Neural prediction of Lagrangian drift trajectories on the sea surface

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

Neural prediction of Lagrangian drift trajectories on the sea surface

Daria Botvynko,[a, b, c] Carlos Granero-Belinchon,[b] Simon van Gennip,[c] Abdesslam Benzinou,[a] Ronan Fablet[b]

[a] *ENIB, Lab-STICC, UMR CNRS 6285, 29238 Brest, France*

[b] *IMT Atlantique, Lab-STICC, UMR CNRS 6285, 29238 Brest, France*

[c] *Mercator Ocean International, 31400 Toulouse, France*

Corresponding author: Daria Botvynko, daria.botvynko@enib.fr

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

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

1. Introduction

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

From a methodological standpoint, we can distinguish three main categories of approaches for Lagrangian drift simulation: model-based approaches [Liubartseva et al. (2018); Zambianchi et al. (2017)], probablistic data-driven schemes [Visser (2008)], and more recent deep learning techniques [Botvynko et al. (2023); Jenkins et al. (2023)]. Model-based approaches involve a sequential advection process based on known sea surface velocity fields [Lange and van Sebille (2017)]. However, small errors in the underlying velocity fields or velocity fields without very fine spatial resolution may lead to unrealistic Lagrangian trajectories [Callies et al. (2021)]. This often impedes their operational usuefulness due to significant uncertainties in the forecasting and reconstruction of sea surface dynamics, especially regarding mesoscale ocean dynamics [Prants et al. (2017)]. Due to their sequential nature, these schemes also face scalability challenges when simulating large ensembles of drift trajectories. By contrast, probabilistic data-driven schemes usually leverage first-order Markovian models and naturally account for uncertainties in the drift process.

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

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

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

2. Problem statement

In the context of Lagrangian fluid dynamics[1], one may describe a flow by following the motion of an ensemble of individual particles [Davis (1991)]. In order to describe the flow in the

[1]From now on, the Lagrangian notations are presented with $\rightarrow$symbol, and the Eulerian notations are in **bold**.

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

$$\frac{\partial \vec{r}(\vec{r_{t_0}},t)}{\partial t} = \vec{v}(\vec{r_{t_0}},t) = \mathbf{u}(\vec{r}(t),t) \tag{1}$$

where $\vec{r}(\vec{r_{t_0}},t)$ and $\vec{v}(\vec{r_{t_0}},t)$ are respectively the position and velocity at time t of the particle situated at $\vec{r_{t_0}}$ at $t = 0$, and $\mathbf{u}(\vec{r}(t),t)$ is the Eulerian velocity of the underlying flow given position $\vec{r}(t)$ and time t. Classical model-driven approaches simulate Lagrangian drift dynamics through explicit time and space integration methods such as Runge-Kutta 4. Thus, the integration of eq.(1) leads to :

$$\vec{r}(\vec{r_{t_0}},t+\Delta t) = \vec{r}(\vec{r_{t_0}},t) + \int_t^{t+\Delta t} \vec{v}(\vec{r_{t_0}},\tau)\mathrm{d}\tau \tag{2}$$

The sequential nature of this equation introduces important propagation of errors when the underlying velocities are not accurately known [Callies et al. (2021)].

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

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

where x represents spatial position in the Eulerian framework, t the time, $p_{\vec{r}}(x,t)$ the PDF of the particle at position $\vec{r}$ and $\mu(x,t)$ a drift field. In our case, this drift term relates to the underlying Eulerian velocity field $\mathbf{u}(x,t)$. Solving eq.(3) relies on a time integration scheme from the initial PDF $p_{\vec{r}}(x,t=0) = p_{\vec{r_{t_0}}}(x,0)$.

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

3. DriftNet for Lagrangian drift simulation

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

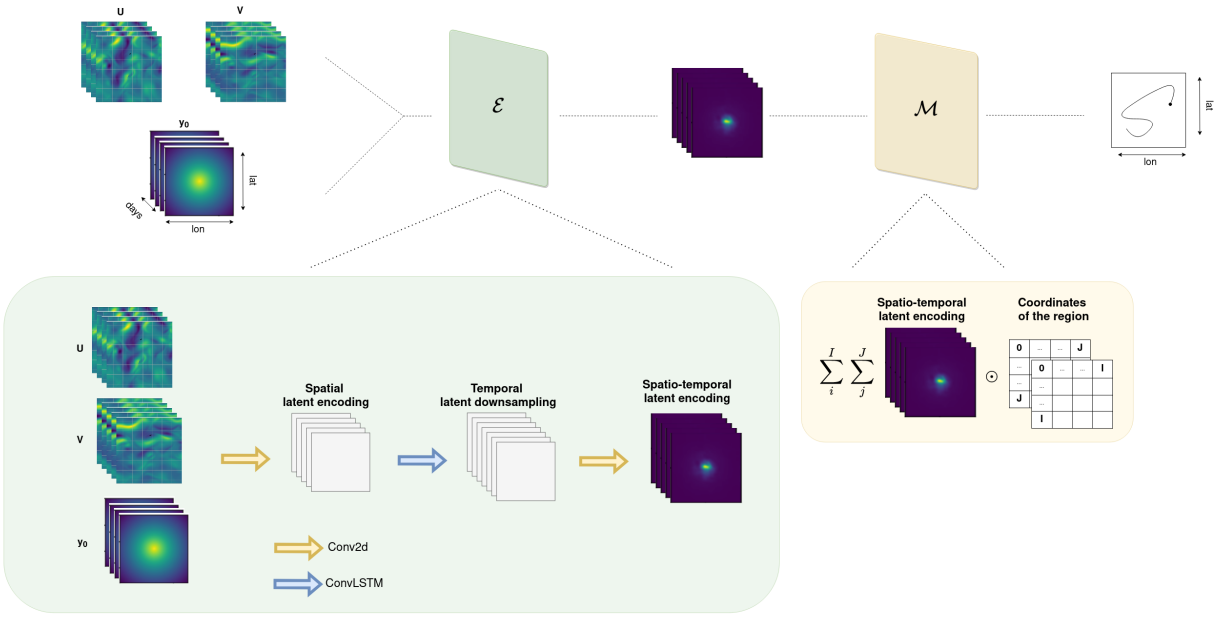

FIG. 1: **DriftNet architecture:** The input geophysical fields over 9-day (here zonal **U** and meridional **V** components of the velocity field) coupled to the initial spatio-temporal positional encoding $\mathbf{y_0}$ are first passed to the encoding block $\mathscr{E}$ and output a spatio-temporal latent encoding. The second block of DriftNet $\mathscr{M}$ maps this latent representation to a 1D trajectory, *i.e.* a time series of positions on the sea surface.

