Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

Atmospheric Lagrangian coherent structures considering unresolved turbulence and forecast uncertainty

CrossMark

Amir E. BozorgMagham *, Shane D. Ross

Engineering Science and Mechanics, Virginia Tech, Blacksburg, VA 24061, USA

ABSTRACT

To obtain more realistic approximations of atmospheric Lagrangian coherent structures, the material surfaces which form a template for the Lagrangian transport, two concepts are considered. First, the effect of unresolved turbulent motion due to finite spatiotemporal resolution of velocity field data is studied and the resulting qualitative changes on the FTLE field and LCSs are observed. Stochastic simulations show that these changes depend on the probabilistic distribution of position of released virtual particles after backward or forward time integration. We find that even with diffusion included, the LCSs play a role in structuring and bifurcating the probability distribution. Second, the uncertainty of the forecast FTLE fields is analyzed using ensemble forecasting. Unavoidable errors of the forecast velocity data due to the chaotic dynamics of the atmosphere is the salient reason for errors of the flow maps from which the FTLE fields are determined. The common practice for uncertainty analysis is to apply ensemble forecasting and here this approach is extended to FTLE field calculations. Previous work has shown an association between LCS passage and fluctuations in microbial populations and we find that ensemble FTLE forecasts are sufficient to predict such passages one day ahead of time with an accuracy of about 2 h.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The notion of hyperbolic Lagrangian coherent structures (LCSs) provide a framework for understanding transport and mixing phenomena especially in the case of passive particles in fluid systems [1–3]. These structures are codimension 1 manifolds (or material surfaces) which effectively separate two regions of fluid with different qualitatively different past histories or fates. Several such critical material surfaces are present in any given geophysical flow and may aid in understanding Lagrangian transport patterns. In the atmosphere, they persist from a few hours to a few days.

The present study is motivated by the role of atmospheric LCSs in aerial transport of microorganisms and the statistical correlation between sudden changes in aerobiota density and passage of LCS features over a fixed location [4–10]. Considering that role, if one can predict the LCSs to a good degree of accuracy and reliability, then important knowledge about the front propagation of microorganisms would be available.

Considering this fact, in a previous study, forecast LCSs were compared with reanalysis (pastcast) results to ascertain the accuracy and reliability of the forecasts [11,12]. Based on that study we infer the need for including more realistic considerations of atmospheric flows such as (i) finite spatiotemporal resolution of the available fluid velocity data, (ii) the accuracy of the fluid velocity data and (iii) the uncertainty of the forecast velocity fields. There have been studies about each of these

---

* Corresponding author.
  E-mail address: aebm@vt.edu (A.E. BozorgMagham).

concerns. For example, Griffa et al. [13] and Ozgokmen et al. [14] studied the predictability of Lagrangian trajectories and Kahl and Samson [15] considered the uncertainty in trajectory calculations due to low resolution data. In addition Wilson and Sawford [16] and Stohl [17] studied different models for generating Lagrangian stochastic trajectories and Palmer [18] and Ehrendorfer [19] reviewed the concepts of predictability and uncertainty of atmospheric forecasts. Also Kalnay [20] represented various approaches for data assimilation, ensemble forecasting and uncertainty analysis. Regarding LCSs, Haller [21] discussed the errors of approximate velocity field and their effect on hyperbolic LCS features and Lermusiaux et al. [22] described the uncertainty of oceanic LCSs and their numerical studies indicated that the more intense FTLE ridges are usually more certain. Moreover, Olcay et al. [23] studied the role of flow field resolution and random errors on LCS identification. Finally, Turbulent dispersion velocity at length scales smaller than wind data grid was considered by Peng and Peterson [24] and it was shown that attracting LCS structures coincide with the regions of high particle concentration. In that study, finite time Lyapunov exponent (FTLE) fields and the associated LCSs were calculated by using deterministic atmospheric flow map and the volcanic ash particles' dispersion was calculated by adding random walk and deposition velocities to the background velocity field.

In this study we connect notions such as unresolved turbulence and uncertainty of flow field to the atmospheric LCSs. For this aim we consider two concepts. First we use a Lagrangian particle dispersion model (LPDM) which approximately represents the effects of unresolved turbulence. Accordingly, we study the FTLE scalar field and the resultant LCSs in the presence of stochastic diffusion by adding the stochastic component of displacement to the deterministic flow map (note, we will often refer to FTLE-derived LCSs as FTLE-LCSs). As a result, the trajectories will be stochastic and non-differentiable instead of deterministic and smooth. We observe significant changes in the (probabilistic) position of particles and the associated FTLE fields. We show that the spatiotemporal dependence of the stochastic velocity component to the time-varying deterministic background velocity field [25–27] plays an important role in the determination of the probabilistic distribution of the end-position of released virtual particles and consequently on the stochastic FTLE fields. We should note that Peng and Peterson [24] just considered the dispersion velocity for the ash particles and the FTLE-LCSs were calculated by deterministic velocity field.

