

Uncertainties in Lagrangian prediction

Sanjeeva Balasuriya

School of Mathematical Sciences, University of Adelaide, Adelaide SA 5005, Australia

December 20, 2019

Abstract

Lagrangian (following the flow) motion of fluid particles is connected to the Eulerian (spatio-temporally defined) velocity field by an ordinary differential equation. Lagrangian motion causes transport (of fluid, as well as of other quantities such as energy and pollutants), and is therefore highly relevant to climate, weather, the environment, industrial mixers, etc. Velocities, on the other hand, are often known in an Eulerian sense from observational, experimental or computational fluid dynamics data. The uncertainties in this Eulerian data will lead to uncertainties in Lagrangian prediction. This short article examines some recent approaches to ascribing uncertainties to Lagrangian-derived information.

1 Lagrangian motion and uncertainty

At all spatial scales (microfluidic, laboratory, oceanic, atmospheric, Jovian), *transport* is governed by where particles move, based on the forces acting on them. This movement strongly influences how chemical species, pollutants, nutrients, heat energy, and biological organisms disperse. Forces give rise to a *velocity* with the relationship given, for example, by the Navier-Stokes equations in the classical fluids context. Velocities are typically unsteady (time-varying), and the rate of change of a particle's position is then instantaneously given by the local velocity. The tracks followed by particles as a result of these velocities is called *Lagrangian* motion in fluid mechanics. In many cases, we have the velocities from observations, experiments, computational fluid dynamics simulations, or via some proxy (e.g., inferring Jupiter's atmospheric motion by performing optical processing on videos, imputing oceanic velocities from satellite measurements of sea-surface heights by using the so-called geostrophic approximation). These are *Eulerian* measurements, meaning that they are given as functions of position and time. Since

the information is moreover typically available on a spatial grid, at discrete values of time, there are uncertainties in the velocity away from gridpoints. In addition to this resolution uncertainty, there are uncertainties *even at the gridpoints*, due to measurement error, modelling (proxy) limitations, and intermittent coverage (e.g., spatio-temporally varying cloud cover when taking satellite measurements). Given this, computed Lagrangian trajectories will certainly inherit these errors.

While earlier work in fluid mechanics typically used Eulerian measurements to (incorrectly) infer transport, seminal work by Haller & Yuan [16] highlighted the necessity of understanding Lagrangian trajectories instead. One aspect is to try to find *coherent* entities which, while moving in time as is inevitable in an unsteady velocity field, retain aspects which might be deemed ‘coherent.’ Some examples of this could be the Gulf stream, an oceanic eddy, a moving hurricane, the Antarctic circumpolar vortex (‘ozone hole’), Jupiter’s Great Red Spot, a blob of fluid in a mixing device which resists mixing with its surroundings, and so on. There is considerable debate as to what ‘coherent’ might mean, as is evidenced by the multitude of diagnostic methods which have—and continue to be—developed in this endeavor. In general nomenclature, these have come to be known as ‘Lagrangian Coherent Structure’ (LCS) methods, for which several recent reviews are available [8, 13, 22, 24] (note that Haller uses the acronym LCS only for a specific class of methods [14]; more generality will be assumed here). As a general rule, coherent structures are detected from LCS methods based on numerical computations from the velocity field, treated as if it had *no* uncertainty whatsoever. The spatial resolution issue is dealt with in the natural way by interpolation between gridpoints, thereby implicitly using a *deterministic* and (sufficiently) *smooth* velocity across the entire domain.

Recently, the issues of uncertainty is coming to the fore. In climate and weather prediction, unresolved subgrid-scale uncertainties are sometimes referred to as the ‘stochastic parametrization problem’ (typically in these situations, the emphasis is not on tracking coherent structures but in determining the impact of this lack of resolution on global properties such as atmospheric carbon-dioxide). From the Lagrangian perspective, there have been some recent studies which examine the impact of uncertainty in some specific ways [3, 7, 11]. Moreover, there has been an increased interest in understanding *diffusion* in the context of fluid flows [15, 18]; this is related because diffusion can be thought of as the macroscopic impact of trajectory uncertainties at the microscopic level.

For an unsteady flow in \mathbb{R}^n (typically, $n = 2$ or 3) from time 0 to T ,

purely deterministic trajectories solve

$$\frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}, t), \quad \text{also written as} \quad d\mathbf{x}_t = \mathbf{u}(\mathbf{x}_t, t) dt. \quad (1)$$

In determining the trajectories \mathbf{x}_t numerically, since \mathbf{u} is only available on a spatio-temporal grid, it is necessary to interpolate \mathbf{u} ; this artificially extends u beyond the gridpoints, while imposing smoothness. A model for including uncertainties in the velocity field would be to consider instead

$$d\mathbf{y}_t = \mathbf{u}(\mathbf{y}_t, t) dt + \varepsilon d\mathbf{W}_t \quad (2)$$

