

Reconstruction of Submesoscale Velocity Field from Surface Drifters[©]

RAFAEL C. GONÇALVES, MOHAMED ISKANDARANI, AND TAMAY ÖZGÖKMEN

Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida

W. CARLISLE THACKER

Miami, Florida

(Manuscript received 6 February 2018, in final form 14 January 2019)

ABSTRACT

The extensive drifter deployment during the Lagrangian Submesoscale Experiment (LASER) provided observations of the surface velocity field in the northern Gulf of Mexico with high resolution in space and time. Here, we estimate the submesoscale velocity field sampled by those drifters using a procedure that statistically interpolates these data both spatially and temporally. Because the spacing of the drifters evolves with the flow, causing the resolution that they provide to vary in space and time, it is important to be able to characterize where and when the estimated velocity field is more or less accurate, which we do by providing fields of interpolation errors. Our interpolation uses a squared-exponential covariance function characterizing correlations in latitude, longitude, and time. Two novelties in our approach are 1) the use of two scales of variation per dimension in the covariance function and 2) allowing the data to determine these scales along with the appropriate amplitude of observational noise at these scales. We present the evolution of the reconstructed velocity field along with maps of relative vorticity, horizontal divergence, and lateral strain rate. The reconstructed velocity field exhibits horizontal length scales of 0.4–3.5 km and time scales of 0.6–3 h, and features with convergence up to 8 times the planetary vorticity f , lateral strain rate up to $10f$, and relative vorticity up to $13f$. Our results point to the existence of a vigorous and substantial ageostrophic circulation in the submesoscale range.

1. Introduction

The Lagrangian Submesoscale Experiment (LASER) was conducted with the arduous and challenging task of observing submesoscale activity in the northern Gulf of Mexico. Submesoscale processes comprise horizontal spatial scales of $\mathcal{O}(0.1\text{--}10)$ km and time scales of $\mathcal{O}(1\text{--}100)$ h, which are too short to be resolved by satellite altimeters, and extremely hard to observe by ship surveys. In LASER, the task was approached by an extensive drifter deployment, which provided measurements of submesoscales at unprecedented high resolution (D'Asaro et al. 2018; Haza et al. 2018; Novelli et al. 2017). The objective of this study is to reconstruct the Eulerian

velocity of an evolving surface flow field sampled by a subset of those drifters, and to present estimates of its differential kinematic properties, namely horizontal divergence, the vertical component of relative vorticity and lateral strain rate, and dominant space–time scales. One specific goal here is to identify submesoscale processes in an Eulerian frame, providing not only statistics about the flow field, but also a visualization of its evolution in time.

Submesoscale features have been known to oceanographers for some time. Spiral eddies with length scales of $\mathcal{O}(1\text{--}10)$ km were first observed in the sun glitter of photographs taken from space by the Apollo mission in the late 1960s. These spiral eddies were later one of the targets of the space shuttle *Challenger* mission 41-G in 1984, whose sun-glitter images revealed a large occurrence of these features embedded in complex fields of shear currents (Scully-Power 1986; Munk et al. 2000). Although those images suggested a potential relevance of the spiral eddies to the ocean circulation, they only provided brief snapshots of these processes and did not

[©] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JPO-D-18-0025.s1>.

Corresponding author: Rafael C. Gonçalves, rgoncalves@miami.edu

yield information on the underlying dynamics (Munk et al. 2000). As mentioned above, the relatively small space and time scales of submesoscale flows raise formidable challenges for traditional ocean observation techniques; these difficulties hindered progress in understanding submesoscale dynamics in the 1980s and 1990s (McWilliams 2016).

The increased computer power over the last decade enabled theoretical studies about submesoscale turbulence based on numerical models, and their results have drawn attention to the potential role of submesoscales on the ocean energy cascade, and on mixing and material transport. The flow regime at these scales is only partially constrained by geostrophy [$\mathcal{O}(1)$ Rossby number] and hydrostatic balance, and represents a transition zone from the quasi-two-dimensional mesoscale flows to smaller-scale and more three-dimensional flows (Thomas and Ferrari 2008; McWilliams 2016). Studies have shown that submesoscale frontal instabilities extract energy from the mesoscale strain field and transfer it to smaller scales (Capet et al. 2008). By providing a forward energy cascade route to dissipation, the submesoscales might play a key role in solving the dilemma of how energy is dissipated from geostrophically balanced flows (Molemaker et al. 2010), as geostrophic turbulence theory predicts an inverse kinetic energy cascade from meso- to large scales, and an enstrophy cascade from meso- to smaller scales (Charney 1971). Also, an active submesoscale field has direct consequences for oceanic transport and mixing as it controls dispersion on local scales (Özgökmen et al. 2012), in contrast to nonlocal dispersion under an enstrophy cascade regime (Bennett 1984).

Observations of submesoscales are still scarce, but they have been increasing in number in recent years and are providing evidence of the importance of submesoscale flows to ocean dynamics and material transport. In Flament et al. (1985), a succession of satellite infrared images of sea surface temperature (SST) off the California shelf presented the evolution of a cold filament under frontal instability on its cyclonic side, and hydrographic measurements showed the presence of frontal subduction. Analyzing a similar feature, Flament and Armi (2000) were able to estimate velocity gradients from clusters of drifters and ship drift over a cyclonic front sharper than 1 km. The authors found horizontal velocity shears larger than $7.5f$ (f is the planetary vorticity), and horizontal convergence larger than $0.8f$. D'Asaro et al. (2011) used a neutrally buoyant Lagrangian float and a towed Triaxus profiling vehicle to take measurements at a frontal zone in the Kuroshio extension. The authors identified enhanced turbulent activity at a sharp front, marked with vigorous vertical

displacements of the float, and verified that the enhanced turbulence was extracting energy from the front itself by a symmetric instability. Poje et al. (2014) used two-particle statistics from about 300 surface drifters to study the structure of velocity fluctuations from 200-m to 50-km scales. Their results show that particle dispersion was controlled by local scales in the submesoscale range, indicating the presence of an energetic submesoscale flow field, and that the underlying Eulerian kinetic energy spectrum at these scales was shallower than the one expected in an enstrophy cascade regime. Some recent studies have identified submesoscale activity by estimates of differential kinematic properties from the velocity field, as velocity gradients with $\mathcal{O}(f)$ are expected at these scales.¹ In Shcherbina et al. (2013), ADCP velocity measurements from two vessels running on parallel tracks were used to estimate the full gradient tensor of the horizontal velocity field in the North Atlantic Mode Water region. Their results show the presence of strong cyclonic vorticity, up to $3f$, over a mainly weak anticyclonic background. In Berta et al. (2016) and Ohlmann et al. (2017), estimates of flow kinematics at multiple scales were computed from clusters of drifters. In both studies, convergence around $\mathcal{O}(1\text{--}10)f$ were found in submesoscale fronts. The latter study also estimated relative vorticity larger than $5f$ over submesoscale eddies. Raschke et al. (2017) used sun-glitter images taken from airplane to observe finescale sea surface roughness, and X-band radar to estimate the velocity gradients over submesoscale fronts, with both datasets collected during LASER. Velocity estimates from the X-band radar in a 500-m-resolution grid showed the presence of an intense front, characterized by across-front convergence and alongfront shear, and with cyclonic vorticity reaching up to $5f$. In the sea surface roughness anomaly estimated from the sun-glitter images of the same area, they measured a 50-m-wide front, and the sharpness of the roughness anomaly indicated velocity gradients on the order of $80f$.

The philosophy of our approach is based on using recent supervised machine-learning methods to infer information about the Eulerian velocity field from Lagrangian data. This information includes estimates of the velocity field at unsampled points, querying the observations for the dominant length and time scales experienced by the drifters and inferring the derived quantities. All information is gleamed from the drifter data, without recourse to outside sources. The technique applied here is Gaussian Process Regression (GPR),

¹ For large and mesoscales, velocity gradients much smaller than f are expected.

which has wide application in geostatistics and problems of forward propagation of uncertainty in numerical models (Kennedy and O'Hagan 2000; Rasmussen and Williams 2006; Thacker et al. 2015; Iskandarani et al. 2016). The analysis is centered on one specific event when 326 drifters were released inside a frontal cyclonic eddy. The dense space–time information enabled us to estimate the time evolution of the Eulerian velocity field, which is presented along with the evolving relative vorticity, divergence and strain rate inside the cyclone. To our knowledge, this is the first time that submesoscales are identified by the reconstruction of an Eulerian velocity field inferred entirely from surface drifters.

The plan for the paper is as follows: section 2 describes the LASER data, highlights evidence of submesoscale activity in the selected observations to be used in the GPR, and presents the GPR procedures for the space–time reconstruction and estimation of the space–time scales. Section 3 presents results and discussion, with the space–time scales estimation, the validation of the reconstruction procedure against withheld data, and the results of the analysis. Finally, section 4 presents a summary of the main findings and conclusion.

2. Material and methods

a. Data

The data used in this study are estimates of positions and velocities of drifters released during LASER, which was conducted in the northern Gulf of Mexico around the DeSoto Canyon. Two ships and two airplanes collected data under El Niño winter conditions from 18 January 2016 to 15 February 2016. The main objective of this experiment was to sample transient submesoscale features in the outflow of the Mississippi river using hundreds of drifters. The biodegradable surface drifters were developed by CARTHE for LASER (Novelli et al. 2017). They were equipped with a small drogue to sample the upper 0.6-m circulation, and they reported their GPS position every 5 min. [For more information on the drifter design, the reader is referred to Novelli et al. (2017).] Before we received the data, the GPS positions had been processed to identify drogue loss, trajectories had been filtered with an acceleration-based filter to eliminate positioning errors, positions had been interpolated into regular 15-min intervals, and horizontal components of velocity were estimated with centered finite differences of the drifters' displacements (Haza et al. 2018). We used only data from drifters with drogues attached.

b. Submesoscale observations

We selected 326 drifters that were deployed on the western side of the DeSoto Canyon, about 100 km from the Mississippi River Delta. Their deployment started at 0200 UTC 7 February 2016 and lasted for about 14 h. At the time of deployment, this area presented strong density gradients induced by the lighter, fresher, and colder water from the Mississippi River plume and the heavier, saltier, and warmer waters from the Gulf of Mexico (D'Asaro et al. 2018). The drifters were released over a strong cyclonic vortex, which can be seen in the satellite images of SST and chlorophyll-a (Figs. 1a,b). The observed SST from an aircraft survey, taken about 6 h before the beginning of the drifter deployment, presents a more detailed picture (Fig. 1c), with the presence of colder and warmer filaments spiraling inside the cyclone. Figure 1c also presents temperatures at 10-m depth measured by moving vessel profiler (MVP) CTD casts right after the end of the deployments (colored dots). The MVP sections present a cold water pattern around 88.5°W and 88.45°W, which could be explained by the waters of a strong cold water filament observed in the aircraft SST, although the former is displaced to the west/northwest in relation to the latter. The offset is probably due to advection of the water masses by a background flow, as the MVP measurements were taken around 20–26 h after the aircraft survey.

All drifters remained inside the eddy for about 40 h after the last drifter deployment, and during this period they sampled the flow field at several scales. The time evolution of the histogram of drifter pairs separation, displayed in Fig. 2, shows that a large number of surface velocity samples were obtained at separation distances from $\mathcal{O}(100)$ m to $\mathcal{O}(10)$ km from 1415 UTC 7 February to 1400 UTC 9 February 2016. Most of the drifters remained within 0.1–15 km from each other for about 40 h as all drifters circled around the cyclonic eddy. After that, a storm disrupted the drifters' trajectories, and several clusters moved away from each other (D'Asaro et al. 2018). This event induced the noisy behavior of the drifters' separations mode after $t = 40$ h (solid line in Fig. 2).