Second, we study the quantification of uncertainty and reliability of the FTLE-LCS when forecast velocity data are used to generate flow maps. The effects of forecast velocity data on the accuracy of forecast FTLE-LCSs have previously been observed [11]. In that paper we compared FTLE-LCSs from forecast velocity data with archive based FTLE-LCSs. Results of that study show the sensitivity of FTLE-LCS forecasting to different parameters such forecast lead time, but because we used a deterministic forecast velocity field (NAM-218 data set) we were not able to measure the uncertainty of the forecast FTLE-LCS results. In the present study we propose a practical approach to measure uncertainty of forecast FTLE-LCSs by using ensemble forecasting concepts and considering the distribution of local FTLE values. In addition, we study the ensemble-based distribution of forecast LCS passage times over a fixed geographical location, which provides a measure of reliability of the LCS forecasts.

2. Effects of unresolved turbulence on the FTLE field; stochastic FTLE field

In this section we show how including the stochastic component of unresolved turbulent velocity changes the FTLE field.

In atmospheric applications, the spatial resolution of archived operational data could vary from the order of 10 km to more than 250 km and the temporal resolution is usually of the order of 3 to 12 h. For example, in the operational model NAM-218, which we use, the spatial resolution is about 12 km and the temporal resolution is 6 h (a short term forecast at each intermediate 3 h is also available). As a result of this coarse resolution, important *unresolved motions* with significant effects on the particle flow map and the resultant FTLE field could exist. To investigate this, we use a particle dispersion model to calculate the stochastic turbulent velocity component of the flow field. These components are functions of turbulent diffusivity [25,26].

Considering the Lagrangian frame, a fluid particle trajectory $\mathbf{x}(t)$ is governed by the stochastic differential equation,

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}[\mathbf{x}(t)], \tag{1}$$

where the velocity of the particle is composed of the deterministic grid scale velocity, $\bar{\mathbf{v}}(\mathbf{x}, t)$, and the stochastic turbulent fluctuation component, $\mathbf{v}_t(\mathbf{x}, \bar{\mathbf{v}}, t)$, respectively,

$$\mathbf{v}(\mathbf{x}, t) = \bar{\mathbf{v}}(\mathbf{x}, t) + \mathbf{v}_t(\mathbf{x}, \bar{\mathbf{v}}, t). \tag{2}$$

Using the Langevin equation and assuming a Markovian process [28,29], one can parameterize the turbulent component of velocity as,

$$dv_{t_i} = a_i(\mathbf{x}, \mathbf{v_t}, t)dt + \sum_{j=1}^{3} b_{ij}(\mathbf{x}, \mathbf{v_t}, t)dW_j, \tag{3}$$

where drift (a_i) and diffusion (b_{ij}) are functions of time, position and turbulent velocity and v_{t_i} represents the ith component of $\mathbf{v}_t(\mathbf{x}, \bar{\mathbf{v}}, t)$ [30,16,17,31]. Moreover, dW is the standard white noise Wiener process.

There are different methods for estimating the drift and diffusion terms with respect to the available data and turbulent regime of the flow field. In this study we follow the work of Thomson [30], Fay et al. [26], Legg and Raupach [27], Stohl et al. [31] and Draxler and Hess [25].

2.1. Grid scales velocity data

For generating the deterministic flow map we use numerical data provided by the National Oceanic and Atmospheric Administration (NOAA) and National Centers for Environmental Prediction's (NCEP) Operational Model Archive and Distribution System (NOMADS) project. We use one of the outputs of this nonhydrostatic mesoscale model, which is the North America Mesoscale, NAM-218, with data given on a grid of 614×428 points spaced at about 12.1 km covering North America. We consider a 6230×4670 km rectangular area as our wind velocity field and we use the smooth extension of the velocity field outside the boundaries whenever a particle leaves this domain [21].