in which \mathbf{y}_t is the random trajectory, where now $\varepsilon d\mathbf{W}_t$ gives the uncertainty in location over an infinitesimal time dt , due to n -dimensional Brownian motion \mathbf{W}_t . As necessary in considering the noise as a perturbative effect on (1), it is required that $|\varepsilon|$ be small in some appropriate sense. A more general formulation which allows for the uncertainties to be spatio-temporally dependent would be to replace $\varepsilon d\mathbf{W}_t$ with $\varepsilon\sigma(\mathbf{y}_t, t)d\mathbf{W}_t$, where σ is an $n \times n$ matrix [3, 7]; its (\mathbf{y}_t, t) -dependence allows, for example, modeling of cloud cover, or certainty at gridpoints. Given a stochastic model such as (2), along with spatio-temporal data \mathbf{u} on a grid, it is a straightforward exercise to evolve trajectories \mathbf{y}_t using, for example, the Euler-Maruyama scheme [19]. For each initial condition, this generates *one realization* of the stochasticity. By taking all initial conditions (consisting of a finite number of discrete points in any numerical situation), one can then pretend that this generates a flow map from time 0 to T , which is the basic building block used in most LCS methods. Thus, various LCS techniques, such as the Finite-Time Lyapunov Exponent (FTLE) can easily be applied to this map [12, 9]. Typically, a function f on the space of initial conditions \mathbf{x}_0 is generated by an LCS method, and based on its structure, coherent structures are identified. In the case of the FTLE, f would be the stretching field: if an infinitesimal n -dimensional sphere were positioned at the initial location \mathbf{x}_0 , it would get mapped to an ellipsoid under the flow, and f is the ratio of the ellipsoid’s longest axis to the original diameter. Indeed in this case, f is equal to the operator norm of the gradient of the flow map, thereby illustrating that f would be in general derived from the flow map in some way. $(n - 1)$ -dimensional ridges of this stretching field (usually scaled using a logarithm as befitting the ‘exponent’ part of the definition, and divided by the time-of-flow T) are important flow separators between coherent structures. Other LCS methods similarly use different diagnostic fields f , and interpret them in diverse ways such as using zero contours of f to define a region (see [8] for more detailed descriptions of many methods).

In doing this for one implementation of (2), it should be noted that the map is random (more properly, there is a different random dynamical system that each initial condition is subject to). Thus computing an f for one such map cannot give us uncertainty information in general. The solution, then, would be to perform this numerous times—thereby using many realizations of the stochasticity—and compute f for each such instance. A function \bar{f} , which is f averaged over all implementations, can then be developed. When using \bar{f} as a diagnostic method instead of f , alternative sets associated with LCSs are obtained: more diffused ridges of the FTLE field [12, 6] or of the stable/unstable manifolds [7].

The question, though, is how an *uncertainty* can be assigned to any features obtained this way. As a computational tool, the standard deviation (with respect to implementations) of f is the obvious way to proceed. Any such spatially-varying uncertainty field shall be denoted s . At each initial location \mathbf{x}_0 , $s(\mathbf{x}_0)$ will provide an indicative value of the uncertainty of $\bar{f}(\mathbf{x}_0)$. Put another way, the true diagnostic field $f(\mathbf{x}_0)$ could then be considered as

$$f(\mathbf{x}_0) = \bar{f}(\mathbf{x}_0) \pm s(\mathbf{x}_0) \quad \text{or} \quad f(\mathbf{x}_0) \in [\bar{f}(\mathbf{x}_0) - s(\mathbf{x}_0), \bar{f}(\mathbf{x}_0) + s(\mathbf{x}_0)] . \quad (3)$$

While this process will numerically generate heuristic, spatially-dependent errors, it remains unsatisfying and computationally expensive. Is it possible to find errors in *theoretical* ways? A recent article [5] does just this, though it is confined to two-dimensional flows. For the computation of any diagnostic function f , the flow map from time 0 to T is required. This flow map acquires an uncertainty, and determining a quantifier for this is therefore necessary for almost any LCS method one considers. For the more general spatio-temporally modulated noise $\varepsilon\sigma(\mathbf{y}_t, t) d\mathbf{W}_t$ situation, the theory in [5] concerns the scaled deviation of the random trajectory of (2) from that of (1) by time T , if they began at the same initial condition \mathbf{x}_0 , i.e.,

$$\mathbf{z}(\mathbf{x}_0) := \frac{\mathbf{y}_T - \mathbf{x}_T}{\varepsilon} \quad \text{where} \quad \mathbf{x}_0 = \mathbf{y}_0 . \quad (4)$$

Under some natural conditions, the expectation $\mathbb{E}(\mathbf{z}(\mathbf{x}_0))$ is shown to be zero for all relevant \mathbf{x}_0 in the limit as $\varepsilon \rightarrow 0$ [5], indicating that to leading-order the expected final location due to (2) is precisely the deterministic one. Here, the expectation is with respect to many realizations of (2). A theoretical expression for the uncertainty is suggested as the *stochastic sensitivity* field S^2 given by [5]

$$S^2(\mathbf{x}_0) := \lim_{\varepsilon \rightarrow 0} \sup_{\theta} \mathbb{V}(\mathbf{z}(\mathbf{x}_0) \cdot \hat{\mathbf{n}}_{\theta}) , \quad (5)$$

where \mathbb{V} is the variance with respect to many realizations of (2), and $\hat{\mathbf{n}}_\theta = (\cos \theta \ \sin \theta)^\top$ is a unit vector pointing in the direction $\theta \in [-\pi/2, \pi/2)$. Thus, the variance is computed for a projected direction θ , and then optimized over all directions in defining S^2 . This can be expressed in terms of trajectories and the evolution of the flow map gradient of the deterministic system (1), as well as the matrix σ , as is given in Appendix A. Once S^2 is known, it is noted that the lengthscale of the uncertainty is to leading-order

$$L(\mathbf{x}_0) = \varepsilon \sqrt{S^2(\mathbf{x}_0)}. \quad (6)$$

If $d\mathbf{W}_t$ is thought of in the standard sense ($d\mathbf{W}_t \sim \sqrt{dt}$ is assumed in terms of units) in (2), a natural way to think of ε is that

$$\varepsilon = \sqrt{v_r L_r} \quad (7)$$

where L_r and v_r are resolution scales of the length (i.e., the grid spacing) and velocity respectively. If so, (6) gives $L(\mathbf{x}_0) = \sqrt{v_r L_r S^2(\mathbf{x}_0)}$, enabling the computation of the lengthscale uncertainty of a trajectory beginning at \mathbf{x}_0 , in terms of scales in the problem. (In this interpretation, S^2 has dimensions of time, and \mathbf{z} in (4) has dimensions of time^{-1/2}.)