We focus our attention on the initial 12 h, when there is a shift in the mode of pair separations distribution from 3.6 to 7.8 km. Prior to this shift, a large number of drifters were moving around the cold filament identified previously. Figures 3a and 3b show the drifters' positions and velocities at 1630 UTC 7 February 2016 (2 h and 30 min after the last drifter deployment) and at 1945 UTC 7 February, respectively, along with MVP temperatures at 10-m depth measured approximately at

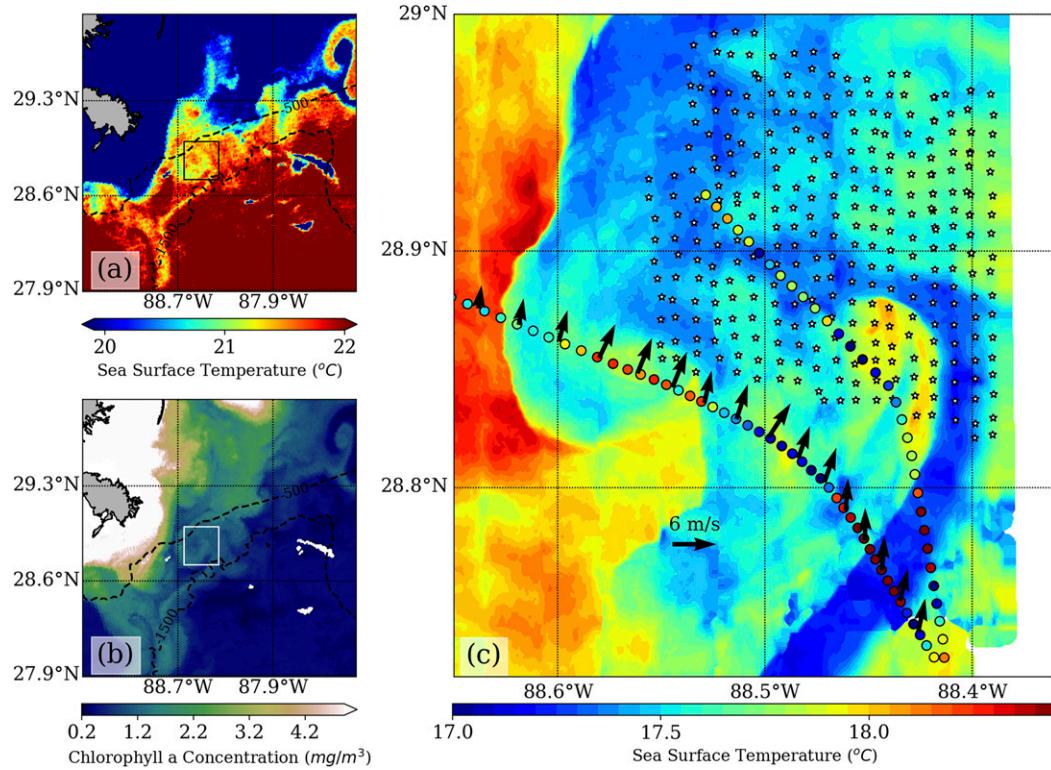


FIG. 1. (a),(b) SST and chlorophyll-*a* concentration from MODIS *Aqua* daytime on 7 Feb 2016 showing a hammerhead feature where the LASER drifters were released. The dashed lines show the 500- and 1500-m isobaths. The black box in (a) and the white box in (b) indicate the boundaries of (c). (c) Initial position of the drifters (white stars), which were deployed between 0200 and 1600 UTC 7 Feb 2016, MVP temperatures at 10-m depth (colored dots) measured on the same day between 1600 and 2200 UTC, aircraft survey of SST taken about 6 h before the beginning of the drifter deployment (colored map), and shipboard wind velocity measurements (vectors). The color scale of the MVP observations is presented in Fig. 3.

the same time (both MVP sections are part of the measurements shown in Fig. 1c). A strong convergence zone can be identified by the presence of near stationary drifters (advection velocity less than 0.1 m s^{-1}), located between $x = 8$ and 12 km , and of drifters to the east moving toward the former. The convergence zone coincides roughly with a temperature gradient zone, as observed from the MVP data (around $x = 12 \text{ km}$ in Fig. 3a, and $x = 8 \text{ km}$ in Fig. 3b). This temperature gradient indicates the transition from the colder, fresher, and lighter waters of the previously mentioned cold filament to warmer and heavier waters trapped inside the eddy. Figure 4 shows the vertical profiles of temperature, salinity, and potential density derived from the MVP sections shown in Fig. 3. For easier reference, the x axis used in the vertical sections is the zonal coordinate axis from Fig. 3, and not distance along the section. The abovementioned convergence zone occurs over a density front that extends from the near surface down to 40–60-m depth. Below 80 m, the potential density increases abruptly with depth indicating the

end of the surface mixed layer (the MVP measurements extend to 130-m depth). The group of drifters with low velocity is located within denser waters, while the drifters to the east are moving within the lighter waters. In the potential density section of Fig. 4f, there is a density high between $x = 4 \text{ km}$ and $x = 8 \text{ km}$ that extends throughout the mixed layer, indicating the presence of a denser filament inside the eddy. This is approximately the area where the drifters are moving slowly in Fig. 3b.

After the shift in the mode of pair separations distribution, the convergence cannot be observed in the drifter velocities anymore. Most of the drifters that were floating within the filament are now circling around what appears to be a small eddy, with diameter between 2 and 3 km (Fig. 3c). The presence of the small eddy can be observed in the pair separation histograms by a secondary mode around 2.5 km (Fig. 2). The presence of another convergence zone is evidenced in Fig. 3c by the meridional alignment of drifters between $x = 0$ and 6 km . The following sections present results

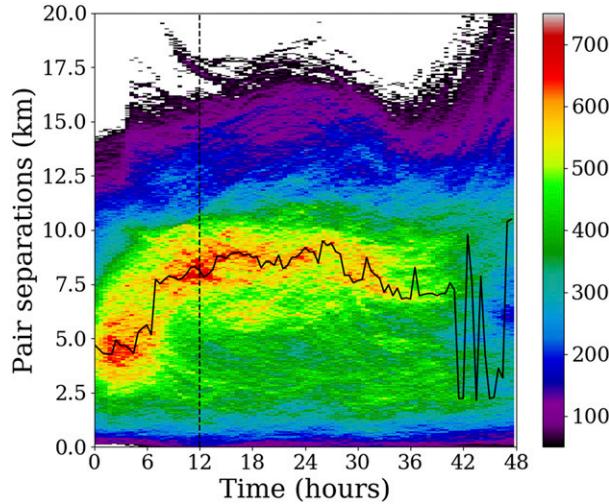


FIG. 2. Two-dimensional histogram of drifter pair separation and time throughout the first 48 h after the last drifter deployment. The size of the bins is 50 m in the pair separations axis and 30 min in the time axis. Only bins with more than 50 samples are shown. The solid black line indicates the mode of the pair separations distribution in each 30-min bin. The vertical dashed line indicates the end of the 12-h period used for the velocity reconstruction.

of the Eulerian velocity field reconstruction during this 12-h period and show the time evolution of horizontal divergence, relative vorticity, and lateral strain rate under the submesoscale features appearing in Fig. 3.

c. Gaussian process regression

Our goal is to estimate the near-surface flow in the vicinity of the drifters by interpolating their positions

and velocities, assuming that the drogues cause the drifters to faithfully follow the water. In doing so, we want to ignore what cannot be resolved and focus on scales allowed by the separations of the drifters. By interpolating to a sufficient number of points, we intend to get a useful picture of the evolving submesoscale flow, in particular time-varying fields of horizontal components of velocity, and from these fields we intend to infer horizontal divergence, relative vorticity, and rate of strain. Our approach is to interpolate the horizontal components u and v independently and filter the unresolved small scales by attributing them to errors in the drifters' velocities. Our method for interpolating is GPR with the scales present in the interpolated fields controlled by a covariance function. Because drifters tend to separate as they are carried by the flow, we use a covariance function with two scale parameters for each of the zonal, meridional, and temporal coordinates so that the interpolation might be tuned to benefit from both smaller and larger separations of neighboring drifters.

The zonal u and meridional v components of surface velocity are a function of space-time $\mathbf{p} = (x, y, t)$ (where x , y , and t are the zonal, meridional, and time coordinates, respectively). Drifter observations are available at n space-time points and are denoted by $(u_i^d, v_i^d) = [u(\mathbf{p}_i^d), v(\mathbf{p}_i^d)]$, with $i = 1, \dots, n$. The superscript d refers to the observational data and the subscript i to the i th such observation. As u and v will be treated as independent scalar fields, the GPR formulation will be presented only for u , and the

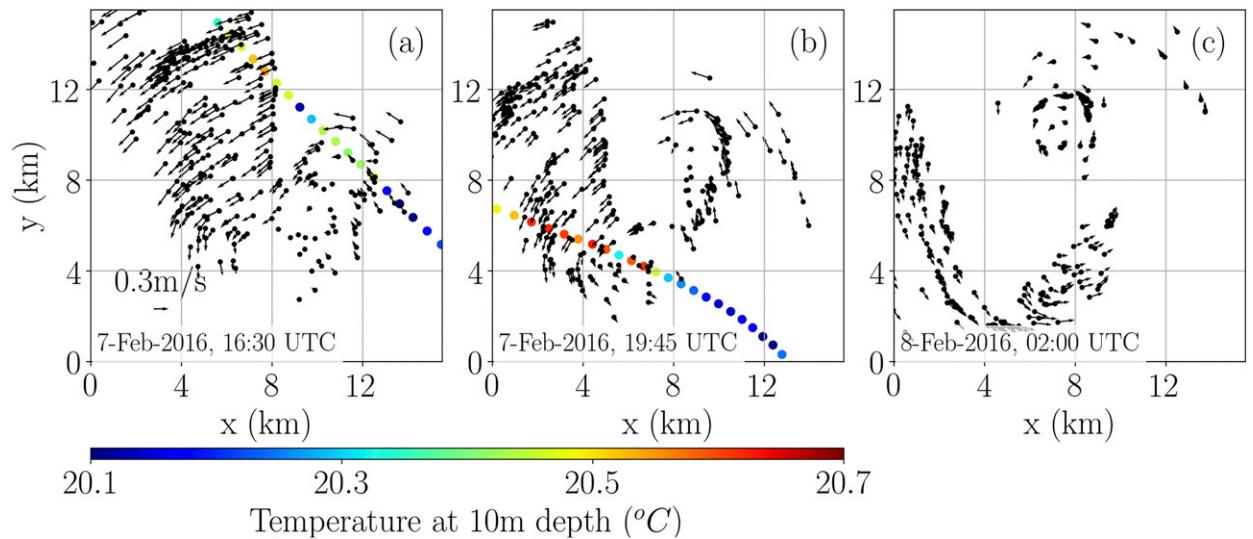


FIG. 3. Position and displacement velocity of the drifters (a) 2, (b) 5, and (c) 12 h after the last drifter deployment. Colored dots are MVP temperatures at 10-m depth, observed on 7 Feb 2016 between 1600 and 1700 UTC in (a) and between 1900 and 2100 UTC in (b).

