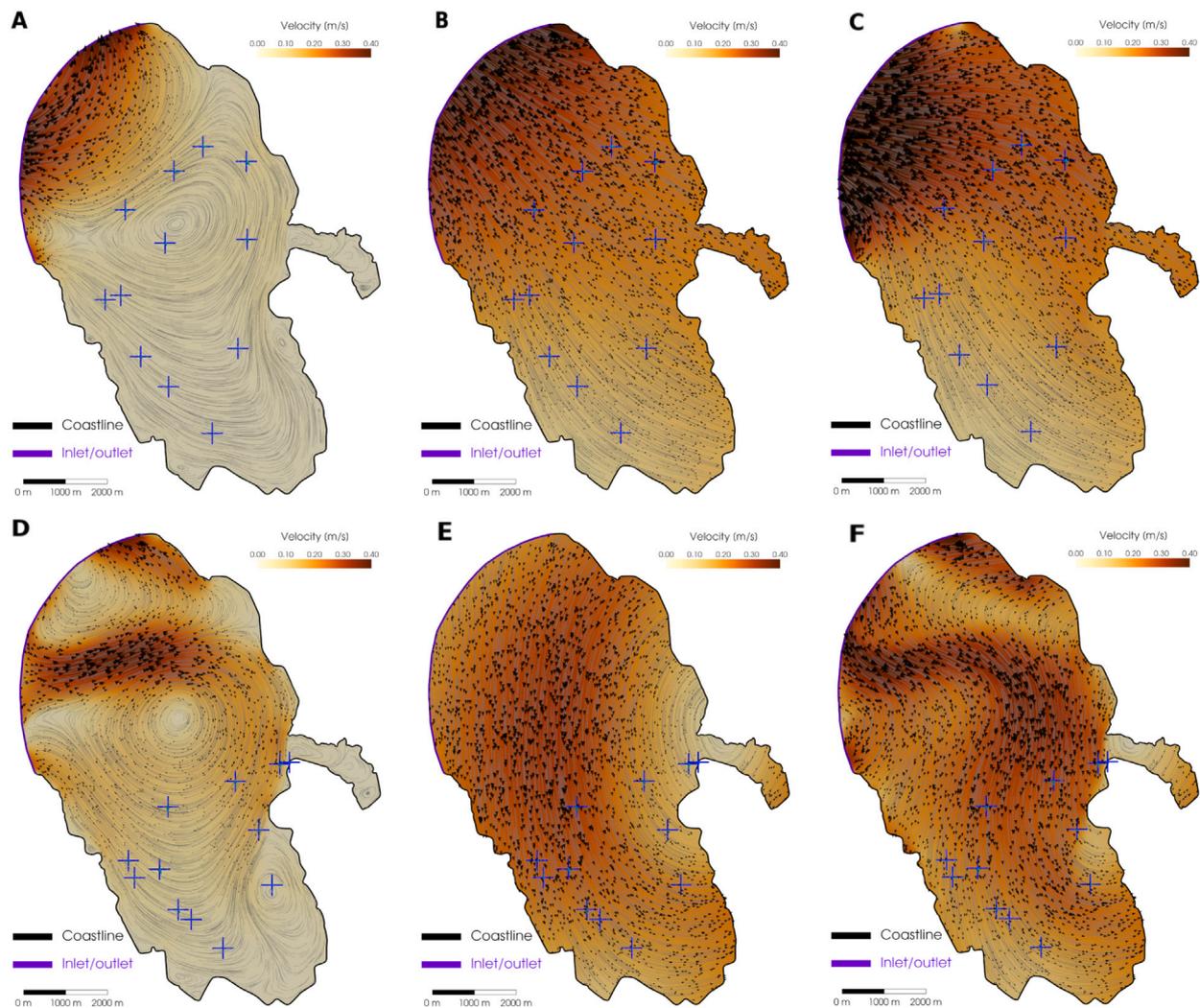


Fig. 10. Figure depicts the flow reconstruction using the fusion model approximation method at the initial and final time step. (A) The bounded domain CFD simulation at $t = 0$ s, with initial drifter locations marked by blue crosses. The simulation shows lower velocities in the southern part of the domain due to its shape, highlighting the limitation of CFD alone in capturing realistic movements in these areas. (B) Usage of open domain allows stronger velocities in the northern region and weaker velocities in the southern region, because of non-uniform open domain flow. (C) Reconstructed flow at $t = 0$ s, obtained by combining the bounded domain simulation with the open domain flow approximation. The resulting velocity field better captures surface dynamics due to the application of a fusion model. (D) The bounded domain CFD simulation of the domain at $t = 7200$ s shows significant changes from $t = 0$ s, as the drifters have moved over time, altering their direction and providing a realistic reference for the time-dependent simulation. (E) Open domain approximation at $t = 7200$ s, where the open domain flow has been adjusted to match the drifter trajectories, demonstrating the adaptability of the approximation model. (F) A notable change in the reconstructed flow compared to the initial (C) and final (F) time steps, highlighting the impact of fusion forcing on the velocity field. The alignment of the reconstructed flow with drifter movement indicates that the approximation effectively captures surface current variations over time.

challenging even with advanced equipment like HF radars or complex oceanic flow simulations due to the dynamic nature of flows. Current methods, while effective in some aspects, often fail to provide detailed insights into the dynamic properties of flow fields and can often be computationally demanding, limiting their use in rapid response scenarios.

In response, we present a data-driven framework that approximates the sea surface velocity field using scattered observation points. The framework utilizes a fused simplified two-dimensional surrogate flow model combined with an optimization approach to adjust boundary conditions, aligning the computed velocity field with measurements. By excluding factors like wind, tides, and temperature fluctuations, the method prioritizes speed, providing an efficient solution for real-world applications where quasi-steady flow assumptions are adequate for predicting advection in cases such as pollution dispersion and object tracking.

From a computational perspective, the use of coarse numerical meshes in our methodology improves efficiency, especially for time-critical applications. However, it is important to recognize that simplifying key environmental factors in the surrogate model may introduce inaccuracies in capturing the full complexity of real-case flow fields. These inaccuracies are acceptable when only an approximate velocity field is needed. Despite these simplifications, our approach offers significant advantages as it can mimic transient behavior and produce quasi-steady flow fields at specific time frames. This also ensures that the optimization can adapt to changing surface flow conditions, as demonstrated in the Cres sea experiment. Moreover, our method provides a reliable velocity field across the entire simulated domain, even in areas without nearby sampling points. This eliminates the need for extensive data, interpolation, or finite differencing, reducing both data collection time and costs. Overall, while challenges remain, the adaptability and efficiency of our methodology make it a valuable tool for real-time monitoring and decision-making in environmental applications.