a. Proposed neural scheme

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

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

$$\left\{ \begin{array}{rcl} \mathbf{y} & = & \mathscr{E}(\mathbf{u}, \mathbf{y}_0) \\ \vec{r} & = & \mathscr{M}(\mathbf{y}) \end{array} \right. \tag{4}$$

where $\mathbf{y} = \{\mathbf{y}_{t_0}, \mathbf{y}_{t_0+\Delta}, \ldots, \mathbf{y}_{t_0+K\Delta}\}$ is a space-time-explicit latent embedding of $\vec{r}$, and $\mathbf{y_0} = \{\mathbf{y}_{0,t_0}, \mathbf{y}_{0,t_0+\Delta}, \ldots, \mathbf{y}_{0,t_0+K\Delta}\}$ some initial encoding of the initial position $\vec{r}_{t_0}$. In particular, we define $\mathbf{y}_{0,t_0}$ by assigning to each spatial pixel of the grid a value corresponding to the normalized distance

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

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

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

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

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

$$\mathscr{M} = x \cdot \mathbf{y}_t \tag{7}$$

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

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

1) a 2D convolutional layer increasing the number of channels from 3 (zonal and meridional components of $\mathbf{u}$ coupled to $\mathbf{y_0}$) to 64, followed by a LeakyReLU activation function.

2) 2D convolutional LSTM block [Shi et al. (2015)].

3) 2D convolutional layer decreasing the number of channels from 64 to 2, followed by 2D Softmax function.

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

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

apply the same neural architecture, just increasing the number of channels of the input fields. The same applies for a prior on the diffusivity from the initial location of a particle in the field, that can be also included by stacking with the temporal evolution of the initial condition $\mathbf{y_0}$.

b. Learning scheme

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

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

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

   where $\vec{r}_R$ and $\vec{r}_S$ are the vector of positions of the reference and simulated trajectories respectively.

2. Liu index between the reference and simulated trajectories [Liu and Weisberg (2011)]

$$\mathscr{L}_{Liu} = \frac{1}{N_T} \sum_{i=1}^{N_T} \frac{\sum_{j=1}^{K} d_{i_j}}{\sum_{j=1}^{K} l_{i_j}} \tag{9}$$

   where

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

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

Overall, the training loss $\mathscr{L}$ is a weighted sum: $\mathscr{L} = \alpha \cdot \mathscr{L}_{MSE} + \beta \cdot \mathscr{L}_{Liu}$. From cross-validaton experiments, we set α and β to 0.6 and 0.4. Using Pytorch[2], our learning setup relies on Adam optimizer with a learning rate of 5e-4 over 750 epochs.

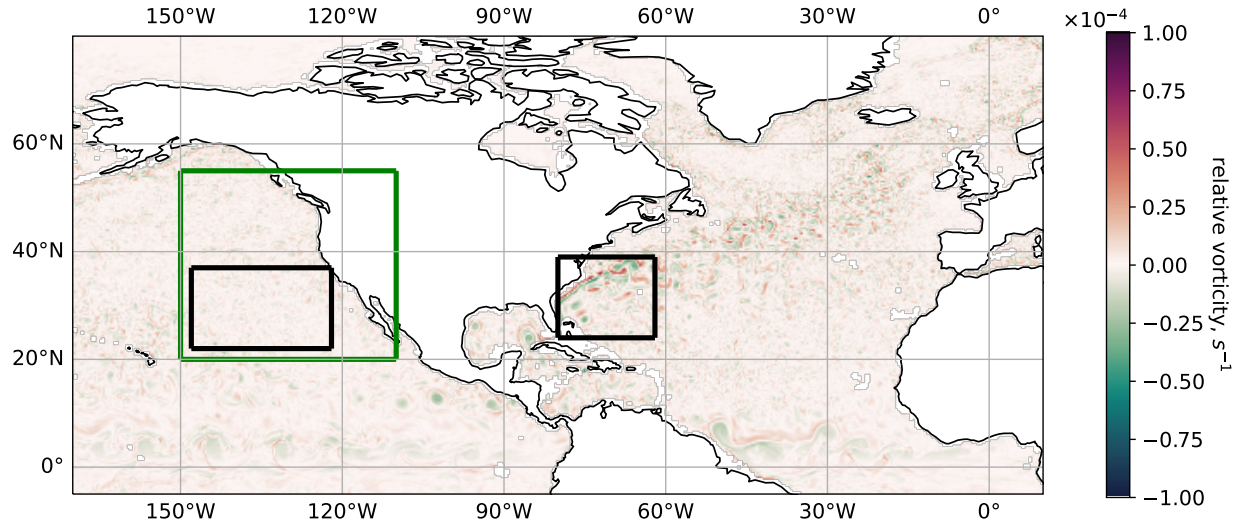

FIG. 2: **Study cases delineation:** Relative vorticity of North East Pacific and North Atlantic oceans from E1 simulation of 01/10/2015. The two black boxes indicate the North East Pacific and Gulf Stream case study regions for benchmarks B1 and B2. The green box defines the North East Pacific region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