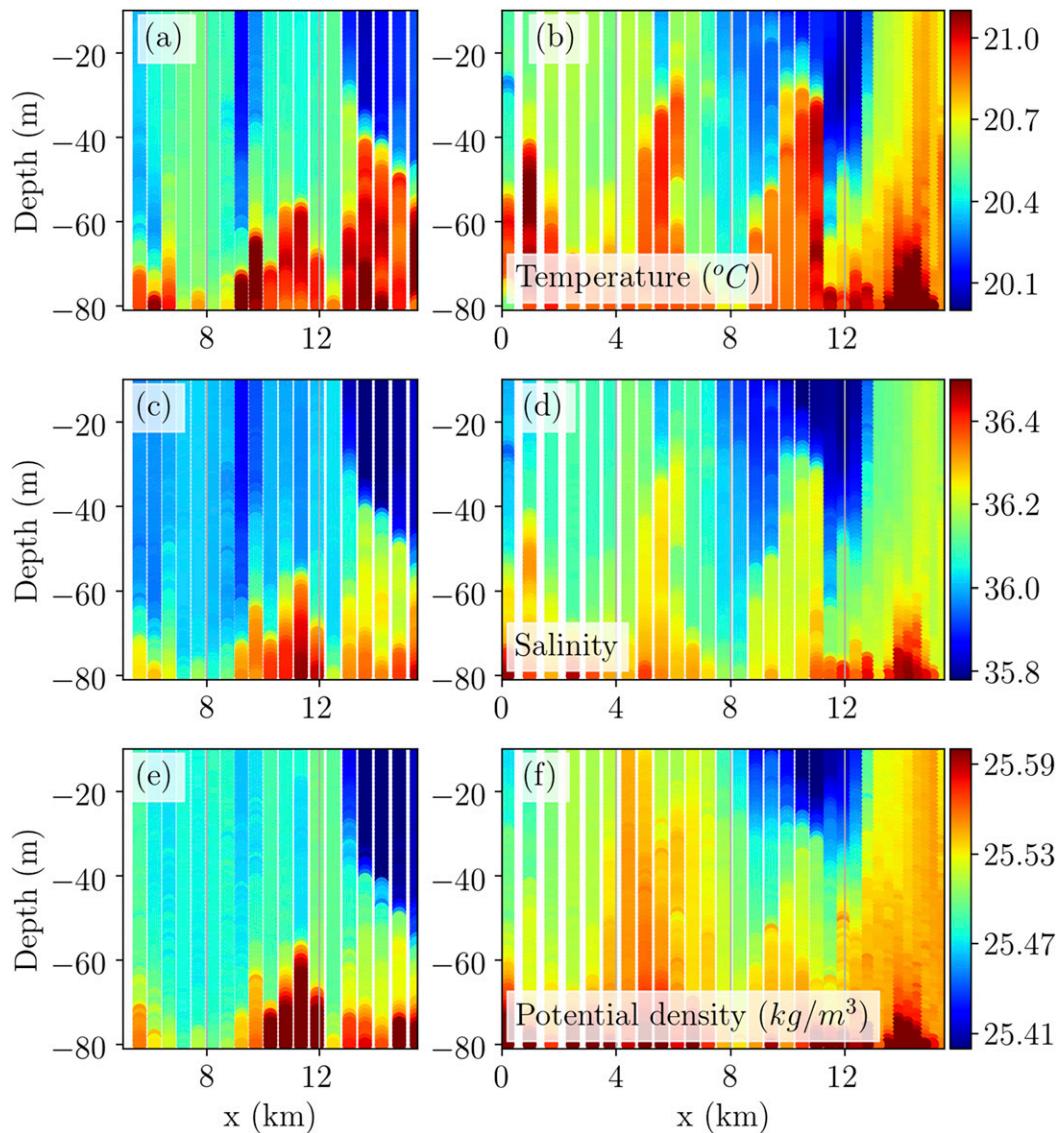


FIG. 4. (a),(b) Temperature; (c),(d) salinity; and (e),(f) potential density from the MVP sections presented in Fig. 3, taken on 7 Feb 2016. (left) Measurements taken between 1600 and 1700 UTC (Fig. 3a) and (right) measurements taken between 1900 and 2100 UTC (Fig. 3b).

estimator of u at a *target* point \mathbf{q} will be denoted as $u^t(\mathbf{q})$.²

The main assumption of GPR is that the quantity of interest, $u(\mathbf{p})$, follows a Gaussian process,³ which is completely specified by its mean function $\bar{u}(\mathbf{p})$ and

covariance function $\mathcal{K}(\mathbf{p}, \mathbf{q})$ (Rasmussen and Williams 2006):

$$u(\mathbf{p}) \sim \mathcal{GP}[\bar{u}(\mathbf{p}), \mathcal{K}(\mathbf{p}, \mathbf{q})]. \quad (1)$$

The Gaussian process (1) characterizes prior beliefs in terms of a prior mean field \bar{u} , which accounts for what we might guess for the velocity field in the absence of data, and \mathcal{K} , which specifies covariances that account for how we expect the velocities at any two points \mathbf{p} and \mathbf{q} might jointly deviate from our guesses at those points. While Eq. (1) presumes \mathbf{p} to represent any and all points in a continuum, in practice it is necessary to work with a

²Because we are interpolating from points where we have drifter data to points where we do not, it is convenient to distinguish which is which notationally with superscripts d for data and t for target.

³A Gaussian process is defined as a collection of normally distributed random variables, from which any finite samples have a joint Gaussian distribution (Rasmussen and Williams 2006).

finite number of points, so Eq. (1) becomes a multivariate normal distribution for the velocity components at these points with the prior mean field replaced by a vector of velocities and the covariance function replaced by a matrix of covariances.

Assuming that \mathbf{u}^d are noisy observations of u , and that the noise is independent, normally distributed and with zero mean and variance σ_N^2 , we can write the multivariate Gaussian prior distribution of \mathbf{u}^t and \mathbf{u}^d following from Eq. (1) as

$$\begin{bmatrix} \mathbf{u}^d \\ \mathbf{u}^t \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \bar{\mathbf{u}}^d \\ \bar{\mathbf{u}}^t \end{bmatrix}, \begin{bmatrix} \mathbf{K}^{dd} + \sigma_N^2 \mathbf{I} & \mathbf{K}^{dt} \\ (\mathbf{K}^{dt})^T & \mathbf{K}^{tt} \end{bmatrix}\right), \quad (2)$$

where $\bar{\mathbf{u}}^d$ and $\bar{\mathbf{u}}^t$ refer to the prior estimate of u at the observation and target points, respectively, \mathbf{K}^{dd} refers to the covariance matrix between pairs of observation points, \mathbf{K}^{tt} to the covariance matrix between pairs of target points, \mathbf{K}^{dt} to the covariance matrix between observation and target points, $(\mathbf{K}^{dt})^T$ is the matrix transpose of \mathbf{K}^{dt} , and \mathbf{I} is the $n \times n$ identity matrix. More specifically, the entries of the different vectors and matrices are given by

$$\begin{aligned} \bar{u}_i^d &= \bar{u}(\mathbf{p}_i^d), & K_{ij}^{dd} &= \mathcal{H}(\mathbf{p}_i^d, \mathbf{p}_j^d), & i, j &= 1, 2, \dots, n, \\ \bar{u}_i^t &= \bar{u}(\mathbf{q}_i), & K_{ij}^{tt} &= \mathcal{H}(\mathbf{q}_i, \mathbf{q}_j), & i, j &= 1, 2, \dots, m, \\ K_{ij}^{dt} &= \mathcal{H}(\mathbf{p}_i^d, \mathbf{q}_j), & i &= 1, 2, \dots, n, & j &= 1, 2, \dots, m. \end{aligned} \quad (3)$$

By conditioning the prior distribution on the observations we obtain the joint Gaussian posterior distribution:

$$\begin{aligned} \mathbf{u}^t | \mathbf{u}^d &\sim \mathcal{N}(\tilde{\mathbf{u}}^t, \mathbf{Q}), \\ \tilde{\mathbf{u}}^t &= \bar{\mathbf{u}}^t + (\mathbf{K}^{dt})^T (\mathbf{K}^{dd} + \sigma_N^2 \mathbf{I})^{-1} (\mathbf{u}^d - \bar{\mathbf{u}}^d), \\ \mathbf{Q} &= \mathbf{K}^{tt} - (\mathbf{K}^{dt})^T (\mathbf{K}^{dd} + \sigma_N^2 \mathbf{I})^{-1} \mathbf{K}^{dt}. \end{aligned} \quad (4)$$

The expressions in Eq. (4) are the same as the ones resulting from optimal interpolation⁴ (Gandin 1965) with $\bar{\mathbf{u}}$ representing the background estimate, \mathbf{K}^{dd} the background error covariance matrix at the observation points, $\sigma_N^2 \mathbf{I}$ the observational error covariance matrix, \mathbf{K}^{tt} the background covariance matrix at the target points, and \mathbf{K}^{dt} as the background covariance matrix between observation and target points. The left-hand side of Eq. (4) is usually referred to as the analysis in data assimilation.

⁴GPR is mathematically equivalent to optimal interpolation. The novelty we brought from the GPR community is the optimization of the covariance function by the marginal likelihood method, described next.

The prior mean $\bar{u}(\mathbf{p})$ and the covariance function $\mathcal{H}(\mathbf{p}, \mathbf{q})$ must be specified in order to carry out the regression. We considered the prior mean to be zero, so that all the information about the velocity field is inferred from the regression. A key aspect of GPR, as of optimum interpolation, is the choice of the covariance function from which the covariance matrices used in the mapping are drawn from. The main requirement is that the covariance matrices must be positive semidefinite (Bretherton et al. 1976; Rasmussen and Williams 2006), and ideally they should reflect the real space–time correlations of the field to be estimated. In applications of optimum interpolation focused on the mapping of meso- and large scales circulation, covariance functions were chosen to resemble empirical estimates of the covariance matrix computed from observations and from climatology (McWilliams 1976; Carter and Robinson 1987; Mariano and Brown 1992; Wilkin et al. 2002). For these scales, geostrophy and nondivergence constraints are imposed in the covariance function by the addition of cross-covariance terms between the horizontal velocity components (Bretherton et al. 1976; McWilliams 1976; Hua et al. 1986). As our goal is to characterize the submesoscale activity, and at this scale range the flow is significantly ageostrophic, with significant horizontal divergence, the abovementioned constraints are not desirable. Given that the LASER observations sampled the flow field at several separation scales (Fig. 2), we would like to explore the possibility of capturing more than one correlation scale per space–time coordinate dimension. Previous applications of multiscale analysis involve low-pass filtering the observations, which are used to build the larger scale. The latter is then used as prior mean for the smaller scale, which is constructed from the high-pass-filtered observations. This process can be implemented sequentially, as in Li et al. (2015), or simultaneously, as in Le Traon (1990) and Le Traon and Hernandez (1992). Our approach differs from the above, as the multiscale is added in the covariance function, and no filtering of the observations is required. Another difference is that, in GPR, the correlation scales and signal to noise ratio are determined directly by an optimization procedure, in which their probability conditioned on the observations is maximized. We define the covariance function as a sum of three-dimensional squared exponential functions (SE):

$$\mathcal{H}(\mathbf{p}, \mathbf{p}') = \sum_{i=1}^M \sigma_i^2 \exp\left[-\frac{(t-t')^2}{2r_{ii}^2} - \frac{(y-y')^2}{2r_{yi}^2} - \frac{(x-x')^2}{2r_{xi}^2}\right], \quad (5)$$

where M is the number of scales per dimension, σ^2 is the signal variance, r_i is the correlation time scale, and

r_x and r_y are the horizontal length scales in the zonal and meridional directions, respectively. The parameters that define the SE covariance function r_{ti} , r_{yi} , r_{xi} , and σ_i^2 , along with the noise variance σ_N^2 , are referred as hyperparameters. We set $M = 2$, so that nine hyperparameters have to be determined by the optimization procedure. As it will be shown, the optimization successfully captured two distinct scales per dimension: a dominant, larger one, from a smaller scale, with lower variance. In tests with $M = 3$ (not shown), the optimization did not result in three distinct scales. The larger scale was basically the same as with $M = 2$, while the signal of the smaller scale was split in two, having similar values of r_{ti} , r_{yi} , r_{xi} , and σ_i^2 .