Having this information on Lagrangian uncertainty can be invaluable in evaluating uncertainties of various LCS measures. The implications on the FTLE field have already been determined, and are presented in another recent article [6]. There are unexpected implications—validated via stochastic simulations as well—that the error increases when the spatial resolution scale L_r is decreased.

An alternative—but connected—way to understand uncertainties in Lagrangian locations is through *diffusion*. Equations such as (2) are directly connected to a Fokker-Planck equation, which governs how densities $\rho(x, t)$ of trajectories evolve. The Brownian term in (2) contributes to a standard diffusive (Laplacian) operator on ρ (it is more complicated if the spatio-temporally modulated form is considered instead). The Fokker-Planck equation is thus an advection-diffusion equation for ρ ; indeed, precisely so if u is area-preserving and σ is the identity. Because of this reason or otherwise, there have been several recent approaches [10, 15, 18] to uncertainty by using an advection-diffusion partial differential equation rather than the stochastic ordinary differential equation (2).

Thus theoretical methods for uncertainty quantification of Lagrangian trajectories is an emerging area, for which the methods mentioned here may set the foundations. In the remainder of this article, some examples related to the discussion above will be presented.

Figure 1: Movie of the RAM flow protocol: (left) deterministic, and (right) stochastic with $\varepsilon = 0.01$. *[provided as a movie]*

2 Rotating Arc Mixer (RAM)

The Rotating Arc Mixer (RAM) is an industrial mixing device, where a thin layer of fluid lies within a circle, and where the circular boundary has a number of arcs which can be moved independently [20, 21]. A three-dimensional version of this has the circle extending outwards to form a cylinder, with the additional potential for the arcs to be in different locations as one proceeds in the extended direction [23]. The idea is to use arc-activation in a sequential way, usually to maximize mixing of fluid within the device. For the standard two-dimensional situation with $\mathbf{x} = (x, y)$, there is an *exact* expression for the velocity within the circular domain $x^2 + y^2 < 1$ when *one* arc is activated [17]. This explicit solution is derived using Stokes’ flow with appropriate boundary conditions, and repeated activation of different arcs can then be implemented by combining these solutions appropriately. The presence of chaotic mixing in the RAM is well-established [21, 23]; hence it is an example in which laminar flow can be used to generate significant mixing.

In the simulations to be presented, a RAM with three equally spaced arcs, each subtending an angle of $\pi/4$, is considered. The flow protocol is that the arcs are rotated counterclockwise with angular frequency 1, in a sequential order which also goes counterclockwise. Four complete cycles of this procedure constitutes the full flow. Fig. 1 shows on the left the impact of the flow on a scatter of fluid particles. The right figure, for comparison, is a stochastically perturbed version; specifically one realization of (2) with $\varepsilon = 0.01$. Here, each particle is subject to its own stochastic perturbation (as opposed to the approach in ‘random dynamical systems,’ in which one stochastic realization affects all initial conditions equally [2]). The arcs are shown in blue, and when active, an arc is highlighted in magenta. Particles near an active arc get pushed strongly counterclockwise, and the impacts are also felt (more weakly) in the interior regions. The RAM system is highly sensitive at the points where the moving arcs end, because the angular velocity on the circular boundary at these points has a discontinuity during the instances in which that particular arc is active: it is 1 on the arc, and 0 just off it. While there is no discontinuity *within* the fluid region, the intense gradients near to these points leads to high uncertainty. The deterministic simulation gives evidence of this; particles apparently ‘disappear’ because in

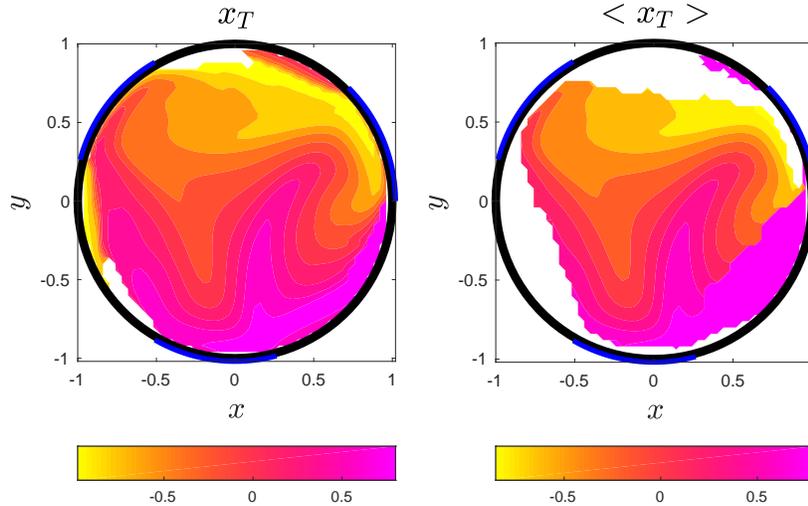


Figure 2: Final x -locations of the particles, shown as a contour plot over initial conditions: (left) deterministic, and (right) average of the stochastic simulations with $N = 500$.

the advection they are sometimes transported across the flow boundary. The stochastic simulation, in which additional velocity uncertainties are imposed on top of the high-gradient issue, is understandably affected substantially more by this issue than the deterministic case.

