



**HAL**  
open science

# Neural prediction of Lagrangian drift trajectories on the sea surface

Daria Botvynko, Carlos Granero-Belinchon, Simon Van Gennip, Abdesslam Benzinou, Ronan Fablet

► **To cite this version:**

Daria Botvynko, Carlos Granero-Belinchon, Simon Van Gennip, Abdesslam Benzinou, Ronan Fablet.  
Neural prediction of Lagrangian drift trajectories on the sea surface. 2024. hal-04569385

**HAL Id: hal-04569385**

**<https://hal.science/hal-04569385>**

Preprint submitted on 6 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



8 ABSTRACT: We simulate Lagrangian drift on the sea surface and investigate deep learning ap-  
9 proaches to address the shortcomings of current model-based and Markovian approaches, particu-  
10 larly concerning error propagation and computational complexity. We present a novel deep learn-  
11 ing framework, referred to as DriftNet, inspired by the Eulerian Fokker-Planck representation of  
12 Lagrangian dynamics. Through numerical experiments for simulated and real drift trajectories on  
13 the sea surface, we illustrate the effectiveness of DriftNet compared to existing state-of-the-art  
14 schemes. We also delve into the influence of diverse geophysical fields, whether derived from  
15 models or observations, used as inputs by DriftNet on drift simulation. Our objective is to assess  
16 the amount of dynamic information required to accurately simulate realistic trajectories.

17 SIGNIFICANCE STATEMENT: We propose a novel Deep Learning model, DriftNet, for con-  
18 ditional generation of Lagrangian trajectories on the sea surface. Our model is based on Eulerian  
19 Fokker-Planck formalism and can be conditioned by multiple geophysical fields. We highlight the  
20 overall over-performance of DriftNet compared to the baseline model-based and learning-based  
21 approaches. We put in evidence the capacity of the proposed method to extract pertinent infor-  
22 mation from various geophysical fields, both from modeled and observed data. We highlight the  
23 significant impact of the observed sea surface height when combined to sea surface currents in the  
24 quality of the generation of Lagrangian trajectories.

## 25 **1. Introduction**

26 The modeling and simulation of Lagrangian drift on the sea surface holds relevance for various  
27 applications such as the tracking of plastics and other debris Maximenko et al. (2012), the study  
28 of algae and plankton dynamics Son et al. (2015), or the prediction of future locations of drifting  
29 objects that is crucial for search and rescue operations Breivik et al. (2013). Furthermore, the  
30 analysis of Lagrangian drifts serves as a valuable tool for diagnosing ocean numerical models in  
31 their capability for reproducing small-scale dynamics Barron et al. (2007). However, reproducing  
32 realistic Lagrangian trajectories on the sea surface poses a significant scientific challenge within  
33 operational oceanography Röhrs et al. (2021).

34 From a methodological standpoint, we can distinguish three main categories of approaches  
35 for Lagrangian drift simulation: model-based approaches Liubartseva et al. (2018); Zambianchi  
36 et al. (2017), probabilistic data-driven schemes Visser (2008), and more recent deep learning tech-  
37 niques Botvynko et al. (2023); Jenkins et al. (2023). Model-based approaches involve a sequen-  
38 tial advection process based on known sea surface velocity fields Lange and van Sebille (2017).  
39 However, small errors in the underlying velocity fields or velocity fields without very fine spa-  
40 tial resolution may lead to unrealistic Lagrangian trajectories Callies et al. (2021). This often  
41 impedes their operational usefulness due to significant uncertainties in the forecasting and recon-  
42 struction of sea surface dynamics, especially regarding mesoscale ocean dynamics Prants et al.  
43 (2017). Due to their sequential nature, these schemes also face scalability challenges when simu-  
44 lating large ensembles of drift trajectories. By contrast, probabilistic data-driven schemes usually  
45 leverage first-order Markovian models and naturally account for uncertainties in the drift process.

46 However, they are primarily suited for relatively coarse space-time resolutions. Extending these  
47 schemes to capture fine-scale patterns poses a challenge Fine et al. (1998). For these two first cat-  
48 egories, the simulation of the Lagrangian drift relies on location-wise velocities at each time step,  
49 which may only be applicable for smooth velocity fields. Deep Learning has recently emerged as  
50 a novel class of numerical tools for the learning-based simulation of movement patterns. We may  
51 cite applications to pedestrians Korbmacher and Tordeux (2022), seabirds Roy et al. (2022), cars  
52 Jiang et al. (2019) or maritime vessels Nguyen and Fablet (2024). These applications leverage  
53 state-of-the-art recurrent neural networks, such as Long Short-Term Memory (LSTM) networks  
54 Yu et al. (2019), as well as deep generative schemes Goodfellow et al. (2020); Roy et al. (2022);  
55 Ma et al. (2019); Gan et al. (2020); Julka et al. (2021); Paz et al. (2021). Most of these studies can  
56 be regarded as neural extensions of probabilistic Markovian schemes.

57 In this context, we present an innovative Deep Learning framework, referred to as DriftNet, for  
58 simulating Lagrangian trajectories at the sea surface. The proposed neural model takes as input  
59 geophysical fields characterizing the ocean dynamics in a spatio-temporal region and produces  
60 the associated Lagrangian trajectory. Our approach is fully-convolutional and features a spatially-  
61 explicit latent representation, drawing inspiration from Eulerian Fokker-Planck representations of  
62 drift processes Botvynko et al. (2023). From numerical simulation datasets, we demonstrate Drift-  
63 Net to outperform state-of-the-art neural schemes for the simulation of Lagrangian trajectories,  
64 including when considering degraded sea surface velocities. The application to real sea surface  
65 drifters’ datasets further supports the relevance of DriftNet to improve the simulation of Lagra-  
66 gian drift on the sea surface compared with model-driven simulations. This last experiment also  
67 illustrates how DriftNet can exploit additional data sources, besides sea surface velocities, here  
68 satellite-derived sea surface height fields.

69 This paper is organized as follows: section 2 focuses on the problem description; we introduce  
70 DriftNet in section 3 and the considered experimental setup in section 4; section 5 details our  
71 results and section 6 discusses key aspects of this study.

## 72 **2. Problem statement**

73 In the context of Lagrangian fluid dynamics<sup>1</sup>, one may describe a flow by following the motion  
74 of an ensemble of individual particles Davis (1991). In order to describe the flow in the Lagrangian

---

<sup>1</sup>From now on, the Lagrangian notations are presented with  $\rightarrow$ symbol, and the Eulerian notations are in **bold**.

75 framework, an ordinary differential equation defines the particle's trajectory :

$$\frac{\partial \vec{r}(\vec{r}_{i_0}, t)}{\partial t} = \vec{v}(\vec{r}_{i_0}, t) = \mathbf{u}(\vec{r}(t), t) \quad (1)$$

76 where  $\vec{r}(\vec{r}_{i_0}, t)$  and  $\vec{v}(\vec{r}_{i_0}, t)$  are respectively the position and velocity at time  $t$  of the particle situated  
 77 at  $\vec{r}_{i_0}$  at  $t = 0$ , and  $\mathbf{u}(\vec{r}(t), t)$  is the Eulerian velocity of the underlying flow given position  $\vec{r}(t)$  and  
 78 time  $t$ . Classical model-driven approaches simulate Lagrangian drift dynamics through explicit  
 79 time and space integration methods such as Runge-Kutta 4. Thus, the integration of eq.(1) leads  
 80 to :

$$\vec{r}(\vec{r}_{i_0}, t + \Delta t) = \vec{r}(\vec{r}_{i_0}, t) + \int_t^{t+\Delta t} \vec{v}(\vec{r}_{i_0}, \tau) d\tau \quad (2)$$

81 The sequential nature of this equation introduces important propagation of errors when the under-  
 82 lying velocities are not accurately known Callies et al. (2021).

83 Interestingly, we can also derive an Eulerian formulation of Lagrangian dynamics. Through  
 84 Fokker-Planck formalism Visser (2008), we can generalize eq.(2) to the time propagation of the  
 85 Probability Density Function (PDF) of the moving particle as follows:

$$\frac{\partial}{\partial t} p_{\vec{r}}(x, t) = -\frac{\partial}{\partial x} [\mu(x, t) p_{\vec{r}}(x, t)] \quad (3)$$

86 where  $x$  represents spatial position in the Eulerian framework,  $t$  the time,  $p_{\vec{r}}(x, t)$  the PDF of the  
 87 particle at position  $\vec{r}$  and  $\mu(x, t)$  a drift field. In our case, this drift term relates to the underlying  
 88 Eulerian velocity field  $\mathbf{u}(x, t)$ . Solving eq.(3) relies on a time integration scheme from the initial  
 89 PDF  $p_{\vec{r}}(x, t = 0) = p_{\vec{r}_{i_0}}(x, 0)$ .

90 Here, we rely on this Eulerian Fokker-Planck representation of Lagrangian dynamics to explore  
 91 convolutional neural architectures for the simulation of individual drift trajectories.

### 92 3. DriftNet for Lagrangian drift simulation

93 This section introduces the mathematical framework and the architecture of the proposed neural  
 94 network model, called DriftNet, along with the considered learning scheme.

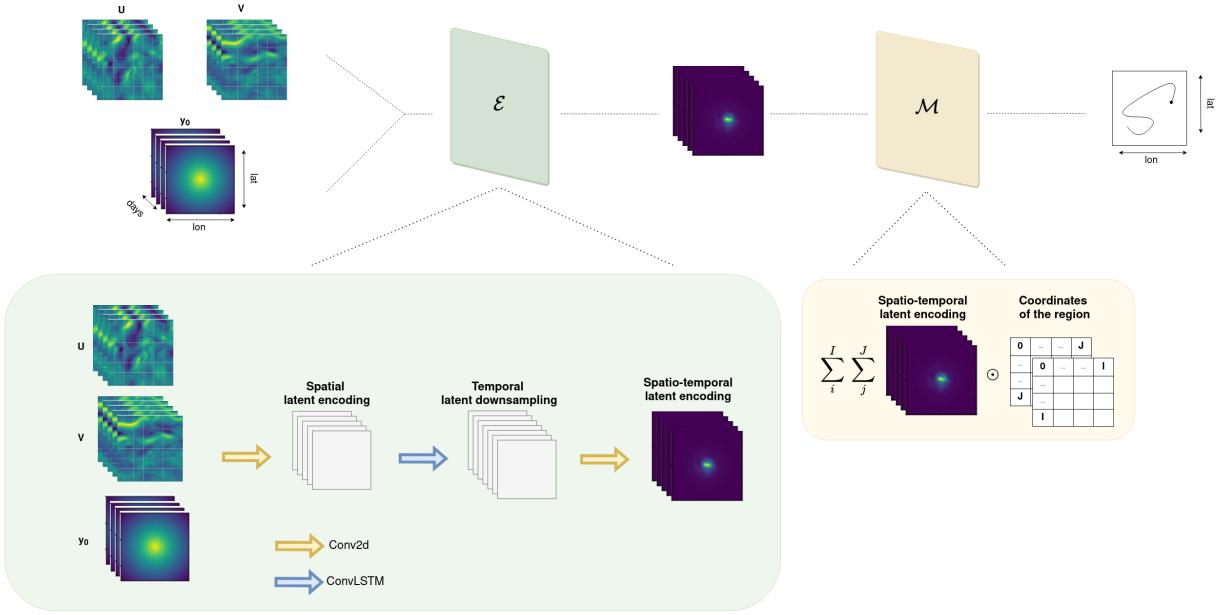


FIG. 1: **DriftNet architecture:** The input geophysical fields over 9-day (here zonal  $\mathbf{U}$  and meridional  $\mathbf{V}$  components of the velocity field) coupled to the initial spatio-temporal positional encoding  $\mathbf{y}_0$  are first passed to the encoding block  $\mathcal{E}$  and output a spatio-temporal latent encoding. The second block of DriftNet  $\mathcal{M}$  maps this latent representation to a 1D trajectory, *i.e.* a time series of positions on the sea surface.

