1   **Applying dynamical systems techniques to real ocean drifters**

2   Irina I. Rypina[1], Timothy Getscher[2], Lawrence Pratt[1], and Tamay Ozgokmen[3]

3   [1]Woods Hole Oceanographic Institution, 266 Woods Hole rd., Woods Hole, MA 02540

4   [2]US Navy

5   [3]Rosenstiel School of Marine and Atmospheric Science, U. of Miami, 4600 Rickenbacker

6   Causeway, Miami, FL 33149

7   Corresponding author: Irina I. Rypina, irypina@whoi.edu

8

9

10

11

12

13

14

15

16

17

18

Abstract

This paper presents the first comprehensive comparison of several different dynamical-systems-based measures of stirring and Lagrangian coherence, computed from real ocean drifters. Seven commonly used methods (finite-time Lyapunov exponent, trajectory path length, trajectory correlation dimension, trajectory encounter volume, Lagrangian-averaged vorticity deviation, dilation, and spectral clustering) were applied to 135 surface drifters in the Gulf of Mexico in order to map out the dominant Lagrangian coherent structures. Among the detected structures were regions of hyperbolic nature resembling stable manifolds from classical examples, divergent and convergent zones, and groups of drifters that moved more coherently and stayed closer together than the rest of the drifters. Many methods highlighted the same structures, but there were differences too. Overall, 5 out of 7 methods provided useful information about the geometry of transport within the domain spanned by the drifters, whereas the path length and correlation dimension methods were less useful than others.

Significance statement

Techniques from the dynamical systems theory have been widely used to study transport in ocean flows. However, they have been typically applied to numerically simulated trajectories of water parcels. This paper applies different dynamical systems techniques to real ocean drifter trajectories from the massive release in the Gulf of Mexico. To our knowledge, this is the first comprehensive comparison of the performance of different dynamical systems techniques with application to real drifters.

1. Introduction

Techniques from the dynamical systems theory can be used to study transport and exchange processes in oceanic flows (Haller, 2015; Samelson and Wiggins, 2006; Balasuriya et al., 2019; Hadjighasem et al., 2017; Filippi et al., 2021a;b; Rypina et al., 2010 and others). In general, they aim to identify the key regions of the flow with qualitatively different Lagrangian behavior and/or to identify boundaries between them. The term Lagrangian Coherent Structures or LCS (Haller and Juan, 2000) has been adopted to refer to both such regions themselves and to their boundaries. Because different methods use different definitions of "different" and "similar," they generally yield different LCS (Balasuriya et al. 2019; Rypina et al. 2011; 2018; Hadjighasem et al., 2017).

Being Lagrangian in nature, most LCS detection methods start with the release of a set of particles or drifters within the domain of interest, and then use observations of their trajectories as the particles are advected by the flow. Obtaining such trajectory datasets is straightforward in applications where the velocity fields are known from either models or observations, and this is exactly the settings in which the dynamical system approach has been used in the past. However, applying the same techniques to real ocean drifters has been a challenge simply because the drifters are rarely released in a manner that adequately spans the domain of interest.

On April 21st 2018, 144 near-surface CARTHE drifters were released nearly simultaneously in a roughly 11 km by 11 km domain in the northern Gulf of Mexico as part of the Submesoscale Processes and Lagrangian Analysis on the Shelf (SPLASH) experiment (Laxague et al., 2018; Solodoch et al., 2020; Lund et al., 2020). The release pattern was a nearly regular, rectangular, 12 x 12 grid with roughly 1 km average spacing between neighboring drifters. The release was

done using 3 boats and took just under 3 hours. The drifters then transmitted their positions every 5 min during the subsequent 5 days. We used all available 144 drifters, and the start time t_{start} for our analysis corresponds to the time when the last drifter was released. The drifter positions at t_{start} and the resulting drifter trajectories are shown in Fig. 1. Such aggressive release strategy is not typical for oceanographic applications due to high costs of vessels and manpower. However, it allowed populating the domain with drifters in a manner most suitable for the dynamical systems applications. Thus, this dataset provided a unique and long-awaited opportunity to try applying the dynamical systems techniques to real, rather than simulated, ocean drifters and to identify the real, rather than simulated, ocean LCS.

In this paper, seven commonly used dynamical systems techniques were applied to the real drifter dataset from the SPLASH experiment: FTLEs, trajectory path length, trajectory correlation dimension, encounter volume, Lagrangian-averaged vorticity deviation, dilation, and spectral clustering. The resulting real ocean LCS were mapped and described and, when possible, parallels were drawn between these observed structures and their more classical counterparts from text-book analytic or numeric examples. The seven techniques were also inter-compared to each other and the similarities/differences were discussed. Our choice of the seven techniques is by no means all-inclusive and was inspired by Hadjighasem et all. (2017) who compared a similar selection of the dynamical systems methods (plus a few more and minus the encounter volume method) in the context of analytical, observed, and numerically-generated flows.

2. Methods

We start with a brief review of the 7 dynamical systems techniques that we will use.

a) FTLEs

One of the most commonly used LCS detection techniques is based upon FTLEs (Haller and Yuan, 2000; Shadden et al., 2005). FTLE is the largest exponential separation rate between a trajectory and its closest neighbors in any direction. Maximizing ridges of FTLE fields can be used as proxies for stable (or unstable for backward-time trajectories) manifolds of hyperbolic trajectories in time-varying fluid flows (with the additional requirement that the fastest separation occurs in the direction normal to the ridge and is caused by the hyperbolic straining rather than shear). Regions with small FTLEs are indicative of slow separation rates between neighboring trajectories and often correspond to eddy cores, ~~although for eddies with a non-solid-body rotating regime, azimuthal shear produces spiraling structures within the core region instead of the uniformly low FTLEs~~. Maps of FTLEs are very visual, and the computation of FTLEs is straightforward, computationally inexpensive, and robust with respect to noise, which makes FTLEs one of the most popular methods in oceanographic studies of transport and mixing. Importantly, FTLEs are also frame-independent and thus give consistent results in any translating or rotating reference frame (Haller 2005; 2015).

For flows where the velocity field is known from either models or observations, FTLEs (λ) can be estimated by releasing dense regularly-spaced orthogonal grids of simulated trajectories (Haller, 2001; 2002~~; Lekien and Ross, 2010; Rypina et al., 2021~~). This method uses 4 (in 2D) closest neighbors to construct the Cauchy-Green tensor $G = \left(\frac{\Delta x_i}{\Delta x_{o,j}}\right)^T \left(\frac{\Delta x_i}{\Delta x_{o,j}}\right)$, whose largest eigenvalue σ is connected to

$$\lambda = \frac{1}{T} \ln \sqrt{\sigma}. \tag{1}$$

123 Here $\Delta x_{0,i}$ and Δx_i are the initial and final distance in the ith-direction between neighboring

124 trajectories. This algorithm requires dense regularly-spaced orthogonal grids of trajectories. For

125 the SPLASH dataset, we manually chose quadruplets of 4 neighboring trajectories that form a

126 near-rectangle, define the local orthogonal coordinate system most strongly aligned with the axes

127 of the near-rectangle, and then estimate FTLEs using eq. (1) for the center of mass of each

128 quadruplet (Fig. 2 shows the quadruplets and their centers of mass locations).

129 A modification for unstructured meshes was described in Lekien and Ross (2010). Rypina et al.

130 (2021) recently used the unstructured grid method to compute FTLEs from a cluster of 6 real

131 drifters in the Alboran Sea. The method estimates FTLEs for each trajectory using its N closest

132 neighbors as

133 $\lambda = \frac{1}{T} \ln \tilde{\sigma}$, (2)

134 where $\tilde{\sigma}$ is the largest ~~positive~~ singular value of a matrix

135 $\widetilde{M} = DX_f \, (DX_0)^T \, (DX_0 (DX_0)^T)^{-1}$

136 which minimizes $\left\| DX_f - M \, DX_0 \right\|$.

137 Here $DX_0 = \begin{pmatrix} x_1^0 - x_i^0 & \dots & x_N^0 - x_i^0 \\ y_1^0 - y_i^0 & \dots & y_N^0 - y_i^0 \end{pmatrix}$ and $DX_f = \begin{pmatrix} x_1^f - x_i^f & \dots & x_N^f - x_i^f \\ y_1^f - y_i^f & \dots & y_N^f - y_i^f \end{pmatrix}$ are matrices of the

138 initial and final displacements between the trajectory and its N neighbors. Because the largest

139 singular value of $\widetilde{M}$ is equal to the square root of the largest eigenvalue of $\tilde{G} = \widetilde{M}^T \widetilde{M}$, eq. (2) is

140 the unstructured-mesh counterpart of eq. (1). We use the Delaunay triangulation partition to

141 define closest neighbors for each drifter (Fig. 2 shows the Delaunay partition for the SPLASH

142 dataset). The FTLE is then estimated using eq. (2) at each drifter's initial position using its

143   Delaunay closest neighbors. When used together, a combination of these two methods – the

144   regular and the unstructured mesh methods – allows estimating FTLEs both at the locations of

145   each drifter and between neighboring quadruplets.

146

147   **b) Trajectory path length**

148   Trajectory path length $L = \int ds = \int_{t_0}^{t_0+T} |\vec{u}(x(t), t)| dt$, where ds is the incremental length of

149   the infinitesimal trajectory segments. For drifter data, summation can be used instead of

150   integration. L has been proposed by Rypina et al. (2011) as one of the "Trajectory Complexity

151   measures" and by Mendoza and Mancho (2010) as the "Lagrangian Descriptor" for identifying

152   LCSs. Curves of near-constant L values with a large ∇L in the perpendicular direction to the

153   curve are indicative of the stable manifolds of hyperbolic trajectories (because trajectories on the

154   manifold approach the hyperbolic trajectory and trajectories slightly off the manifold are repelled

155   from it). This method is less mathematically rigorous than FTLEs and is frame-dependence, but

156   it is commonly used due to its simplicity.

157

158   **c) Trajectory correlation dimension**

159   Trajectory correlation dimension (CD) is a measure of space occupied by a trajectory. In 2D, it

160   varies from 0 for a point, to 1 for a curve, to 2 for a trajectory that densely fills an area. C can be

161   estimated using a box counting algorithm, where the entire trajectory data set is first mapped

162   onto a unit square, and the unit square is then repeatedly split into 2^{-2m}, $m = 0, 1, \ldots, M$ adjacent

163   square boxes with side length $s = 2^{-m}$ (we use $M = 12$ in this paper). A distribution function is

164   then computed for each trajectory as $F_i(s) = \frac{1}{N_i^2} \Sigma \left(N_i^j(s) \right)^2$ where N_i is the total number of

165   points in the ith trajectory and N_i^j is the number of points in the ith trajectory that fall inside the jth

166   box for a given s. The trajectory correlation dimension CD_i for the ith trajectory can then be

167   estimated as the slope of $F_i(s)$ vs. s in log-log coordinates. Just like trajectory path length, CD is

168   another measure of "Trajectory Complexity" and has been proposed by Rypina et al. (2011) as a

169   means for LCS identification. Similar to L, level curves of near-constant CD with a large ∇CD in

170   the perpendicular direction to the curve are indicative of the stable manifolds of hyperbolic

171   trajectories. CD is a more sensitive measure of "Trajectory Complexity" than L but is more

172   computationally expensive. Just like L, CD is also frame dependent. Note also that for flows in

173   the state of chaotic advection, CD (and L) could also be used to highlight slowly-moving

174   coherent eddy-like features (regular islands), embedded into vigorously-stirring regions (chaotic

175   sea). Islands would have less complex trajectories with lower C than trajectories within the

176   chaotic sea. Similarly, although CD was not designed to identify convergence, trajectories

177   converging rapidly into a nearly-stationary convergence zone would have smaller CD than those

178   free to wonder over the entire domain.