The expectation is that if the final location is averaged over many stochastic simulations, the deterministic final location is obtained. This is investigated in Fig. 2, where the final x -locations are averaged from the $N = 500$ simulations, and illustrated here as a contour plot as a function of the initial locations \mathbf{x}_0 . Thus, the value of x_T is the indicative f function considered in this instance. There is clear spatial correlation with x_T of the deterministic advection (shown on the left). The blank areas in this figure correspond to particles for which information has been lost due to the effect described in the previous paragraph. This effect can be reduced by choosing more refined grids locally and—in the case of the stochastic simulations—by performing

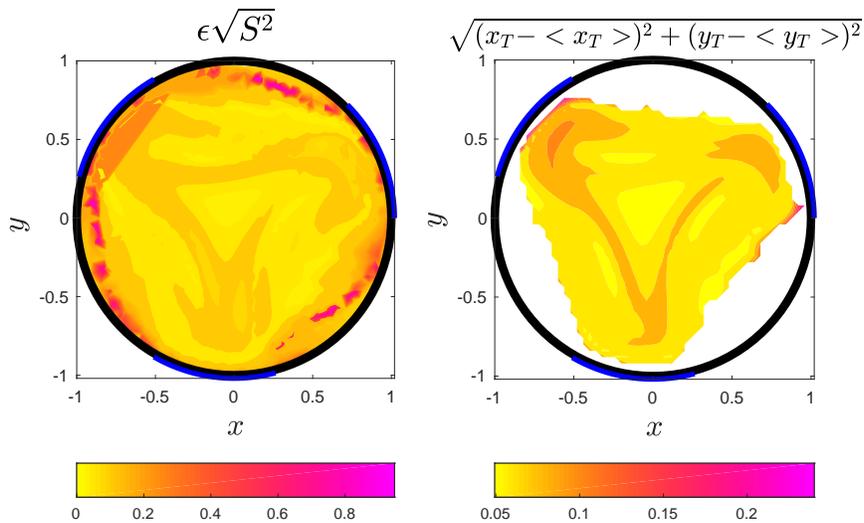


Figure 3: Uncertainty in final location of the RAM flow, expressed as a lengthscale: (left) using stochastic sensitivity theory (5), and (right) using a distance measure from the deterministic location, based on $N = 500$ stochastic simulations.

many more stochastic realizations.

The stochastic sensitivity (5) is a measure of the uncertainty in the final location. More specifically, it can be converted to a lengthscale according to (6). By computing S^2 using the expressions and methodology outlined in [5] (and summarized in Appendix A), the lengthscale of the uncertainty in the final location can be computed. This is displayed in the left figure of Fig. 3. Strong uncertainty ‘ridges’ are seen to emanate from the ends of the arcs, which is expected because of prior discussions. There is a nonuniform distribution of the uncertainty (which is related to the presence of chaotic regions and islands in the flow; however, these concepts only make sense in reality for *infinitely many* iterations of the flow). The figure on the right shows a proxy for the uncertainty in final location as the distance from the deterministic to the averaged stochastic final location, based on the 500 stochastic

Figure 4: Movie of the evolution of the FTLE field (in units of day^{-1}) in the northern Atlantic. A time-of-flow of $T = 20$ days is used throughout, and the FTLE field is at t_0 (the number of days from 1 June 2008) in each frame of the movie. *[provided as a movie]*

simulations. While the two concepts are not identical (compare with the definition (5)), it is clear that the spatial structures of the uncertainties are somewhat correlated. The two figures in Fig. 3 thus provide a theoretical and a numerical estimate for the Lagrangian uncertainty, as fields $s(\mathbf{x}_0)$.

This example has been chosen to highlight how stiff velocity gradients can cause considerable difficulties, as evidenced here by ‘white’ regions in the plots where information is unavailable. In this case, this generates uncertainties such that even deterministic advection—with an explicitly known velocity field—leads to problems. Basically, numerical uncertainties in integration are enough to generate this issue. When performing stochastic simulations, in which the velocity uncertainties are more explicitly included in the numerics, computations are affected in even larger regions. Most examples will not have as extreme a velocity gradient as this, and in such cases computations can often be performed without such clear evidence of them being wrong. Nevertheless, they *will* carry some errors. The stochastic approaches described in this article can help access information on these.

3 Oceanic velocities

Next, oceanic velocity data (Unidata Dataset Discovery V1.0) from the European Union’s Copernicus Marine Environment Monitoring Service is briefly analyzed. This dataset has zonal and meridional geostrophic velocities (on the ocean surface), derived from multiple satellite measurements of sea-surface height of the global ocean. In this case, daily data from from 1 June 2008 for a 3380 day period is available for the region in the northern Atlantic bounded by longitudes 77.125°W to 58.875°W and latitudes 30.125°N to 50.125°N . (Note: x will be represented as longitude eastwards, and y as latitude northwards; time will be measured in days and distances in degrees.) Gridpoints are separated by 0.25 degrees in both directions, and hence the resolution lengthscale is $L_r = 0.25$ deg.