95 *a. Proposed neural scheme*

96 Given a spatial domain of interest  $\mathcal{D}$  defined as a regularly-gridded domain of size  $I \times J$ , we  
 97 define the associated sequence of velocity fields  $\mathbf{u} = \{\mathbf{u}_{t_0}, \mathbf{u}_{t_0+\Delta}, \dots, \mathbf{u}_{t_0+K\Delta}\}$  from time  $t_0$  to  $t_0 +$   
 98  $K\Delta$ , where  $\Delta$  is the time sampling and  $K$  the number of time steps, and the trajectory of a particle as  
 99 a time series of locations  $\vec{r} = \{\vec{r}_{t_0}, \vec{r}_{t_0+\Delta}, \dots, \vec{r}_{t_0+K\Delta}\}$  in  $\mathcal{D}$  from  $t_0$  to  $t_0 + K\Delta$ , where  $\vec{r}_{t_0}$  corresponds  
 100 to the initial position of the particle.

101 Inspired by Fokker-Plank eq.(3), we consider the following latent representation for a simulated  
 102 Lagrangian drift  $\vec{r}$ :

$$\begin{cases} \mathbf{y} = \mathcal{E}(\mathbf{u}, \mathbf{y}_0) \\ \vec{r} = \mathcal{M}(\mathbf{y}) \end{cases} \quad (4)$$

103 where  $\mathbf{y} = \{\mathbf{y}_{t_0}, \mathbf{y}_{t_0+\Delta}, \dots, \mathbf{y}_{t_0+K\Delta}\}$  is a space-time-explicit latent embedding of  $\vec{r}$ , and  $\mathbf{y}_0 =$   
 104  $\{\mathbf{y}_{0,t_0}, \mathbf{y}_{0,t_0+\Delta}, \dots, \mathbf{y}_{0,t_0+K\Delta}\}$  some initial encoding of the initial position  $\vec{r}_{t_0}$ . In particular, we define  
 105  $\mathbf{y}_{0,t_0}$  by assigning to each spatial pixel of the grid a value corresponding to the normalized distance

106 to the initial position of the particle eq.(5).

$$d = -\sqrt{\sum_{x \in \mathcal{D}} (\vec{r}_{t_0} - x)^2} \quad (5)$$

$$\mathbf{y}_{0,t_0} = \frac{d + |\min d|}{\max(d + |\min d|)} \quad (6)$$

107 while the other time steps of  $\mathbf{y}_0$  are initialized as all-ones matrix of the same dimension as  $\mathcal{D}$ .

108 The operator  $\mathcal{E}$  computes a latent embedding  $\mathbf{y}$  given velocity conditions  $\mathbf{u}$  and initial repre-  
 109 sentation  $\mathbf{y}_0$ , as introduced in eq.(3). And, the operator  $\mathcal{M}$  maps latent representation  $\mathbf{y}$  to the  
 110 targeted Lagrangian drift  $\vec{r}$ . In other words, it maps a multi-dimensional tensor to a time series of  
 111 locations in  $\mathcal{D}$ . The analytical expression of the operator  $\mathcal{M}$  for a given time step  $t$  grounds on the  
 112 Hadamard product between the spatially-explicit encoding  $x$  of the space  $\mathcal{D}$  and  $\mathbf{y}_t$  field:

$$\mathcal{M} = x \cdot \mathbf{y}_t \quad (7)$$

113 where  $\mathbf{y}_t \in \mathbf{y}$ , a sample of  $\mathbf{y}$  at a time step  $t \in [t_0, \dots, t_{0+K\Delta}]$ .

114 Regarding the operator  $\mathcal{E}$ , we consider the architecture sketched in Figure 1 and composed of:

- 115 1) a 2D convolutional layer increasing the number of channels from 3 (zonal and meridional  
 116 components of  $\mathbf{u}$  coupled to  $\mathbf{y}_0$ ) to 64, followed by a LeakyReLU activation function.
- 117 2) 2D convolutional LSTM block Shi et al. (2015).
- 118 3) 2D convolutional layer decreasing the number of channels from 64 to 2, followed by 2D  
 119 Softmax function.

120 Since our model is mainly based on 2D convolutional layers, we expect the resulting architecture  
 121 to capture relevant information at different space-time scales, and not only in a point-wise manner,  
 122 as it would be the case for the straightforward implementation of Fokker-Plank representation  
 123 eq.(3). Thus, our network can leverage non-local physical information explaining the particle's  
 124 motion and processes it to propagate the latent representation of particle's positions through time.

125 Interestingly, the proposed architecture easily generalizes to multi-source input data. Gridded  
 126 fields such as SSH fields with the same space-time resolution as the sea surface velocities convey  
 127 relevant information on sea surface dynamics. Stacking SSH and sea surface velocities, we may

128 apply the same neural architecture, just increasing the number of channels of the input fields. The  
 129 same applies for a prior on the diffusivity from the initial location of a particle in the field, that can  
 130 be also included by stacking with the temporal evolution of the initial condition  $\mathbf{y}_0$ .

131 *b. Learning scheme*

132 We consider a supervised training of the proposed neural architecture according to the following  
 133 two losses:

- 134 1. the Mean Square Error (MSE) between the reference and simulated trajectories for the  $K$ -step  
 135 simulation of  $N_T$  particles

$$\mathcal{L}_{MSE} = \frac{1}{N_T K} \sum_{i=0}^{N_T} \sum_{j=0}^K (\vec{r}_R(\vec{r}_{0,i}, j\Delta) - \vec{r}_S(\vec{r}_{0,i}, j\Delta))^2 \quad (8)$$

136 where  $\vec{r}_R$  and  $\vec{r}_S$  are the vector of positions of the reference and simulated trajectories respec-  
 137 tively.

- 138 2. Liu index between the reference and simulated trajectories Liu and Weisberg (2011)

$$\mathcal{L}_{Liu} = \frac{1}{N_T} \sum_{i=1}^{N_T} \frac{\sum_{j=1}^K d_{ij}}{\sum_{j=1}^K l_{ij}} \quad (9)$$

139 where

$$d_{ij} = \sqrt{(\vec{r}_R(\vec{r}_{0,i}, j\Delta) - \vec{r}_S(\vec{r}_{0,i}, j\Delta))^2} \quad (10)$$

140 is the Euclidean distance between the reference and the simulated trajectories number  $i$  at  
 141 time step  $j$  and  $l_{ij}$  is the length of reference trajectory  $i$  between the initial position and the  
 142 position at time step  $j$ .

143 Overall, the training loss  $\mathcal{L}$  is a weighted sum:  $\mathcal{L}_{Liu} = \alpha \cdot \mathcal{L}_{MSE} + \beta \cdot \mathcal{L}_{Liu}$ . From cross-validaton  
 144 experiments, we set  $\alpha$  and  $\beta$  to 0.2 and 0.8. Using Pytorch<sup>2</sup>, our learning setup relies on Adam  
 145 optimizer with a learning rate of 5e-3 over 750 epochs.

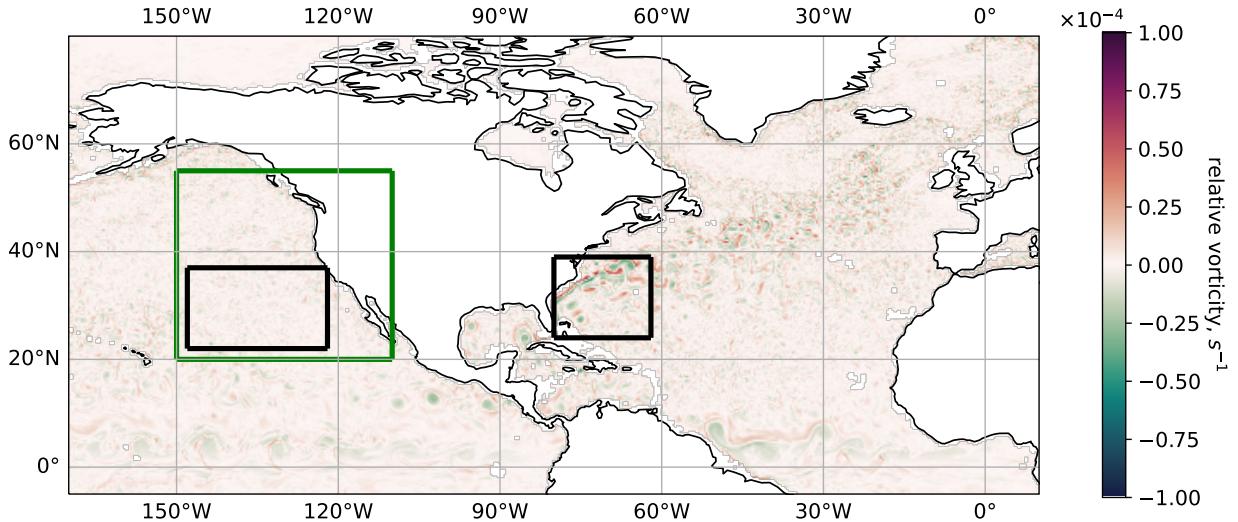


FIG. 2: **Study cases delineation:** Relative vorticity of North East Pacific and North Atlantic oceans from E1 simulation of 01/10/2015. The two black boxes indicate the North East Pacific and Gulf Stream case study regions for benchmarks B1 and B2. The green box defines the North East Pacific region of study for benchmark B3.

#### 146 4. Experimental workflow

147 Section 4.a details the two case-study regions considered because of their contrasting turbulent  
 148 activity: the Gulf Stream and the North-East Pacific region, see Figure 2. We describe the different  
 149 datasets used in our experiments in section 4.b, see Table 1, and the metrics used to evaluate  
 150 performance in section 4.c. In section 4.d we present the benchmarked deep learning models. We  
 151 perform three different experiments: two relying on numerical simulation datasets such that the  
 152 full sea surface velocity field is known, and a last experiment dealing with real drifters in the sea  
 153 surface. These three benchmarks are explained in section 4.e.