OPTIMIZATION PROCEDURE

To optimize the hyperparameters of the SE covariance function, we applied the marginal likelihood method, described in section 5.4.1 of Rasmussen and Williams (2006). Considering the set of hyperparameters $\boldsymbol{\theta} = \{(r_{ti}, r_{yi}, r_{xi}, \sigma_i, \sigma_N) | i = 1, 2\}$ that specify the covariance function $\mathcal{K}(\boldsymbol{\theta})$, the optimization procedure seeks the maximization of the probability of the hyperparameters conditioned on the observations $p(\boldsymbol{\theta}|\mathbf{u}_d)$. Using Bayes's theorem, this is equivalent to maximizing the log marginal likelihood:

$$\log p(\mathbf{u}^d|\boldsymbol{\theta}) = -\frac{1}{2}(\mathbf{u}^d)^T \mathbf{B}^{-1} \mathbf{u}^d - \frac{1}{2} \log |\mathbf{B}| - \frac{n}{2} \log 2\pi, \quad (6)$$

where $\mathbf{B} = (\mathbf{K}^{dd} + \sigma_N^2 \mathbf{I})$ and n is the number of observations. The partial derivative of the log marginal likelihood function with respect to a hyperparameter $\theta_j \in \boldsymbol{\theta}$ is

$$\begin{aligned} \frac{\partial}{\partial \theta_j} \log p(\mathbf{u}^d|\boldsymbol{\theta}) &= \frac{1}{2}(\mathbf{u}^d)^T \mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial \theta_j} \mathbf{B}^{-1} \mathbf{u}^d - \frac{1}{2} \text{Tr} \left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial \theta_j} \right) \\ &= \frac{1}{2} \text{Tr} \left[(\boldsymbol{\beta} \boldsymbol{\beta}^T - \mathbf{B}^{-1}) \frac{\partial \mathbf{B}}{\partial \theta_j} \right], \end{aligned} \quad (7)$$

where $\boldsymbol{\beta} = \mathbf{B}^{-1} \mathbf{u}^d$, and Tr is the trace of a square matrix. Equation (7) can be applied with any gradient based optimization algorithm to find the values of $\boldsymbol{\theta}$ that maximize the log marginal likelihood. Here, we used the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm (Byrd et al. 1995).

3. Results and discussion

a. Optimization of the hyperparameters

We start our analysis by optimizing the covariance function to find the values of σ_1^2 , r_{t1} , r_{y1} , r_{x1} , σ_2^2 , r_{t2} , r_{y2} , r_{x2} , and σ_N^2 for each velocity component. A total of 16 343

TABLE 1. Optimized hyperparameters of the covariance function.

Hyperparameter	u	v
σ_N (m s^{-1})	0.009	0.01
σ_1 (m s^{-1})	0.13	0.14
r_{t1} (h)	2.99	3.24
r_{y1} (km)	1.84	3.54
r_{x1} (km)	2.58	1.42
σ_2 (m s^{-1})	0.028	0.031
r_{t2} (h)	0.64	0.65
r_{y2} (km)	0.4	0.44
r_{x2} (km)	0.39	0.39

data points were selected in the 12-h time window from 1415 UTC 7 February to 0215 UTC 8 February 2016, and the drifter positions (x , y) are defined in Cartesian coordinate system with origin at 28°N and 88°W . Table 1 shows the optimized hyperparameters of the covariance function for the two velocity components. The optimized noise level for u and v was $\sigma_N \approx 0.009 \text{ m s}^{-1}$, and both components presented a dominant larger scale, with signal standard deviation $\sigma_1 \approx 0.1 \text{ m s}^{-1}$, and a smaller scale with signal standard deviation $\sigma_2 \approx 0.03 \text{ m s}^{-1}$. In both velocity components, the larger scale presented a correlation time $r_{t1} \approx 3 \text{ h}$, with u having a larger correlation scale in the zonal direction ($r_{x1} = 2.5 \text{ km}$ and $r_{y1} = 1.8 \text{ km}$), and v presenting larger correlation in the meridional direction ($r_{x1} = 1.4 \text{ km}$ and $r_{y1} = 3.5 \text{ km}$). The smaller scale presented a correlation time $r_{t2} \approx 40 \text{ min}$ and length scales around 400 m.

Each dimension of the optimized SE covariance function is plotted in Fig. 5 along with empirically estimated space and time correlations of the drifters' velocity components. The spatial correlation of the velocity components as a function of radial distances between observations (blue line in Figs. 5a,b) was estimated in 200-m spatial bins, and averaged over the 12-h period used in the analysis. These correlations assume a statistically isotropic and homogeneous flow, but even if that is not the case, they still provide a crude estimate of the dominant scales of motion (Davis 1985). In general, the spatial part of the SE (solid and dotted red lines in Figs. 5a,b) is in good agreement with the empirically estimated spatial correlations, presenting correlation decay of the same order of magnitude. The most striking difference is that the former do not have negative values, which are induced by the cyclonic circulation of the eddy where the drifters were released. The SE can only capture the correlations in one-half of the eddy. Another difference is that the spatial correlation of v appears to be a combination between the broader meridional SE and the narrow

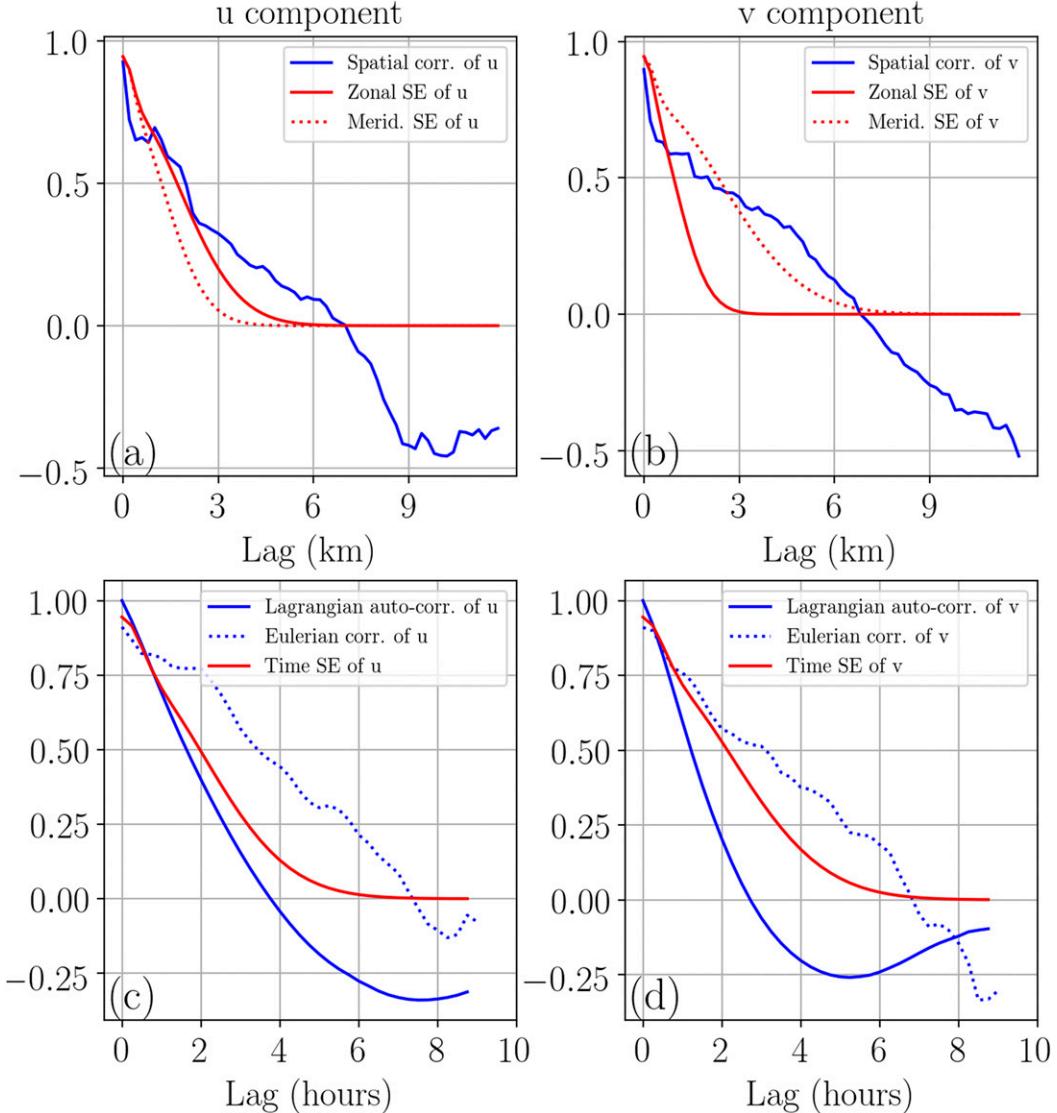


FIG. 5. Comparison of correlations estimated from the observations and the optimized SEs used in the covariance function. The SEs were normalized by their respective total variance ($\sigma_1 + \sigma_2 + \sigma_N$; see Table 1). Right (left) panels contain comparisons for the u (v) component. (a),(b) Simultaneous correlation as a function of drifter separations computed using spatial bins of 200 m (blue) and SE for the zonal (solid red) and meridional (dotted red) directions. (c),(d) Lagrangian and Eulerian lagged autocorrelation (solid and dotted blue, respectively) and SE for the time dimension (red).

zonal SE. This might actually be the case if the flow is anisotropic. The decay in spatial correlations at shorter separations is more abrupt than provided by the SE, particularly for u . Figures 5c and 5d present two estimates of the time correlation as a function of time lag; the Lagrangian autocorrelation, which is averaged over all drifters (solid blue), and an estimate of the Eulerian correlation (dotted blue), which was obtained from the zero spatial lag of the space-time lagged correlations. For the latter, spatial bins of 400 m were considered, and for both Lagrangian and

Eulerian correlations, a time lag interval of 15 min was used. The optimized time SE for each velocity component is presented in red. As with the spatial correlations, the negative lagged autocorrelations are caused by the cyclonic flow of the eddy, which is not captured by the SE. In spite of that, the three curves have similar correlation decay, with the SE curve staying between the Lagrangian and Eulerian correlations curves. The fact that all drifters are trapped inside the eddy explains the proximity between the Lagrangian and Eulerian correlation estimates.

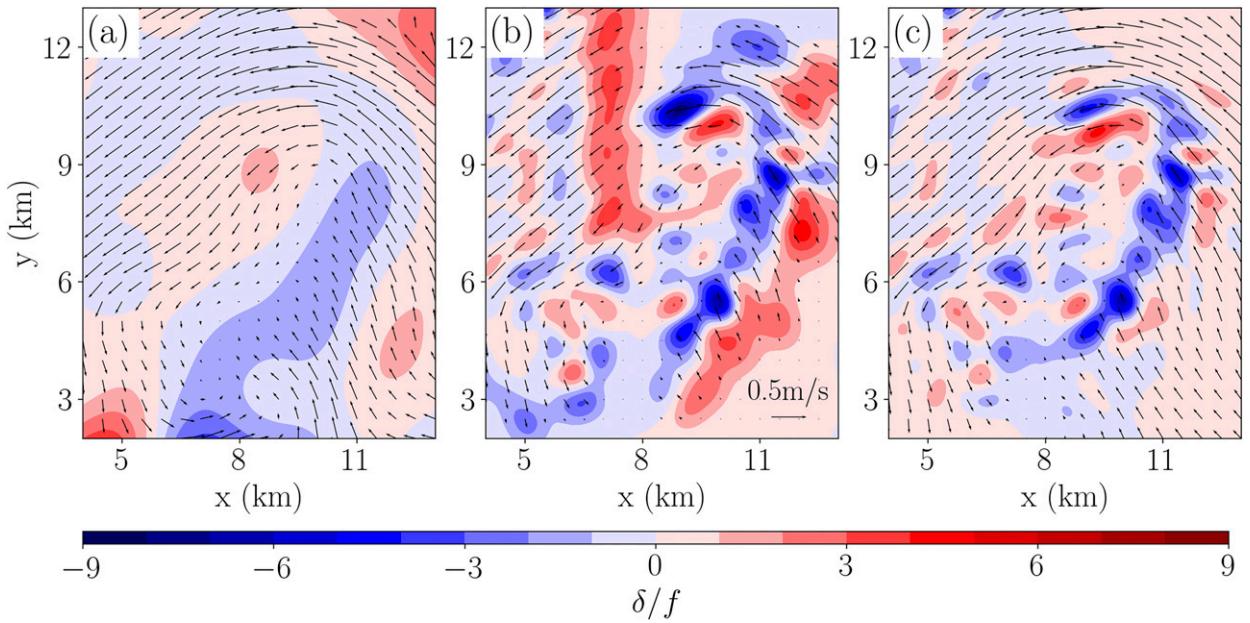


FIG. 6. Horizontal divergence normalized by planetary vorticity and velocity vectors of reconstructed velocity fields in a 50-m-resolution grid (velocity vectors sampled every 15 grid points). (a),(b) Reconstruction using covariance function with only one scale per dimension for each component, with $\sigma = \sigma_1$, $r_t = r_{t,1}$, $r_y = r_{y,1}$, and $r_x = r_{x,1}$ in (a) and $\sigma = \sigma_2$, $r_t = r_{t,2}$, $r_y = r_{y,2}$, and $r_x = r_{x,2}$ in (b) (see Table 1 for the values of each hyperparameter). (c) The covariance function was defined as the sum of the two covariance functions used in (a) and (b).