A particularly significant issue in this dataset is to be noted. While oceanic water is incompressible, this is with respect to three-dimensional flow. Thus, the horizontal velocity components u (eastward) and v (north-

ward) of ocean surface velocities need not necessarily obey (two-dimensional) incompressibility. In other words, it is not necessarily true that $u_x + v_y = 0$, and indeed this data set indicates strong violations of two-dimensional incompressibility, particularly so when close to shorelines where strong downwelling may occur.

In this example, rather than viewing the final locations, a particular field associated with these locations will be examined. This is the Finite-Time Lyapunov Exponent (FTLE) field, which essentially measures the stretching of infinitesimal fluid parcels over a given time duration. More explicitly, this is given by

$$\text{FTLE}(\mathbf{x}_0) = \frac{1}{T} \ln \tau(\mathbf{x}_0) \quad (8)$$

where the stretching field τ intuitively positions a tiny disk (say diameter δ) at \mathbf{x}_0 , allows it to evolve with the deterministic flow (1) till time T to form an ellipse, and then takes the ratio of the longest extended axis of this ellipse with δ . Computationally, $\tau(\mathbf{x}_0)$ is exactly the operator norm of the gradient of the flow map from time 0 to T , and thus after advecting particles using (1), numerical differentiations for estimating flow map derivatives are necessary. The logarithm and then division by the time-of-flow in (8) converts the FTLE to an exponential rate value, in units of inverse time. It is common in oceanography to use the Finite-Size Lyapunov Exponent (FSLE) instead of the FTLE; this uses the same formula as (8), but stops the calculation once the stretching has gone beyond a specified threshold and then uses the T at that instance in the formula. The FTLE is preferred from the dynamical systems perspective, because it ensures that all initial conditions are subject to the identical dynamical system (namely, (1) flowing from time 0 to T), which is not true for an FSLE computation. In any case, FTLEs/FSLEs are probably the most commonly used diagnostic technique for evaluating regions of coherence.

Thus, the FTLE on 1 June 2008 (the first day of data availability), if considering the flow for time $T = 20$ days into the future, will give a field. This will be indexed by stating that the initial day is $t_0 = 0$. Similarly, it is possible to compute an FTLE field on 2 June 2008, again by considering a flow over a time of $T = 20$ days, but this will then be indexed by $t_0 = 1$. In general, the time t_0 will be taken to be the number of days after 1 June 2008. With this understanding, a movie of the evolution of the FTLE field as t_0 changes (but the time-of-flow remains at 20 in all cases), is shown in Fig. 4. A two-year period is shown in the movie, and the white region is of course the northern American continental landmass. Large FTLE values indicate regions which will stretch considerably, while low FTLE value regions will

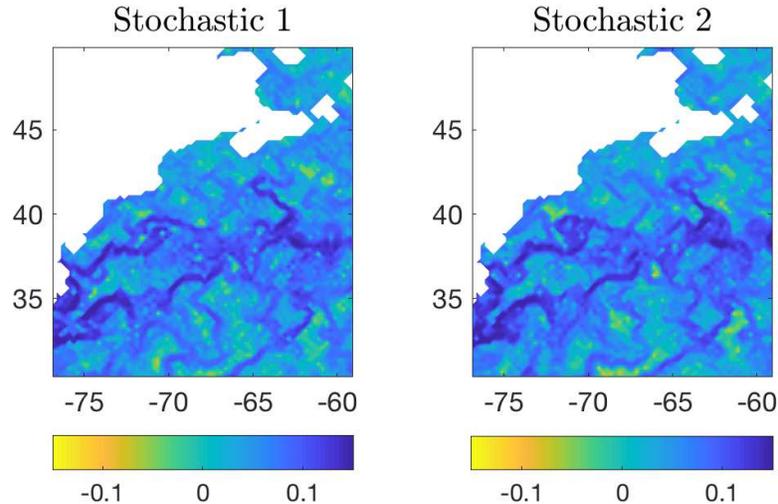


Figure 5: FTLE computed from two stochastic simulations of the ocean data from 11 November 2013 for 20 days using $v_r = 0.02$ deg/day.

not. Hence, the usual argument goes, the former regions would be associated with high levels of flow/shear/mixing (e.g., the flanges of the Gulf Stream), while the latter regions have lower values (e.g., coherent eddies). It should be noted that these conclusions must be considered relative to the flow duration (20 days into the future) considered; choosing a different time-of-flow will give information on that particular time-window. (FTLE/FSLE fields such as these are used often to reach various conclusions about coherent regions of flows; ‘ridges’ have received significant focus in analogy to stable/unstable manifolds of infinite-time flows. However, stable and unstable manifolds cannot be theoretically defined in general for finite-time flows such as this.)

The velocities which are used as input here have significant uncertainties: spatial resolution, measurement error, modelling error (the geostrophic approximation is often used, and other processing also occurs). Thus, any computation of eventual fluid parcel location will have errors. Put another way, there are errors in the flow map derived from the Eulerian velocities.