179

180   **d) Trajectory encounter number and trajectory encounter volume**

181   Trajectory encounter volume V_{en} for a particular trajectory is a volume of fluid that gets in

182   contact with a particular water parcel over a time interval T (Rypina and Pratt, 2017; Rypina et

183   al., 2018). This is a frame-independent quantity. It quantifies the mixing potential of a flow and

184   is related to the eddy or turbulent flow diffusivity κ (Rypina et al., 2018). The larger V_{en}, the

185   more opportunities exist for a parcel to exchange properties with surrounding fluid. Smallest V_{en}

186   occur in isolated secluded regions of the flow such as eddy cores, and largest V_{en} occur in

187   hyperbolic regions and along the stable manifolds of hyperbolic trajectories leading into

188   hyperbolic regions. Thus, V_{en} can be used to characterize both elliptic and hyperbolic LCSs.

189   For data sets containing a finite number of particle trajectories, encounter volume for a particular

190   trajectory can be approximated by assigning small volumes δV_j to all trajectories and summing

191   over those trajectories that come close to the particular trajectory: $V_{en} \approx \sum \delta V_j$. For regular grids

192    $\delta V_j = \delta V = const$ and $V_{en} = \delta V N_{en}$ where N_{en} is the encounter number – the number of

193   trajectories that come close (i.e., within a small radius R) to the particular trajectory. In our

194   calculations, we use $R = 1$ km and $\delta V \approx 1$ km^2, which is the square of the mean distance

195   between the drifters' release locations.

196   Note that the interpretation of the encounter volume in the context of limited trajectories

197   deployed in a small part of a flow domain, such as our SPLASH drifters, differs from the case

198   where drifters are seeded over the entire domain. Only for a domain-wide deployment, encounter

199   volume is representative of the mixing potential of the flow. For a small deployment, encounter

200   volume merely measures the amount of encounters within the dataset. This undersampling issue

201   leads to important consequences in both hyperbolic and elliptic regions. While for a domain-

202   wide deployment a lot of encounters occur in hyperbolic regions (as discussed above), these are

203   also the exact same regions where initially-nearby trajectories separate rapidly from each other,

204   yielding low encounter values in the case of a small deployment. Similarly, whereas coherent

205   eddy cores produce fewer encounters than hyperbolic regions for a domain-wide drifter release,

206   these regions trap drifters allowing them to encounter many of their neighbors deployed within

207   the same eddy, which produces large values in the case of small deployment. Thus, encounter

208   volume might be a poor measure of the mixing potential of a flow in the case of a small

209   deployment (but because this metric is still sensitive to differences between hyperbolic/elliptic

210   behaviors even for a small deployment, it might still be able to highlight regions with different

211   transport characteristics, so we go ahead and apply it to SPLASH drifters in the next section).

212

213   **e) Dilation**

214   Dilation rate (with units of inverse time) is the velocity divergence averaged along a particle's

215   trajectory, $D = \frac{1}{T} \int_{t_0}^{t_0+T} div\big(u(x(t), t)\big)$. This frame-independent quantity was proposed by

216   Huntley et al. (2015) as a method for identifying clusters of material at the ocean surface. We

217   will refer to D simply as "dilation" for brevity. Trajectories with the largest positive/negative D

218   experience the strongest divergence/convergence and thus repel/accumulate buoyant floating

219   surface tracers (including drifters). D can be used to identify convergence-type LCS marked by

220   the extrema of D. For drifter data, summation can be used instead of integration, and the Linear

221   Least Squares method of Molinari and Kirwan (1975) can be used to estimate $div(u)$ at each

222   point along each trajectory.

223

224   **f) Lagrangian-Averaged Vorticity Deviation (LAVD)**

225    $LAVD$ is the vorticity deviation with respect to the domain-averaged instantaneous vorticity,

226   averaged along a particle trajectory, $LAVD = \frac{1}{T} \int_{t_0}^{t_0+T} |\omega(x(x_0, t) - \overline{\omega(t)}| \, dt$. It was introduced

227   by Haller et al. (2016) as a frame-independent metric for identifying rotationally coherent

228   Lagrangian eddies, which correspond to a region contained within the outermost closed convex

229   level surface of $LAVD$ surrounding an isolated maximum. For drifter data, we again use

230   summation instead of integration and estimate vorticity using a Linear Least Squares method.

231 Note that $LAVD$ would only be able to identify those rotationally-coherent Lagrangian eddies

232 that are smaller than, and lay entirely within, the domain seeded with drifters.

233

234 **g) Spectral Clustering**

235 The last method for identifying the LCSs that we will be testing using drifter data is the

236 optimized-parameter Spectral Clustering described in Filippi et al., 2021a;b (see also Shi and

237 Malik (2000), Hadjighasem et al. (2016) and references therein). This was originally a data

238 science technique that was adopted by the dynamical systems community. This method aims at

239 identifying, within a given dataset of trajectories, clusters of trajectories that are most similar to

240 each other and, at the same time, most dissimilar from trajectories in other clusters. A direct

241 connection between spectral clusters and elliptic/hyperbolic/convergence-type LCS from other

242 methods is not always straightforward, although some of the identified spectral clusters often

243 coincide with elliptic regions, regions of strong convergence, or regions delineated by segments

244 of hyperbolic LCS. The method starts with the construction of a matrix of weights

245 $w_{ij} = \begin{cases} \dfrac{1}{r_{ij}} \\ w_{diag} \end{cases}$, where r_{ij} is the time-average distance between the ith and j^{th} trajectories, and

246 w_{diag} is a large constant offset value (we use $w_{diag} = max(w_{ij}) \times 10^7$). Based on this matrix,

247 the method used ideas from machine learning theory, specifically, N-cut matrix partitioning and

248 K-means clustering algorithms, to identify the spectral clusters with the largest/smallest degree

249 of intra-/inter-cluster similarity. Importantly, the optimized-parameter version of the Spectral

250 Clustering method (Filippi et al., 2021a;b) that we are using automatically detects both the

251 optimal number of clusters and the cluster sizes (based on the normalized eigengap between the

252   eigenvalues of the generalized normalized Laplacian, as described in Filippi et al. (2021a)).

253   Being based on the distances between trajectories, spectral clustering is frame-independent.

254

255   **h) Linear Least Squares (LLS) method for estimating drifter-based divergence and**

256   **vorticity**

257   In order to estimate divergence and vorticity from drifters, we follow the approach of Rypina et

258   al. (2021), where we first compute horizontal velocities from drifter positions using a centered

259   finite-difference scheme and then apply the linear least squares (LLS) method of Molinari and

260   Kirwan (1975) to estimate horizontal velocity gradients. LLS method is based on the Taylor

261   expansion of velocity, $U = DA$, where $U = [u_1, \dots u_N]^T$ is a (known) vector containing the

262   u −velocity at a given time t for each of the N drifters, $D = \begin{pmatrix} 1 & x_1 - \bar{x} & y_1 - \bar{y} \\ \vdots & \vdots & \vdots \\ 1 & x_N - \bar{x} & y_N - \bar{y} \end{pmatrix}$ is a known

263   distance matrix containing instantaneous distances from each drifter to the center of mass of the

264   drifter distribution at time t, and $A = \left[\bar{u}, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right]^T$ is the vector containing the unknown velocity

265   derivatives at time t that can be estimated using the Moore-Penrose pseudo-inverse as $A =$

266    $(D^T D)^{-1} D^T U$ (and similarly for the v-component).

267   Note that methods other than LLS can also be used to compute divergence and vorticity.

268   Specifically, divergence can be estimated as a rate of change of the area spanned by the drifter

269   polygon, and both divergence and vorticity can be estimated using Green's theorem as,

270   respectively, the circulation around and total flux through the drifter polygon. Rypina et al.

271   (2021) compared all three techniques in detail using both real and simulated drifters deployed in

272   the Alboran Sea at similar inter-drifter distances as the SPLASH drifters, and observed good

273   correspondence between all three techniques for clusters of 6 drifters, as long as the drifters

stayed within a few km of each other and the aspect ratio was reasonably small (≤5). For larger aspect ratios all methods started to deteriorate. Essink et al. (2022) also investigated the optimal way of computing velocity gradients, divergence, and vorticity from drifters. By quantifying the uncertainty in the velocity gradient calculation for different methods and different drifter configurations in a high-resolution submesoscale-resolving ocean circulation model, they concluded that the LLS was the most robust among the three methods, and that the accuracy of the LLS estimates grew linearly with the increasing number of drifters, and decreased logarithmically with the increasing aspect ratio of the drifter polygon (i.e., LLS works best for tight equidistant polygons with many drifters). Based on their analysis, they favored LLS over the area-rate-of-change and Green's theorem methods as their preferred method, and proposed 6 drifters with a polygon length scale of about 10 km and an aspect ratio of less than 10 as an optimal parameter range for reliable estimation of velocity gradients. They then successfully used LLS with these parameter criteria for estimating divergence and vorticity from the drifters in the Bay of Bengal.

Guided by recommendations of Rypina et al. (2022) and Essink et al. (2021), in this paper we will rely on the LLS method for estimating velocity gradients, and will refer to the LLS estimates of divergence and vorticity as trustworthy (and mark them by colored circles) if there are ≥ 6 drifters within a $3 -$km radius, the center of mass of the drifter distribution is located within the polygon, and the polygon aspect ratio is ≤ 6. If only the aspect ratio condition is not satisfied (but the number of drifters, the distance, and the center of mass conditions are), we will still compute LLS estimates but we will refer to them as less trustworthy (and mark them by colored diamonds). In all other cases, we do not produce estimates of divergence and vorticity.

3. Results

We start by qualitatively separating the motion of drifters into three stages. For about a day after deployment, all drifters started moving together in an anticyclonic fashion to the north and then northeast towards the coast (Fig. 1) – this is what we will refer to as the initial stage of motion. Upon approaching the shelf, the drifters halted their on-shore motion and split into two groups, a smaller northern group that headed northward along the coast and a larger southern group that moved southward. This splitting behavior was reminiscent of a hyperbolic motion in the vicinity of a hyperbolic trajectory, with a stable manifold emanating from a hyperbolic trajectory in the off-shore direction, and two unstable manifolds northward and southward from it in the along shore direction. As a result, a long and narrow filament roughly aligned with the coast is quickly formed just after 1 day. This filament contains about one third of all the drifters, with the rest of the drifters forming a less elongated and more compact blob just south-southwest of the filament. Some clustering temporarily occurs at about 1 day near the southeastern corner of the drifter configuration but goes away later. The slow-down of the on-shore movement, the splitting into the north-south groups, and the formation of the elongated along-shore filament constitute the second stage of motion, which lasted from about 0.9 to about 1.25 days after the deployment. Finally, during the third stage of motion, the drifters started moving off-shore to the southwest. As they progress further from the coast, trajectories started exhibiting more looping and the drifters dispersed further apart from one another, although they still remained in an elongated filament configuration (not anymore aligned with the coast) all the way until day 5, which is the end time of this dataset.