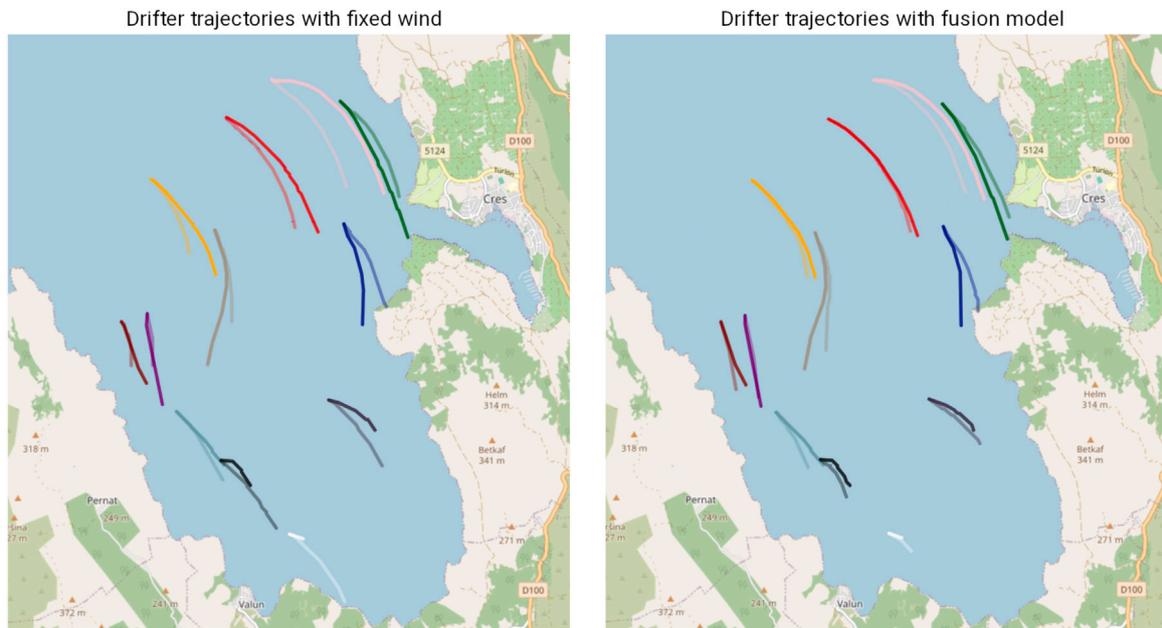


Fig. 11. The figure illustrates the difference in reconstructed experimental drifter trajectories using fixed wind and fusion model, with new measurements taken every 300 s over a 2 h period. Darker colors represent the experimental drifter trajectories, while the reconstructed trajectories are displayed in the same color with increased transparency. It is evident that the usage of fusion model better captures the realistic curvature of the trajectories, with several of the reconstructed trajectories showing high similarity to the experimental ones.

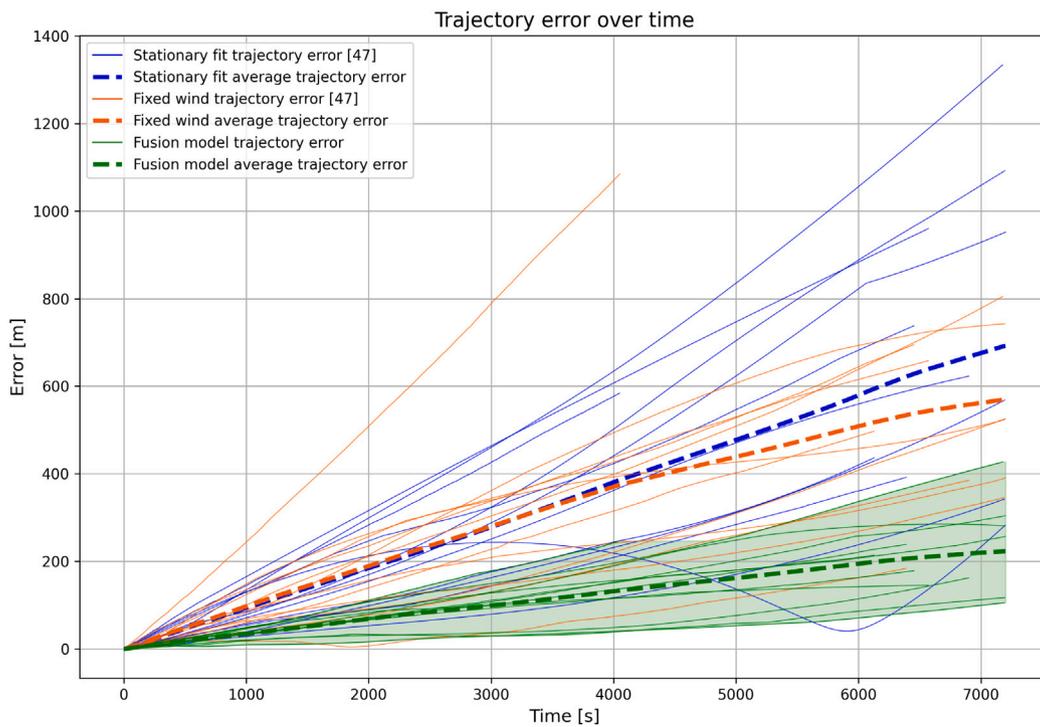


Fig. 12. The blue color on the graph represents trajectory reconstructions based on a stationary flow reconstruction with the fixed wind approximation, where the initial drifter locations are advected solely using the reconstructed flow from the initial time step. The orange color illustrates the approach using transient fit with fixed wind with updated measurements every 300 s in order to better guide optimization. This method shows improving results, particularly visible in prolonged periods, as it adapts to new measurements, accounting for changes in the surface flow over time. The green color represents the use of the fusion model, combined with updated measurements every 300 s. The difference in trajectory reconstruction accuracy is evident, as the fusion model approximation results in a significant reduction in the average trajectory reconstruction error, offering a more precise match to the experimental data.