4. Experimental workflow

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

a. Case-study Regions

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

[2]DriftNet code is available at `https://github.com/CIA-Oceanix/DriftNet`

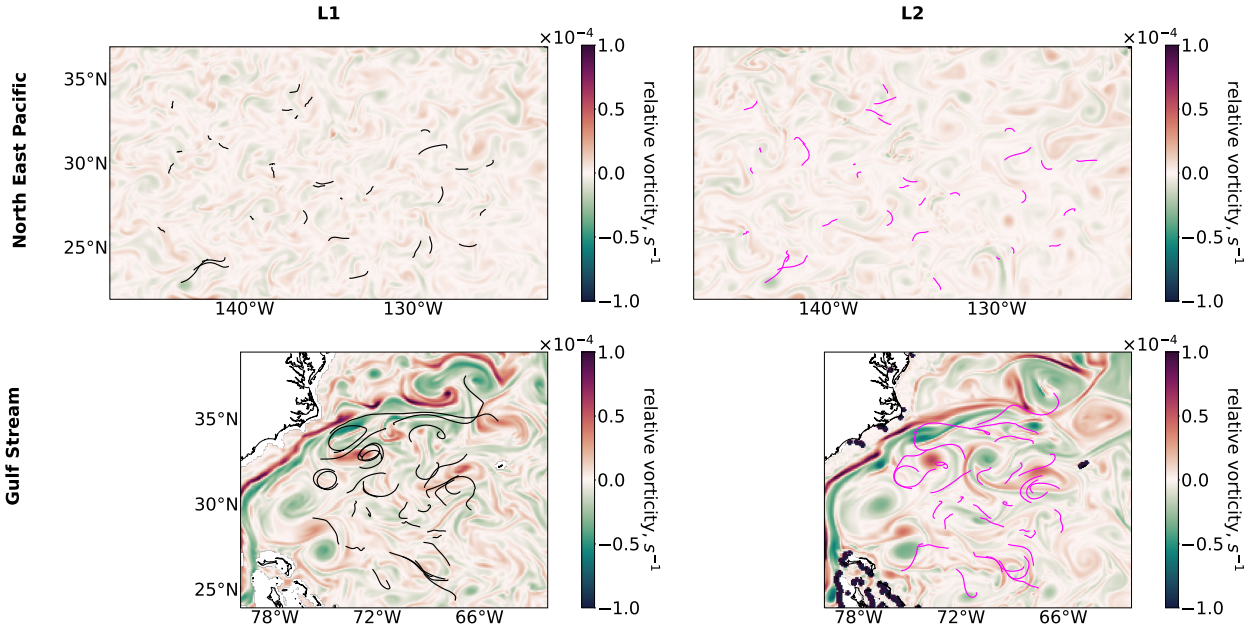

FIG. 3: **Examples of trajectories:** Trajectories from dataset L1 and dataset L2 superimposed respectively to the E1 and E2 relative vorticity fields of 2015/01/01.

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

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

b. Datasets

Our study involves three different Eulerian datasets used as input data for the benchmarked 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.

Name	*Source*	*Variable*	*Period*	*Spatial resolution*	*Temporal resolution*
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

- **Eulerian dataset E1:** This Eulerian dataset contains the sea surface current (SSC) velocity and sea surface height (SSH) fields. Both fields are obtained from a high-resolution free simulation of the NEMO ocean model without any data assimilation [Benkiran et al. (2021b)]. Those are regularly-gridded products with a horizontal spatial resolution of $1/12^\circ$ and a daily time resolution. This study involves the entire year of 2015.

- **Eulerian dataset E2** This dataset contains the SSC fields from a data assimilated run, which dynamically reconstructs the ocean state using pseudo observations from E1 which mimic the current observational network (satellite altimetry tracks). Those fields are at the same $1/12^\circ$ and daily space-time resolution as the E1 dataset. As in the previous E1 dataset, we equivalently use the complete 2015 year.

- **Eulerian Real-world dataset E3**: this real-world dataset combines the SSC from the operational ocean reanalysis GLORYS12 [Lellouche et al. (2021)] and the optimally-interpolated altimetry-derived SSH product DUACS [Pujol et al. (2016)]. The GLORYS12 velocity fields are provided on the regular grid with spatial resolution of 1/12° and temporal resolution of 1 day. We re-interpolate the SSH fields on the same 1/12° regular grid at daily temporal resolution.

For each individual trajectory, the considered spatial domain of the Eulerian conditional fields is 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 the domains corresponds to the furthest displacement from a given particle's deployment location in each case study region.

Our benchmarks involve four different Lagrangian datasets:

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

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

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

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

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

c. Performance metrics

As evaluation metrics, we consider

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

$$D = \frac{1}{N_T} \cdot \sum_{i=0}^{N_T} d_{i_9} \tag{11}$$

  with N_T the number of trajectories and d_{i_9} defined in eq.(10) with $j = 9$ days.

- the mean Liu index $\mathscr{L}_{Liu}$ without units, see eq.(9).

- the mean error between autocorrelation functions, $R_{\vec{u}}(\tau)$ and $R_{\vec{v}}(\tau)$ respectively for zonal and meridional components of trajectories' velocities [Krauß and Böning (1987); Kang et al. (2005); Wunsch (1999)]:

$$\Delta R_{\vec{u}} = \langle |\overline{R}_{\vec{u}_R}(\tau) - \overline{R}_{\vec{u}_S}(\tau)| \rangle \tag{12}$$

  where $\langle\rangle$ is the average across temporal dimension, $\overline{R}_{\vec{u}_R}(\tau)$ and $\overline{R}_{\vec{u}_S}(\tau)$ are respectively the ensemble mean of the reference and simulated normalized autocorrelation functions of the Lagrangian velocity, defined as: $R_{\vec{u}}(\tau) = \frac{1}{9 \cdot \Delta} \cdot \sum_{t=0}^{9} \vec{u}(t\Delta + \tau)\vec{u}(t\Delta)$. The same definition is used for the meridional velocity.

The mean euclidean distance and the mean Liu index are computed between the reference and generated trajectories and averaged over the whole ensemble of studied trajectories.

d. Benchmarked deep learning models

To evaluate the performance of DriftNet, we compare it to the following CNN and LSTM baselines inspired by [Zheng et al. (2022); Ma et al. (2019)]:

- Convolutional Neural Network (CNN): a 2D convolutional block varying the number of channels 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 $2{\cdot}35$. 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 $\mathscr{E}$;
- DriftNet without the CNN block in the encoding operator $\mathscr{E}$;
- a DriftNet-U-Net architecture, referred to as DriftNet$_U$: in this architecture, we replace the CNN block in the encoding operator $\mathscr{E}$ 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%) datasets. DriftNet is fed with E1 velocity fields and trained to reproduce reference trajectories from the L1 dataset. For benchmarking purposes, we consider the CNN and LSTM baselines trained following the same strategy as DriftNet.

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

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

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

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

5. Results

This section reports our numerical experiments for the three benchmarks introduced in the previous section, namely benchmark B1 (section 5.a), benchmark B2 (section 5.b) and benchmark B3 (section 5.c).

a. Results for Benchmark B1

DriftNet presents the best performance for all metrics and significantly outperforms the CNN and LSTM baselines, see Table 2. For the North East Pacific, the separation distance of 5.3km after 9 days is about one order of magnitude smaller than the baseline CNN and LSTM models. Similar performances are noted for Liu index, 0.06, and for the autocorrelation functions errors in both zonal and meridional components, 0.0103 and 0.0145. For the Gulf Stream, the overall accuracy is lower (separation distance after 9 days of 91.1km, Liu index 0.26 and autocorrelation

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 ($\mathscr{L}_{Liu}$) and error between autocorrelation functions ($\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$). Bold indicates the best performance. We refer the reader to section 4.d for the details of the benchmarked architectures.

	North East Pacific				**Gulf Stream**			
Method	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$
DriftNet	**5.3**	**0.06**	**0.0103**	**0.0145**	**91.1**	**0.26**	**0.0405**	**0.0348**
CNN	43	0.5	0.1076	0.1379	151.85	0.63	0.229	0.2173
LSTM	30.6	0.4	0.2136	0.2283	188	0.77	0.2244	0.2123

errors 0.0405 and 0.0348), outperforming 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 its 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 ConvLSTM blocks from the DriftNet architecture impacts the simulation performance, see Table 3. The separation distance increases from 5.3km to 36.2km (resp. 57.1km) for the North East Pacific region when we remove the CNN (resp. ConvLSTM) 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 U-Net configuration. Same qualitative results are found in the Gulf Stream region.