Having split the drifter movement into 3 stages, we next apply our Lagrangian methods to trajectory segments from $t_{start} = 0$ days until $t_{end} = 0.5, 1,$ and 3 days, respectively (top,

15

middle, and bottom row of panels in Figs. 3-9). The resulting fields highlight the dominant LCSs

that existed at the time of the drifter deployment (i.e., at $t_{start} = 0$) and that governed the

movement of drifters during the subsequent 0.5, 1, and 3 days, respectively. Since all 7

identifiers map out LCS at the start time of trajectory, using $t_{start} = 0$ allowed making the best

use of the nearly-regular deployment pattern. Fields computed for other time intervals, for

example, [0.5, 1] day or [1, 3] days, would need to be mapped to the location of trajectories at

0.5 days and 1 day, respectively, when drifters already stretched into highly elongated filaments,

thus losing the advantage of the regular deployment grid.

FTLEs (Fig. 3): During the initial stage of motion ($t_{start} = 0$ days and $t_{end} = 0.5$ days), the

FTLE field did not show any clear coherent structures, neither hyperbolic (maximizing ridges)

nor elliptic (isolated regions with significantly lower FTLEs). During the intermediate stage

($t_{start} = 0$ days and $t_{end} = 1$ day), the largest FTLEs were observed along the northwestern

edge of the release domain, containing drifters that split north-south upon approaching the coast

and formed an elongated along-shelf filament. FTLEs were negative for drifters released near the

middle of the northeastern edge of the release domain, which converged into a tight cluster in the

southeastern corner of the drifter distribution at 1 day. This feature was transient and

disappeared as the drifters moved offshore. The rest of the release domain has small positive

FTLE values; these were the drifters which did not experience strong along-shore alignment and

formed a more compact group in the southern part of the drifter distribution at 1 day. Finally,

during the third stage of motion ($t_{start} = 0$ days and $t_{end} = 3$ days), as the drifters moved

offshore and re-shaped into a northwest-southeast configuration, the only distinguishing feature

of the FTLE field was the blue cluster near the central part of the release domain. This cluster

contained trajectories that either remained together, or separated and then came back together

343   (since some of these data points are marked by yellow in the top row). When mapped to the

344   current positions of the drifters at *3* days, these smallest blue FTLEs corresponded to a group of

345   drifters in the western part of the distribution, i.e., a cluster of blue dots in the lower middle and

346   right panels of Fig. 4. (Note that the northwest-southeast configuration at 3 days was mostly

347   formed from the drifters located in the southern part of the distribution at 1 day, and so is

348   different from the along-shelf "tail".)

349   To summarize, although no clear coherent sets were distinguishable at early stage, the

350   characteristic patterns became clearer at later stages. Largest FTLEs indicated regions of strong

351   drifter separation that, during the intermediate stage of motion, were reminiscent of stable

352   manifolds of hyperbolic trajectories. Smallest FTLEs highlighted groups of drifters that stayed

353   closer together compared to their neighbors. Transient negative FTLE regions were also present

354   and highlighted groups of drifters temporarily converging into tight clusters (before spreading

355   apart again later on). The FTLEs varied significantly with the increasing duration of trajectories,

356   i.e., increasing t_{end}, suggesting that different flow features governed the movement of drifters

357   during different stages of motion. The calculation of FTLEs was straightforward and

358   computationally inexpensive, and by combining the structured and unstructured grid methods,

359   we were able to obtain FTLE values at both drifter release positions and in between them,

360   providing twice higher resolution compared to other methods.

361

362   *L (Fig. 4):* Trajectory path length *L* showed a~~n monotonic~~n increase in values with increasing

363   latitude across the release domain at all times, with the largest/smallest values in the

364   northwest/southeast. This large-scale gradient in *L* was dominated by the faster anticyclonic

365   motion of the northwestern drifters at early times. This was reminiscent of a solid body rotation,

17

366 where the northwestern drifters that were located further from the center of rotation than their

367 southeastern neighbors moved at a faster speed and thus covered a longer path length over a

368 given time interval. (This effect could presumably be removed by recalculation of L in an

369 appropriate rotating frame of reference, an operation that would not change the values of the

370 FTLEs. Thus it is perhaps not surprising that the distributions of the two metrics differ in

371 significant ways.) All other characteristic features, such as the splitting of trajectories into the

372 northern and southern group at about 1 day, the formation of an elongated along-shelf filament,

373 the transient convergence region, and the reshaping of the drifter configuration as it progressed

374 further offshore had only minor effects on the resulting path length fields. Specifically, we tried

375 looking for hyperbolic LCS, which would show up as level sets of L with the highest gradient in

376 the perpendicular direction, and for slow-moving elliptic regular regions which should be

377 characterized by a uniformly low L with a high gradient toward large L at the periphery, but we

378 did not find any. Thus, despite being easy to compute and straightforward to interpret, the path

379 length L was only marginally useful in identifying the dominant LCSs.

380

381 *CD (Fig. 5):* Results for the trajectory correlation dimension CD were generally similar to those

382 for the trajectory path length, in that CD was also dominated by the across domain gradient from

383 northwest to southeast, and the distribution of CD did not change dramatically in time. Although

384 CD is a more sensitive, and also more computationally expensive, measure of trajectory

385 complexity, it was still not able to identify the LCSs responsible for either the formation of the

386 elongated filament at 1 day, or the transient convergence zones just after 1 day, nor the

387 suppressed separation between trajectories coming from the central part of the domain at 3 days.

388   Overall, CD was no more useful than L in identifying the LCS, and, like L, had the same frame

389   dependence issues.

390

391   V_{en} ($Fig.\,6$): The encounter volume V_{en} ~~showed a lot of variability across the domain and~~ was

392   able to successfully highlight several different flow features governing the movement of drifters

393   at different stages of motion. During the initial stage, V_{en} had largest values in the southern part

394   of the release domain. From the top middle panel (the map of V_{en} at the current position of the

395   drifters) we observed that these enhanced values were caused by the tighter clustering of drifters

396   (so that they were able to meet more neighbors). During the intermediate stage of motion, the

397   distribution of V_{en} changed, and the largest values migrated to the northeastern edge of the

398   release domain. This was associated with the transient convergence zone (that we also observed

399   in the FTLE fields); trajectories released in that area converged into a tight cluster located at the

400   southeastern corner of the drifter distribution at *1* day (2nd row, middle panel). The elongated

401   along-shore filament seen at *1* day contained smallest V_{en} since trajectories in the filament

402   separated rapidly from their nearby neighbors and thus did not encounter many SPLASH

403   trajectories. This is likely a consequence of undersampling in hyperbolic regions (note that the

404   same region was marked by largest FTLEs indicative of hyperbolic behavior). Since SPLASH

405   drifters were only seeded over a small O($10\ km^2$) domain, the resulting V_{en} characterizes

406   encounters within this limited dataset, rather that with all trajectories in the entire domain,

407   leading to smallest V_{en} in this hyperbolic region instead of largest V_{en}, as would likely have been

408   the case for a domain-wide trajectory deployment. Trajectories that headed north after

409   approaching the coast at *1* day never caught up with the rest of the distribution, always staying

410   behind, i.e., to the north from the rest of the drifters. Thus, these drifters experienced the least

411 amount of encounters and, during the third stage of motion, had the smallest V_{en} values. Apart

412 from this low-encounter-number group, there were no other pronounced features in the V_{en} field

413 during the third stage of motion.

414 It is interesting to compare and contrast V_{en} with FTLEs, which became sort of a benchmark for

415 the LCS detection problems, being frame independent, commonly used, and easy to compute.

416 There are significant differences between the distributions of the two metrics, reflecting

417 differences in what the two are actually measuring. While both FTLEs and V_{en} are sensitive to

418 flow convergence/divergence, trajectory clustering, and hyperbolic behavior, one of the key

419 differences between them is that V_{en} is a time-integrated measure that depends on the behavior of

420 trajectories over the entire time interval between the initial and final times, whereas FTLEs only

421 depend on the initial and final positions of drifters (i.e., FTLEs do not care how trajectories got

422 to their final positions, whereas V_{en} does). For example, even though trajectories comprising the

423 low-FTLE blue cluster in the western part of the distribution at 3 days have come close together

424 at that time, over a time frame of 3 days they experienced no more trajectory encounters than

425 many other trajectories outside of that blue FTLE cluster (and thus were not standing out in the

426 V_{en} field). V_{en} is also more susceptible to undersampling issues than FTLEs, since the number of

427 encounters within a limited dataset is not necessarily representative of that with trajectories

428 seeded over the entire domain. For SPLASH drifters, undersampling led to smallest V_{en} along the

429 northwestern edge of the release domain during the 2nd stage of motion, where large FTLEs

430 indicated the presence of a stable manifold of a hyperbolic trajectory that was responsible for the

431 formation of an elongated along-shore filament at 1 day.

432   Overall, despite some challenges with undersampling, the encounter volume V_{en} proved to be an

433   interesting frame-independent diagnostic that was sensitive to both enhanced clustering,

434   hyperbolic behavior, and flow convergence, and was complementary to FTLEs.

435

436   *D (Fig. 7):* The challenge with computing dilation D (as well as $LAVD$) for real drifters is the

437   inability to reliably estimate divergence (vorticity) for isolated drifters and drifters forming

438   strongly elongated polygons. This was not a problem for SPLASH drifters during the early stage

439   of motion but became an issue as the drifters started to spread apart and formed elongated

440   filaments. During the initial stage of motion (top row), the most pronounced feature of the D

441   field was the negative cluster in the southern corner of the release domain, which contained

442   drifters that converged more than their neighbors. A similar feature has been identified by V_{en} as

443   the high-encounter-volume region. The rest of the domain had near-zero dilation. During the

444   second stage of motion (middle row), the negative dilation in the south diminished, and another

445   convergent negative-D region appeared along the northeastern edge and eastern corner of the

446   release domain. This is reminiscent of the negative-FTLE / high-V_{en} region in the middle rows of

447   Figs. 3 and 6. Trajectories released there converged into the southeastern corner of the drifter

448   distribution at *1* day. Around this time, an increasing number of trajectories started having

449   unreliable divergence values; for example, divergence and thus dilation, could not anymore be

450   reliably computed for the northern group of trajectories, which became too few and too sparse.

451   During the third stage of motion, this problem became even more important and by day *3*, the

452   dilation field was undefined for about half of the trajectories. The resulting D field was noisy and

453   did not exhibit any pronounced features.

454   Overall, dilation D was useful in highlighting the convergence zones during the first two stages

455   of motion, but numerical difficulties associated with reliably estimating divergence for sparse

456   datasets and elongated drifter configurations made it challenging to compute D over long time

457   intervals from real drifters.

458

459   *LAVD (Fig. 8):* During the first stage of motion, the strongest feature in the $LAVD$ map was the

460   red large-$LAVD$ region near the southern corner of the release domain. This area coincided

461   roughly with the negative-D and large-V_{en} in Figs. 6-7. During the second stage of motion, this

462   feature diminished in intensity and a second high-$LAVD$ region appeared near the eastern corner

463   of the domain. Again, a similar region has been highlighted by low FTLEs, high V_{en}, and

464   negative D, although $LAVD$ emphasized the eastern corner rather than the entire northeastern

465   edge of the release domain. Trajectories starting there converged into a tight cluster near the

466   southeastern corner of the drifter distribution at *1* day. It is interesting that $LAVD$ identified

467   similar regions as FTLEs, V_{en}, and D, despite the fact that clustering behavior and flow

468   convergence do not necessarily need to be associated with increased vorticity deviation. In our

469   case, clustering and convergence did coincide with increased vorticity deviation, suggesting that

470   perhaps a small-scale eddy or recirculation that was affecting this particular cluster of drifters

471   might have been responsible for all of these effects. (Note that interpreting the vorticity deviation

472   as vorticity is only possible when the domain-averaged background vorticity, $\bar{\omega}$, is small, which

473   was not always the case for the SPLASH drifters.) Finally, during the third stage of motion

474   (bottom row), the map of $LAVD$ became gappy (because, similar to the challenges with dilation,

475   here we could not reliably estimate $LAVD$ for about half of the drifters) and showed no

476   distinguished regions. However, when mapped to the current position of the drifters (lower

477 middle panel), the cluster in the middle of the drifter distribution showed larger *LAVD* values

478 than clusters to the northwest and southeast (but since trajectories forming the middle cluster

479 came from different parts of the release domain, this feature did not stand out in the left panel).