CRediT authorship contribution statement

Karlo Jakac: Writing – original draft, Validation, Methodology, Data curation, Visualization, Software, Investigation, Conceptualization. **Luka Lanča:** Methodology, Writing – review & editing, Investigation. **Ante Sikirica:** Validation, Investigation, Writing – review & editing, Methodology. **Stefan Ivić:** Writing – review & editing, Validation, Methodology, Conceptualization, Supervision, Investigation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Karlo Jakac reports financial support was provided by Croatian Science Foundation. Luka Lanča reports financial support was provided by Croatian Science Foundation. Ante Sikirica reports financial support was provided by Croatian Science Foundation. Stefan Ivić reports financial support was provided by Croatian Science Foundation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This publication is supported by the Croatian Science Foundation, Croatian Science Foundation under the project UIP-2020-02-5090.

Data availability

All parameters required to reproduce the study are provided in the manuscript. The data necessary for reproducing the presented method and cases can be accessed in The Open Science Framework repository: <https://osf.io/wjsb2/>. The Python code for reproducing this research is available upon request.

References

- Allan, A.A., Plourde, J.V., 1999. Review of Leeway: Field Experiments and Implementation. Tech. Rep.
- Bellomo, L., Griffa, A., Cosoli, S., Falco, P., Gerin, R., Iermano, I., Kalampokis, A., Kokkini, Z., Lana, A., Magaldi, M., et al., 2015. Toward an integrated hf radar network in the Mediterranean Sea to improve search and rescue and oil spill response: The toscia project experience. *J. Oper. Ocean.* 8 (2), 95–107.
- Bennett, J.R., Clites, A.H., 1987. Accuracy of trajectory calculation in a finite-difference circulation model. *J. Comput. Phys.* 68 (2), 272–282.
- Berta, M., Ursella, L., Nencioli, F., Doglioli, A.M., Petrenko, A.A., Cosoli, S., 2014. Surface transport in the northeastern Adriatic Sea from fsl analysis of hf radar measurements. *Cont. Shelf Res.* 77, 14–23.
- Bolaños, R., Sørensen, J.V.T., Benetazzo, A., Carniel, S., Scavo, M., 2014. Modelling ocean currents in the northern Adriatic Sea. *Cont. Shelf Res.* 87, 54–72.
- Breivik, Ø., Allen, A.A., 2008. An operational search and rescue model for the Norwegian Sea and the North Sea. *J. Mar. Syst.* 69 (1–2), 99–113.
- Carson, R.T., Mitchell, R.C., Hanemann, M., Kopp, R.J., Presser, S., Ruud, P.A., 2003. Contingent valuation and lost passive use: Damages from the Exxon Valdez oil spill. *Environ. Resour. Econ.* 25, 257–286.
- Chaturvedi, S., Yadav, B.P., Siddiqui, N.A., Chaturvedi, S.K., 2020. Mathematical modelling and analysis of plastic waste pollution and its impact on the ocean surface. *J. Ocean. Eng. Sci.* 5 (2), 136–163.
- Chen, J., Di, Z., Shi, J., Shu, Y., Wan, Z., Song, L., Zhang, W., 2020. Marine oil spill pollution causes and governance: A case study of Sanchi Tanker collision and explosion. *J. Clean. Prod.* 273, 122978.
- Cho, K., Reid, R.O., Nowlin, Jr., W.D., 1998. Objectively mapped stream function fields on the Texas–Louisiana shelf based on 32 months of moored current meter data. *J. Geophys. Res.: Ocean.* 103 (C5), 10377–10390.
- Cosoli, S., Ličer, M., Vodopivec, M., Malačič, V., 2013. Surface circulation in the gulf of trieste (northern Adriatic Sea) from radar, model, and adcp comparisons. *J. Geophys. Res.: Ocean.* 118 (11), 6183–6200.
- Elliott, A., Jones, B., 2000. The need for operational forecasting during oil spill response. *Marine Poll. Bull.* 40 (2), 110–121.
- Eremeev, V., Ivanov, L., Kirwan, Jr., A., 1992. Reconstruction of oceanic flow characteristics from quasi-Lagrangian data: 1. Approach and mathematical methods. *J. Geophys. Res.: Ocean.* 97 (C6), 9733–9742.