##### 154 a. Case-study Regions

155 **The Gulf Stream** is a Western boundary current system characterized by a fast flowing jet that  
 156 originates in the Gulf of Mexico and flows into the North Atlantic Gula et al. (2015); Dewar  
 157 and Bane (1989), becoming a part of it’s anticyclonic subtropical gyre Talley (2011). Surface  
 158 velocity can exceed  $2 \text{ m} \cdot \text{s}^{-1}$  and the current is characterized by high eddy variability. Exhibiting  
 159 strong velocities for hundreds of kilometers, the Gulf Stream transports warm saline water from  
 160 lower latitudes north-eastward. The Gulf Stream plays a crucial role in redistributing heat and

<sup>2</sup>DriftNet code is available at <https://github.com/CIA-Oceanix/DriftNet>

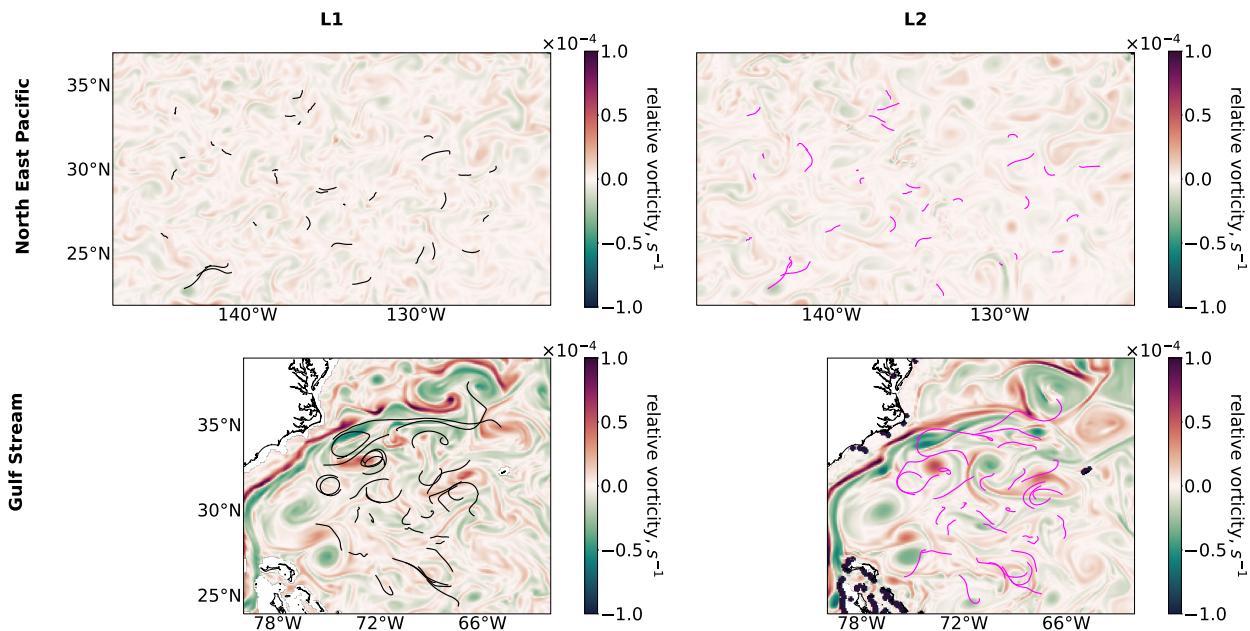


FIG. 3: **Examples of trajectories:** Trajectories from dataset L1 and dataset L2 superimposed respectively to the E1 and E2 relative vorticity fields of 2015/01/01.

161 influencing climate patterns and the surrounding marine ecosystem Gula et al. (2015), Wills et al.  
 162 (2016).

163 **The North East Pacific** is a vast oceanic region extending from the western coast of North  
 164 America to the central Pacific. It is influenced by various ocean currents, including the California  
 165 Current and North Pacific Current. The California Current System (CCS) is an eastern boundary  
 166 of the slow ( $0.1 \text{ m} \cdot \text{s}^{-1}$ ) eastward North Pacific Current. It is a cold, nutrient-rich current that  
 167 flows southward along the Western coast of North America Checkley and Barth (2009); Aua  
 168 et al. (2011), impacting marine ecosystems and fisheries. The CCS is an upwelling region due to  
 169 Ekman transport, spinning up eddies westward offshore. It has a mean surface velocity of  $0.4 - 0.8$   
 170  $\text{m} \cdot \text{s}^{-1}$ . Due to the baroclinic instability of the coastal upwelling current, the CCS contains high  
 171 mesoscale eddy activity, associated with the highly variable sea-surface height Talley (2011).

## 172 *b. Datasets*

173 Our study involves three different Eulerian datasets used as input data for the benchmarked  
 174 models:

TABLE 1: **Summary description of the Eulerian and Lagrangian datasets used in the current study:** *Source* makes reference to the used numerical or satellite product, *Variable* indicates the physical quantity of interest, *Period* the studied years. The spatial and temporal resolutions of each product are also indicated. SSC - sea surface currents. SSH - sea surface height.

<i>Name</i>	<i>Source</i>	<i>Variable</i>	<i>Period</i>	<i>Spatial resolution</i>	<i>Temporal resolution</i>
Eulerian					
E1	Nature Run	SSC and SSH	2015	1/12°	1 day
E2	OSSE-based assimilated free run	SSC	2015	1/12°	1 day
E3	Operational assimilated reanalysis GLORYS12	SSC	1992 - 2020	1/12°	1 day
E3	DUACS	SSH	1992 - 2020	1/4°	1 day
Lagrangian					
L1	Ocean Parcels simulation on E1 SSC	Position	2015	-	6 hours
L2	Ocean Parcels simulation on E2 SSC	Position	2015	-	6 hours
L3	CMEMS drifters	Position	1992 - 2020	-	6 hours
L4	Ocean Parcels simulation on E3 SSC	Position	1992 - 2020	-	6 hours

- 175 • **Eulerian dataset E1:** This Eulerian dataset contains the sea surface current (SSC) velocity  
176 and sea surface height (SSH) fields. Both fields are obtained from a high-resolution free  
177 simulation of the NEMO ocean model without any data assimilation Benkiran et al. (2021b).  
178 Those are regularly-gridded products with a horizontal spatial resolution of 1/12° and a daily  
179 time resolution. This study involves the entire year of 2015.
- 180 • **Eulerian dataset E2** This dataset contains the SSC fields from a data assimilated run, which  
181 dynamically reconstructs the ocean state using pseudo observations from E1 which mimic  
182 the current observational network (satellite altimetry tracks). Those fields are at the same  
183 1/12° and daily space-time resolution as the E1 dataset. As in the previous E1 dataset, we  
184 equivalently use the complete 2015 year.
- 185 • **Eulerian Real-world dataset E3:** this real-world dataset combines the SSC from the oper-  
186 ational ocean reanalysis GLORYS12 Lellouche et al. (2021) and the optimally-interpolated  
187 altimetry-derived SSH product DUACS Pujol et al. (2016). The GLORYS12 velocity fields  
188 are provided on the regular grid with spatial resolution of 1/12° and temporal resolution of  
189 1 day. We re-interpolate the SSH fields on the same 1/12° regular grid at daily temporal  
190 resolution.

191 For each individual trajectory, the considered spatial domain of the Eulerian conditional fields is  
192 equal to,  $4^\circ \times 4^\circ$  window for the North East Pacific and  $10^\circ \times 10^\circ$  for the Gulf Stream. The size of  
193 the domains corresponds to the furthest displacement from a given particle's deployment location  
194 in each case study region.

195 Our benchmarks involve four different Lagrangian datasets:

196 • **Lagrangian Dataset L1:** this dataset is composed of 9-day trajectories of virtual particles  
197 advected by the sea surface velocities from dataset E1 using Ocean Parcels, a state-of-the-art  
198 tool for model-based Lagrangian drift simulation Lange and van Sebille (2017). Around 20  
199 particles were randomly seeded every day across the domain and the studied period, amount-  
200 ing to a total of 7900 particles.

201 • **Lagrangian Dataset L2:** this dataset contains 9-day trajectories of virtual particles advected  
202 by the SSC of the E2 dataset with Ocean Parcels. It contains 7900 particles deployed at the  
203 exact same locations and times as in the Lagrangian Dataset L1 described above, see Figure 3.

204 • **Lagrangian Dataset L3** This dataset contains trajectories from CMEMS drifters Etienne  
205 et al. (2023) in the North East Pacific region, from 1992 to 2020. Those drifters are a part  
206 of the Surface Velocity Program (SVP) and are built of 1-m long floating buoy attached to a  
207 “drogue” deployed at 15 meters depth in order to reduce the windage slip Lumpkin et al.  
208 (2017). Drifters positions are subject to a quality control and editing procedures before  
209 being operational. Subsequently, trajectory positions undergo regular reinterpolation using  
210 the Kriging technique at 6-hour intervals Etienne et al. (2023). For this study, we refer to  
211 drogued-only drifters which are representative of the ocean currents at 15 meters depth and  
212 are widely used in various application domains Lumpkin et al. (2012); Koszalka et al. (2011);  
213 Lumpkin and Flament (2013); Kaplan et al. (2005); Lumpkin and Garzoli (2011).

214 This dataset contains 12.785 9-day trajectories in the North East Pacific case-study region,  
215 green box in Figure 2. These trajectories are obtained by dividing the whole drifter trajectories  
216 of the region in 9-day long non-overlapping segments.

217 • **Lagrangian Dataset L4:** this dataset contains 9-day trajectories of virtual particles advected  
218 by the velocity fields from E3 with Ocean Parcels. We use the same initial deployment posi-  
219 tion and time of the trajectories as in dataset L3.

220 *c. Performance metrics*

221 As evaluation metrics, we consider

- 222 • the mean Euclidean distance at the last time step (9-th day) in kilometers, that we denote  $D$ :

$$D = \frac{1}{N_T} \cdot \sum_{i=0}^{N_T} d_{i_{9\Delta}} \quad (11)$$

223 with  $N_T$  the number of trajectories and  $d_{i_{9\Delta}}$  defined in eq.(10) with  $j = 9$  days.