480 Overall, during the first two stages of motion, *LAVD* highlighted two regions with enhanced

481 *LAVD* values. While large *LAVD* does not generally indicate convergence, in our case both

482 regions were strongly convergent. At later times, vorticity estimation became less reliable, and it

483 became harder to distinguish coherent features in the sparse and noisy map of *LAVD*. Note that

484 our high-*LAVD* regions differed from the classical examples of rotationally-coherent Lagrangian

485 eddies. Our regions were not circular, did not have a single maximum, and were too noisy to

486 identify the outermost convex contour level, which marks the outer edge of the coherent

487 rotational eddies in the standard application of the *LAVD* technique. Thus we cannot call these

488 high-*LAVD* features rotationally-coherent Lagrangian eddies. It is interesting that even though

489 trajectories exhibited clear anticyclonic rotation during the first 12 hours, *LAVD* did not identify

490 this anticyclonic eddy. We think this might be because the SPLASH release domain was too

491 small and was located entirely within this vortex structure.

492 *Spectral Clustering (Fig 9):* At early times, the number of coherent clusters identified by the SC

493 algorithm was quite large (12), although some clusters only contained a few drifters. (Recall that

494 the optimized-parameter SC is able to autonomously identify the optimal number and optimal

495 size of the clusters, without input from the user). Among the detected clusters, the yellow cluster

496 located in the south-southwest of the release domain is perhaps the most noteworthy because it

497 resembled the low-FTLE / large-V_{en}/ negative-*D*/ large-*LAVD* region that contained trajectories

498 that stayed close together during the initial stage of motion. As the drifters entered the second

23

499   stage of motion, the number of identified coherent clusters decreased to 6. Most of the release

500   domain was split between two large clusters – the cyan cluster in the north-northeast containing

501   drifters attracted by the convergence region (i.e., drifters that converged/came close to the

502   southeastern corner of the drifter distribution at 1 day), and the green cluster in the south of the

503   release domain containing drifters that did not feel the pull of that convergence zone. The

504   remainder of the domain, i.e., the northwestern edge of the domain that mostly contained the

505   trajectories forming an elongated along-shore filament, was split into 4 more clusters. Finally, at

506   the third stage of motion, the drifters were split into 8 clusters, and the grouping was most

507   straightforward to interpret by looking at the lower middle panel. All trajectories in the western

508   cluster were blue (these trajectories came from the central and southern portion of the domain in

509   the bottom left panel), with the yellow cluster to the southeast of it (these trajectories came from

510   around the periphery of the blue cluster in the bottom left panel), and with the orange group

511   further to the southeast of the yellow cluster (most orange trajectories originate from the

512   northeastern edge of the release domain in the bottom left panel). The remaining 5 clusters only

513   contained 1 or 2 trajectories.

514   Overall, the spectral clustering algorithm seems to have identified physically-meaningful and

515   intuitively-clear coherent clusters; the movement was similar for drifters within each cluster and

516   dissimilar between the clusters. There were also good correspondences between the spectral

517   clusters and coherent featured highlighted by other methods.

518

519   **4. Summary and Discussion**

520   SPLASH drifter experiment provided the long-awaited opportunity to test the performance of

521   different dynamical systems techniques with real, rather than simulated, ocean drifters. Although

522   many other drifter data sets are available for various regions of the World Ocean, drifters are

523   typically released by a handful here and there, and the resulting data is typically inadequate for

524   mapping out the LCS. For example, NOAA's Global Drifter Program data set contains several

525   thousands of near-surface drifter trajectories released between 1971 and today, but the density of

526   the drifter distribution at any given time is only about 1 per 5-by-5 deg box, which is too sparse

527   to identify even mesoscale LCSs.

528   Three qualitatively-different stages of motion were evident in the SPLASH drifter data. During

529   the first stage, all drifters moved anticyclonically toward the coast. During the second stage, the

530   drifters halted their on-shore motion, split north-south, and formed an elongated along-shelf

531   filament. During the third stage, the drifters moved off-shore, rearranging themselves into a

532   northwest-southeast configuration. As the character of drifter movement changed with time, the

533   maps of the Lagrangian metrics and the resulting LCSs that they highlighted changed as well. In

534   order to capture this time-dependence, we have applied the Lagrangian metrics to segments of

535   trajectories from fixed $t_{start} = 0$ days to variable $t_{end} = 0.5, 1,$ and 3 days. When the

536   Lagrangian metrics were mapped back to the initial positions of drifters at t_{start}, the resulting

537   maps highlighted the dominant LCS (such as the hyperbolic-type LCS responsible for the

538   formation of the along-shore filament at 1 day, the convergence-type LCS attracting drifters into

539   the southeastern corner at 1 day, and the elliptic-type LCS forming during the 3rd stage of off-

540   shore motion) which existed at the time of the deployment within the deployment domain, and

541   which govern the subsequent motion of drifters over the corresponding time interval. The fact

542   that the results for any particular measure differed between the three time intervals is consistent

543   with submesoscale dynamics, where fronts, small eddies, and filaments form, evolve, and

544   disappear on time scales of days or less.

545   The Lagrangian techniques we have examined include FTLEs, trajectory path length, trajectory

546   correlation dimension, trajectory encounter number, dilation, $LAVD$, and optimized-parameter

547   spectral clustering. This list was motivated by Hadjighasem et all. (2017) and is by no means

548   exhaustive, but it includes a variety of commonly-used methods that are based on different

549   properties of trajectories, make use of the different definitions of coherence, and thus aim to

550   identify different types of LCSs. Interestingly, despite the differences in their underlying

551   principles and methodologies, many of these methods identified similar features within the

552   SPLASH drifter data set.

553   Among the most prominent features that were highlighted by multiple methods were: 1) the

554   region near the northwestern edge of the release domain (large FTLEs, small V_{en}, yellow/orange

555   clusters), which contained trajectories that split north-south upon approaching the shelf and

556   formed an elongated along-shelf filament at about 1 day; 2) the very strong but transient

557   convergence region located near the northeastern edge of the release domain (negative FTLEs,

558   large V_{en}, strongly negative dilation, cyan spectral cluster), which contained trajectories that

559   converged into a tight cluster at about 1 day; and 3) the region in the central/southern part of the

560   release domain (small FTLEs, blue spectral cluster), which contained trajectories that remained

561   close to each other starting from 2.5 days and onward.

562   Although all of the identified structures were noisier and more complex than the classical elliptic

563   and hyperbolic LCSs in textbook examples, some of the features bore resemblance to their

564   classical counterparts. For example, the north-south splitting of trajectories starting within the

565   red FTLE region near the northwestern edge of the domain was qualitatively similar to the

566   behavior of trajectories near a hyperbolic region, where particles approach the hyperbolic

567   trajectory along a stable manifold and then split and move away from the hyperbolic trajectory

568 along the two unstable directions. The detected large-FTLE region near the northwestern edge of

569 the release domain might thus possibly indicate the presence of a stable manifold in this region.

570 From the standpoint of numerical efficiency, FTLEs and L were the least computationally

571 expensive, whereas CD, V_{en}, and Spectral Clustering were the most computationally expensive.

572 However, with only 135 trajectories, the differences in the amount of time required to apply each

573 technique were not critical. More importantly, FTLEs had the advantage of providing values at

574 the positions of each drifter as well as between the neighboring drifters, effectively yielding

575 output fields with twice the resolution of the other methods. FTLEs were also less affected by the

576 gaps in GPS transmissions along trajectories, because the estimation of FTLEs at a particular

577 time only required knowing the initial and the current positions of the drifters, rather than

578 requiring the information about the entire trajectory up to that time, as in the case of all other

579 methods - path length, correlation dimension, encounter number, dilation, $LAVD$, and spectral

580 clusters.

581 One challenge with dilation and $LAVD$ is the loss of accuracy at longer times, when the drifters

582 form elongated polygons. The deterioration of the velocity gradients estimates (that are required

583 for estimating dilation and $LAVD$) with the increasing aspect ratio of the drifter polygon is

584 intuitively clear. As the polygon elongates, the information about the velocity gradient in the

585 perpendicular direction diminishes and is lost when the polygon approaches a one dimensional

586 line. This is true for all methods of estimation, not just for LLS, and presents a fundamental

587 challenge for estimating dilation and $LAVD$ from drifters, which tend to naturally form elongated

588 filaments in oceanic flows.

It is interesting to note that the two frame-dependent methods – L and CD, which were dominated by the large-scale gradient across the entire release domain and did not highlight any submesoscale features – were the least useful in identifying LCSs. For SPLASH drifters, this dominant large-scale gradient developed during the initial anticyclonic phase of motion, when drifters deployed closer/further from the center of rotation were shorter/longer and less/more complex. This overpowering trend could potentially be removed by moving into a co-rotating reference frame (i.e., a natural frame of reference), but identifying such natural reference frame is non-trivial in the absence of additional information about the flow.

Massive drifter releases such as SPLASH are extremely useful for improving our understanding of the transport and exchange processes at submesoscale. Specifically, data from the SPLASH and other similar experiments have been used for estimating diffusivity and studying particle spreading regimes at submesoscale (Poje et al., 2014; Beron-Vera and LaCasce, 2016). We have shown that a simultaneous release of about 100 drifters provides a glimpse of the dominant Lagrangian Coherent Structures that govern the transport of water and the movement of drifters.

The SPLASH experiment was not specifically focused on identifying LCSs, so the drifter release locations and timings were not optimized for capturing the underlying LCSs. Our analysis suggested that, luckily, a stable manifold of a hyperbolic trajectory was likely present in the northwestern edge of the domain spanned by the drifters at the time of their release and persisted for at least the first 1-1.5 days of the experiment. As explained above, this feature manifested itself as a high-FTLE region and was characterized by the north-south splitting of trajectories around day 1. However, no clear elliptic LCSs (i.e., coherent eddy cores) were identified by any of the methods, even though an anticyclone was likely present near the SPLASH release site at the time of deployment (based on the numerical model simulations and the clockwise movement

of drifters during the first day after release). Note that even the *LAVD* method, which was specifically designed to identify rotationally-coherent Lagrangian eddies, was also not able to highlight this anticyclone, possibly because *LAVD* is a wrong tool for identifying an eddy from a small trajectory set located entirely within an eddy. It is also possible that this anticyclone did not possess a Lagrangian core, or the core was located outside of the drifter release domain and/or was not properly resolved by the SPLASH drifters.

In the future, it would be interesting to repeat the experiment with the drifter deployment site and the release pattern optimized for capturing specific LCSs whose presence could have been predicted based on a model or satellite data.