To illustrate the effect of the different scales in the regression, Fig. 6 presents a snapshot at $t = 4$ h of the velocity vectors and the divergence field from reconstructions using covariance functions considering only the larger scale (Fig. 6a), only the smaller scale (Fig. 6b), and both scales (Fig. 6c). The two-scale reconstruction shows a cyclonic circulation around a sharp convergence zone. While the larger scale captures the cyclonic circulation, the sharp velocity gradients in the reconstructed velocity field are resolved by the smaller scale.

b. GPR validation

To validate the results of the GPR velocity reconstruction, we randomly selected half of the observations to build an estimator for u and v [Eq. (4)], while the other half was set as verification data. Using the optimized hyperparameters from Table 1, we estimated the velocity components at the space–time locations of the verification data, and used their respective observations to compute the absolute error:

$$\begin{aligned} \text{Err}(v) &= |\tilde{v}_i^t - v_i^d|, \\ \text{Err}(u) &= |\tilde{u}_i^t - u_i^d|, \end{aligned} \quad (8)$$

where $\text{Err}(u)$ and $\text{Err}(v)$ are the absolute errors of the estimates of u and v , respectively, (u_i^d, v_i^d) are the withheld velocity observations, and $(\tilde{u}_i^t, \tilde{v}_i^t)$ are their corresponding GPR estimates. The normalized cumulative

histograms of $\text{Err}(v)$ and $\text{Err}(u)$ are presented in Fig. 7a, and the absolute errors normalized by the observed value are shown in Fig. 7b. The reconstruction of both components presented satisfactory results, with more than 90% of the errors staying below 0.02 m s^{-1} , and almost 99% below 0.04 m s^{-1} . The distribution of $\text{Err}(v)$ had a mean of 0.008 m s^{-1} , and standard deviation of 0.013 m s^{-1} , while the distribution of $\text{Err}(u)$ had a mean of 0.007 m s^{-1} , and standard deviation of 0.011 m s^{-1} . More than 85% of $\text{Err}(v)$ and about 80% of $\text{Err}(u)$ were below 10% of the observed values.

The square root of the diagonal of the posterior covariance matrix \mathbf{Q} [Eq. (4)] provides error estimates for the GPR, which we will use to bound our Eulerian velocity reconstructions. We compared these errors estimates to the absolute errors presented before. Figures 7c and 7d present histograms of the difference between the posterior covariance error estimates (Err_Q) and the absolute errors (Err). Negative values indicate that Err_Q underestimated the actual error, while for positive values, Err_Q overestimated Err . On average, Err_Q overestimated the actual error by 0.01 m s^{-1} for both velocity components, indicating that the former is a good metric to bound the velocity reconstructions.

c. Velocity maps and kinematics

Next, we evaluate the evolution of the velocity field sampled by the drifters over the 12-h period between

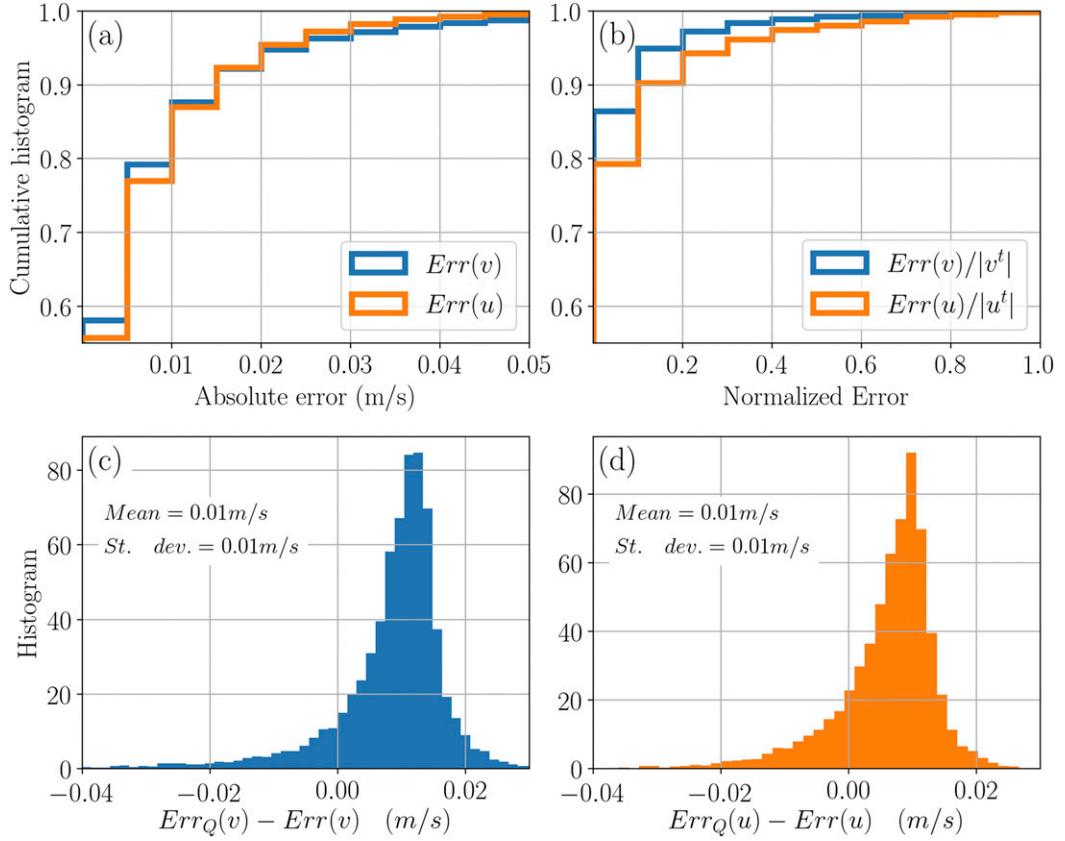


FIG. 7. (a) Cumulative histograms of absolute error of v and u reconstructions at verification points. (b) Cumulative histogram of the absolute error normalized by the absolute observed value. (c),(d) Difference between error estimates from the posterior covariance (Err_Q) and the actual errors computed from the observations for v and u , respectively. A total of 8171 verification points were used in these histograms.

1415 UTC 7 February and 0215 UTC 8 February 2016. We set a grid with 50-m resolution, and the reconstructions were carried out every 15 min. For this step all observations available were used. Figures 8–11 present the evolution of speed, horizontal divergence ($\delta = \partial u / \partial x + \partial v / \partial y$), relative vorticity ($\zeta = \partial v / \partial x - \partial u / \partial y$), and lateral strain rate ($\alpha = [(\partial u / \partial x - \partial v / \partial y)^2 + (\partial v / \partial x + \partial u / \partial y)^2]^{1/2}$), respectively. Only grid points with $\text{Err}_Q(u)$ and $\text{Err}_Q(v)$ smaller than 0.04 m s^{-1} are shown, and the differential kinematic properties were normalized by the planetary vorticity f . These figures also show vectors of the reconstructed velocity, along with vectors of the drifters' velocities.

The velocity vectors present a large cyclonic pattern, with radius on the order of 10 km, which is consistent with the eddy observed in the satellite and aircraft images (Fig. 1). The density front between the cold lighter filament and warmer heavier waters inside the eddy shown in section 2b has a clear imprint in the reconstructed velocity and its derived kinematics. The maps of speed present a low speed zone elongated in

the south-southwest–north-northeast direction between $x = 4$ and 12 km, and $y = 2$ –12 km, which gets narrower from $t = 2$ to 8 h. The snapshot at $t = 6$ h (Fig. 8c) is inside the time period where the MVP section in Fig. 3b was taken, and the low speed zone coincides with the area of high-density waters west of the cold filament (Fig. 4f), between $x = 4$ km and $x = 8$ km. The differential kinematic properties over the density gradient area have characteristics of a submesoscale front (Munk et al. 2000), with intense horizontal convergence, reaching up to $8f$, and strong cyclonic vorticity, with values up to $13f$. As the low speed area gets narrower, the strain rate and relative vorticity under the convergence zone intensify, indicating the occurrence of strain driven frontogenesis (Mahadevan and Tandon 2006). The shipboard wind measurements presented in Fig. 1c show that, at the time the MVP sections were taken, the wind was nearly in the alongfront direction, considering that the denser waters were in the western side of the front. This is a favorable condition for wind induced frontogenesis, as the Ekman transport could be expected to drive denser