- 224 • the Mean Liu index  $\mathcal{L}_{Liu}$  without units, see eq.(9).
- 225 • the Mean absolute error between Lagrangian Time Scales  $\Delta\mathcal{T}$  Krauß and Böning (1987), in
- 226 days:

$$\Delta\mathcal{T} = \frac{1}{N_T} \cdot \sum_{i=0}^{N_T} |\mathcal{T}_{R_i} - \mathcal{T}_{S_i}| \quad (12)$$

227 where  $\mathcal{T}_{R_i}$  and  $\mathcal{T}_{S_i}$  are respectively the Lagrangian time of the reference and simulated  $i$ -th  
228 trajectory defined as:

$$\mathcal{T} = \int_{t_0}^{t_0+9\Delta} R_{\vec{v}}(\tau) d\tau \quad (13)$$

229 where  $R_{\vec{v}}(\tau) = \frac{1}{9\Delta} \cdot \sum_{t=0}^9 \vec{v}(t\Delta + \tau)\vec{v}(t\Delta)$  is the autocorrelation function of the Lagrangian  
230 velocity of a given particle.

231 These three metrics are computed between the reference and generated trajectories and averaged  
232 over the whole ensemble of studied trajectories.

233 *d. Benchmarked deep learning models*

234 To evaluate the performance of DriftNet, we compare it to the following CNN and LSTM base-  
235 lines inspired by Zheng et al. (2022); Ma et al. (2019):

- 236 • Convolutional Neural Network (CNN): a 2D convolutional block varying the number of chan-  
237 nels from 27 ( $\mathbf{u}$  and  $\mathbf{y}_0$  concatenated along the time dimension) to  $8 \rightarrow 16 \rightarrow 32 \rightarrow 35$ . Each

convolution layer has a kernel of size 5 and is followed by a LeakyReLU activation function with a slope of 0.2 for negative inputs. The output of this convolutional block is introduced by a reshape, where the two spatial dimensions of initial 3d matrix are flattened together, and the output is passed to a fully connected layer with 256 and 128 neurons decreasing the mentioned spatial dimension to 2.

- Long Short-Term Memory (LSTM): a 5-layer LSTM block with hidden size 96, followed by a ReLU activation function, then two fully connected layers with a ReLU in between. Each dense layer contains 864, and 128 neurons respectively, reducing the latent space to 2·35. This output is then reshaped so to contain the spatial dimension 2 and temporal one 35.

We also explore different configurations of DriftNet. In order to investigate the impact of each component of the model on the evaluation metrics, we modify its inner architecture, see section 3. a and Figure 1, and assess the performance of the following three architectures:

- DriftNet without the ConvLSTM block in the encoding operator  $\mathcal{E}$ ;
- DriftNet without the CNN block in the encoding operator  $\mathcal{E}$ ;
- a DriftNet-Unet architecture, referred to as DriftNet<sub>U</sub>: in this architecture, we replace the CNN block in the encoding operator  $\mathcal{E}$  by a U-net block. The latter is composed of three 2d Conv-MaxPool-ReLU layers with kernel size 3 and channels evolving from 27 to 256 and three 2d ConvTranspose-Upsample-ReLU layers with kernel size 3 and channels evolving from 256 to 27.

### *e. Benchmarks*

This section presents the experimental setups that we employ to assess the performance of DriftNet in both virtual and real-case scenarios.

**Benchmark B1** This benchmark B1 aims to assess the reliability and accuracy of learning-based schemes in reproducing Lagrangian drift trajectories given fully-known underlying dynamics. This benchmark uses the Eulerian velocity field extracted from the dataset E1 and the trajectories from the Lagrangian Dataset L1. The resulting dataset is randomly splitted into train (80%), validation (10%) and test (10%) sets. DriftNet is fed with E1 velocity fields and trained to

265 reproduce reference trajectories from the L1 dataset. For benchmarking purposes, we consider the  
266 CNN and LSTM baselines trained following the same strategy as DriftNet.

267 **Benchmark B2** This benchmark aims to assess from simulation data the ability to predict accu-  
268 rately Lagrangian drift trajectories when the sea surface velocities are imperfectly known. We use  
269 E2 velocity fields and L1 trajectories as reference. Similarly to Benchmark B1, we randomly split  
270 the resulting dataset into train (80%) validation (10%) and test (10%) datasets. For benchmarking  
271 purposes, we consider the state-of-the-art model-based method corresponding to trajectory dataset  
272 L2 as well as the CNN and LSTM baselines trained following the same strategy as DriftNet.

273 Benchmark B2 also assesses the ability of DriftNet to exploit multi-source geophysical data.  
274 Here, we complement the data with the Sea Surface Height (SSH) from Dataset E1. In this con-  
275 figuration, Driftnet is fed with E2 velocity fields and E1 SSH fields to reproduce L1 trajectories.

276 **Benchmark B3** This benchmark addresses a real-case scenario with real drifters’ trajectories  
277 and operational ocean reanalyses. We use E3 velocity and SSH fields and real trajectories from L3  
278 dataset as reference. Similarly to Benchmark B1, we randomly split the resulting dataset into train  
279 (70%) validation (10%) and test (20%) datasets.

280 For benchmarking purposes, we consider the state-of-the-art model-based method correspond-  
281 ing to trajectory dataset L4 as well as the CNN and LSTM baselines trained following the same  
282 strategy as DriftNet. This benchmark assesses a DriftNet using solely E3 velocity fields and a  
283 DriftNet using jointly E3 velocity and SSH fields.

## 284 5. Results

285 This section reports our numerical experiments for the three benchmarks introduced in the pre-  
286 vious section, namely benchmark B1 (section 5.a), benchmark B2 (section 5.b) and benchmark  
287 B3 (section 5.c).

### 288 *a. Results for Benchmark B1*

289 DriftNet presents the best performance for all metrics and significantly outperforms the CNN  
290 and LSTM baselines, see Table 2. For the North East Pacific, the separation distance of 5.3km after  
291 9 days is of around one order of magnitude smaller than the baseline CNN and LSTM models.  
292 Similar performances are noted for Liu index, 0.06, and the error between Lagrangian time scales,

TABLE 2: **Performance of DriftNet for Benchmark B1:** we report the performance metrics for DriftNet and classical learning-based approaches for benchmark B1 in the two study regions: North East Pacific and Gulf Stream. The evaluation metrics are the last time step mean separation distance ( $D$ ), mean Liu index ( $\mathcal{L}_{Liu}$ ) and mean absolute error between Lagrangian time scales ( $\Delta\mathcal{T}$ ). Bold indicates the best performance. We refer the reader to section 4.d for the details of the benchmarked architectures.

Method	North East Pacific			Gulf Stream		
	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days
<b>DriftNet</b>	<b>5.3</b>	<b>0.06</b>	<b>0.14</b>	<b>91.1</b>	<b>0.26</b>	<b>1.17</b>
CNN	43	0.5	1.46	151.85	0.63	4.03
LSTM	30.6	0.4	3.36	188	0.77	4.42

0.14. For the Gulf Stream, the overall accuracy is lower (separation distance after 9 days of 91.1km, Liu index 0.26 and Lagrangian time scale error 1.17), but still DriftNet outperforms the baselines. These experiments support the proposed Eulerian Fokker-Plank-inspired architecture.

DriftNet simulation error is larger for the Gulf Stream region. This is consistent with the resolution error in encoding initial positions, as evidenced in Callies et al. (2021). Here, Lagrangian trajectories simulated with an error corresponding to the resolution of the grid (*i.e.*,  $1/12^\circ$ ) leads to a separation distance after 9 days of 20km (resp. 94km) for the North East Pacific region (resp. the Gulf Stream region). The larger sea surface velocities and sharper associated spatial gradients in the Gulf Stream region explains the larger separation distance after 9 days.

- **Impact of DriftNet components** we compare DriftNet in it’s original configuration to three variations presented in section 4.d namely a configuration combining two commonly used U-Net models separated by one ConvLSTM block, a second one containing the ConvLSTM block only, and finally one composed of 2D convolutional layers only.

Removing the CNN or LSTM blocks from the DriftNet architecture impacts the simulation performance, see Table 3. The separation distance increases from 5.3km to 42.1km (resp. 36.2km) for the North East Pacific region when we remove the LSTM (resp. CNN) block. Similarly, the U-Net version of the DriftNet does not lead to a better simulation performance. This likely relates to some overfitting in our training schemes due to the greater complexity of this UNet configuration. Same qualitative results are found in the Gulf Stream region.

TABLE 3: **Ablation study of DriftNet**: we compare the metrics of different configurations of DriftNet for benchmark B1: namely, the reference DriftNet, DriftNet<sub>U</sub> with a U-net block instead of CNN, a DriftNet without the CNN block and DriftNet without ConvLSTM. The evaluation metrics are the last time step mean separation distance ( $D$ ), mean Liu index ( $\mathcal{L}_{Liu}$ ) and mean absolute error between Lagrangian time scales ( $\Delta\mathcal{T}$ ). Bold indicates best performance.

<i>Method</i>	North East Pacific			Gulf Stream		
	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days
<b>DriftNet</b>	<b>5.3</b>	<b>0.06</b>	<b>0.14</b>	<b>91.1</b>	<b>0.26</b>	1.17
DriftNet <sub>U</sub>	16.1	0.21	1.79	96	0.32	<b>0.9</b>
DriftNet without CNN	36.2	0.39	0.32	163	0.68	1.7
DriftNet without ConvLSTM	42.1	0.37	1.45	177	0.73	2.84

TABLE 4: **Impact of spatial resolution on DriftNet performance** Performance of DriftNet in function of the spatial resolution of the input Eulerian velocity field for benchmark B1 in both study regions: North East Pacific and Gulf Stream. Each line corresponds to a different spatial resolution of the input E1 field. The evaluation metrics are the last time step mean separation distance ( $D$ ), mean Liu index ( $\mathcal{L}_{Liu}$ ) and mean absolute error between Lagrangian time scales ( $\Delta\mathcal{T}$ ). Bold indicates best performance.

<i>Resolution</i>	North East Pacific			Gulf Stream		
	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days
$1/12^\circ \times 1/12^\circ$	<b>5.3</b>	<b>0.06</b>	<b>0.14</b>	<b>91.1</b>	<b>0.26</b>	<b>1.17</b>
$1/6^\circ \times 1/6^\circ$	8.7	0.15	0.17	102.3	0.35	1.23
$1/3^\circ \times 1/3^\circ$	17.9	0.27	0.28	109.4	0.47	0.91
$1/2^\circ \times 1/2^\circ$	28.2	0.40	0.99	115.9	0.46	1.05

- **Impact of the resolution of the conditioning velocity fields.** Coarsening the resolution of the conditioning fields leads to degrading the performance, see Table 4. For both regions, the impact is more significant in the mean separation distance and the Liu index, while Lagrangian time scale estimation does not seem to be impacted by the degradation in the underlying resolution. Resolution is thus an important characteristic of the input fields.

### b. Results for Benchmark B2

Table 5 illustrates the performance of the simulation of Lagrangian trajectories using degraded velocity fields instead of the true velocity fields. Similarly to Benchmark B1, we compare DriftNet to CNN and LSTM baselines. Here we also benchmark DriftNet with respect to the Lagrangian simulations with Ocean Parcels Lange and van Sebille (2017).