The very rapid nature of evolution at submesoscales may cause an evenly spaced array of drifters to rapidly collect into filaments, making it difficult to continue to accurately compute certain Lagrangian measures. In the SPLASH experiment, for example, the nearly-rectangular deployment mesh of drifters (which took quite a bit of effort to achieve) eroded into an elongated filament over a time scale of about a day. Note, however, that it is precisely this rapid filamentation process and the rapid deformation of the initial mesh that gives rise to the strong, pronounced, and detectable LCSs. A related challenge that complicates the understanding of the flow from the Lagrangian analysis presented here is that the features that are found (for example, a hyperbolic region) say something about local kinematic features of the flow but do not allow one to say much about how the flow looks like on broader spatial scales, or over time scales longer than just a few days. The rapid filamentation experienced by the drifters prevents mapping out the structures at later times and over regions other than the original deployment domain. This might be one of the important things that we have learned about the flow, and about sampling through massive drifter releases.

Finally, in order to investigate the ~~robustness and~~ reliability of the real-drifter-derived LCS, we have simulated the SPLASH drifter dataset in a model and then compared the resulting SPLASH-like drifter-based LCS to those computed using dense regular orthogonal grids of trajectories (we refer to the latter as dense-grid simulations). We used the operational data-assimilative Navy Coastal Ocean Model (NCOM) forecasting model for this purposes (https://data.gulfresearchinitiative.org/data/R4.x265.245:0002).

Comparison between SPLASH-like and dense-grid model simulations showed reasonably good agreement for many, although not all, metrics and times, suggesting that many, although not all, SPLASH fields were reliable (see Supplementary Material). Specifically, SPLASH-like FTLEs were most reliable at shorter times and still meaningful at longer times in regions with strong hyperbolic-type LCS located far enough away from each other to be resolved by the deployment grid. L and CD were reliable at all times, but since they did not identify any hyperbolic, elliptic, or convergence-type LCS for SPLASH drifters, they were perhaps least useful among the 7 methods. In contrast to FTLEs, V_{en} was not reliable at short times but improved its reliability at longer times. D and $LAVD$ worked well at short times when drifters were still relatively close and didn't form elongated filaments, but deteriorated at longer times due to the rapid filamentation of the drifter distribution. Finally, SPLASH-like and dense-grid SC both identified large numbers of clusters within the SPLASH domain at short times and fewer clusters at longer times. At short time, the clusters were different between the SPLASH-like and dense-grid simulations; at longer times, there was a number of similarities in the identified clusters (longitudinal split along same longitude at intermediate time, and assignment of most of the domain to 1 cluster at later time), but the details of the cluster configurations were different, especially near the edges of the domain.

Comparing observations to simulations, Lagrangian metrics were of similar magnitude for the real and simulated SPLASH drifter. The actual range of values in simulations and observations matched for $FTLEs$ and D, as well as L/CD at 0.5 days and 3 days, and V_{en} at 0.5 days. $LAVD$ was 2 to 3 times larger in observations, V_{en} was 2 to 3 times larger in simulations at the intermediate and late stages, and L/CD were slightly larger in simulations at the intermediate stage (note, however, that we used 1 day/2 days as a characteristic time for the intermediate stage in observations/simulations). Hyperbolic- and convergence-type LCS were present in both observations and simulations, and no clear elliptic-type LCS were seen in either model or observations. The model fields were generally significantly less noisy, exhibited a larger degree of coherence, and at early times had more positive dilation, compared to mostly near-zero and negative in observations. Detailed comparisons can be found in Supplementary Material.

Data Availability Statement

Data from the Submesoscale Processes and Lagrangian Analysis on the Shelf (SPLASH) surface drifters used in this paper is available from:

https://data.gulfresearchinitiative.org/data/R4.x265.000:0074.

679 The NCOM model output fields used in this paper are available from:

680 https://data.gulfresearchinitiative.org/data/R4.x265.245:0002.

681 **Author Contribution Statement:** IR led the overall effort and primarily wrote the manuscript,

682 TG performed estimation of most Lagrangian metrics, LP and TO contributed to the

683 interpretation of the results and editing of the manuscript.

684 **Competing interests Statement:** no competing interests

685 **References:**

686

687 Balasuriya, S., Ouellette, N.T. and Rypina, I.I.: Generalized Lagrangian coherent

688 structures. *Physica D: Nonlinear Phenomena*, *372*, pp.31-51, 2018.

689

690 Beron-Vera, F.J. and LaCasce, J.H.: Statistics of simulated and observed pair separations in the

691 Gulf of Mexico. *Journal of Physical Oceanography*, *46*(7), pp.2183-2199, 2016.

692

693 Essink, S., Hormann, V., Centurioni, L.R. and Mahadevan, A., 2022. On characterizing ocean

694 kinematics from surface drifters. *Journal of Atmospheric and Oceanic Technology.*

695

696 Filippi, M., Rypina, I.I., Hadjighasem, A. and Peacock, T. (a): An Optimized-Parameter Spectral

697 Clustering Approach to Coherent Structure Detection in Geophysical Flows. *Fluids*, *6*(1), p.39,

698 2021.

699

700   Filippi, M., Hadjighasem, A., Rayson, M., Rypina, I.I., Ivey, G., Lowe, R., Gilmour, J. and

701   Peacock, T. (b): Investigating transport in a tidally driven coral atoll flow using Lagrangian

702   coherent structures. *Limnology and Oceanography*, *66*(11), pp.4017-4027, 2021.

703

704   Jacobs, G.A., Bartels, B.P., Bogucki, D.J., Beron-Vera, F.J., Chen, S.S., Coelho, E.F., Curcic,

705   M., Griffa, A., Gough, M., Haus, B.K. and Haza, A.C.: Data assimilation considerations for

706   improved ocean predictability during the Gulf of Mexico Grand Lagrangian Deployment

707   (GLAD). *Ocean Modelling*, *83*, pp.98-117, 2014.

708

709   Hadjighasem, A., Karrasch, D., Teramoto, H. and Haller, G.: Spectral-clustering approach to

710   Lagrangian vortex detection. *Physical Review E*, *93*(6), p.063107, 2016.

711

712   Hadjighasem, A., Farazmand, M., Blazevski, D., Froyland, G. and Haller, G.: A critical

713   comparison of Lagrangian methods for coherent structure detection. *Chaos: An Interdisciplinary*

714   *Journal of Nonlinear Science*, *27*(5), p.053104, 2017.

715

716   Haller G. 2001. Distinguished material surfaces and coherent structures in three-dimensional

717   fluid flows. *Physica D* 149:248–77

718

719   Haller, G.: Lagrangian coherent structures from approximate velocity data. *Physics of*

720   *fluids*, *14*(6), pp.1851-1861, 2002.

721

722   Haller, G.: An objective definition of a vortex. *Journal of fluid mechanics*, *525*, pp.1-26, 2005.

723

724   Haller, G.: Lagrangian coherent structures. *Annual Review of Fluid Mechanics*, *47*, pp.137-162,

725   2015.

726

727   Haller, G. and Yuan, G.: Lagrangian coherent structures and mixing in two-dimensional

728   turbulence. *Physica D: Nonlinear Phenomena*, *147*(3-4), pp.352-370, 2000.

729

730   Haller, G., Hadjighasem, A., Farazmand, M. and Huhn, F.: Defining coherent vortices

731   objectively from the vorticity. *Journal of Fluid Mechanics*, *795*, pp.136-173, 2016.

732

733   Huntley, H.S., Lipphardt Jr, B.L., Jacobs, G. and Kirwan Jr, A.D.: Clusters, deformation, and

734   dilation: Diagnostics for material accumulation regions. *Journal of Geophysical Research:*

735   *Oceans*, *120*(10), pp.6622-6636, 2015.

736

737   Lekien, F. and Ross, S.D.: The computation of finite-time Lyapunov exponents on unstructured

738   meshes and for non-Euclidean manifolds. *Chaos: An Interdisciplinary Journal of Nonlinear*

739   *Science*, *20*(1), p.017505, ., 2010.

740

741   Laxague, N., Özgökmen, T. M., Haus, B. K., Novelli,G., Shcherbina, A., Sutherland, P.,

742   Guigand, C., Lund, B., Mehta, S., Alday, M., Molemaker, J.: Observations of near-surface

743   current shear help describe oceanic oil and plastic transport. *Geophysical Research Letters*,

744   45(1), 245-249, 2018.

745

746 Lund, B., Haus, B.K., Graber, H.C., Horstmann, J., Carrasco, R., Novelli, G., Guigand, C.,

747 Mehta, S., Laxague, N., and Özgökmen, T.M.: Marine X-Band Radar Currents and Bathymetry:

748 An Argument for a Wave Number-Dependent Retrieval Method. JGR-Oceans, 125(2),

749 e2019JC015618, 2020.

750

751 Mendoza, C. and Mancho, A.M., 2010. Hidden geometry of ocean flows. *Physical review*

752 *letters, 105*(3), p.038501.

753

754 Molinari, R. and Kirwan Jr, A.D.: Calculations of differential kinematic properties from

755 Lagrangian observations in the western Caribbean Sea. *Journal of Physical Oceanography*, *5*(3),

756 pp.483-491, 1975.

757

758 Özgökmen, T.M., Bracco, A., Chassignet, E.P., Chang, H., Chen, S.C., D'Asaro, E., Fox-

759 Kemper, B., Haza, A.C., Jacobs, G., Novelli, G. and Poje, A.: Basin-Scale and Near-Surface

760 Circulation in the Gulf of Mexico. In *International Oil Spill Conference* (Vol. 2021, No. 1, p.

761 686903), 2021.

762

763 Poje, A.C., Özgökmen, T.M., Lipphardt, B.L., Haus, B.K., Ryan, E.H., Haza, A.C., Jacobs, G.A.,

764 Reniers, A.J.H.M., Olascoaga, M.J., Novelli, G. and Griffa, A.: Submesoscale dispersion in the

765 vicinity of the Deepwater Horizon spill. *Proceedings of the National Academy of*

766 *Sciences*, *111*(35), pp.12693-12698, 2014.

767

768   Rypina, I.I., Pratt, L.J., Pullen, J., Levin, J. and Gordon, A.L.: Chaotic advection in an

769   archipelago. *Journal of Physical Oceanography*, *40*(9), pp.1988-2006, 2010.

770

771   Rypina, I.I., Scott, S.E., Pratt, L.J. and Brown, M.G.: Investigating the connection between

772   complexity of isolated trajectories and Lagrangian coherent structures. *Nonlinear Processes in*

773   *Geophysics*, *18*(6), pp.977-987, 2011.

774

775   Rypina, I.I. and Pratt, L.J.: Trajectory encounter volume as a diagnostic of mixing potential in

776   fluid flows. *Nonlinear Processes in Geophysics*, *24*(2), pp.189-202, 2017.

777

778   Rypina, I.I., Llewellyn Smith, S.G. and Pratt, L.J.: Connection between encounter volume and

779   diffusivity in geophysical flows. *Nonlinear Processes in Geophysics*, *25*(2), pp.267-278, 2018.

780

781   Rypina, I.I., Getscher, T.R., Pratt, L.J. and Mourre, B.: Observing and quantifying ocean flow

782   properties using drifters with drogues at different depths. *Journal of Physical*

783   *Oceanography*, *51*(8), pp.2463-2482, 2021.

784

785   Samelson, R.M. and Wiggins, S.: *Lagrangian transport in geophysical jets and waves: The*

786   *dynamical systems approach* (Vol. 31). Springer Science & Business Media, 2006.

787

788   Shadden, S.C., Lekien, F. and Marsden, J.E.: Definition and properties of Lagrangian coherent

789   structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D:*

790   *Nonlinear Phenomena*, *212*(3-4), pp.271-304, 2005..