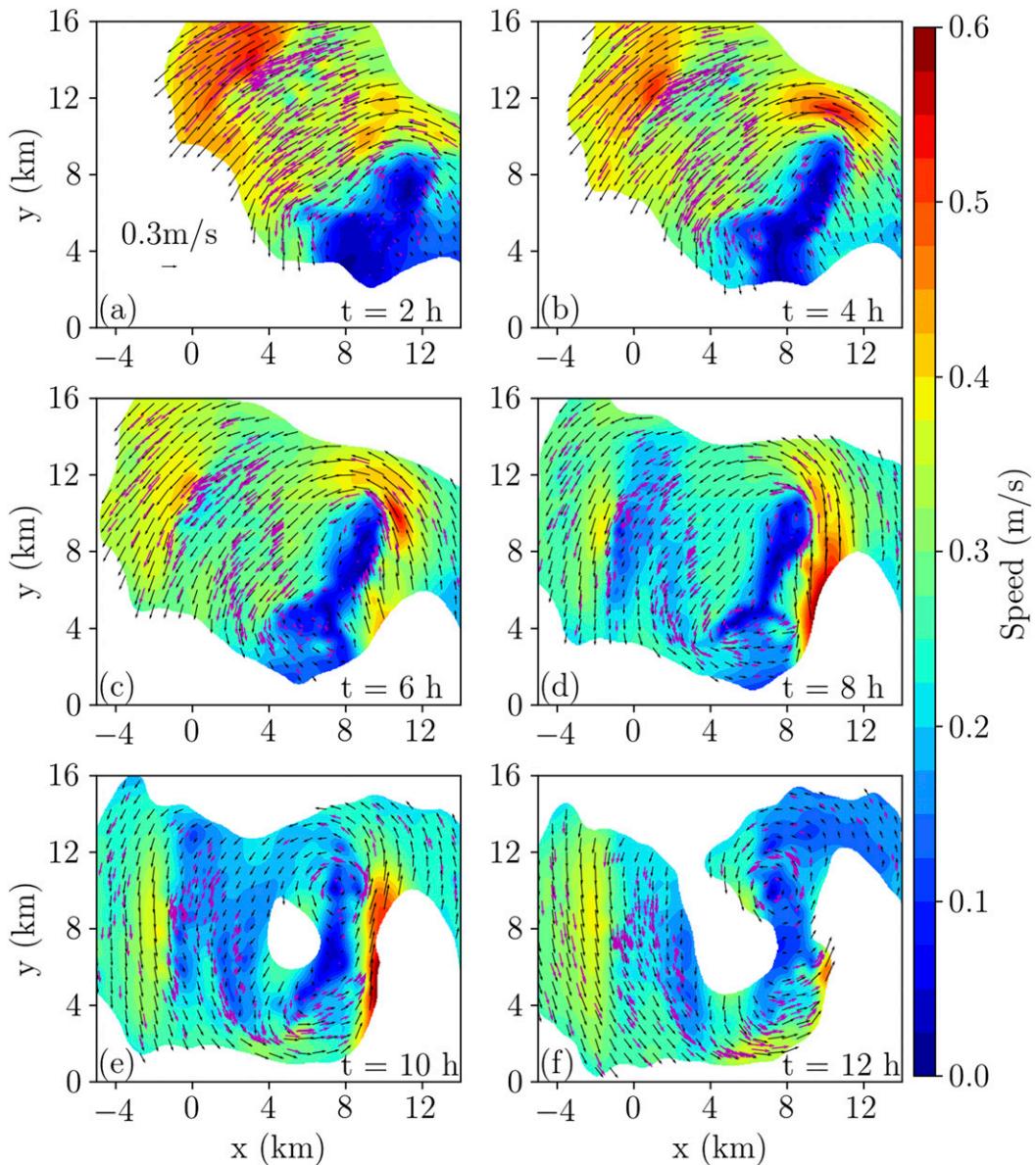


FIG. 8. Snapshots of the reconstructed velocity field (black vectors) sampled every 15 grid points superimposed over speed (m s^{-1}). The time of each snapshot is referenced at 1415 UTC 7 Feb 2016. The magenta vectors are the drifters' velocities at the time of each snapshot.

waters toward lighter waters (Thomas and Lee 2005; Thomas and Ferrari 2008; Capet et al. 2008), and could be contributing to the strong convergence observed in the reconstructed velocity.

At $t = 10$ h, the convergence zone is broken as the drifters in the northern edge distance themselves from the ones in the south, and the northern edge has turned into the small eddy, with radius of $\mathcal{O}(1)$ km, described in section 2b (Fig. 3b). Other frontal features can be observed throughout the 12 h period in the western side of the larger cyclone, between $x = -4$ and 4 km, with

enhanced convergence, cyclonic vorticity and strain rate. The online supplemental material contains animations of the velocity field and of maps of relative vorticity, horizontal divergence and lateral strain rate showing the time evolution of the frontal features.

The reconstructed differential kinematic quantities document the presence of fronts that exhibit strong positive relative vorticity and strongly negative horizontal divergence on short space-time scales. These features are typical of mixed layer submesoscale turbulence (McWilliams 2016) and were captured by the

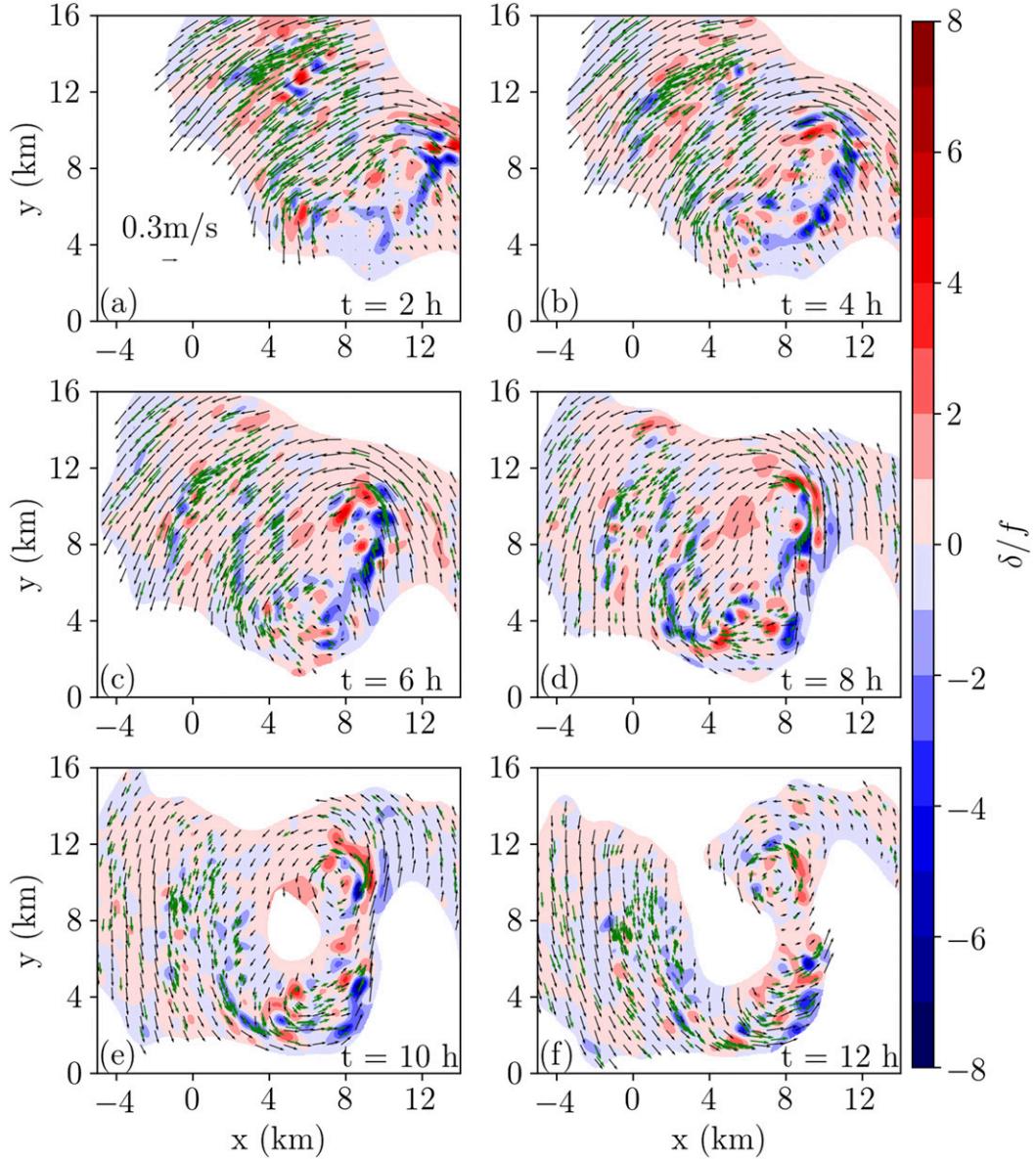


FIG. 9. Horizontal divergence of the reconstructed velocity field normalized by planetary vorticity. The time of each snapshot is referenced at 1415 UTC 7 Feb 2016. The green vectors are the drifters' velocities at the time of each snapshot.

Lagrangian observations as they tend to cluster in and align over convergence zones (D'Asaro et al. 2018). Figure 12 presents histograms of relative vorticity, horizontal divergence, and lateral strain rate computed over the whole 12-h period considered. Our estimates of relative vorticity shows a distribution with mean of $0.83f$, standard deviation of $1.57f$, and skewness of 1.73, while the horizontal divergence distribution has zero mean, standard deviation of $0.92f$, and skewness of -0.91 , and the distribution of lateral strain rate has a mean of $1.42f$, standard deviation of $1.12f$, and skewness of 2.2.

The histograms of relative vorticity and horizontal divergence are similar to the ones reported by Ohlmann et al. (2017) over submesoscale fronts off the California coast. Ohlmann et al. (2017) used clusters of four drifters to estimate the velocity gradients, and obtained a positively skewed relative vorticity distribution, with mean of $0.68f$ and standard deviation of $1.13f$ (the values of skewness were not presented), and a divergence distribution with negative skewness, mean of $-0.13f$, and standard deviation of $0.95f$. Shcherbina et al. (2013) estimated the differential kinematics from ADCP velocity measurements obtained

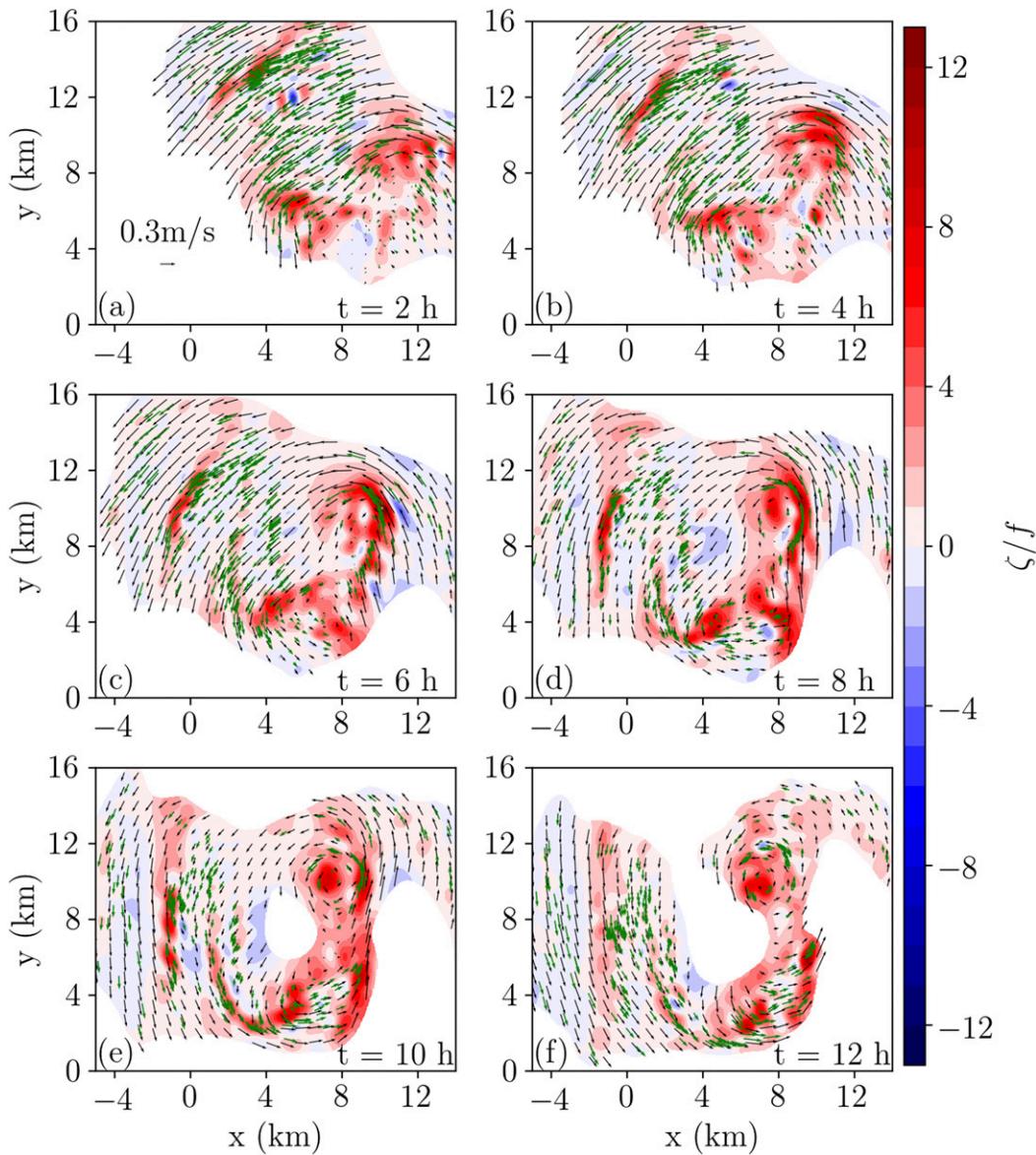


FIG. 10. Vertical relative vorticity of the reconstructed velocity field normalized by planetary vorticity. The time of each snapshot is referenced at 1415 UTC 7 Feb 2016. The green vectors are the drifters' velocities at the time of each snapshot.

by two vessels sailing in parallel tracks; the two vessels covered a 500-km area of intense submesoscale turbulence in the North Atlantic Mode Water region. They found a positively skewed relative vorticity distribution throughout the mixed layer (skewness = 2.5 between 0 and 50 m), with a slightly negative mean of $-0.07f$, and median of $-0.32f$. Values of high strain rate matching high cyclonic vorticity estimates indicated the presence of submesoscale fronts, which were embedded in an anticyclonic background, as suggested by the weak negative mean and median. Their distribution of divergence was nearly Gaussian, with zero mean, and skewness slightly negative (-0.2).