- Ferziger, J., Peric, M., 2012. Computational Methods for Fluid Dynamics. Springer, Berlin Heidelberg. URL <https://books.google.hr/books?id=BZnvCAAQBAJ>.
- Foundation, T.O., 2016. Openfoam user guide. <https://www.openfoam.com/documentation/user-guide>. [Online; (Accessed 04 September 2024)].
- García-Garrido, V., Ramos, A., Mancho, A., Coca, J., Wiggins, S., 2016. A dynamical systems perspective for a real-time response to a marine oil spill. *Marine Poll. Bull.* 112 (1–2), 201–210.
- Ghalambaz, S., Abbaszadeh, M., Sadrehaghghi, I., Younis, O., Ghalambaz, M., Ghalambaz, M., 2024. A forty years scientometric investigation of artificial intelligence for fluid-flow and heat-transfer (aifh) during 1982 and 2022. *Eng. Appl. Artif. Intell.* 127, 107334.
- Grossi, M.D., Kubat, M., Özgökmen, T.M., 2020. Predicting particle trajectories in oceanic flows using artificial neural networks. *Ocean. Model.* 156, 101707.
- Gucel, M.U., Sakalli, A., 2024. Analyzing the Mediterranean Sea's dynamic current system and modeling of renewable current energy potential. *J. Mar. Sci. Eng.* 12 (4), 671.
- Gulakaram, V.S., Vissa, N.K., Bhaskaran, P.K., 2018. Role of mesoscale eddies on atmospheric convection during summer monsoon season over the Bay of Bengal: A case study. *J. Ocean. Eng. Sci.* 3 (4), 343–354.
- Gunzburger, M.D., 2012. Finite Element Methods for Viscous Incompressible Flows: A Guide to Theory, Practice, and Algorithms. Elsevier.
- Haza, A.C., D'Asaro, E., Chang, H., Chen, S., Curic, M., Guigand, C., Huntley, H., Jacobs, G., Novelli, G., Özgökmen, T., et al., 2018. Drogue-loss detection for surface drifters during the lagrangian submesoscale experiment (laser). *J. Atmos. Ocean. Technol.* 35 (4), 705–725.
- Hernandez, F., Le Traon, P.-Y., Morrow, R., 1995. Mapping mesoscale variability of the azores current using topex/poseidon and ers 1 altimetry, together with hydrographic and Lagrangian measurements. *J. Geophys. Res.: Ocean.* 100 (C12), 24995–25006.
- Jakac, K., Lanča, L., Sikirica, A., Ivić, S., 2024. Approximation of sea surface velocity field by fitting surrogate two-dimensional flow to scattered measurements. arXiv preprint arXiv:2401.12746.
- Juretic, F., 2015. cfmesh User Guide. Creative Fields, Ltd, p. 1.
- Kirby, M.F., Law, R.J., 2010. Accidental spills at sea—risk, impact, mitigation and the need for co-ordinated post-incident monitoring. *Marine Poll. Bull.* 60 (6), 797–803.
- Kovačević, V., Gačić, M., Mosquera, I.M., Mazzoldi, A., Marinetti, S., 2004. Hf radar observations in the northern Adriatic: Surface current field in front of the Venetian Lagoon. *J. Mar. Syst.* 51 (1–4), 95–122.
- Kronbichler, M., 2009. Numerical Methods for the Navier–Stokes Equations Applied to Turbulent Flow and to Multi-Phase Flow (Ph.D. thesis). Uppsala University.