Two types of data are used in our FTLE-LCSs calculations, archive data and forecast data. *Archive data* is the state of the atmospheric system determined by a combination of numerous measurements from, e.g., weather stations, weather balloons, satellites and any available atmospheric observations through data assimilation methods which use large-scale oceanic-atmospheric geophysical fluid models. By *forecast data*, we mean the output of geophysical models which use the most current outputs of the data assimilation process (archive data) as input to estimate the state of the system for (near-) future times [20,32,33]. In the NAM-218 data set, the temporal resolution of archived data is 6 h corresponding to 00:00, 06:00, 12:00 and 18:00 UTC of each day. Accordingly, the forecast chain can start based on each of these archive time snapshots. Temporal resolution for forecast data is 3 h and at each moment, the maximum available forecast data is for 84 h lead time, i.e., ahead of the most recent archive data.

In this paper we will show the results for 2D flows over constant pressure surfaces which could be a simplified representation of the full 3D cases. Isobaric calculation is common in Lagrangian studies and it generates a first approximation to trajectories while avoiding the complexity of three-dimensional flow field integrations [17,34,35]. In addition, these results can be regarded as examples of what is expected if more realistic conditions are considered for FTLE field computations. Finally, from Branicki et al. [36] and Sulman et al. [37] we know that 3D LCS features are nearly vertical within a layer whose thickness is sufficiently smaller than the ratio of the average horizontal velocities to their average vertical gradients. In that layer we may approximate the 3D structures with 2D features with a satisfactory accuracy. In our case of study that ratio can be estimated by the horizontal approximation of the wind velocity profile [38] showing the upper and lower limits of that ratio ($\bar{u}/(\overline{\frac{\partial u}{\partial z}})$) are between 250 and 1500 m. So, we may conclude that within a layer centered at 850 mb pressure surface, whose thickness is in the range of 25 to 150 m (10% of that ratio), the 2D LCS features are acceptable approximation of the true 3D structures. More careful investigation of this approximation is left for future work.

2.2. Unresolved turbulent velocity component

By considering essential hypotheses such as the Kolmogorov similarity hypothesis [39] and the criteria described by Thomson [30], various solutions have been suggested for the unresolved turbulent velocity term. We use a model which is implementable with respect to the available data [26,25]. This model is the same as the stochastic Lagrangian particle dispersion model which is implemented by the HYSPLIT (Hybrid Single Particle Lagrangian Integrated Trajectory) model[1] maintained by the NOAA Air Resources Laboratory (ARL). We note that there exist other models such as STILT model[2] which also consider the problem of Lagrangian particle dispersion model for their specific aims. We choose the HYSPLIT model since it is the most widely used for aerobiological applications [40–47]. Our goal is to show how Lagrangian structures obtained from HYSPLIT or other models might change if more realistic conditions such as unresolved turbulence (stochastic diffusion) are considered.

In our implemented method, components of $\mathbf{v}_t(\mathbf{x}, \bar{\mathbf{v}}, t)$, which are shown as v_{t_i}, are estimated by a Markov-chain formulation as,

$$v_{(t+\delta t)_i} = R_{\delta t} v_{t_i} + \left(1 - R_{\delta t}^2\right)^{0.5} \xi, \tag{4}$$

where $R_{\delta t}$ is the turbulent velocity Markov autocorrelation coefficient. This coefficient is approximated from the Lagrangian time scale of the flow system as,

$$R_{\delta t} = \exp\left(-\delta t / T_L\right), \tag{5}$$

where δt is the integration time step and the Lagrangian time scale (T_L) is considered to be constant over the pressure surface, assuming δt to be sufficiently smaller than T_L. In our calculations δt is between 5 to 15 min and T_L is 180 min according to Draxler and Hess [25]. T_L describes how long turbulent fluctuations are 'remembered' by the particle. The Lagrangian time scale (T_L) is estimated with respect to the averaged turbulent kinetic energy of the flow field, and is approximated by the length scale of eddies and their velocity scale [48].

For calculation of the second term of (4), only lateral diffusion within constant pressure surfaces is considered and density variations over those surfaces are neglected [49]. The ξ term is estimated with respect to the standard deviation of the turbulent velocity σ, as,

---

[1] http://ready.arl.noaa.gov/HYSPLIT.php
[2] http://www.stilt-model.org

$$\xi = \mathcal{N}(0,1)\sigma, \tag{6}$$

where $\mathcal{N}(0,1)$ is the standard normal distribution and σ at each point is estimated as,

$$\sigma = \sqrt{\kappa/T_L}, \tag{7}$$

where the sub-grid scale horizontal mixing coefficient, κ, is computed from the velocity deformation tensor as,

$$\kappa = 2^{-0.5}(c\chi)^2 \left[\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right)^2 + \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right)^2 \right]^{0.5}, \tag{8}$$

where $\bar{\mathbf{v}} = (u, v)$, χ is the meteorological data grid size and c is an empirical constant equal to 0.14 [25].