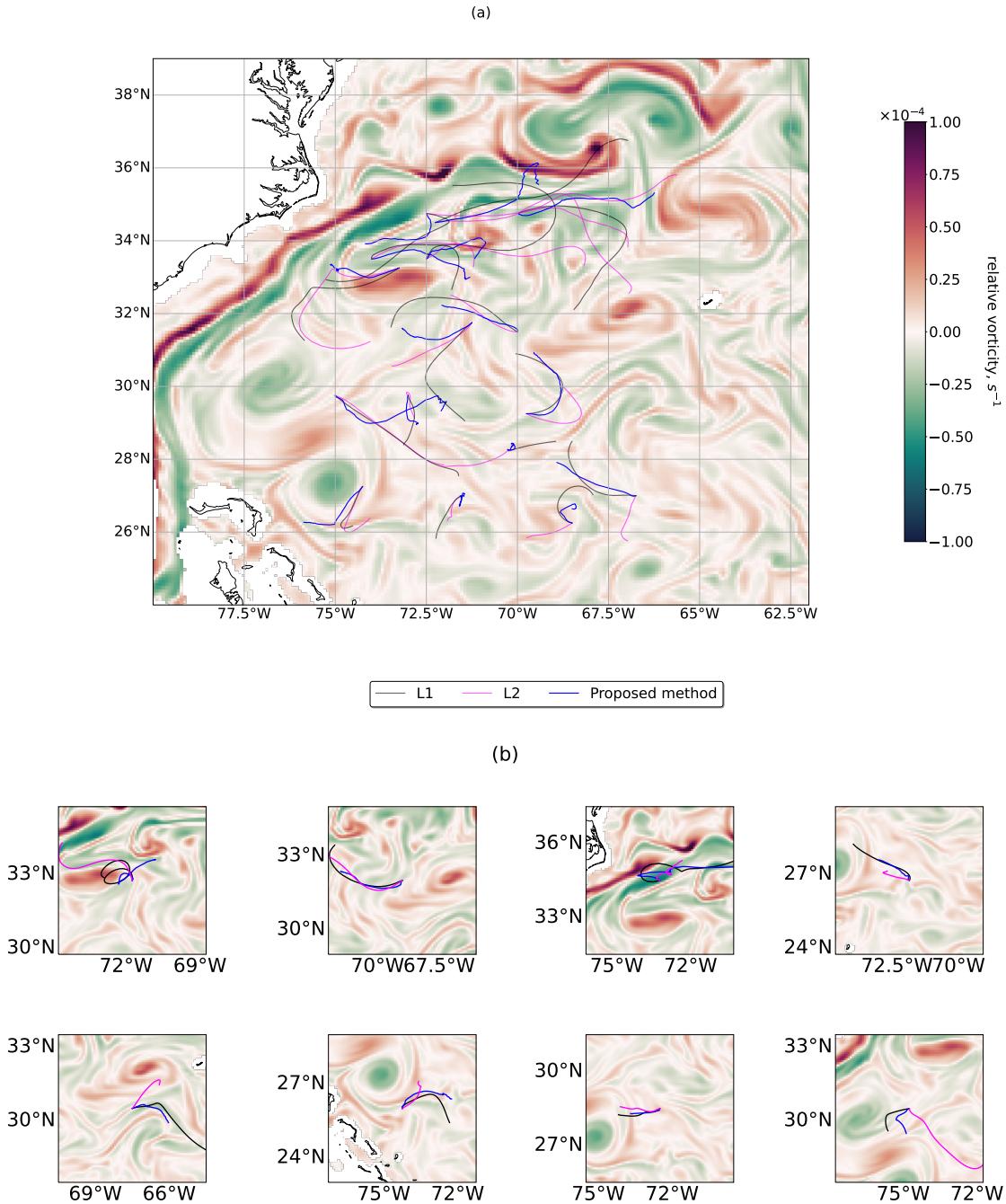
TABLE 5: **Performance metrics for Benchmark B2 experimental framework.** We report the performance metrics for DriftNet and baseline models, all supplied with velocity fields from E2 and trained to generate trajectories from reference Dataset L1 in the two study regions: North East Pacific and Gulf Stream. The evaluation metrics are the last time step mean separation distance ( $D$ ), mean Liu index ( $\mathcal{L}_{Liu}$ ) and mean absolute error between Lagrangian time scales ( $\Delta\mathcal{T}$ ). Bold indicates the best performance.

<i>Method</i>	North East Pacific			Gulf Stream		
	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days
Ocean Parcels	53.6	0.87	<b>0.1</b>	168.3	0.75	<b>0.04</b>
DriftNet	<b>45.4</b>	0.7	1.96	<b>154.6</b>	<b>0.63</b>	1.82
CNN	51.7	0.72	2.12	184.1	0.67	2.48
LSTM	47.9	<b>0.69</b>	4.35	163.2	0.67	2.1

322 DriftNet presents smaller  $D$  45.4 km (154.6 km) and  $\Delta\mathcal{T}$  than the other learning-based models,  
 323 and smaller  $D$ ,  $\mathcal{L}_{Liu}$  than Ocean Parcels, thus consistently outperforming the baseline approaches.  
 324 The  $\mathcal{L}_{Liu}$  for the North East Pacific region is of 0.7 and is smaller than the most baseline ap-  
 325 proaches, except the LSTM model (0.69) slightly outperforming the proposed method. While in  
 326 the Gulf Stream the  $\mathcal{L}_{Liu}$  (0.63) is smaller than all the other approaches. The Lagrangian time  
 327 scale error is larger for the proposed method than for the baseline Ocean Parcels.

328 **Impact of auxiliary variables** Table 6 assesses how the availability of the SSH fields could  
 329 improve the simulation of Lagrangian drift trajectories. Whereas Lagrangian model eq.(1) solely  
 330 depends on the sea surface velocities, deep learning schemes can easily extend to multiple condi-  
 331 tioning fields. Among the observed satellite-derived tracers on the sea surface, the SSH explicitly  
 332 seems appealing as it informs the geostrophic component of sea surface currents Taqi et al. (2019);  
 333 Ballarotta et al. (2022). In an idealized setting, we evaluate the potential added value of the true  
 334 SSH fields provided as inputs to DrifNet either in place of E2 velocity fields or to complement  
 335 them, see Table 6.

336 These results emphasize the synergistic potential of SSH observations and E2 velocity fields  
 337 compared with using only one of these two input data. In the North East Pacific region,  $D$  is  
 338 reduced by over 55%,  $\mathcal{L}_{Liu}$  by over 60% and  $\Delta\mathcal{T}$  by over 45%. Similarly in the Gulf Stream,  
 339 we report a relative gain greater than 25% for  $D$ , 30% for  $\mathcal{L}_{Liu}$  and 20% for  $\Delta\mathcal{T}$ . These re-  
 340 sults are in agreement with the ability of observation-based SSH products to capture geostrophic  
 341 dynamics Ballarotta et al. (2022), while data-assimilation-based velocity fields can reveal the



**FIG. 4: Trajectories for Benchmark B2 in the Gulf Stream** Lagrangian trajectories from Dataset L2 are depicted in magenta, trajectories simulated with DriftNet conditioned with SSC from E2 and SSH from E1 in blue, and trajectories from Dataset L1 in black. For better visual appreciation panel a) shows 16 randomly picked trajectories for the whole region while b) shows 8 random trajectories zoomed. All superimposed to the mean relative vorticity field of E1.

<sup>342</sup> ageostrophic sea surface dynamics. The latter are expected to be significant for horizontal scales

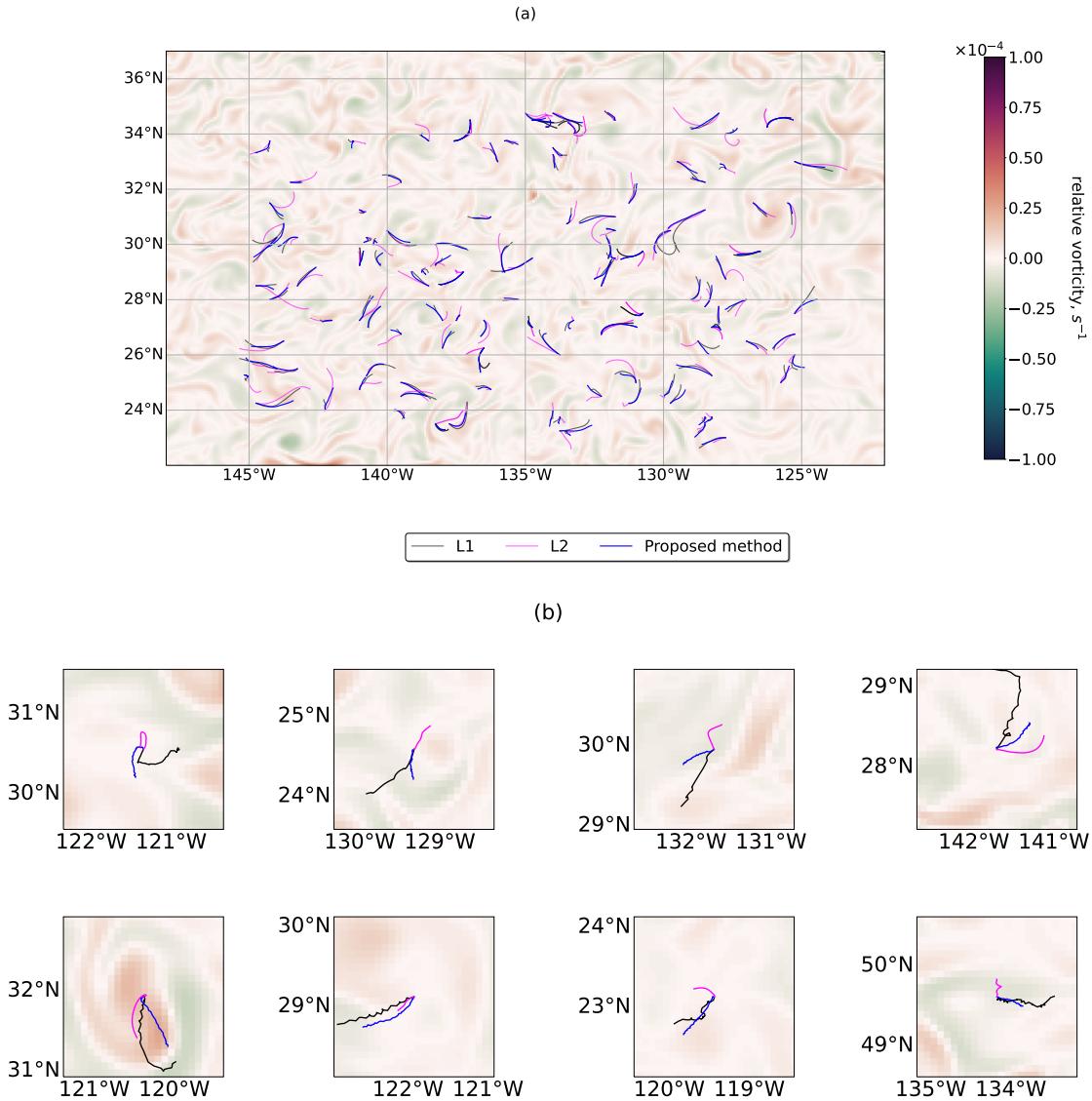


FIG. 5: **Trajectories for Benchmark B2 in the North East Pacific** Lagrangian trajectories from Dataset L2 are depicted in magenta, trajectories simulated with DriftNet conditioned with SSC from E2 and SSH from E1 in blue, and trajectories from Dataset L1 in black. a) shows 128 randomly picked trajectories for the whole region while b) shows 8 random trajectories zoomed. All superimposed to the mean relative vorticity field of E1.