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 significant in the mean separation distance and the Liu index. Interestingly, the error between

TABLE 3: **Ablation study of DriftNet**: we compare the metrics of different configurations of DriftNet for benchmark B1: namely, the reference DriftNet, DriftNet$_U$ 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 ($\mathscr{L}_{Liu}$) and error between autocorrelation functions ($\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$). Bold indicates best performance.

	North East Pacific				**Gulf Stream**			
Method	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$
DriftNet	**5.3**	**0.06**	**0.0103**	**0.0145**	**91.1**	**0.26**	**0.0405**	**0.0348**
DriftNet$_U$	16.1	0.21	0.1552	0.1694	96.4	0.32	0.0573	0.0432
DriftNet without CNN	36.2	0.39	0.0299	0.0379	163	0.68	0.0834	0.1035
DriftNet without ConvLSTM	57.1	0.72	0.0846	0.1093	177	0.73	0.1568	0.1592

TABLE 4: **Impact of spatial resolution on DriftNet performance** Performance of DrifNet 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 ($\mathscr{L}_{Liu}$) and error between autocorrelation functions ($\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$). Bold indicates best performance.

	North East Pacific				**Gulf Stream**			
Resolution	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$
$1/12^\circ \times 1/12^\circ$	**5.3**	**0.06**	**0.0103**	**0.0145**	**91.1**	**0.26**	0.0405	0.0348
$1/6^\circ \times 1/6^\circ$	8.7	0.15	0.0162	0.0152	102.3	0.35	0.0377	**0.0255**
$1/3^\circ \times 1/3^\circ$	17.9	0.27	0.0272	0.034	109.4	0.47	**0.0376**	0.0313
$1/2^\circ \times 1/2^\circ$	28.2	0.40	0.0768	0.0995	115.9	0.46	0.048	0.0448

the autocorrelation functions in the Pacific region increases when the input resolution is coarsen, while in the Gulf Stream a lower input resolution allows to slightly improve the estimation of the zonal component for $1/3^\circ \times 1/3^\circ$ and the meridional component for $1/6^\circ \times 1/6^\circ$. 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 Lagragian 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 ($\mathscr{L}_{Liu}$) and error between autocorrelation functions($\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$). Bold indicates the best performance.

	North East Pacific				**Gulf Stream**			
Method	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$
Ocean Parcels	53.6	0.87	**0.0086**	**0.016**	168.3	0.75	**0.0084**	**0.0212**
DriftNet	**45.4**	0.7	0.1212	0.1202	**154.6**	**0.63**	0.1623	0.135
CNN	51.7	0.72	0.1492	0.1807	184.1	0.67	0.1603	0.1574
LSTM	47.9	**0.69**	0.2281	0.237	163.2	0.67	0.1742	0.157

DriftNet presents smaller D 45.4 km (154.6 km), $\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$ than the other learning-based models, and smaller D, $\mathscr{L}_{Liu}$ than Ocean Parcels, thus consistently outperforming the baseline approaches. The $\mathscr{L}_{Liu}$ for the North East Pacific region is of 0.7 and is smaller than the most baseline approaches, except the LSTM model (0.69) slightly outperforming the proposed method. While in the Gulf Stream the $\mathscr{L}_{Liu}$ (0.63) is smaller than all the other approaches. The autocorrelation errors is smaller for the proposed method that the other learning-based approaches in the North East Pacific region. In the Gulf Stream, the autocorrelation error in the meridional component is smaller than the other learning-based methods. For both regions, the proposed method's autocorrelation errors are larger than for the Ocean Parcels.

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

These results emphasize the synergistic potential of SSH observations and E2 velocity fields compared with using only one of these two input data. In the North East Pacific region, D is reduced by over 55%, $\mathscr{L}_{Liu}$ by over 60%, $\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$ by over 60% and 50%. Similarly in