791

792 Shi, J. and Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on pattern*

793 *analysis and machine intelligence*, *22*(8), pp.888-905, 2000.

794

795 Solodoch, A., Molemaker, J. M., Srinivasan, K., Berta, M., Marie, L., Jagannathan, A.:

796 Observations of Shoaling Density Current Regime Changes in Internal Wave Interactions. J.

797 Phys. Oceanogr., 50(6), 1733-1751, 2020.

798

799 Von Luxburg, U.: A tutorial on spectral clustering. *Statistics and computing*, *17*(4), pp.395-416,

800 2007.

801

802 Yaremchuk, M. and Coelho, E.F., 2014. Filtering drifter trajectories sampled at submesoscale

803 resolution. *IEEE Journal of Oceanic Engineering*, *40*(3), pp.497-505.

804

805

806

807

808

809
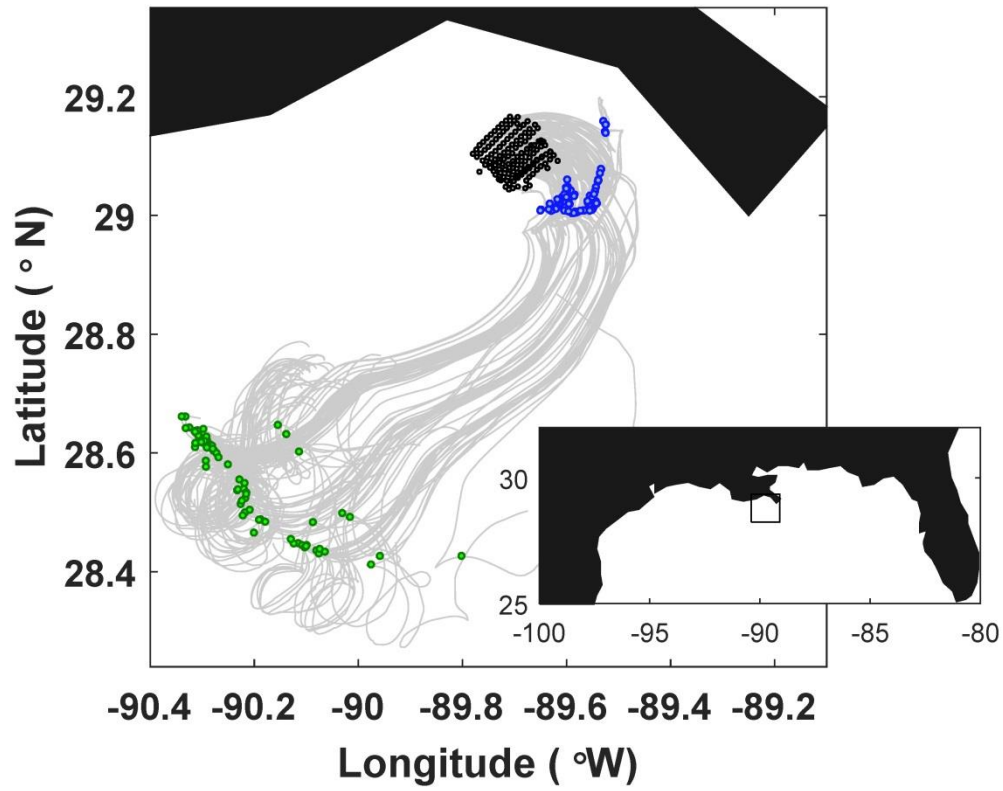
810 **Figure 1. Trajectories of the SPLASH drifters, with their positions at the release time, 1 day, and 3 days shown by black, blue,**

811 **and green dots, respectively. The inset shows the geographical location of the experiment site, with black box indicating the**

812 **domain shown in the main panel.**

813

814

815

816

817

818

819

820
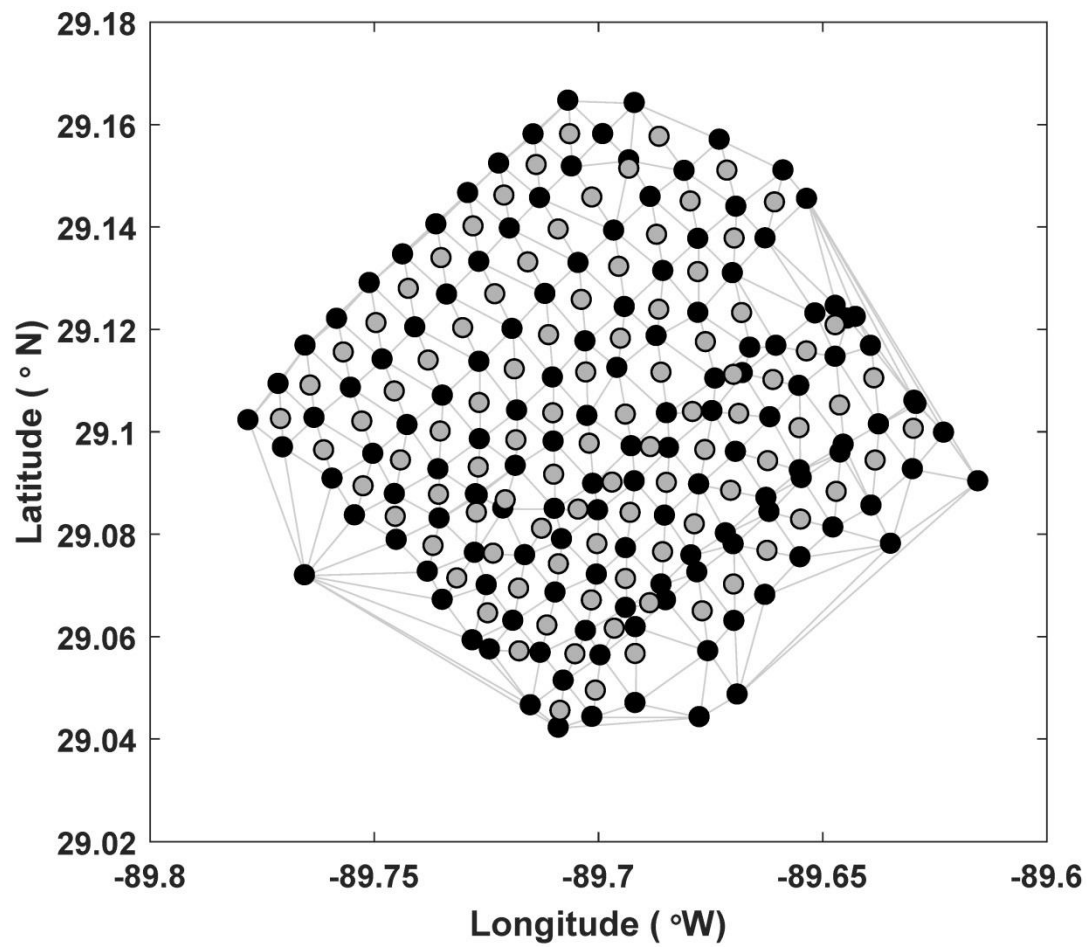
Figure 2. Release locations of the SPLASH drifters (black dots) with the Delaunay delineation (grey lines) used to define the closest neighbors for estimating FTLEs at the drifter positions using the unstructured grid method. Grey circles show locations between the drifters, at which FTLEs were estimated via the structured grid method (using a quadruplet of black drifters around each grey dot).

821

822

823

824

825

826

827

828

829

830

831
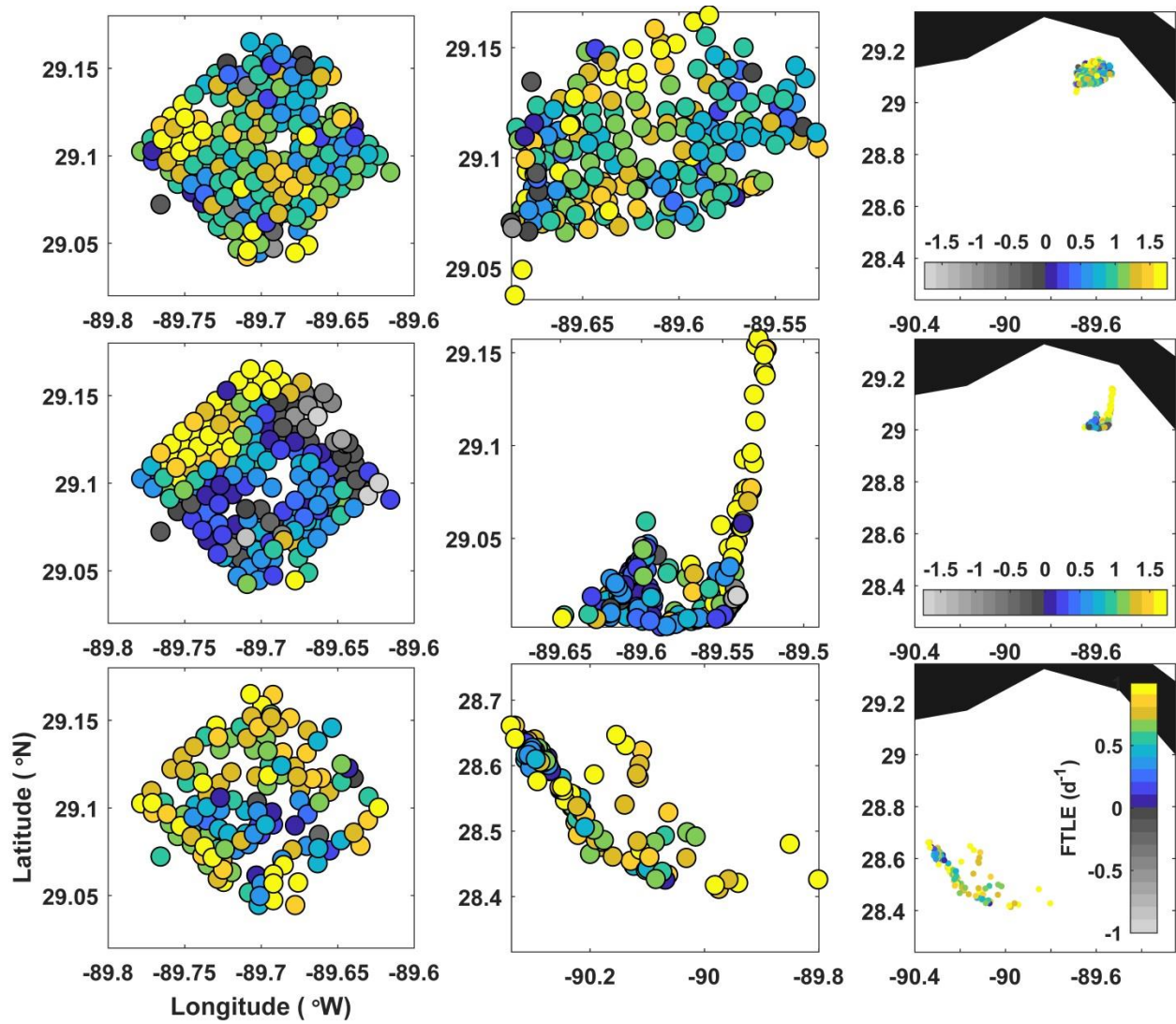
Figure 3. Real-drifter-based FTLEs at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left) and current

(middle and right) positions of the drifters.