343 below a few hundreds of kilometers, especially in Western Boundary currents such as the Gulf  
 344 Stream Gula et al. (2015).

345 For both regions the trajectories are visualised at the Figure 4 and Figure 5 which show that  
 346 DriftNet generates trajectories closer to the ground truth from L1 than the baseline L2.

TABLE 6: **Synergistic use of SSH observations and data-assimilation-based velocity fields in Benchmark B2.** We benchmark DriftNet schemes using as inputs data-assimilation-based velocity fields from dataset E2 and idealized perfectly-observed Sea Surface Height (SSH) fields (i.e., SSH fields from dataset E1). We illustrate performance in the two study regions: North East Pacific and Gulf Stream. The evaluation metrics are the last time step mean separation distance ( $D$ ), mean Liu index ( $\mathcal{L}_{Liu}$ ) and mean absolute error between Lagrangian time scales ( $\Delta\mathcal{T}$ ). Bold indicates the best performance.

Method( <i>Variable</i> )	North East Pacific			Gulf Stream		
	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days
DriftNet (U, V)	45.4	0.71	1.96	157	0.73	1.84
DriftNet (SSH)	31.2	0.32	1.32	160.3	0.7	<b>1.41</b>
<b>DriftNet (U, V, SSH)</b>	<b>22.2</b>	<b>0.29</b>	<b>0.73</b>	<b>127</b>	<b>0.52</b>	1.44

TABLE 7: **Performance metrics for Benchmark B3.** Performance of different Lagrangian simulation schemes for Benchmark B3 which involves real data. For each scheme, we detail the considered input data: namely the velocity fields U and V from the GLORYS12 reanalysis Lellouche et al. (2018) possibly complemented by satellite-derived optimally-interpolated SSH fields Pujol et al. (2016). We report the performance metrics in the two study regions: North East Pacific and Gulf Stream. The evaluation metrics are the last time step mean separation distance ( $D$ ), mean Liu index ( $\mathcal{L}_{Liu}$ ) and mean absolute error between Lagrangian time scales ( $\Delta\mathcal{T}$ ). Bold indicates the best performance.

North East Pacific			
Method( <i>Variable</i> )	$D$ , km	$\mathcal{L}_{Liu}$	$\Delta\mathcal{T}$ , days
Ocean Parcels (U, V)	73.77	0.65	2.2
DriftNet (U, V)	62.9	0.55	0.71
CNN (U, V)	68.64	0.57	0.32
LSTM (U, V)	64.3	0.55	2.1
<b>DriftNet (U, V, SSH)</b>	<b>60.2</b>	<b>0.52</b>	<b>0.04</b>

### 347 *c. Results for Benchmark B3*

348 In this real case situation, DriftNet outperforms the classical advection method, see Table 7  
349 and Figure 6. Compared to the model-based baseline, the mean separation distance of 62.9km  
350 is reduced by around 15%, the Liu index of 0.55 by 15%. Unlike previous B2 experiment, the  
351 Lagrangian time scale error of 0.71 is significantly reduced. As in benchmark B2, we observe a  
352 significant impact of the observed SSH fields in the quality of the simulation of the Lagrangian  
353 trajectories. The combination of GLORYS12 velocities and satellite-derived SSH fields leads to  
354 a reduction of the mean separation distance by over 4% compared to the baseline DriftNet and

355 by over 18% compared to the baseline model-based advection method. The Liu index is also  
356 reduced by more than 5% compared to the DriftNet baseline and by over 20% compared to the  
357 baseline advection method. The Lagrangian time scale error is significantly reduced, by over  
358 94% compared to the DriftNet baseline and by over 98% compared to the baseline model-based  
359 approach.

## 360 **6. Conclusions and Discussion**

361 In this study, we presented a novel deep learning framework, called DriftNet, for the condi-  
362 tional simulation of Lagrangian drift trajectories on the sea surface. We drew inspiration from the  
363 Fokker-Planck equation to explore a spatially-explicit Eulerian latent representation of the trajec-  
364 tories in the proposed neural architecture. We evaluate DriftNet on three different benchmarks  
365 using synthetic and real-world datasets. Our experiments support the relevance of DriftNet to ad-  
366 vance Lagrangian drift simulations on the sea surface and better exploit available reanalysis and  
367 observation datasets.

368 DriftNet consistently demonstrated a better performance than state-of-the-art model-based and  
369 learning-based approaches, with respect to the following Lagrangian metrics: mean separation  
370 distance, Liu index and Lagrangian time scale error between the ground-truth trajectories and  
371 the simulated ones. When DriftNet is provided with fully known velocity fields from Nature  
372 Run simulation (Benchmark B1), it reduces the overall error by more than 80% compared to the  
373 state-of-the-art learning-based approaches. Moreover, when DriftNet is provided with degraded  
374 velocity fields from OSSE-based assimilated simulation (Benchmark B2), it reduces the overall  
375 error by over 15% compared to conventional point-wise schemes. These results are consistent for  
376 the two case-study regions, even if they are characterized by notably different dynamical regimes.  
377 Interestingly, we draw similar conclusions between synthetic and real datasets.

378 The performance of DriftNet exhibits variabilities across the two different regions. It likely un-  
379 derscores the importance of region-specific adaptations in trajectory simulation models. As such,  
380 the deployment of DriftNet on a global scale may require retraining the model to accommodate  
381 the diverse dynamics in various oceanic regions Miao et al. (2023). While we expect the proposed  
382 neural architecture to be generic, fine-tuning its parameters to align with the characteristics of a  
383 given region seems crucial for the relevance of the simulations. This approach allows for the incor-

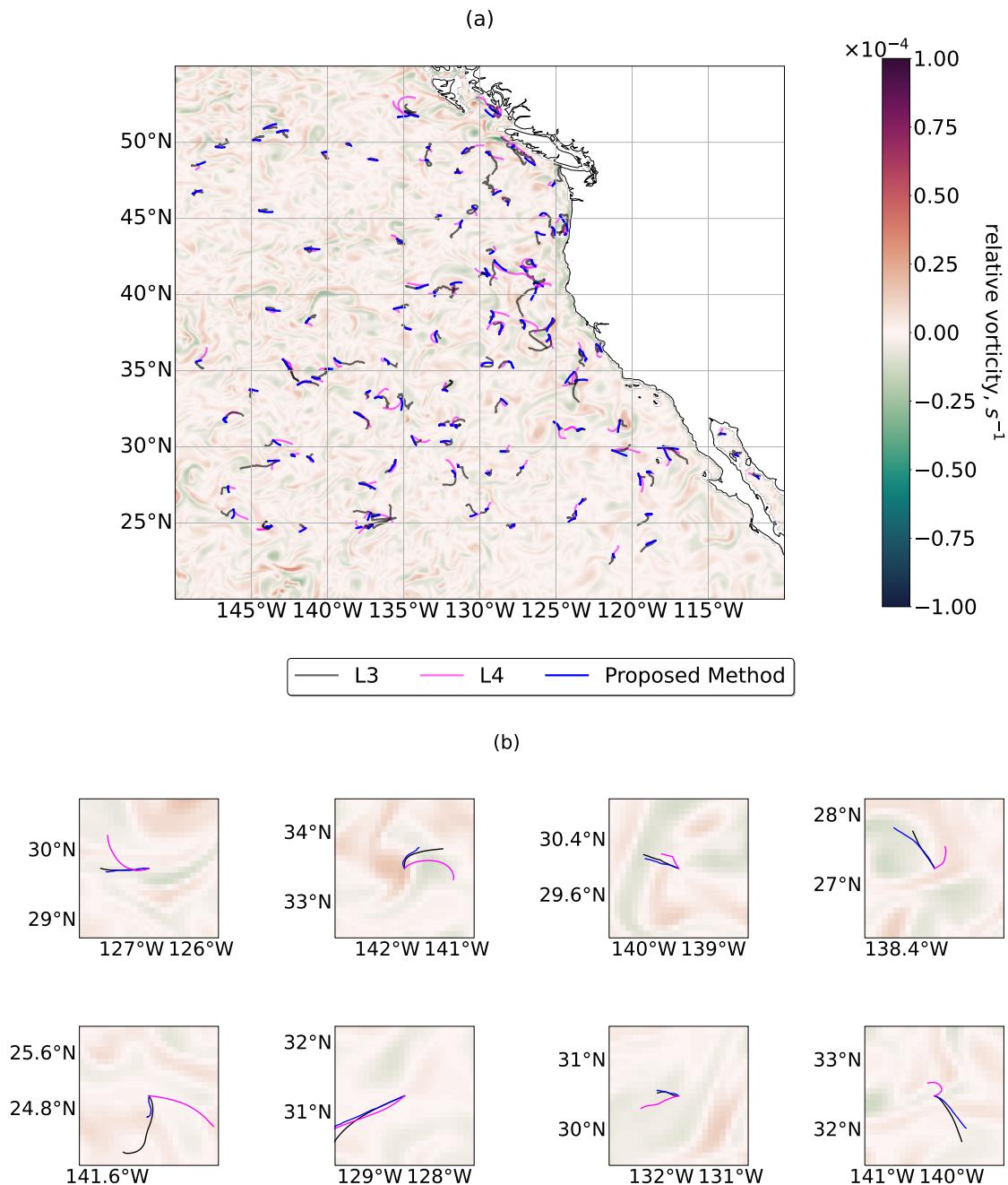


FIG. 6: **Examples of simulated trajectories for Benchmark B3:** we depict real drifters trajectories in black (L3), trajectories simulated with Ocean Parcels using GLORYS12 velocity fields in magenta (L4) and trajectories simulated with DriftNet using both GLORYS12 velocity fields and satellite-derived optimally-interpolated SSH fields in blue. Panel (a) illustrates trajectories from the whole region. Panel (b) zoom on a subset of randomly-selected trajectories. All superimposed to the mean relative vorticity of the GLORYS12 velocity fields.