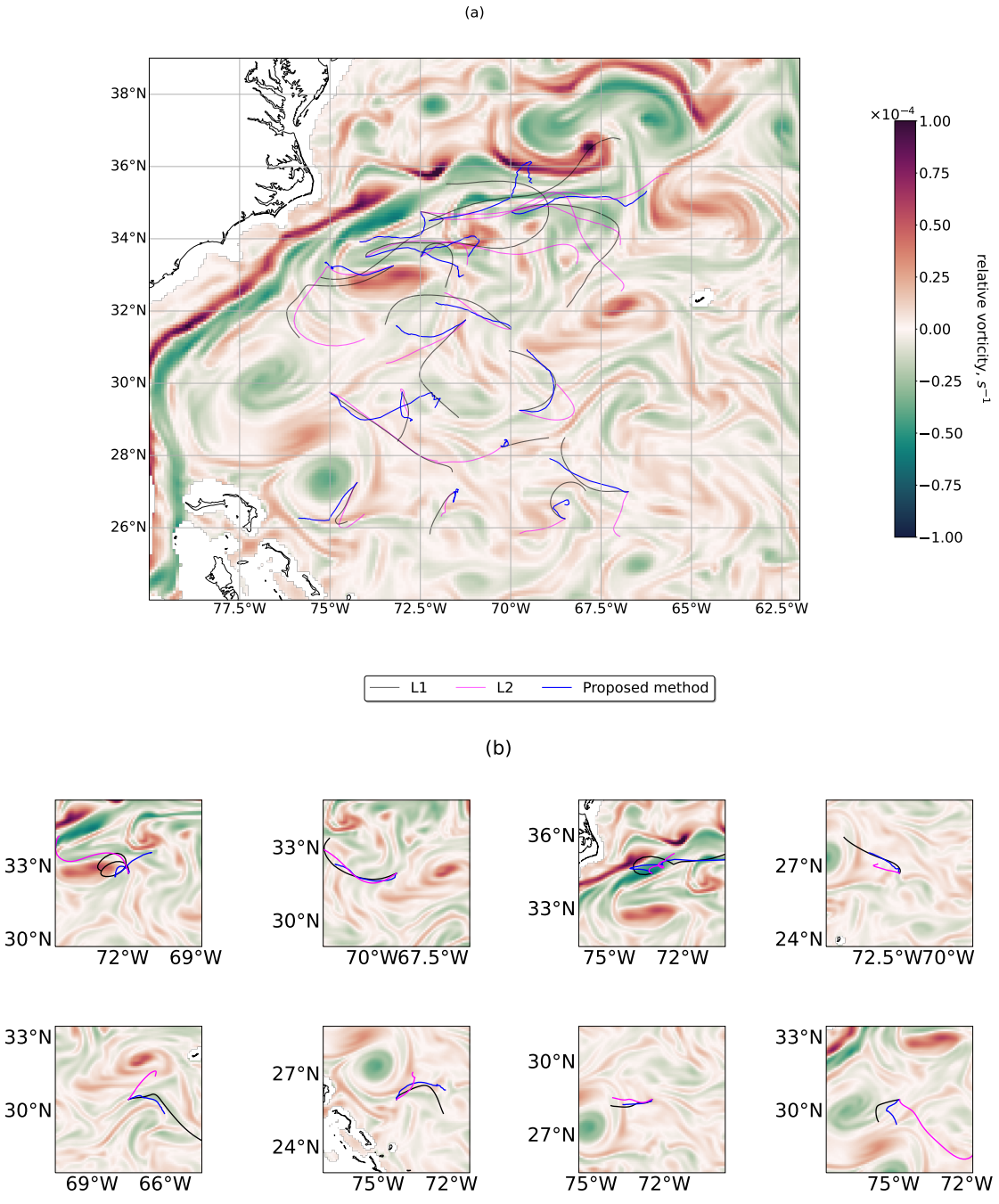

FIG. 4: **Trajectories for Benchmark B2 in the Gulf Stream** Lagrangian trajectories from Dataset L2 are depicted in magenta, trajectories simulated with DriftNet conditioned with SSC from E2 and SSH from E1 in blue, and trajectories from Dataset L1 in black. For better visual appreciation panel a) shows 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.

the Gulf Stream, we report a relative gain greater than 25% for D, 30% for $\mathscr{L}_{Liu}$. The $\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$ however are reduced by over 56% and 57% for the case of SSH fields input only. These

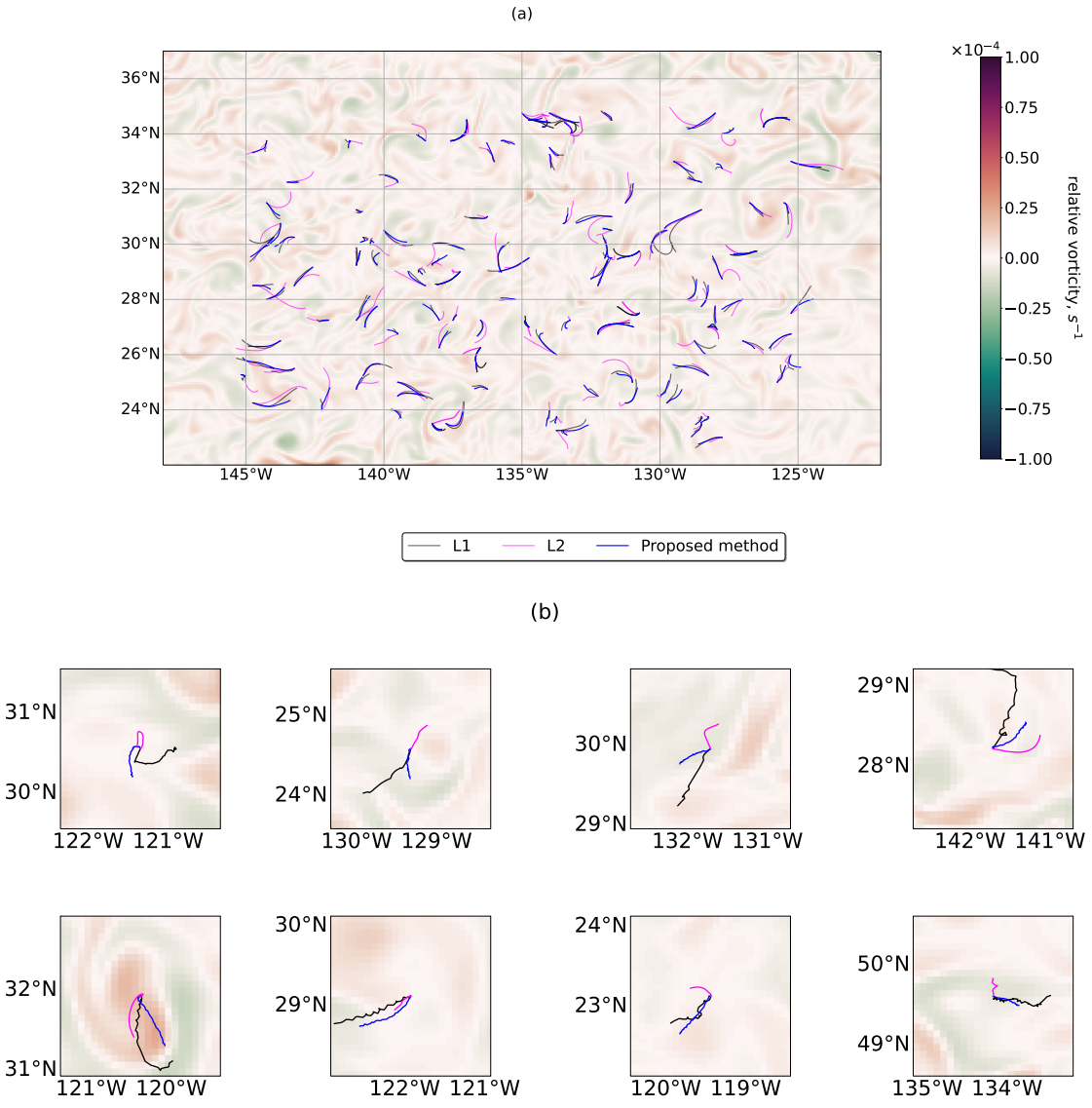

FIG. 5: **Trajectories for Benchmark B2 in the North East Pacific** Lagrangian trajectories from Dataset L2 are depicted in magenta, trajectories simulated with DriftNet conditioned with SSC from E2 and SSH from E1 in blue, and trajectories from Dataset L1 in black. a) shows 128 randomly picked trajectories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

results are in agreement with the ability of observation-based SSH products to capture geostrophic dynamics [Ballarotta et al. (2022)], while data-assimilation-based velocity fields can reveal the ageostrophic sea surface dynamics. The latter are expected to be significant for horizontal scales below a few hundreds of kilometers, especially in Western Boundary currents such as the Gulf Stream [Gula et al. (2015)].

TABLE 6: **Synergistic use of SSH observations and data-assimilation-based velocity fields in Benchmark B2**. We becnhmark 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 ($\mathscr{L}_{Liu}$) and error between autocorrelation functions ($\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$). Bold indicates the best performance.