The set of Eqs. (1)–(8) provides all one needs to calculate the stochastic velocity and stochastic trajectory of particles with respect to the background 2D velocity field. This process can be repeated for 3D flows by considering the vertical component of the velocity [31] as well.

2.3. Stochastic FTLE field (SFTLE)

Similar to the deterministic case, the essential step for defining the stochastic FTLE (SFTLE) field is to define the stochastic flow map as,

$$\Phi_{t_0}^t : \mathbf{x}_0 \rightarrow \mathbf{X}(\mathbf{x}_0, t_0, t), \tag{9}$$

where conventionally deterministic (sure) values are written in lower case and random (stochastic) variables are represented by capital letters. Regarding the stochastic trajectory definition, the stochastic FTLE field is *defined* for the random variable $\Phi_{t_0}^t(\mathbf{x}_0)$ as,

$$\Sigma_{t_0}^t(\mathbf{x}_0) = \frac{1}{|t - t_0|} \log \left\| D\Phi_{t_0}^t(\mathbf{x}_0) \right\| = \frac{1}{|t - t_0|} \log \sqrt{\Lambda_{max} \left[\mathfrak{C}_{t_0}^t(\mathbf{x}_0) \right]}, \tag{10}$$

where the right Cauchy–Green strain tensor, $\mathfrak{C}_{t_0}^t(\mathbf{x}_0)$, and the deformation gradient (Jacobian), $D\Phi_{t_0}^t$, are defined with respect to the stochastic map $\Phi_{t_0}^t$ and consequently they are random variables. The notation $\Lambda_{max}[\cdot]$ denotes the maximum eigenvalue.

Assuming that the initial position of a particle, $\mathbf{x}_0$, is known with infinite precision, the main difference between deterministic and stochastic FTLE fields is that when the deterministic velocity field is considered, the source or destination (referring to backward and forward integration respectively) of a released particle would be a point, but when unresolved stochastic turbulence is considered that single point becomes a probabilistic source or destination distribution.

For calculation of the probabilistic distribution of the end-position of trajectories, we divide the domain into small square boxes $\{B_1, B_2, \ldots\}$ of side length d. Then by using the Monte Carlo method in forward or backward integration, the number of final landings in each box is counted. We re-index all the boxes by number of landings per box $\left\{ B^{(1)}, B^{(2)}, \ldots \right\}$ such that $n_{B^{(1)}} \geqslant n_{B^{(2)}} \geqslant \cdots$, where $n_{B^{(i)}}$ is the number of landings in box i. For $\alpha \in (0, 1)$, we define the probabilistic α-source (or destination) area as,

$$A_\alpha(\mathbf{x_0}, t_0, t) = \bigcup_{i=1}^k B^{(i)}, \quad \text{where} \quad k = \inf \left\{ j \in \mathbb{N} | \sum_{i=1}^j n_{B^{(j)}}/N \geqslant \alpha \right\}, \tag{11}$$

where N is the total number of released virtual particles from each reference point $\mathbf{x_0}$ (number of realizations of the Monte Carlo method). The interpretation is that $A_\alpha(\mathbf{x_0}, t_0, t)$ contains a fraction α of the ensemble of particles, starting with the most dense boxes. In other words, the probabilistic source or destination regions, $A_\alpha(\mathbf{x_0}, t_0, t)$, are represented by the collection of small boxes which contain α percent of the sampled or released virtual particles. Our simulations show that area($A_\alpha(\mathbf{x_0}, t_0, t)$) converges to a fixed number for each case by choosing sufficient number of realizations and considering small size for boxes (i.e., we have convergence as $d \rightarrow 0$). For computational purposes, we see reliable convergence for $d = 4$ km, and we use this value throughout the computations shown. We note this is smaller than the spatial resolution of our model (12 km).

For brevity, we will often refer to $A_\alpha(\mathbf{x_0}, t_0, t)$ as simply A_α, where the dependence on $\mathbf{x_0}$, t_0, and t is understood. It is sometimes more convenient to refer to the integration time $T = t - t_0$, in which case A_α can be considered dependent on $\mathbf{x_0}$, t_0, and T.

For $T < 0$ (backward integration) we call A_α the "probabilistic α-source area" and for $T > 0$ (forward integration), A_α is called the "probabilistic α-destination area". Note that for both cases, N particles are released from each grid point of the domain ($\mathbf{x_0}$) at an initial time t_0, and after time integration of duration T, the end positions are calculated.