384 poration of region-specific features, thereby enhancing the model’s predictive capabilities across  
385 diverse geographical regions. Furthermore, leveraging transfer learning techniques to capitalize  
386 on the knowledge gained from training in one region to inform model adjustments in others holds  
387 promising avenue. By iteratively refining DriftNet through region-specific training and fine-tuning  
388 iterations, we hope to develop a robust trajectory simulation tool capable of capturing the intricacies of oceanic dynamics on a global scale.

390 Our benchmarks highlight the relevance of the SSH fields to improve the prediction of Lagrangian drifts on the sea surface both in synthetic and real case-studies, *i.e.* Benchmarks B2  
391 and B3. We report a significant improvement for all the considered metrics when we complement reanalysis velocity fields with SSH fields. The SSH informs the geostrophic component of  
392 the sea surface currents, which is important for the two case-study regions for horizontal scales  
393 above 100km Checkley and Barth (2009); Thomas and Joyce (2010); Johns et al. (1989). As ocean  
394 reanalysis datasets can usually retrieve sea surface currents for horizontal scales above a few hundred of kilometers Benkiran et al. (2021a); Lellouche et al. (2021), the learning schemes likely  
395 benefit from the SSH fields to recover a more robust representation of the sea surface dynamics and improve the simulation of Lagrangian drift on the sea surface. The advent of wide-swath  
396 satellite altimeters Benkiran et al. (2021b); Tchonang et al. (2021), possibly combined with neural  
397 mapping schemes Martin et al. (2023); Beauchamp et al. (2022); Fablet et al. (2023), will lead  
398 to a better reconstruction of satellite-derived SSH fields, which could in turn result in improved  
399 learning-based Lagrangian simulations for real drifters on the sea surface.

404 Beyond the exploraton of satellite-derived SSH fields, the use of additional geophysical variables to enhance trajectory generation offers another interesting approach to improve the simulation  
405 performance. Future studies could investigate other variables such as Sea Surface Temperature Ciani et al. (2020), Ocean Color Yang et al. (2015); Liu et al. (2017), Winds Cucco et al. (2016);  
406 Solanlki et al. (2001) or Waves Tang et al. (2007) to refine trajectory simulations and gain deeper  
407 insights into oceanic processes. By leveraging these variables, the proposed method could not only  
408 improve trajectory accuracy but also facilitate comprehensive analysis of the underlying ocean  
409 dynamics.

412 Inspired by established initiatives such as WeatherBench Rasp et al. (2020) and OceanBench  
413 Johnson et al. (2024), we propose a new benchmark for a standardized assessing and comparing

414 Lagrangian trajectory simulation models <https://github.com/CIA-Oceanix/DriftNet>. Through pro-  
415 posed datasets and evaluation metrics, this framework seeks to foster collaborative frameworks  
416 involving ocean scientists and ML practitioners. In this context, future work could explore new  
417 evaluation metrics and extend the proposed benchmarks to a broader array of oceanic regions.

418 From a methodological point of view, it seems appealing to extend the DriftNet to a probabilistic  
419 framework with a view to sampling ensembles of realistic drift trajectories. Conditional Genera-  
420 tive Adversarial Networks (GAN) Roy et al. (2022) naturally arise among the promising solutions.  
421 Other generative models such as diffusion models Tashiro et al. (2021) could also be explored. Re-  
422 garding real-time operational applications, such as search and rescue operations, or iceberg drifts,  
423 such generative schemes would also naturally apply to the short-term forecasting of Lagrangian  
424 trajectories on the sea surface using sea surface conditions issued from short-term forecasts. This  
425 research direction would likely benefit from the development of neural ocean forecasts Wang et al.  
426 (2024); Xiong et al. (2023) to advance model-based operational forecasts Drévilion et al. (2008);  
427 Benkiran et al. (2024).

428 *Acknowledgments.* We gratefully acknowledge the Mercator Ocean International and Coperni-  
429 cus Marine Environment Monitoring Service (CMEMS) for providing access to the comprehen-  
430 sive datasets used in this study, which was instrumental in our research efforts. Special thanks  
431 to Mounir Benkiran for generously providing Nature Run and OSSE data, which significantly  
432 enriched our analysis and insights. This work was supported by LEFE program (LEFE MANU  
433 and IMAGO projects IA-OAC), CNES (OSTST DUACS-HR and SWOT ST DIEGO) and ANR  
434 Projects Melody (ANR-19-CE46-0011) and OceaniX (ANR-19-CHIA-0016). It benefited from  
435 HPC and GPU resources from Azure (Microsoft Azure grant) and from GENCI-IDRIS (Grant  
436 2021-101030). The authors declare no conflict of interest.

437 *Data availability statement.* The data used in this study, including ocean currents and sea surface  
438 height are openly accessible through the Copernicus Marine Environment Monitoring Service  
439 (CMEMS). The data from the B1 and B2 benchmarks are subject to access restriction, one should  
440 refer to Mercator Ocean International in order to claim the access rights. Additionally, the code  
441 implemented for the evaluation of the methods discussed in this article is available as an open-  
442 source project on GitHub: <https://github.com/CIA-Oceanix/DriftNet>.

## 443 **References**

- 444 Auad, G., D. Roemmich, and J. Gilson, 2011: The california current system in relation to the  
445 northeast pacific ocean circulation. *Progress in Oceanography*, **91** (4), 576–592.
- 446 Ballarotta, M., and Coauthors, 2022: Improved global sea surface height and currents maps from  
447 remote sensing and in situ observations. *Earth System Science Data Discussions*, **2022**, 1–32.
- 448 Barron, C. N., L. F. Smedstad, J. M. Dastugue, and O. M. Smedstad, 2007: Evaluation of ocean  
449 models using observed and simulated drifter trajectories: Impact of sea surface height on syn-  
450 thetic profiles for data assimilation. *Journal of Geophysical Research: Oceans*, **112** (C7).
- 451 Beauchamp, M., Q. Febvre, H. Georgenthum, and R. Fablet, 2022: 4dvarnet-ssh: end-to-end  
452 learning of variational interpolation schemes for nadir and wide-swath satellite altimetry. *Geo-  
453 scientific Model Development Discussions*, **2022**, 1–37.

- 454 Benkiran, M., P.-Y. Le Traon, and G. Dibarboure, 2021a: Contribution of a constellation of two  
455 wide-swath altimetry missions to global ocean analysis and forecasting. *Ocean Science Discus-*  
456 *sions*, **2021**, 1–25.
- 457 Benkiran, M., P.-Y. Le Traon, E. Rémy, and Y. Drillet, 2024: Impact of two high resolution al-  
458 timetry mission concepts for ocean forecasting. *EGUsphere*, **2024**, 1–19.
- 459 Benkiran, M., and Coauthors, 2021b: Assessing the impact of the assimilation of swot observa-  
460 tions in a global high-resolution analysis and forecasting system part 1: Methods. *Frontiers in*  
461 *Marine Science*, **8**, 691 955.
- 462 Botvynko, D., C. Granero-Belinchon, S. van Gennip, A. Benzinou, and R. Fablet, 2023: Deep  
463 learning for lagrangian drift simulation at the sea surface. *ICASSP 2023 - 2023 IEEE Interna-*  
464 *tional Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.
- 465 Breivik, O., A. A. Allen, C. Maisondieu, and M. Olagnon, 2013: Advances in search and rescue  
466 at sea. *Ocean Dynamics*, **63**, 83–88.
- 467 Callies, U., M. Kreuz, W. Petersen, and Y. G. Voynova, 2021: On using lagrangian drift simulations  
468 to aid interpretation of in situ monitoring data. *Frontiers in Marine Science*, **8**, 666 653.
- 469 Checkley, D. M., and J. A. Barth, 2009: Patterns and processes in the california current system.  
470 *Progress in Oceanography*, **83 (1-4)**, 49–64.
- 471 Ciani, D., M.-H. Rio, B. B. Nardelli, H. Etienne, and R. Santoleri, 2020: Improving the altimeter-  
472 derived surface currents using sea surface temperature (SST) data: A sensitivity study to SST  
473 products. *Remote Sensing*, **12 (10)**, 1601.
- 474 Cucco, A., G. Quattrocchi, A. Satta, F. Antognarelli, F. De Biasio, E. Cadau, G. Umgiesser, and  
475 S. Zecchetto, 2016: Predictability of wind-induced sea surface transport in coastal areas. *Jour-*  
476 *nal of Geophysical Research: Oceans*, **121 (8)**, 5847–5871.
- 477 Davis, R. E., 1991: Lagrangian ocean studies. *Annual Review of Fluid Mechanics*, **23 (1)**, 43–64.
- 478 Dewar, W. K., and J. M. Bane, 1989: Gulf stream dynamics. pad II: Eddy energetics at 73 w.  
479 *Journal of Physical Oceanography*, **19 (10)**, 1574–1587.

480 Dréville, M., and Coauthors, 2008: The godae/mercator-ocean global ocean forecasting system:  
481 results, applications and prospects. *Journal of Operational Oceanography*, **1** (1), 51–57.

482 Etienne, H., and Coauthors, 2023: Quality information document: Global ocean-delayed mode  
483 in-situ observations of surface (drifters and hfr) and sub-surface (vessel-mounted adcps) water  
484 velocity.

485 Fablet, R., Q. Febvre, and B. Chapron, 2023: Multimodal 4dvarnets for the reconstruction of  
486 sea surface dynamics from SST-SSH synergies. *IEEE Transactions on Geoscience and Remote  
487 Sensing*.

488 Fine, S., Y. Singer, and N. Tishby, 1998: The hierarchical hidden markov model: Analysis and  
489 applications. *Machine Learning*, **32**, 41–62.

490 Gan, J., P. Liu, and R. K. Chakrabarty, 2020: Deep learning enabled lagrangian particle trajectory  
491 simulation. *Journal of Aerosol Science*, **139**, 105 468.

492 Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and  
493 Y. Bengio, 2020: Generative adversarial networks. *Communications of the ACM*, **63** (11), 139–  
494 144.

495 Gula, J., M. J. Molemaker, and J. C. McWilliams, 2015: Gulf stream dynamics along the south-  
496 eastern US seaboard. *Journal of Physical Oceanography*, **45** (3), 690–715.

497 Jenkins, J., A. Paiement, Y. Ourmières, J. Le Sommer, J. Verron, C. Ubelmann, and H. Glotin,  
498 2023: A DNN framework for learning lagrangian drift with uncertainty. *Applied Intelligence*,  
499 **53** (20), 23 729–23 739.

500 Jiang, H., L. Chang, Q. Li, and D. Chen, 2019: Trajectory prediction of vehicles based on deep  
501 learning. *2019 4th International Conference on Intelligent Transportation Engineering(ICITE)*,  
502 IEEE, 190–195.

503 Johns, E., D. R. Watts, and H. T. Rossby, 1989: A test of geostrophy in the gulf stream. *Journal of  
504 Geophysical Research: Oceans*, **94** (C3), 3211–3222.