	North East Pacific				**Gulf Stream**			
Method(*Variable*)	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$
DriftNet	45.4	0.7	0.1212	0.1202	154.6	0.63	0.1623	0.135
DriftNet (SSH)	31.2	0.32	0.059	0.0672	160.3	0.7	**0.0709**	**0.057**
DriftNet (U, V, SSH)	**22.2**	**0.29**	**0.0456**	**0.0579**	**127**	**0.52**	0.0833	0.0616

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 ($\mathscr{L}_{Liu}$) and mean absolute error between autocorrelation functions ($\Delta R_{\vec{u}}$ and $\Delta R_{\vec{v}}$). Bold indicates the best performance.

	North East Pacific			
Method(*Variable*)	D, km	$\mathscr{L}_{Liu}$	$\Delta R_{\vec{u}}$	$\Delta R_{\vec{v}}$
Ocean Parcels (U, V)	73.77	0.65	0.1374	0.1471
DriftNet (U, V)	62.4	0.55	0.0522	**0.0315**
CNN (U, V)	68.64	0.57	0.0466	0.0621
LSTM (U, V)	64.3	0.55	0.0982	0.0893
DriftNet (U, V, SSH)	**60.2**	**0.52**	**0.0296**	0.034

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

c. Results for Benchmark B3

In this real case situation, DriftNet outperforms the classical advection method, see Table 7 and Figure 6. Compared to the model-based baseline, the mean separation distance is reduced by around 15% to 62.4km, the Liu index by 15% to 0.55. Unlike previous B2 experiment, the autocorrelation errors are significantly reduced to 0.0522 and 0.0315 compared to Ocean Parcels. As

in benchmark B2, we observe a significant impact of the observed SSH fields in the quality of the simulation of the Lagrangian trajectories. The combination of GLORYS12 velocities and satellite-derived SSH fields leads to a reduction of the mean separation distance by almost 4% compared to the baseline DriftNet and by over 18% compared to the baseline model-based advection method. The Liu index is also reduced by more than 5% compared to the DriftNet baseline and by over 20% compared to the baseline advection method. The autocorrelation error in the zonal component is reduced by more than 43% compared to the DriftNet baseline and by over 78% compared to the baseline model-based approach.

6. Conclusions and Discussion

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

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

The performance of DriftNet exhibits variabilities across the two different regions. It likely underscores the importance of region-specific adaptations in trajectory simulation models. As such, the deployment of DriftNet on a global scale may require retraining the model to accommodate

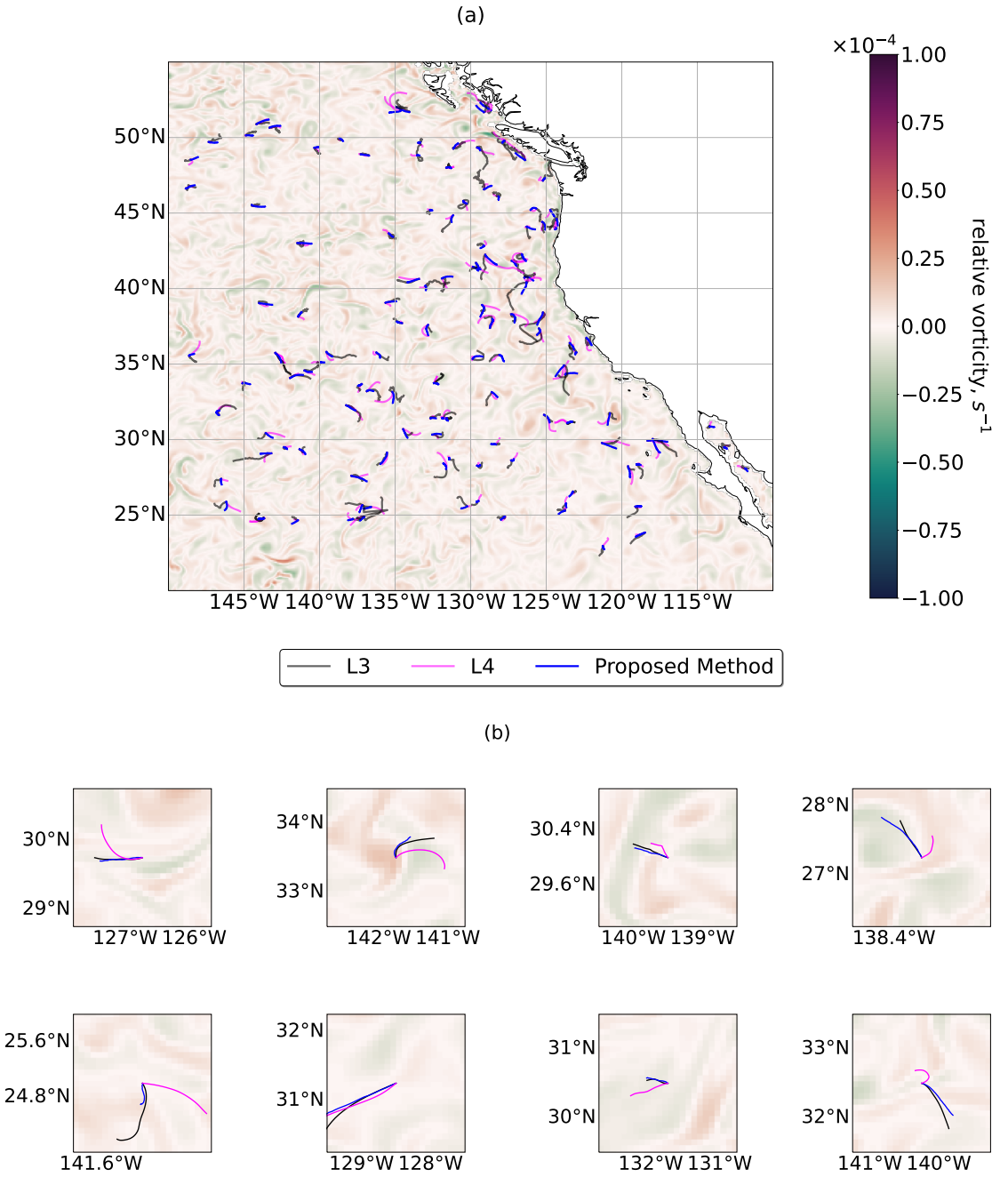

FIG. 6: **Examples of simulated trajectories for Benchmark B3**: we depict real drifters trajectories in black (L3), trajectories simulated with Ocean Parcels using GLORYS12 velocity fields in magenta (L4) and trajectories simulated with DriftNet using both GLORYS12 velocity fields and satellite-derived 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.

the diverse dynamics in various oceanic regions [Miao et al. (2023)]. While we expect the proposed neural architecture to be generic, fine-tuning its parameters to align with the characteristics of a given region seems crucial for the relevance of the simulations. This approach allows for the incorporation of region-specific features, thereby enhancing the model's predictive capabilities across diverse geographical regions. Furthermore, leveraging transfer learning techniques to capitalize on the knowledge gained from training in one region to inform model adjustments in others holds promising avenue. By iteratively refining DriftNet through region-specific training and fine-tuning iterations, we hope to develop a robust trajectory simulation tool capable of capturing the intricacies of oceanic dynamics on a global scale.