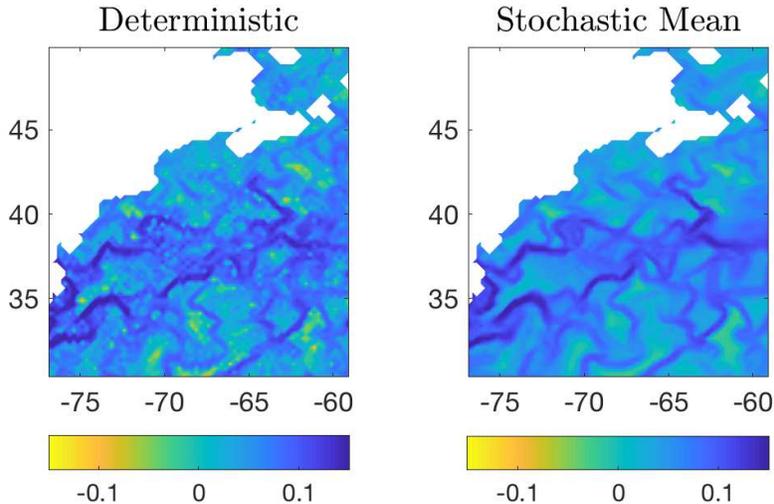


Figure 6: FTLE computed from the ocean data from 11 November 2013 for 20 days using $v_r = 0.02$ deg/day: (left) deterministic, and (right) average of 200 stochastic simulations.

Since this flow map is often used in computing various diagnostics f of the flow (case in point: the FTLE field), it is instructive to try to put some error-bars on any fields f computed. For $f = \text{FTLE}$, there is recent work [6] which does exactly this, generating an error field $s(\mathbf{x}_0)$ which depends on estimates of the velocity uncertainty scale v_r and the spatial resolution scale L_r according to [6]

$$s(\mathbf{x}_0) = \frac{2\sqrt{v_r S^2(\mathbf{x}_0)}}{e^{\text{FTLE}(\mathbf{x}_0)|T|}\sqrt{L_r}|T|}. \quad (9)$$

$\text{FTLE}(\mathbf{x}_0)$ is the FTLE field computed from (8) assuming that no uncertainties are present. The stochastic sensitivity field S^2 enters into this because the uncertainty in the final location is what leads to uncertainties in the FTLE field. In the general sense, it is noted that $s \sim \sqrt{v_r/L_r}$ in terms of the two uncertainty scales.

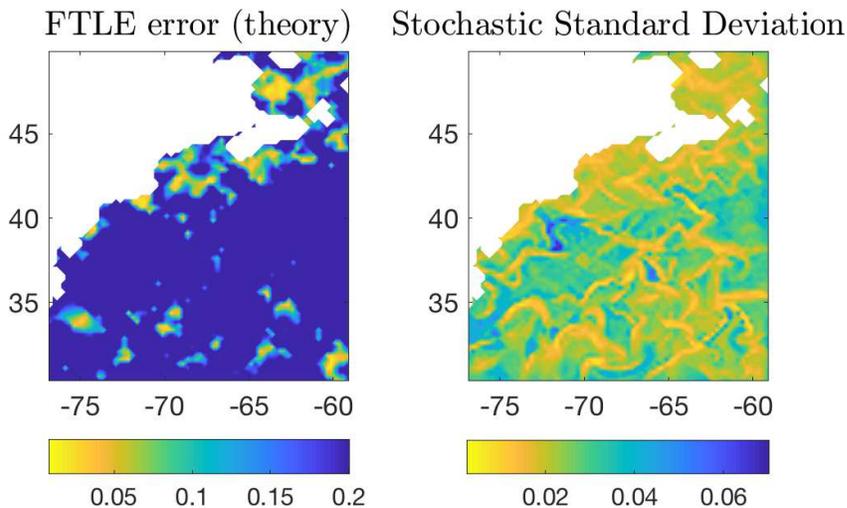


Figure 7: Error in FTLE computed from the ocean data from 11 November 2013 for 20 days using $v_r = 0.02$ deg/day: (left) theoretical bound using (9), and (right) standard deviation of FTLE from 200 stochastic simulations.

To examine the impact of uncertainty, let us focus on the FTLE field on 21 November 2013, using once again a 20-day period into the future. The resolution lengthscale $L_r = 0.25$ deg for this data set, and a velocity uncertainty of $v_r = 0.02$ deg/day will be assumed; this is very small considering that the velocities are typically of the order 1 deg/day. Thus from (7), $\varepsilon = 0.0707$ deg day $^{-1/2}$. Now, (2) can be implemented for this duration of time, and pretending that the final flow map is accurate, the FTLE field can be computed from (8). Fig. 5 displays the result of two such simulations. Some differences are to be observed, as expected. Thus, even with about 2% uncertainty in the velocity field, there can be errors in computed FTLEs. Next, following recent ideas that *averaging stochastic simulations* give a better approximation of the truth [12, 9, 6], the averaged FTLE field for 200 such simulations is shown on the right in Fig. 6. The left figure is of the FTLE field based on the *deterministic* data, pretending that there

are no uncertainties. On a gross level, the figures are very similar, indicating a level of robustness of the deterministic FTLE prediction to this level of noise. However, there are differences in the two figures in particular with respect to finer details. Clearly, the high-resolution detail from the deterministic FTLE field is unreasonable to use when uncertainties are considered. The averaged version of many stochastic realizations, on the other hand, smoothens out these finer (and unjustifiable) details.