834
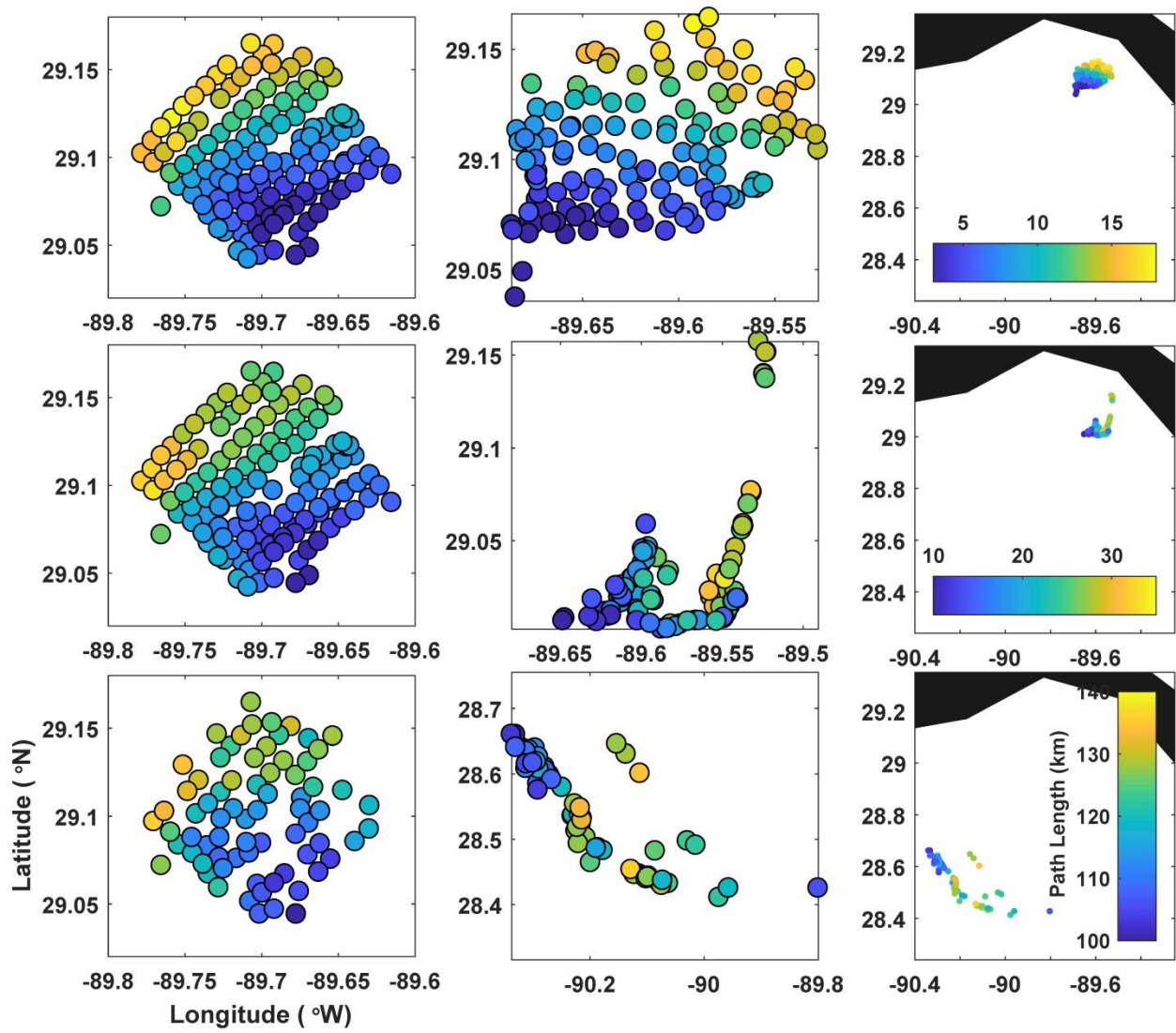
Figure 4. Real-drifter-based path length at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left) and current

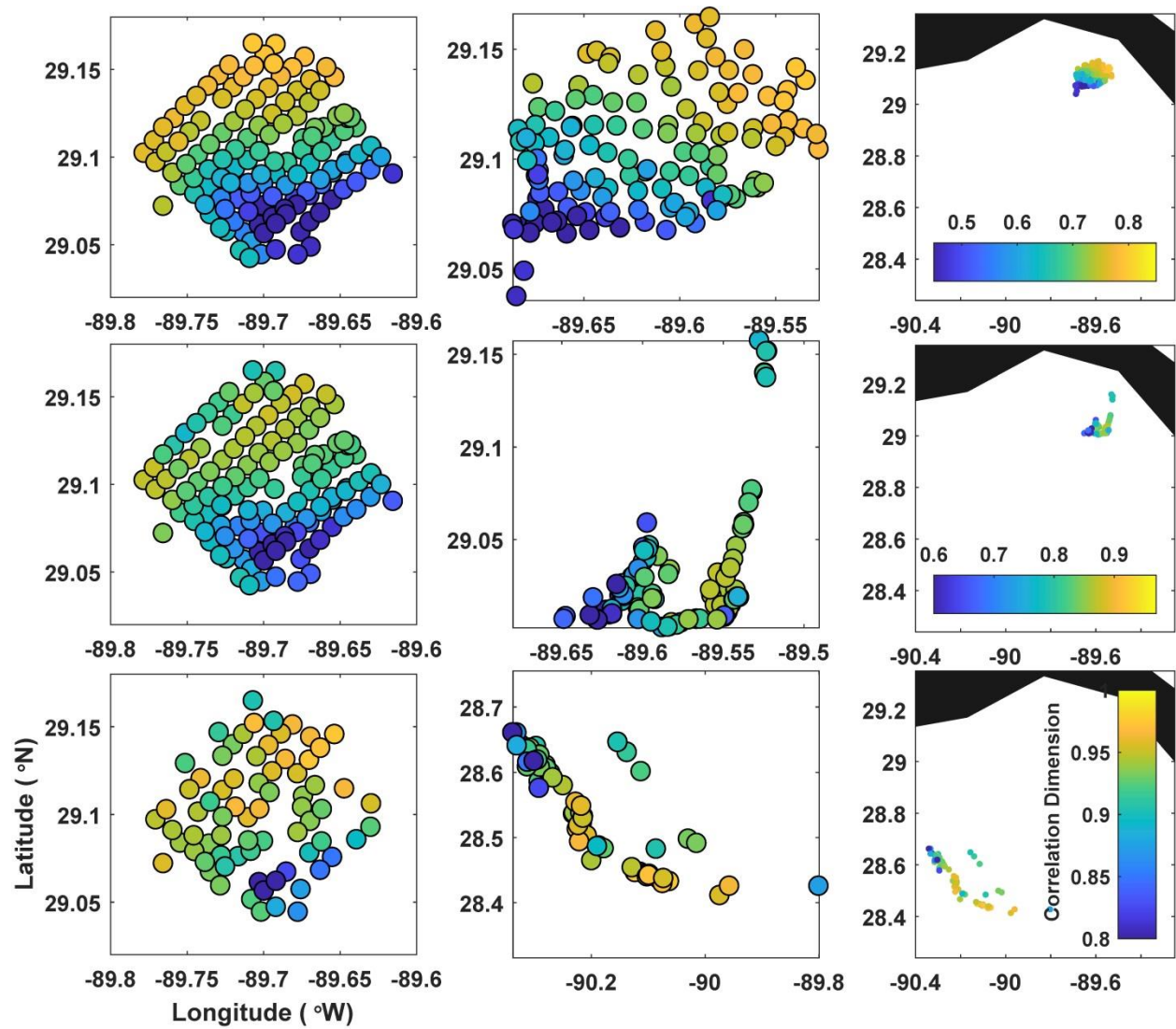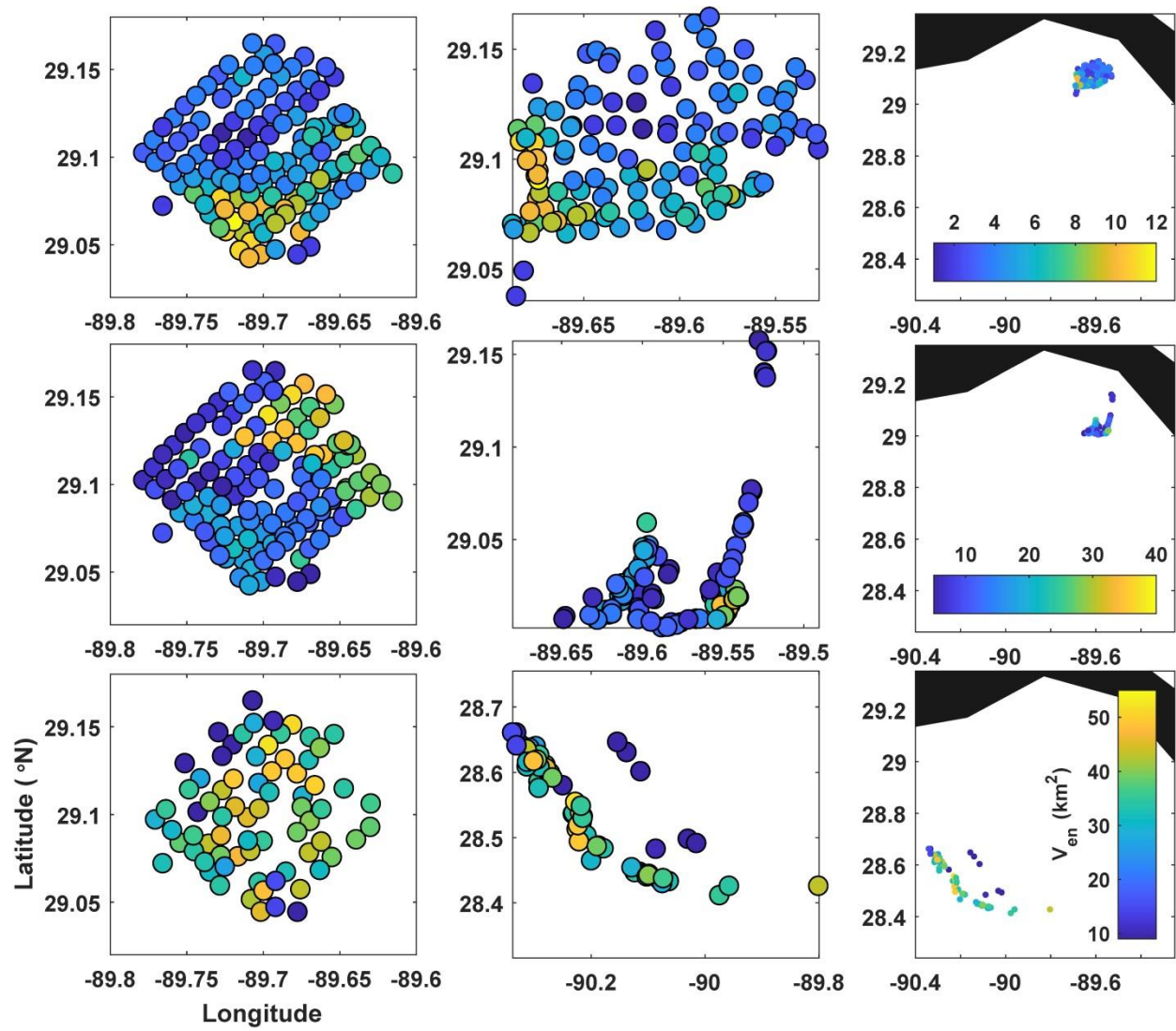(middle and right) positions of drifters.

837

838

Figure 5. Real-drifter-based correlation dimension at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left)

840   and current (middle and right) positions of drifters.