- Li, Y., Yu, H., Wang, Z.-y., Li, Y., Pan, Q.-q., Meng, S.-j., Yang, Y.-q., Lu, W., Guo, K.-x., 2019. The forecasting and analysis of oil spill drift trajectory during the Sanchi collision accident, east China Sea. *Ocean Eng.* 187, 106231.
- Lions, P.-L., 1996. Mathematical Topics in Fluid Mechanics: Volume 2: Compressible Models. In: Oxford Lecture Mathematics, vol. 2.
- Lorente, P., Piedracoba, S., Sotillo, M., Aznar, R., Amo-Balandron, A., Pascual, A., Soto-Navarro, J., Álvarez-Fanjul, E., 2016. Characterizing the surface circulation in ebro delta (nw Mediterranean) with hf radar and modeled current data. *J. Mar. Syst.* 163, 61–79.
- Mariano, A., Kourafalou, V., Srinivasan, A., Kang, H., Halliwell, G., Ryan, E., Roffer, M., 2010. On the modeling of the 2010 gulf of Mexico oil spill. *Dyn. Atmos. Ocean.* 52 (1–2), 322–340.
- Marine, A., 2024. Portable ais, tb560. <https://www.alltekmarine.com/products/portable-ais/tb560>. [Online;(Accessed 18 December 2024)].
- Marmain, J., Molcard, A., Forget, P., Barth, A., Ourmieres, Y., 2014. Assimilation of hf radar surface currents to optimize forcing in the northwestern Mediterranean sea. *Nonlinear Process. Geophys.* 21 (3), 659–675.
- Menter, F.R., 1992. Improved Two-Equation K-Omega Turbulence Models for Aerodynamic Flows. Tech. Rep., NASA Ames Research Center Moffett Field, CA, United States.
- Molcard, A., Piterbarg, L.I., Griffa, A., Özgökmen, T.M., Mariano, A.J., 2003. Assimilation of drifter observations for the reconstruction of the Eulerian circulation field. *J. Geophys. Res.: Ocean.* 108 (C3).
- Notarstefano, G., Poulain, P.-M., Mauri, E., 2008. Estimation of surface currents in the Adriatic Sea from sequential infrared satellite images. *J. Atmos. Ocean. Technol.* 25 (2), 271–285.
- Olascoaga, M.J., Haller, G., 2012. Forecasting sudden changes in environmental pollution patterns. *Proc. Natl. Acad. Sci.* 109 (13), 4738–4743.
- Paduan, J.D., Washburn, L., 2013. High-frequency radar observations of ocean surface currents. *Annu. Rev. Mar. Sci.* 5 (1), 115–136.
- Patankar, S.V., Spalding, D.B., 1983. A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows. In: Numerical Prediction of Flow, Heat Transfer, Turbulence and Combustion. Elsevier, pp. 54–73.
- Poulain, P.-M., 2001. Adriatic Sea surface circulation as derived from drifter data between 1990 and 1999. *J. Mar. Syst.* 29 (1–4), 3–32.
- Qazi, W.A., Emery, W.J., Fox-Kemper, B., 2014. Computing ocean surface currents over the coastal California current system using 30-min-lag sequential sar images. *IEEE Trans. Geosci. Remote Sens.* 52 (12), 7559–7580.
- Rao, D.B., Schwab, D.J., 1981. A method of objective analysis for currents in a lake with application to Lake Ontario. *J. Phys. Oceanogr.* 11 (5), 739–750.