Fig. 7 displays two different ways of attempting to capture the spatially-varying errors in the deterministically computed FTLE field. The left figure is computed using (9). Note its scale in comparison to Fig. 6. The error values are large, and a prime factor in this is that the data turns out to not obey area-preservation, presumably due to upwelling and downwelling. When computing S^2 (see equation (10) in Appendix A) for the usage of (9), the term $\nabla \cdot \mathbf{u}$ turns out to contribute significantly. The implications from the theory are that the deterministic FTLE field is therefore *highly* uncertain. However, it should be noted that the computation of this FTLE error field is difficult in this instance because of the irregular boundary (the coastline), which in particular renders the S^2 -field computation difficult. (Several levels of interpolation is needed; the errors that can arise in nonregularly scattered data points is well-known.) The right figure is obtained from the 200 stochastic simulations. The standard deviation (across the 200 values) is computed for the FTLE value at each point to generate this field. It should be noted that in performing stochastic simulations here, the simple strategy of excising particles when they cross to land has been done. If using sufficiently many stochastic simulations, there will be sufficient relevant data left, in order to do this calculations effectively. However, boundaries *are* a vexing problem when using stochastic simulations, since it is not always clear what the physically appropriate boundary conditions should be. Another difficulty with using the standard deviation of stochastic simulations is the computational cost: many simulations need to be performed. Theoretical estimates are therefore highly desirable. In any case, the yellow regions (lower theoretical errors) in the left figure do appear to be in regions where the standard derivations (right figure) are small. While both approaches (left and right) give methods for estimating errors in the FTLE field, there is clearly scope for improvement of these ideas because of issues such as boundary conditions, as well as improving computational efficiencies.

It should be noted that a conservative error scale for the velocity uncertainty was used here: $v_r = 0.02$ deg/day corresponds to an error of about 2% in the velocity measurements. The true errors are likely to be larger, which means that understanding the uncertainties in eventual predictions is

crucial.

4 Concluding remarks

This short article has highlighted the emerging issue that uncertainties in velocities are relevant when inferring Lagrangian motion. Of the two examples shown, the first was mainly concerned about the uncertainty in the final location, whereas the second focussed on a quantity *derived from the location*. Uncertainties play a role in either situation. Moreover, the first was related to a patented [20] industrial device, whereas the second was from observational data. It is for such instances—not merely in toy models—that knowing the uncertainties has a strong practical value.

In general, the main point made in this article is that there needs to be more ways of ascribing uncertainties to any predictions made using spatio-temporally available data. One approach—using stochasticity to model the uncertainty in the context of the governing ordinary differential equation—is the prime focus here, and in this case, glimpses of both numerical and theoretical methods have been provided. A closely related approach is to use the Fokker-Planck equation [10, 15, 4] corresponding to the stochastic ordinary differential equation (2); this encapsulates the macroscopic impact of the microscopic Brownian motions, though not per se capturing uncertainties in Lagrangian trajectories. A continuing development of uncertainty quantification in this realm—building possibly on the methods discussed here or those from different viewpoints [1, 11]—is essential in interpreting Lagrangian results correctly.

Acknowledgements: Partial support from the Australian Research Council via grant DP200101764 is gratefully acknowledged. I also express my thanks to Michel Speetjens for allowing me to mine his expertise in the RAM device, and to Irina Rypina for providing me with the Copernicus data.

A Stochastic sensitivity

This section gives the formula for the stochastic sensitivity field associated with the stochastic differential equation (2), as developed in [5]. The formula to be stated is only valid for two-dimensional situations. Extensions to three dimensions are in progress. Let \mathbf{F}_ξ^η be the flow map of the deterministic flow from time ξ to η . Thus, $\mathbf{x}_T = \mathbf{F}_0^T(\mathbf{x}_0)$ is the location to which the particle

beginning at \mathbf{x}_0 at time 0 gets mapped by the flow till time T . For $t \in [0, T]$, define the 2×2 matrix function

$$H(\mathbf{x}_T, t) := \exp \left[\int_t^T [\nabla \cdot \mathbf{u}] \left(\mathbf{F}_T^\xi(\mathbf{x}_T), \xi \right) d\xi \right] \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \nabla \mathbf{F}_T^t(\mathbf{x}_T), \quad (10)$$

where the gradient in the final term is with respect to locations \mathbf{x}_T . In terms of the components of H , let

$$\begin{aligned} P(\mathbf{x}_T) &= \left| \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^2 \int_0^T H_{ij}^2(\mathbf{x}_T, t) dt \right|, \\ L(\mathbf{x}_T) &= \frac{1}{2} \int_0^T \left[\sum_{i=1}^2 H_{i2}^2(\mathbf{x}_T, t) - \sum_{i=1}^2 H_{i1}^2(\mathbf{x}_T, t) \right] dt, \\ M(\mathbf{x}_T) &= \int_0^T \sum_{i=1}^2 [H_{i1}(\mathbf{x}_T, t) H_{i2}(\mathbf{x}_T, t)] dt, \quad \text{and} \\ N(\mathbf{x}_T) &= \sqrt{L^2(\mathbf{x}_T) + M^2(\mathbf{x}_T)}. \end{aligned}$$