841

842

Figure 6. Real-drifter-based encounter volume at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left) and

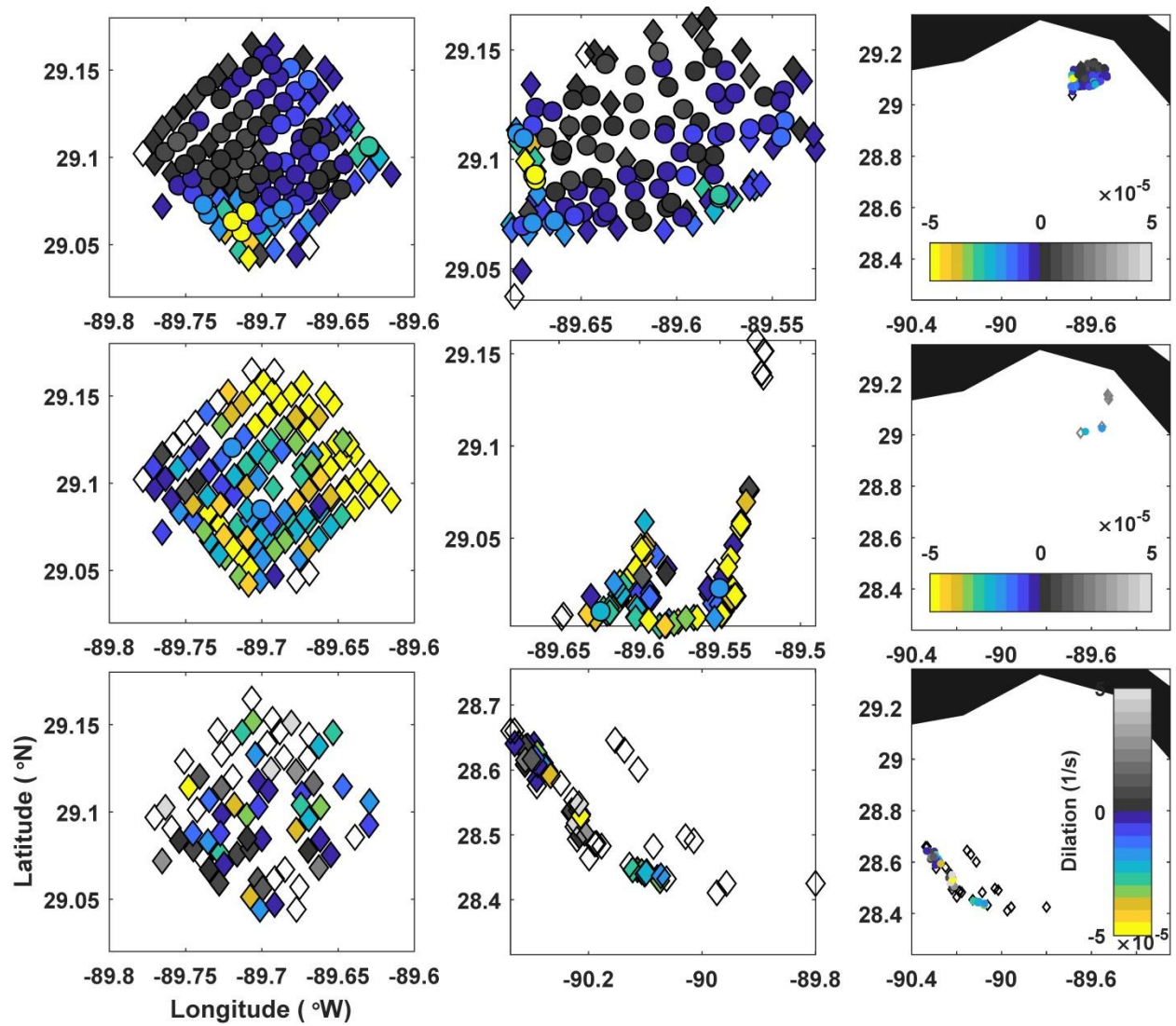**current (middle and right) positions of drifters.**

Figure 7. Real-drifter-based dilation at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left) and current (middle and right) positions of drifters. Colored circles/colored diamonds/empty (white) diamonds mark drifters for which calculation of gradients was most reliable /less reliable/not reliable, according to the LLS reliability criteria described in the methods section.
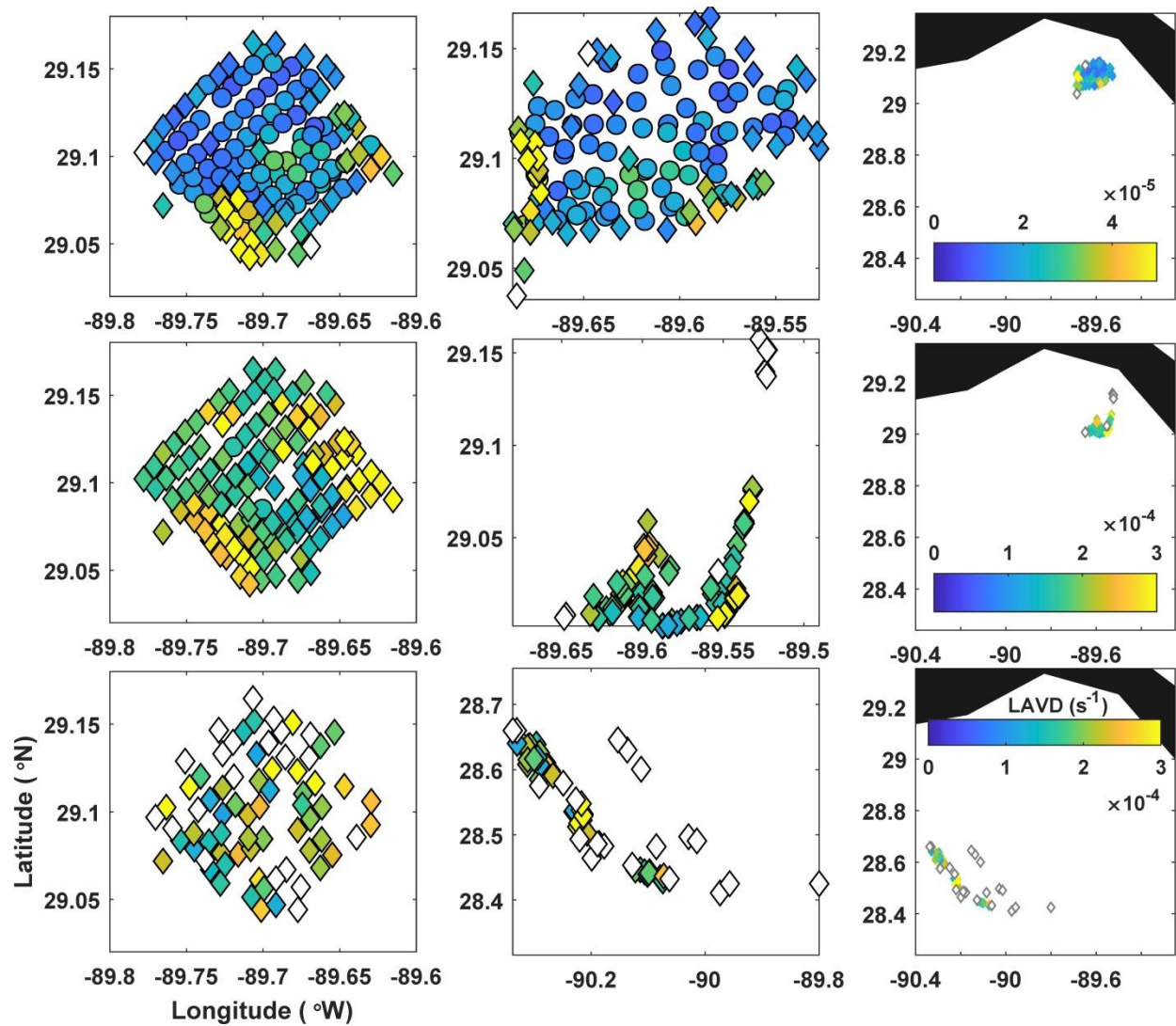
852

Figure 8. Real-drifter-based LAVD at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left) and current

(middle and right) positions of drifters. Colored circles/colored diamonds/empty (white) diamonds mark drifters for which

calculation of gradients was most reliable /less reliable/not reliable, according to the LLS reliability criteria described in the
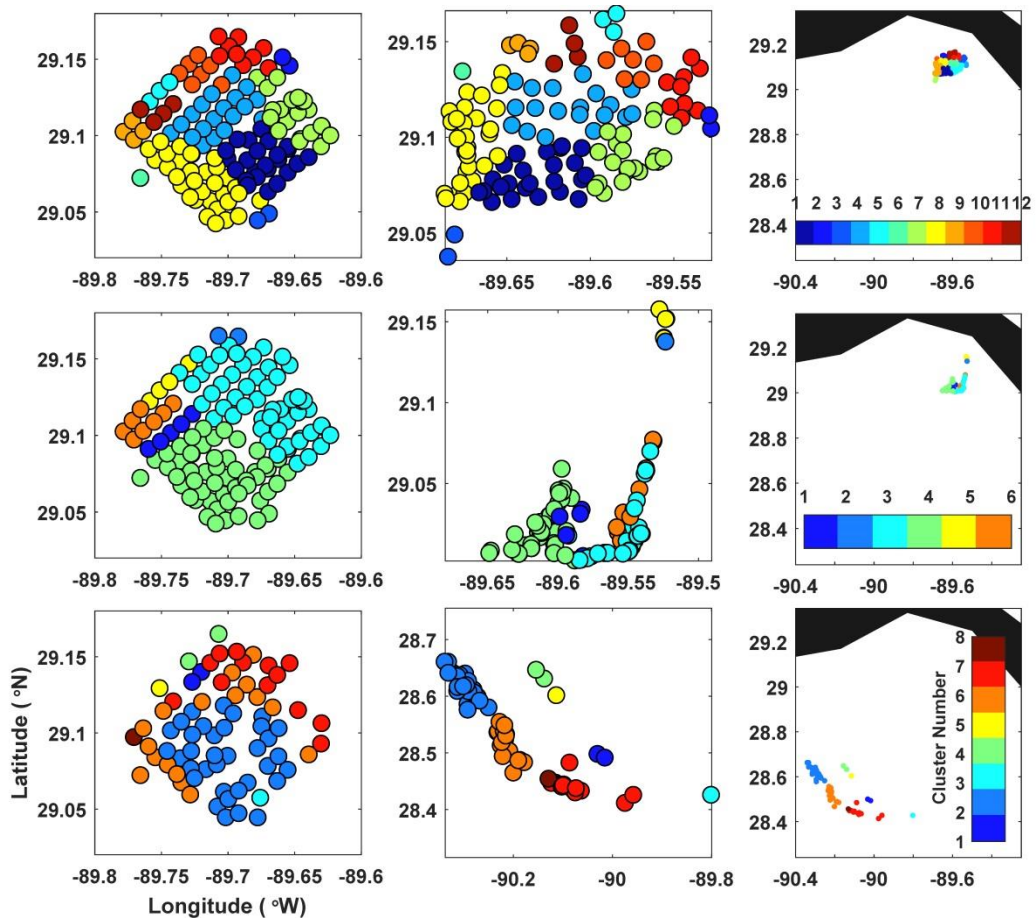
methods section.

857

858

859

Figure 9. Real-drifter-based spectral clusters at (top) 0.5, (middle) 1, and (bottom) 3 days, mapped to the initial (left) and current (middle and right) positions of drifters.

862

863

864

865

866

867