- 505 Johnson, J. E., Q. Febvre, A. Gorbunova, S. Metref, M. Ballarotta, and J. e. a. Le Sommer, 2024:  
506 Oceanbench: The sea surface height edition. *Advances in Neural Information Processing Sys-*  
507 *tems*, **36**.
- 508 Julka, S., V. Sowrirajan, J. Schloetterer, and M. Granitzer, 2021: Conditional generative adversar-  
509 ial networks for speed control in trajectory simulation. *International Conference on Machine*  
510 *Learning, Optimization, and Data Science*, Springer, 436–450.
- 511 Kaplan, D. M., J. Largier, and L. W. Botsford, 2005: HF radar observations of surface circulation  
512 off bodega bay (northern california, USA). *Journal of Geophysical Research: Oceans*, **110**.
- 513 Korbmacher, R., and A. Tordeux, 2022: Review of pedestrian trajectory prediction methods: Com-  
514 paring deep learning and knowledge-based approaches. *IEEE Transactions on Intelligent Trans-*  
515 *portation Systems*, **23 (12)**, 24 126–24 144.
- 516 Koszalka, I., J. H. LaCasce, M. Andersson, K. A. Orvik, and C. Mauritzen, 2011: Surface cir-  
517 culation in the nordic seas from clustered drifters. *Deep Sea Research Part I: Oceanographic*  
518 *Research Papers*, **58 (4)**, 468–485.
- 519 Krauß, W., and C. W. Böning, 1987: Lagrangian properties of eddy fields in the northern north  
520 atlantic as deduced from satellite-tracked buoys. *Journal of Marine Research*, **45 (2)**, 259–291.
- 521 Lange, M., and E. van Sebille, 2017: Parcels v0. 9: prototyping a lagrangian ocean analysis  
522 framework for the petascale age. *Geoscientific Model Development*, **10 (11)**, 4175–4186.
- 523 Lellouche, J. M., E. Greiner, R. Bourdalle-Badie, G. Garric, A. Melet, M. Drévilion, C. Bricaud,  
524 and et al., 2021: The copernicus global 1/12 oceanic and sea ice GLORYS12 reanalysis. *Fron-*  
525 *tiers in Earth Science*, 698876.
- 526 Lellouche, J. M., and Coauthors, 2018: Recent upyears to the copernicus marine service global  
527 ocean monitoring and forecasting real-time 1/ 12 high-resolution system. *Ocean Science*, **14 (5)**,  
528 1093–1126.
- 529 Liu, J., W. J. Emery, X. Wu, M. Li, C. Li, and L. Zhang, 2017: Computing ocean surface cur-  
530 rents from GOCI ocean color satellite imagery. *IEEE Transactions on Geoscience and Remote*  
531 *Sensing*, **55 (12)**, 7113–7125.

532 Liu, Y., and R. H. Weisberg, 2011: Evaluation of trajectory modeling in different dynamic regions  
533 using normalized cumulative lagrangian separation. *Journal of Geophysical Research: Oceans*,  
534 **116 (C9)**.

535 Liubartseva, S., G. Coppini, R. Lecci, and E. Clementi, 2018: Tracking plastics in the mediter-  
536 ranean: 2d lagrangian model. *Marine pollution bulletin*, **129 (1)**, 151–162.

537 Lumpkin, R., and P. J. Flament, 2013: Extent and energetics of the hawaiian lee countercurrent.  
538 *Oceanography*, **26 (1)**, 58–65.

539 Lumpkin, R., and S. Garzoli, 2011: Interannual to decadal changes in the western south atlantic’s  
540 surface circulation. *Journal of Geophysical Research: Oceans*, **116**.

541 Lumpkin, R., N. Maximenko, and M. Pazos, 2012: Evaluating where and why drifters die. *Journal*  
542 *of Atmospheric and Oceanic Technology*, **29 (2)**, 300–308.

543 Lumpkin, R., T. Özgökmen, and L. Centurioni, 2017: Advances in the application of surface  
544 drifters. *Annual Review of Marine Science*, **9 (1)**, 59–81.

545 Ma, Y., X. Zhu, S. Zhang, R. Yang, W. Wang, and D. Manocha, 2019: Trafficpredict: Trajectory  
546 prediction for heterogeneous traffic-agents. *Proceedings of the AAAI conference on artificial*  
547 *intelligence*, Vol. 33, 6120–6127.

548 Martin, S. A., G. E. Manucharyan, and P. Klein, 2023: Synthesizing sea surface temperature and  
549 satellite altimetry observations using deep learning improves the accuracy and resolution of  
550 gridded sea surface height anomalies. *Journal of Advances in Modeling Earth Systems*, **15 (5)**,  
551 e2022MS003 589.

552 Maximenko, N., J. Hafner, and P. Niiler, 2012: Pathways of marine debris derived from trajectories  
553 of lagrangian drifters. *Marine pollution bulletin*, **65 (1-3)**, 51–62.

554 Miao, Y., X. Zhang, Y. Li, L. Zhang, and D. Zhang, 2023: Monthly extended ocean predictions  
555 based on a convolutional neural network via the transfer learning method. *Frontiers in Marine*  
556 *Science*, **9**, 1073 377.

557 Nguyen, D., and R. Fablet, 2024: A transformer network with sparse augmented data representa-  
558 tion and cross entropy loss for AIS-based vessel trajectory prediction. *IEEE Access*.

- 559 Paz, D., H. Zhang, and H. I. Christensen, 2021: Tridentnet: A conditional generative model for  
560 dynamic trajectory generation. *International Conference on Intelligent Autonomous Systems*,  
561 Springer, 403–416.
- 562 Prants, S. V., M. Y. Uleysky, and M. V. Budyansky, 2017: *Lagrangian oceanography: large-scale*  
563 *transport and mixing in the ocean*. Springer.
- 564 Pujol, M.-I., Y. Faugère, G. Taburet, S. Dupuy, C. Pelloquin, M. Ablain, and N. Picot, 2016:  
565 DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years. *Ocean*  
566 *Science*, **12** (5), 1067–1090.
- 567 Rasp, S., P. D. Dueben, S. Scher, J. A. Weyn, S. Mouatadid, and N. Thuerey, 2020: Weatherbench:  
568 a benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling*  
569 *Earth Systems*, **12** (11), e2020MS002 203.
- 570 Roy, A., R. Fablet, and S. L. Bertrand, 2022: Using generative adversarial networks (GAN) to  
571 simulate central-place foraging trajectories. *Methods in Ecology and Evolution*, **13** (6), 1275–  
572 1287.
- 573 Röhrs, J., and Coauthors, 2021: Surface currents in operational oceanography: Key applications,  
574 mechanisms, and methods. *Journal of Operational Oceanography*, **16** (1), 60–88.
- 575 Shi, X., Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, 2015: Convolutional  
576 LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural*  
577 *information processing systems*, **28**.
- 578 Solanki, H. U., R. M. Dwivedi, S. R. Nayak, J. V. Jadeja, D. B. Thakar, H. B. Dave, and M. I. Patel,  
579 2001: Application of ocean colour monitor chlorophyll and AVHRR SST for fishery forecast:  
580 Preliminary validation results off gujarat coast, northwest coast of india. *Indian Journal of Geo-*  
581 *Marine Sciences*, **30**(3), 132–138.
- 582 Son, Y. B., B.-J. Choi, Y. H. Kim, and Y.-G. Park, 2015: Tracing floating green algae blooms in the  
583 yellow sea and the east china sea using GOCI satellite data and lagrangian transport simulations.  
584 *Remote Sensing of Environment*, **156**, 21–33.
- 585 Talley, L. D., 2011: *Descriptive physical oceanography: an introduction*. Academic press.

- 586 Tang, C. L., W. Perrie, A. D. Jenkins, B. M. DeTracey, Y. Hu, B. Toulany, and P. C. Smith, 2007:  
587 Observation and modeling of surface currents on the grand banks: A study of the wave effects  
588 on surface currents. *Journal of Geophysical Research: Oceans*, **112**.
- 589 Taqi, A. M., A. M. Al-Subhi, M. A. Alsaafani, and C. P. Abdulla, 2019: Estimation of geostrophic  
590 current in the red sea based on sea level anomalies derived from extended satellite altimetry  
591 data. *Ocean Science*, **15 (3)**, 477–488.
- 592 Tashiro, Y., J. Song, Y. Song, and S. Ermon, 2021: CSDI: Conditional score-based diffusion models  
593 for probabilistic time series imputation. *Advances in Neural Information Processing Systems*,  
594 **34**, 24 804–24 816.
- 595 Tchonang, B. C., M. Benkiran, P.-Y. Le Traon, S. Jan van Gennip, J. M. Lellouche, and G. Rug-  
596 giero, 2021: Assessing the impact of the assimilation of swot observations in a global high-  
597 resolution analysis and forecasting system—part 2: Results. *Frontiers in Marine Science*, **8**,  
598 687 414.
- 599 Thomas, L. N., and T. M. Joyce, 2010: Subduction on the northern and southern flanks of the gulf  
600 stream. *Journal of physical oceanography*, **40 (2)**, 429–438.
- 601 Visser, A. W., 2008: Lagrangian modelling of plankton motion: From deceptively simple random  
602 walks to fokker–planck and back again. *Journal of Marine Systems*, **70 (3-4)**, 287–299.
- 603 Wang, X., and Coauthors, 2024: Xihe: A data-driven model for global ocean eddy-resolving  
604 forecasting. *arXiv preprint arXiv:2402.02995*.
- 605 Wills, S. M., D. W. J. Thompson, and L. M. Ciasto, 2016: On the observed relationships between  
606 variability in gulf stream sea surface temperatures and the atmospheric circulation over the north  
607 atlantic. *Journal of Climate*, **29 (10)**, 3719–3730.
- 608 Xiong, W., Y. Xiang, H. Wu, S. Zhou, Y. Sun, M. Ma, and X. Huang, 2023: Ai-goms: Large  
609 ai-driven global ocean modeling system. *arXiv preprint arXiv:2308.03152*.
- 610 Yang, H., R. Arnone, and J. Jolliff, 2015: Estimating advective near-surface currents from ocean  
611 color satellite images. *Remote Sensing of Environment*, **158**, 1–14.

- 612 Yu, Y., X. Si, C. Hu, and J. Zhang, 2019: A review of recurrent neural networks: LSTM cells and  
613 network architectures. *Neural computation*, **31** (7), 1235–1270.
- 614 Zambianchi, E., M. Trani, and P. Falco, 2017: Lagrangian transport of marine litter in the mediter-  
615 ranean sea. *Frontiers in Environmental Science*, **5**, 5.
- 616 Zheng, X., X. Peng, J. Zhao, and X. Wang, 2022: Trajectory prediction of marine moving target  
617 using deep neural networks with trajectory data. *Applied Sciences*, **12** (23), 11 905.