Each of these quantities can be computed quite easily for a given final location \mathbf{x}_T . The complication is that the flow map gradient needs to be evaluated along backwards trajectories starting at \mathbf{x}_T , until they reach \mathbf{x}_0 . At each instance in time, the gradient can be evaluated in the standard way (e.g., as is normally done in finite-time Lyapunov exponent computations). Now, the stochastic sensitivity—as defined in (5)—at \mathbf{x}_0 is then given by [5]

$$S^2(\mathbf{x}_0) = P(\mathbf{x}_T) + N(\mathbf{x}_T). \quad (11)$$

A field across all \mathbf{x}_0 can thus be constructed numerically. The proof of this result is based in stochastic calculus methods (Itô lemma and isometry), coupled with some techniques more familiar to classical ordinary differential equations [5]. Moreover, a more general version associated with the spatio-temporally varying stochasticity $\varepsilon\sigma(\mathbf{y}_t, t)d\mathbf{W}_t$ is also available [5]. Numerically applying this technique does have its difficulties, because one needs to map the values of S^2 (which is computed at \mathbf{x}_T by this process) to $\mathbf{x}_0 = \mathbf{F}_T^0(\mathbf{x}_T)$ to generate a field at the time instance 0. If a regular grid were chosen for the distribution of \mathbf{x}_T , since the \mathbf{x}_0 s will not be regularly distributed, it will be necessary to extend S^2 from these scattered points to the relevant domain. Depending on the scatter, this process can be error-prone (see [5] for more details).

References

- [1] F. Antown, D. Dagićević, and G. Froyland. Optimal linear responses for Markov chains and stochastically perturbed dynamical systems. *J. Statis. Phys.*, 170:1051–1087, 2018.
- [2] L. Arnold. *Random Dynamical Systems*. Springer-Verlag, Berlin, 1998.
- [3] S. Balasuriya. Stochastic uncertainty of advected curves in finite-time unsteady flows. *Phys. Rev. E*, 95:062201, 2017.
- [4] S. Balasuriya. Stochastic approaches to Lagrangian coherent structures. *Adv. Stud. Pure. Math.*, accepted, 2019.
- [5] S. Balasuriya. Stochastic sensitivity: a computable uncertainty measure for unsteady flows. *SIAM Review*, accepted, 2020.
- [6] S. Balasuriya. Uncertainty in finite-time Lyapunov exponent calculations. *J. Comp. Dyn.*, under review, 2020.
- [7] S. Balasuriya and G. Gottwald. Estimating stable and unstable sets and their role as transport barriers in stochastic flows. *Phys. Rev. E*, 98:013106, 2018.
- [8] S. Balasuriya, N. Ouellette, and I. Rypina. Generalized Lagrangian coherent structures. *Physica D*, 372:31–51, 2018.
- [9] F. Balibrea-Iniesta, C. Lopesino, S. Wiggins, and A. Mancho. Lagrangian descriptors for stochastic differential equations: a tool for revealing the phase portrait of stochastic dynamical systems. *Intern. J. Bifurc. Chaos*, 26:1630036, 2016.
- [10] A. Denner, O. Junge, and D. Matthes. Computing coherent sets using the Fokker–Planck equation. *J. Comp. Dyn.*, 3:163–177, 2016.
- [11] G. Froyland, C. González-Tokman, and A. Quas. Stochastic stability of Lyapunov exponents and Oseledets splitting for semi-invertible matrix cocycles. *Commun. Pure Appl. Math.*, 68:2052–2081, 2015.
- [12] H. Guo, W. He, T. Peterka, H.-W. Shen, S. Collis, and J. Helmus. Finite-time Lyapunov exponents and Lagrangian coherent structures in uncertain unsteady flows. *IEEE Trans. Visual. Comp. Graphics*, 22:1672–1682, 2016.

- [13] A. Hadjighasem, M. Farazmand, D. Blazeovski, G. Froyland, and G. Haller. A critical comparison of Lagrangian methods for coherent structure detection. *Chaos*, 27:053104, 2017.
- [14] G. Haller. Lagrangian Coherent Structures. *Annu. Rev. Fluid Mech.*, 47:137–162, 2015.
- [15] G. Haller, D. Karrasch, and F. Kogelbauer. Material barriers to diffusive and stochastic transport. *Proc. Nat. Acad. Sci.*, 115/37:9074–9079, 2018.
- [16] G. Haller and G.-C. Yuan. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Phys. D*, 147:352–370, 2000.
- [17] T.-Y. Hwu, D.-L. Young, and Y.-Y. Chen. Chaotic advection for Stokes flow in circular cavity. *J. Engin. Mech.*, 123:774–782, 1997.
- [18] D. Karrasch and J. Keller. A geometric heat-flow theory of Lagrangian coherent structures. *arXiv:1608.05598*, under review, 2018.
- [19] P. Kloeden, E. Platen, and H. Schurz. *Numerical solution of SDE through computer experiments*. Universitext. Springer-Verlag, Berlin, 1994.
- [20] G. Metcalfe and M. Rudman. Fluid mixer using viscous drag. *U.S. Patent*, 7,121,714, 2006.
- [21] G. Metcalfe, M. Rudman, A. Bryden, L. Graham, and R. Hamilton. Composing chaos: An experimental and numerical study of an open duct mixing flow. *AIChE J.*, 52:9–28, 2006.
- [22] S. Shadden. Lagrangian coherent structures. In R. Grigoriev, editor, *Transport and mixing in laminar flows: from microfluidics to oceanic currents*. Wiley, 2011.
- [23] M. Speetjens, E. Demissie, G. Metcalfe, and H. Clercx. Lagrangian transport characteristics of a class of three-dimensional inline mixing flows with fluid inertia. *Phys. Fluids*, 26:11360, 2014.
- [24] M. Speetjens, G. Metcalfe, and M. Rudman. Lagrangian transport and chaotic advection in three-dimensional laminar flows. *arXiv:1904.07580*, under review, 2019.