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 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 the sea surface currents, which is important for the two case-study regions for horizontal scales above 100km [Checkley and Barth (2009); Thomas and Joyce (2010); Johns et al. (1989)]. As ocean 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 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 satellite altimeters [Benkiran et al. (2021b); Tchonang et al. (2021)], possibly combined with neural mapping schemes [Martin et al. (2023); Beauchamp et al. (2022); Fablet et al. (2023)], will lead to a better reconstruction of satellite-derived SSH fields, which could in turn result in improved learning-based Lagrangian simulations for real drifters on the sea surface.

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 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); Solanlki et al. (2001)] or Waves Tang et al. (2007) to refine trajectory simulations and gain deeper insights into oceanic processes. By leveraging these variables, the proposed method could

not only improve trajectory accuracy but also facilitate comprehensive analysis of the underlying ocean dynamics.

Inspired by established initiatives such as WeatherBench [Rasp et al. (2020)] and OceanBench [Johnson et al. (2024)], we propose a new benchmark for a standardized assessing and comparing Lagrangian trajectory simulation models https://github.com/CIA-Oceanix/DriftNet. Through proposed datasets and evaluation metrics, this framework seeks to foster collaborative frameworks involving ocean scientists and ML practitioners. In this context, future work could explore new evaluation metrics and extend the proposed benchmarks to a broader array of oceanic regions.

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

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

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

References

Auad, G., D. Roemmich, and J. Gilson, 2011: The california current system in relation to the northeast pacific ocean circulation. *Progress in Oceanography*, **91 (4)**, 576–592.

Ballarotta, M., and Coauthors, 2022: Improved global sea surface height and currents maps from remote sensing and in situ observations. *Earth System Science Data Discussions*, **2022**, 1–32.

Barron, C. N., L. F. Smedstad, J. M. Dastugue, and O. M. Smedstad, 2007: Evaluation of ocean models using observed and simulated drifter trajectories: Impact of sea surface height on synthetic profiles for data assimilation. *Journal of Geophysical Research: Oceans*, **112 (C7)**.

Beauchamp, M., Q. Febvre, H. Georgenthum, and R. Fablet, 2022: 4dvarnet-ssh: end-to-end learning of variational interpolation schemes for nadir and wide-swath satellite altimetry. *Geoscientific Model Development Discussions*, **2022**, 1–37.

Benkiran, M., P.-Y. Le Traon, and G. Dibarboure, 2021a: Contribution of a constellation of two wide-swath altimetry missions to global ocean analysis and forecasting. *Ocean Science Discussions*, **2021**, 1–25.

Benkiran, M., P.-Y. Le Traon, E. Rémy, and Y. Drillet, 2024: Impact of two high resolution altimetry mission concepts for ocean forecasting. *EGUsphere*, **2024**, 1–19.

Benkiran, M., and Coauthors, 2021b: Assessing the impact of the assimilation of swot observations in a global high-resolution analysis and forecasting system part 1: Methods. *Frontiers in Marine Science*, **8**, 691 955.

Botvynko, D., C. Granero-Belinchon, S. van Gennip, A. Benzinou, and R. Fablet, 2023: Deep learning for lagrangian drift simulation at the sea surface. *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.

Breivik, O., A. A. Allen, C. Maisondieu, and M. Olagnon, 2013: Advances in search and rescue at sea. *Ocean Dynamics*, **63**, 83–88.

Callies, U., M. Kreus, W. Petersen, and Y. G. Voynova, 2021: On using lagrangian drift simulations to aid interpretation of in situ monitoring data. *Frontiers in Marine Science*, **8**, 666 653.

Checkley, D. M., and J. A. Barth, 2009: Patterns and processes in the california current system. *Progress in Oceanography*, **83 (1-4)**, 49–64.

Ciani, D., M.-H. Rio, B. B. Nardelli, H. Etienne, and R. Santoleri, 2020: Improving the altimeter-derived surface currents using sea surface temperature (SST) data: A sensitivity study to SST products. *Remote Sensing*, **12 (10)**, 1601.

Cucco, A., G. Quattrocchi, A. Satta, F. Antognarelli, F. De Biasio, E. Cadau, G. Umgiesser, and S. Zecchetto, 2016: Predictability of wind-induced sea surface transport in coastal areas. *Journal of Geophysical Research: Oceans*, **121 (8)**, 5847–5871.

Davis, R. E., 1991: Lagrangian ocean studies. *Annual Review of Fluid Mechanics*, **23 (1)**, 43–64.

Dewar, W. K., and J. M. Bane, 1989: Gulf stream dynamics. pad II: Eddy energetics at 73 w. *Journal of Physical Oceanography*, **19 (10)**, 1574–1587.

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

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

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

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

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

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

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

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

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

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

Johnson, J. E., Q. Febvre, A. Gorbunova, S. Metref, M. Ballarotta, and J. e. a. Le Sommer, 2024: Oceanbench: The sea surface height edition. *Advances in Neural Information Processing Systems*, **36**.

Julka, S., V. Sowrirajan, J. Schloetterer, and M. Granitzer, 2021: Conditional generative adversarial networks for speed control in trajectory simulation. *International Conference on Machine Learning, Optimization, and Data Science*, Springer, 436–450.

Kang, Y., K. Morooka, and H. Nagahashi, 2005: Scale invariant texture analysis using multi-scale local autocorrelation features. *Scale Space and PDE Methods in Computer Vision: 5th International Conference, Scale-Space 2005, Hofgeismar, Germany, April 7-9, 2005. Proceedings 5*, Springer, 363–373.

Kaplan, D. M., J. Largier, and L. W. Botsford, 2005: HF radar observations of surface circulation off bodega bay (northern california, USA). *Journal of Geophysical Research: Oceans*, **110**.

Korbmacher, R., and A. Tordeux, 2022: Review of pedestrian trajectory prediction methods: Comparing deep learning and knowledge-based approaches. *IEEE Transactions on Intelligent Transportation Systems*, **23 (12)**, 24 126–24 144.