The maximum values of velocity gradients reported here are about 3 times larger than those found by Shcherbina et al. (2013), which could be explained by the relatively coarser space resolution of their observations, as the two vessels were 1 km apart. On the other hand, Raschke et al. (2017) estimated velocity gradients up to $80f$ over a frontal feature from finescale sea surface roughness anomalies. The sea surface roughness anomalies were obtained from high-resolution (0.5–6 m) sun-glint images from airplane surveys during LASER, and the observations covered the same cyclonic eddy as the drifters trajectories used here, although 4 days later.

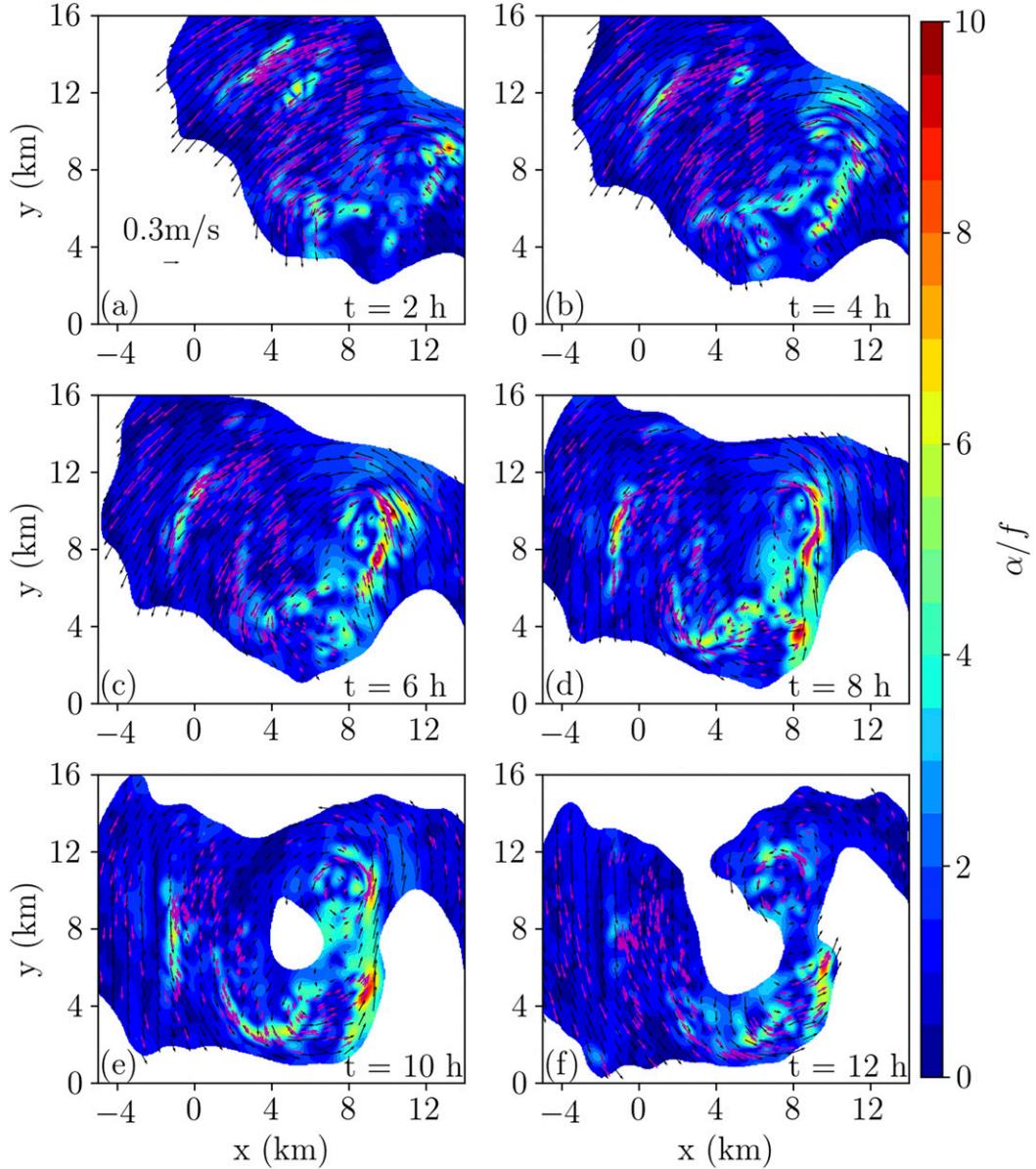


FIG. 11. Lateral strain rate of the reconstructed velocity field normalized by planetary vorticity. The time of each snapshot is referenced at 1415 UTC 7 Feb 2016. The magenta vectors are the drifters' velocities at the time of each snapshot.

The space–time scales captured in the present analysis are only a subset of the scales of the true flow; this is due to the finite resolution of the drifter sampling and of the limitation of the current GPR approach. The observed scales are set by the spatial distribution of the drifters, and hence dictated by the flow being sampled, and by the frequency of observations. Furthermore, the current GPR approach uses only a limited number of stationary space–time scales to represent a flow field which might be energetic at several scales of motion and which might be nonstationary. Increasing the grid resolution of our

Eulerian velocity field would not change the intensity of the velocity gradients, as these are dictated by the smoothness of a covariance function whose shortest optimal correlation scales were found to be around 400 m. The front in [Raschle et al. \(2017\)](#) was 50 m wide as seen from airplane observations.

4. Summary and conclusions

Gaussian process regression has been applied to obtain high-resolution velocity maps from a dense array of

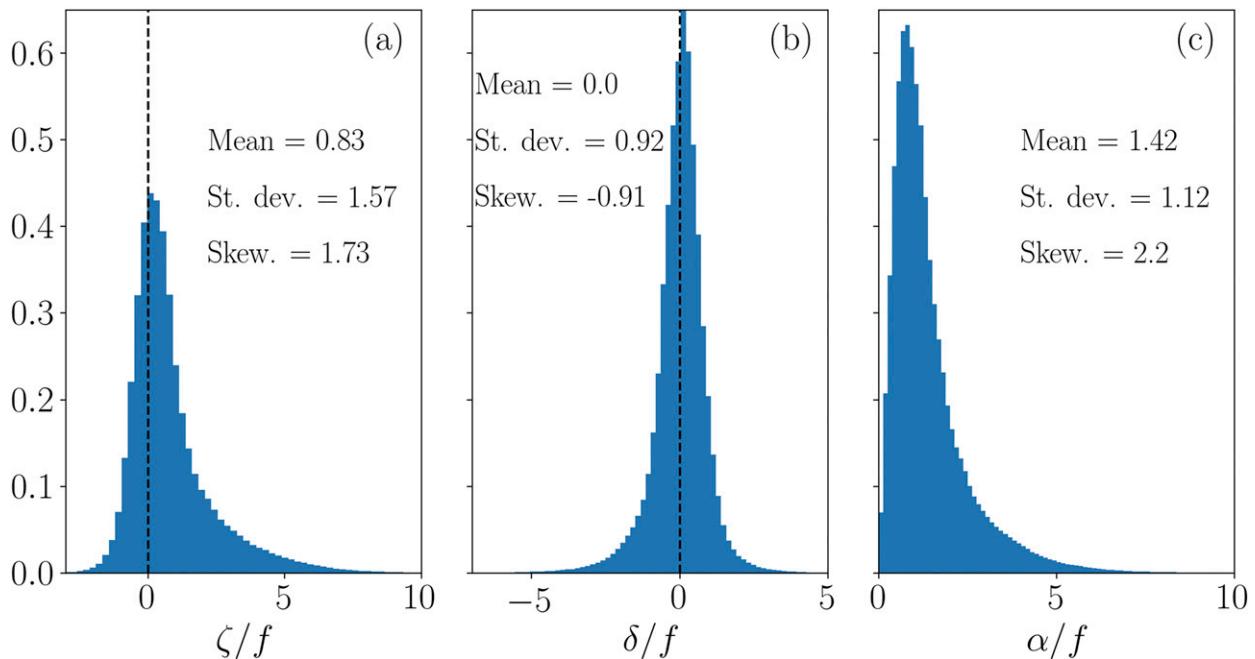


FIG. 12. Normalized histograms of (a) relative vorticity, (b) horizontal divergence, and (c) lateral strain rate, all quantities normalized by f . The histograms were built from velocity reconstructions taken in 15-min intervals from 1415 UTC 7 Feb to 0215 UTC 8 Feb 2016. Only grid points with $\text{Err}_Q(u)$ and $\text{Err}_Q(v)$ smaller than 0.04 m s^{-1} were considered.

drifter observations. Data from 326 drifters with time resolution of 15 min were used in a 12-h window, when all drifters remained inside a cyclonic frontal eddy. The space and time correlation scales of the covariance function, as well as the data noise level, were obtained from the drifter observations themselves via a Bayesian optimization procedure. The covariance function was defined with two scales per dimension, and the optimization procedure resulted in a dominant larger scale, which captured the cyclonic circulation inside the eddy, and a smaller scale that represented the fast evolving sharp velocity gradients inside the eddy. The reconstructed Eulerian velocity field presents the time evolution of strong submesoscale features displaying significant ageostrophic circulation. Frontal features with horizontal convergence reaching $8f$, lateral strain rate up to $10f$, and cyclonic relative vorticity up to $13f$ were observed.

It is important to point out that as all information about the velocity field were provided by the drifters, the sharp velocity gradients only manifested where (and when) there were enough samples. As drifters move away from each other, beyond the correlation scales from the covariance function, the velocity gradients are smoothed out. The empirically estimated covariances in Fig. 5 display features that were not present in the optimized SE covariance functions: the steep correlation decay at short separation scales, and negative correlations at large separations. The SE, however, yielded acceptable results,

as evidenced by the error verification in section 3b, even though no attempt was made to optimize the type of covariance function or to account for the varying space-time scales of the flow. Model selection algorithms, including combination of different types of covariance function, can be implemented in a similar fashion as the optimization procedure presented for the hyperparameters (Rasmussen and Williams 2006). Likewise, more complex and nonstationary covariance functions (Paciorek and Schervish 2006) can be used to capture the changes in the dominant dynamical scales of heterogeneous and nonstationary flows. These improvements could provide a better representation of the flow field and should be the subject of future research.

Nevertheless, our approach enabled us to not only identify submesoscale activity, but also follow its time evolution for 12 h. This aspect represents an advantage to other approaches used to identify submesoscales, like point estimates of differential kinematic quantities, which do not provide a visualization of the processes, or static images from vessel or aircraft surveys, as these do not provide the time evolution of an observed process.