- Reed, M., Johansen, Ø., Brandvik, P.J., Daling, P., Lewis, A., Fiocco, R., Mackay, D., Prentki, R., 1999. Oil spill modeling towards the close of the 20th century: Overview of the state of the art. *Spill Sci. Technol. Bull.* 5 (1), 3–16.
- Roache, P.J., 1998. *Verification and Validation in Computational Science and Engineering*, vol. 895, Hermosa Albuquerque, NM.
- Sayol, J.M., Orfila, A., Simarro, G., Conti, D., Renault, L., Molcard, A., 2014. A lagrangian model for tracking surface spills and sar operations in the ocean. *Environ. Model. Softw.* 52, 74–82.
- Solano, M., Canals, M., Leonardi, S., 2018. Development and validation of a coastal ocean forecasting system for Puerto Rico and the us virgin islands. *J. Ocean. Eng. Sci.* 3 (3), 223–236.
- Song, M., Hu, W., Liu, S., Chen, S., Fu, X., Zhang, J., Li, W., Xu, Y., 2024. Developing an artificial intelligence-based method for predicting the trajectory of surface drifting buoys using a hybrid multi-layer neural network model. *J. Mar. Sci. Eng.* 12 (6), 958.
- Sun, L., Gao, H., Pan, S., Wang, J.-X., 2020. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. *Comput. Methods Appl. Mech. Engrg.* 361, 112732.
- Tang, M., Liu, Y., Durlófsky, L.J., 2021. Deep-learning-based surrogate flow modeling and geological parameterization for data assimilation in 3d subsurface flow. *Comput. Methods Appl. Mech. Engrg.* 376, 113636.
- Tkalich, P., 2006. A cfd solution of oil spill problems. *Environ. Model. Softw.* 21 (2), 271–282.
- Ursella, L., Poulain, P.-M., Signell, R.P., 2006. Surface drifter derived circulation in the northern and middle Adriatic Sea: Response to wind regime and season. *J. Geophys. Res.: Ocean.* 111 (C3).
- Wang, J., Shen, Y., 2010. Modeling oil spills transportation in seas based on unstructured grid, finite-volume, wave-ocean model. *Ocean. Model.* 35 (4), 332–344.