Koszalka, I., J. H. LaCasce, M. Andersson, K. A. Orvik, and C. Mauritzen, 2011: Surface circulation in the nordic seas from clustered drifters. *Deep Sea Research Part I: Oceanographic Research Papers*, **58 (4)**, 468–485.

Krauß, W., and C. W. Böning, 1987: Lagrangian properties of eddy fields in the northern north atlantic as deduced from satellite-tracked buoys. *Journal of Marine Research*, **45 (2)**, 259–291.

Lange, M., and E. van Sebille, 2017: Parcels v0. 9: prototyping a lagrangian ocean analysis framework for the petascale age. *Geoscientific Model Development*, **10 (11)**, 4175–4186.

Lellouche, J. M., E. Greiner, R. Bourdalle-Badie, G. Garric, A. Melet, M. Drévillon, C. Bricaud, and et al., 2021: The copernicus global 1/12 oceanic and sea ice GLORYS12 reanalysis. *Frontiers in Earth Science*, 698876.

Lellouche, J. M., and Coauthors, 2018: Recent upyears to the copernicus marine service global ocean monitoring and forecasting real-time 1/ 12 high-resolution system. *Ocean Science*, **14 (5)**, 1093–1126.

Liu, J., W. J. Emery, X. Wu, M. Li, C. Li, and L. Zhang, 2017: Computing ocean surface currents from GOCI ocean color satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **55 (12)**, 7113–7125.

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

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

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

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

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

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

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

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

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

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

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

Paz, D., H. Zhang, and H. I. Christensen, 2021: Tridentnet: A conditional generative model for dynamic trajectory generation. *International Conference on Intelligent Autonomous Systems*, Springer, 403–416.

Prants, S. V., M. Y. Uleysky, and M. V. Budyansky, 2017: *Lagrangian oceanography: large-scale transport and mixing in the ocean*. Springer.

Pujol, M.-I., Y. Faugère, G. Taburet, S. Dupuy, C. Pelloquin, M. Ablain, and N. Picot, 2016: DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years. *Ocean Science*, **12 (5)**, 1067–1090.

Rasp, S., P. D. Dueben, S. Scher, J. A. Weyn, S. Mouatadid, and N. Thuerey, 2020: Weatherbench: a benchmark data set for data-driven weather forecasting. *Journal of Advances in Modeling Earth Systems*, **12 (11)**, e2020MS002 203.

Roy, A., R. Fablet, and S. L. Bertrand, 2022: Using generative adversarial networks (GAN) to simulate central-place foraging trajectories. *Methods in Ecology and Evolution*, **13 (6)**, 1275–1287.

Röhrs, J., and Coauthors, 2021: Surface currents in operational oceanography: Key applications, mechanisms, and methods. *Journal of Operational Oceanography*, **16 (1)**, 60–88.

Shi, X., Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, 2015: Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, **28**.

Solanlki, H. U., R. M. Dwivedi, S. R. Nayak, J. V. Jadeja, D. B. Thakar, H. B. Dave, and M. I. Patel, 2001: Application of ocean colour monitor chlorophyll and AVHRR SST for fishery forecast:

Preliminary validation results off gujarat coast, northwest coast of india. *Indian Journal of Geo-Marine Sciences*, **30(3)**, 132–138.

Son, Y. B., B.-J. Choi, Y. H. Kim, and Y.-G. Park, 2015: Tracing floating green algae blooms in the yellow sea and the east china sea using GOCI satellite data and lagrangian transport simulations. *Remote Sensing of Environment*, **156**, 21–33.

Talley, L. D., 2011: *Descriptive physical oceanography: an introduction*. Academic press.

Tang, C. L., W. Perrie, A. D. Jenkins, B. M. DeTracey, Y. Hu, B. Toulany, and P. C. Smith, 2007: Observation and modeling of surface currents on the grand banks: A study of the wave effects on surface currents. *Journal of Geophysical Research: Oceans*, **112**.

Taqi, A. M., A. M. Al-Subhi, M. A. Alsaafani, and C. P. Abdulla, 2019: Estimation of geostrophic current in the red sea based on sea level anomalies derived from extended satellite altimetry data. *Ocean Science*, **15 (3)**, 477–488.

Tashiro, Y., J. Song, Y. Song, and S. Ermon, 2021: Csdi: Conditional score-based diffusion models for probabilistic time series imputation. *Advances in Neural Information Processing Systems*, **34**, 24 804–24 816.

Tchonang, B. C., M. Benkiran, P.-Y. Le Traon, S. Jan van Gennip, J. M. Lellouche, and G. Ruggiero, 2021: Assessing the impact of the assimilation of swot observations in a global high-resolution analysis and forecasting system–part 2: Results. *Frontiers in Marine Science*, **8**, 687 414.

Thomas, L. N., and T. M. Joyce, 2010: Subduction on the northern and southern flanks of the gulf stream. *Journal of physical oceanography*, **40 (2)**, 429–438.

Visser, A. W., 2008: Lagrangian modelling of plankton motion: From deceptively simple random walks to fokker–planck and back again. *Journal of Marine Systems*, **70 (3-4)**, 287–299.

Wang, X., and Coauthors, 2024: Xihe: A data-driven model for global ocean eddy-resolving forecasting. *arXiv preprint arXiv:2402.02995*.

Wills, S. M., D. W. J. Thompson, and L. M. Ciasto, 2016: On the observed relationships between variability in gulf stream sea surface temperatures and the atmospheric circulation over the north atlantic. *Journal of Climate*, **29 (10)**, 3719–3730.

Wunsch, C., 1999: The interpretation of short climate records, with comments on the north atlantic and southern oscillations. *Bulletin of the american meteorological society*, **80 (2)**, 245–256.

Xiong, W., Y. Xiang, H. Wu, S. Zhou, Y. Sun, M. Ma, and X. Huang, 2023: Ai-goms: Large ai-driven global ocean modeling system. *arXiv preprint arXiv:2308.03152*.

Yang, H., R. Arnone, and J. Jolliff, 2015: Estimating advective near-surface currents from ocean color satellite images. *Remote Sensing of Environment*, **158**, 1–14.

Yu, Y., X. Si, C. Hu, and J. Zhang, 2019: A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, **31 (7)**, 1235–1270.

Zambianchi, E., M. Trani, and P. Falco, 2017: Lagrangian transport of marine litter in the mediterranean sea. *Frontiers in Environmental Science*, **5**, 5.

Zheng, X., X. Peng, J. Zhao, and X. Wang, 2022: Trajectory prediction of marine moving target using deep neural networks with trajectory data. *Applied Sciences*, **12 (23)**, 11 905.