Acknowledgments. This research was made possible in part by a grant from BP/The Gulf of Mexico Research Initiative to the CARTHE Consortium. R. Gonçalves acknowledges support by the Brazilian Ministry of Science, Technology and Innovation (CNPq—Council for

Scientific and Technological Development) through a Ph.D. scholarship from the Science Without Borders program, Grant 202263/2012-6. Dr. Iskandarani was partially supported by an EarthCube NSF Grant ICER1639722. This research was conducted in collaboration with and using the resources of the University of Miami Center for Computational Science. The LASER data are publicly available through the Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC) at <https://data.gulfresearchinitiative>. Data doi: 10.7266/N7W0940J (drifters), 10.7266/N7H130FC (MVP CTD), and 10.7266/N7S75DRP (shipboard wind measurements). The MODIS *Aqua* data displayed in Fig. 1 are publicly available at NASA's Jet Propulsion Laboratory PODAAC website at <https://podaac.jpl.nasa.gov/dataaccess>. The Gaussian process regression and the optimization of the hyperparameters were carried out using the Python package GPy (GPy 2012). We thank the two anonymous reviewers whose comments and suggestions substantially improved the manuscript.

REFERENCES

- Bennett, A. F., 1984: Relative dispersion: Local and nonlocal dynamics. *J. Atmos. Sci.*, **41**, 1881–1886, [https://doi.org/10.1175/1520-0469\(1984\)041<1881:RDLAND>2.0.CO;2](https://doi.org/10.1175/1520-0469(1984)041<1881:RDLAND>2.0.CO;2).
- Berta, M., A. Griffa, T. M. Özgökmen, and A. C. Poje, 2016: Submesoscale evolution of surface drifter triads in the Gulf of Mexico. *Geophys. Res. Lett.*, **43**, 11 751–11 759, <https://doi.org/10.1002/2016GL070357>.
- Bretherton, F. P., R. E. Davis, and C. B. Fandry, 1976: A technique for objective analysis and design of oceanographic experiment applied to MODE-73. *Deep-Sea Res. Oceanogr. Abstr.*, **23**, 559–582, [https://doi.org/10.1016/0011-7471\(76\)90001-2](https://doi.org/10.1016/0011-7471(76)90001-2).
- Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu, 1995: A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.*, **16**, 1190–1208, <https://doi.org/10.1137/0916069>.
- Capet, X., J. C. McWilliams, M. J. Molemaker, and F. Shchepetkin, 2008: Mesoscale to submesoscale transition in the California Current System. Part II: Frontal processes. *J. Phys. Oceanogr.*, **38**, 44–64, <https://doi.org/10.1175/2007JPO3672.1>.
- Carter, E. F., and A. R. Robinson, 1987: Analysis models for the estimation of oceanic fields. *J. Atmos. Oceanic Technol.*, **4**, 49–74, [https://doi.org/10.1175/1520-0426\(1987\)004<0049:AMTEO>2.0.CO;2](https://doi.org/10.1175/1520-0426(1987)004<0049:AMTEO>2.0.CO;2).
- Charney, J. G., 1971: Geostrophic turbulence. *J. Atmos. Sci.*, **28**, 1087–1095, [https://doi.org/10.1175/1520-0469\(1971\)028<1087:GT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1971)028<1087:GT>2.0.CO;2).
- D'Asaro, E., C. Lee, L. Rainville, R. Harcourt, and L. Thomas, 2011: Enhanced turbulence and energy dissipation at ocean fronts. *Science*, **332**, 318–322, <https://doi.org/10.1126/science.1201515>.
- , and Coauthors, 2018: Ocean convergence and the dispersion of flotsam. *Proc. Natl. Acad. Sci. USA*, **115**, 1162–1167, <https://doi.org/10.1073/pnas.1718453115>.
- Davis, R. E., 1985: Drifter observations of coastal surface currents during CODE: The statistical and dynamical views. *J. Geophys. Res.*, **90**, 4756–4772, <https://doi.org/10.1029/JC090iC03p04756>.
- Flament, P., and L. Armi, 2000: The shear, convergence, and thermohaline structure of a front. *J. Phys. Oceanogr.*, **30**, 51–66, [https://doi.org/10.1175/1520-0485\(2000\)030<0051:TSCATS>2.0.CO;2](https://doi.org/10.1175/1520-0485(2000)030<0051:TSCATS>2.0.CO;2).
- , —, and L. Washburn, 1985: The evolving structure of an upwelling filament. *J. Geophys. Res.*, **90**, 11 765–11 778, <https://doi.org/10.1029/JC090iC06p11765>.
- Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translations, 242 pp.
- GPy, 2012: GPy: A Gaussian process framework in python. Accessed 13 February 2019, <http://github.com/SheffieldML/GPy>.
- Haza, A., and Coauthors, 2018: Drogue-loss detection for surface drifters during the Lagrangian submesoscale experiment (laser). *J. Atmos. Oceanic Technol.*, **35**, 705–725, <https://doi.org/10.1175/JTECH-D-17-0143.1>.
- Hua, B. L., J. C. McWilliams, and W. B. Owens, 1986: An objective analysis of the POLYMODE local dynamics experiment. Part II: Streamfunction and potential vorticity fields during the intensive period. *J. Phys. Oceanogr.*, **16**, 506–522, [https://doi.org/10.1175/1520-0485\(1986\)016<0506:AOAOTP>2.0.CO;2](https://doi.org/10.1175/1520-0485(1986)016<0506:AOAOTP>2.0.CO;2).
- Iskandarani, M., S. Wang, A. Srinivasan, W. Carlisle Thacker, J. Winokur, and O. M. Knio, 2016: An overview of uncertainty quantification techniques with application to oceanic and oil-spill simulations. *J. Geophys. Res. Oceans*, **121**, 2789–2808, <https://doi.org/10.1002/2015JC011366>.
- Kennedy, M. C., and A. O'Hagan, 2000: Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, **87**, 1–13, <https://doi.org/10.1093/biomet/87.1.1>.
- Le Traon, P. Y., 1990: A method for optimal analysis of fields with spatially variable mean. *J. Geophys. Res.*, **95**, 13 543–13 547, <https://doi.org/10.1029/JC095iC08p13543>.
- , and F. Hernandez, 1992: Mapping the oceanic mesoscale circulation: Validation of satellite altimetry using surface drifters. *J. Atmos. Oceanic Technol.*, **9**, 687–698, [https://doi.org/10.1175/1520-0426\(1992\)009<0687:MTOMCV>2.0.CO;2](https://doi.org/10.1175/1520-0426(1992)009<0687:MTOMCV>2.0.CO;2).
- Li, Z., J. C. McWilliams, K. Ide, and J. D. Farrara, 2015: A multi-scale variational data assimilation scheme: Formulation and illustration. *Mon. Wea. Rev.*, **143**, 3804–3822, <https://doi.org/10.1175/MWR-D-14-00384.1>.
- Mahadevan, A., and A. Tandon, 2006: An analysis of mechanisms for submesoscale vertical motion at ocean fronts. *Ocean Modell.*, **14**, 241–256, <https://doi.org/10.1016/j.ocemod.2006.05.006>.
- Mariano, A. J., and O. B. Brown, 1992: Efficient objective analysis of dynamically heterogeneous and nonstationary fields via the parameter matrix. *Deep-Sea Re.*, **39A**, 1255–1271, [https://doi.org/10.1016/0198-0149\(92\)90068-5](https://doi.org/10.1016/0198-0149(92)90068-5).
- McWilliams, J. C., 1976: Maps from the Mid-Ocean Dynamics Experiment: Part I. Geostrophic streamfunction. *J. Phys. Oceanogr.*, **6**, 810–827, [https://doi.org/10.1175/1520-0485\(1976\)006<0810:MOMD>2.0.CO;2](https://doi.org/10.1175/1520-0485(1976)006<0810:MOMD>2.0.CO;2).
- , 2016: Submesoscale currents in the ocean. *Proc. Roy. Soc. London*, **472A**, 20160117, <https://doi.org/10.1098/rspa.2016.0117>.
- Molemaker, M. J., J. C. McWilliams, and X. Capet, 2010: 2010: Balanced and unbalanced routes to dissipation in an equilibrated Eady flow. *J. Fluid Mech.*, **654**, 35–63, <https://doi.org/10.1017/S002211200993272>.
- Munk, W., L. Armi, K. Fischer, and F. Zachariassen, 2000: Spirals on the sea. *Proc. Roy. Soc. London*, **456A**, 1217–1280, <https://doi.org/10.1098/rspa.2000.0560>.
- Novelli, G., C. M. Guigand, C. Cousin, E. H. Ryan, N. J. Laxague, H. Dai, B. K. Haus, and T. M. Özgökmen, 2017: A biodegradable surface drifter for ocean sampling on a massive scale. *J. Atmos. Oceanic Technol.*, **34**, 2509–2532, <https://doi.org/10.1175/JTECH-D-17-0055.1>.

- Ohlmann, J. C., M. J. Molemaker, B. Baschek, B. Holt, G. Marmorino, and G. Smith, 2017: Drifter observations of submesoscale flow kinematics in the coastal ocean. *Geophys. Res. Lett.*, **44**, 330–337, <https://doi.org/10.1002/2016GL071537>.
- Özgökmen, T. M., A. C. Poje, P. F. Fischer, H. Childs, H. Krishnan, C. Garth, A. C. Haza, and E. Ryan, 2012: On multi-scale dispersion under the influence of surface mixed layer instabilities and deep flows. *Ocean Model.*, **56**, 16–30, <https://doi.org/10.1016/j.ocemod.2012.07.004>.
- Paciorek, C. J., and M. J. Schervish, 2006: Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, **17**, 483–506, <https://doi.org/10.1002/env.785>.
- Poje, A. C., and Coauthors, 2014: Submesoscale dispersion in the vicinity of the Deepwater Horizon spill. *Proc. Natl. Acad. Sci. USA*, **111**, 12 693–12 698, <https://doi.org/10.1073/pnas.1402452111>.
- Rasclle, N., J. Molemaker, L. Marié, F. Nouguier, B. Chapron, B. Lund, and A. Mouche, 2017: Intense deformation field at oceanic front inferred from directional sea surface roughness observations. *Geophys. Res. Lett.*, **44**, 5599–5608, <https://doi.org/10.1002/2017GL073473>.
- Rasmussen, C. E., and C. K. I. Williams, 2006: *Gaussian Processes for Machine Learning*. MIT Press, 266 pp.
- Scully-Power, P., 1986: Navy oceanographer shuttle observations STS 41-G mission report. NUSC Tech. Doc. 7611, Naval Underwater Systems Center, 117 pp., <https://apps.dtic.mil/dtic/tr/fulltext/u2/a233578.pdf>.
- Shcherbina, A. Y., E. A. D'Asaro, C. M. Lee, J. M. Klymak, M. J. Molemaker, and J. C. McWilliams, 2013: Statistics of vertical vorticity, divergence, and strain in a developed submesoscale turbulence field. *Geophys. Res. Lett.*, **40**, 4706–4711, <https://doi.org/10.1002/grl.50919>.
- Thacker, W. C., M. Iskandarani, R. C. Gonçalves, A. Srinivasan, and O. M. Knio, 2015: Pragmatic aspects of uncertainty propagation: A conceptual review. *Ocean Model.*, **95**, 25–36, <https://doi.org/10.1016/j.ocemod.2015.09.001>.
- Thomas, L. N., and C. M. Lee, 2005: Intensification of ocean fronts by down-front winds. *J. Phys. Oceanogr.*, **35**, 1086–1102, <https://doi.org/10.1175/JPO2737.1>.
- , and R. Ferrari, 2008: Friction, frontogenesis, and the stratification of the surface mixed layer. *J. Phys. Oceanogr.*, **38**, 2501–2518, <https://doi.org/10.1175/2008JPO3797.1>.
- Wilkin, J. L., M. M. Bowen, and W. J. Emery, 2002: Mapping mesoscale currents by optimal interpolation of satellite radiometer and altimeter data. *Ocean Dyn.*, **52**, 95–103, <https://doi.org/10.1007/s10236-001-0011-2>.