Fig. 1 shows the backward solutions starting from the reference point (shown by the black circle) for two cases. The dark collection of points shows the probable source points for integration time $T = -12$ h and the cyan region shows the probabilistic region for $T = -24$ h. Notice that the distance between the deterministic solutions (red circles) and the centroid of the stochastic realizations (red diamonds) increases with time. Also, since the probable source areas are stretched, the deterministic solution or the centroid of the stochastic solutions are not good representatives for those regions. In Section 2.4, the qualitative effects of probable source regions on stochastic FTLE fields are discussed for different cases.
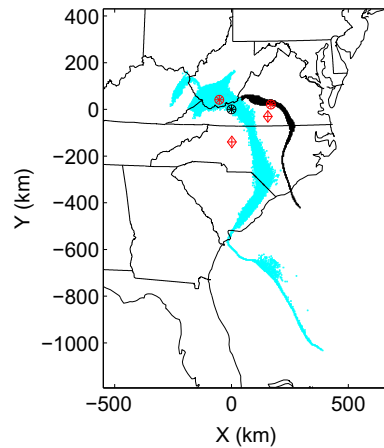
Fig. 1. Collections of the probabilistic source points ($A_{90\%}$), backward integration times $T = -12$ h (black) and -24 h (cyan), black circle: reference point at (37°11′N latitude and 80°35′W longitude), red circles: deterministic solutions, red diamonds: centroid of the stochastic solutions. For this case 10^6 virtual particles are released from the reference point on the 850 mb pressure surface at time $t_0 = 14$:37 UTC 29 Sep 2010. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In general, the final probabilistic distribution of arbitrary individual particles after integration time T are not Gaussian since drift and diffusion terms are functions of position, velocity and time. As a result, the deterministic final position of a particle after the integration time is offset from the centroid position of the stochastic realizations. In general, to find the exact probability distribution one needs to solve the Fokker–Planck equation, however as a practical approach one can use the Monte Carlo method to make an estimate of that distribution [50,51].

2.4. Results; stochastic FTLE fields

As a first case, we consider the backward FTLE field for 12:00 UTC 29 Sep 2010 over the eastern part of the United States. We consider quasi-2D flow on pressure surfaces and all the presented results correspond to the 850 mb surface.

In Fig. 2(a), the deterministic FTLE is shown and we see sharp ridges which are candidates for hyperbolic attracting LCS features. This FTLE field was generated with archived data for $T = -24$ h. One could extract the hyperbolic LCS features for the deterministic velocity field over the time-frame given and for domain shown, but our focus in this paper is on the comparison of FTLE and SFTLE fields (for more information about atmospheric hyperbolic LCSs, see [52,11]).

Fig. 2(b) shows *one particular* realization of the stochastic FTLE field when the stochastic component of the unresolved turbulent velocity field is taken into account. To generate this field we evolve a grid of initial positions, with added stochastic
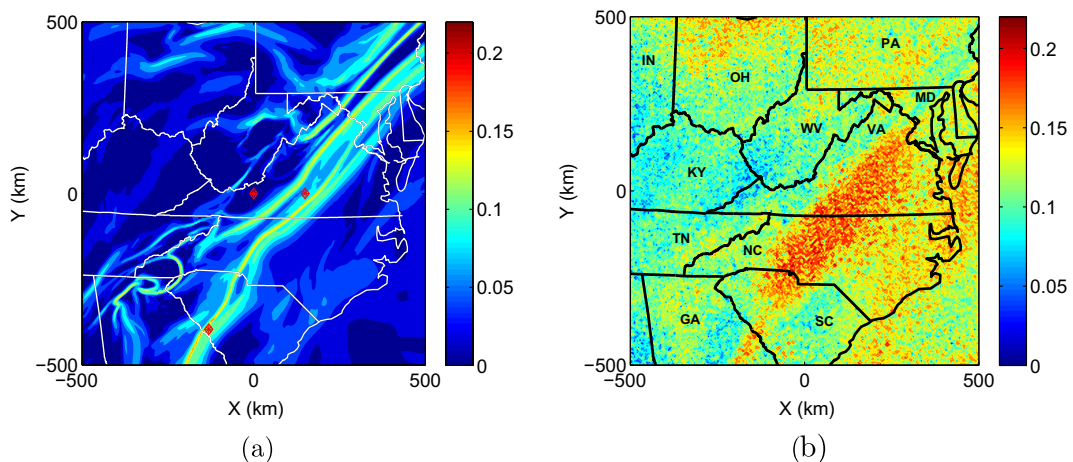

Fig. 2. A typical comparison between deterministic and stochastic FTLE field, (a) deterministic backward FTLE field (b) one realization of the stochastic backward FTLE field; integration time $T = -24$ h, $t_0 = 12$:00 UTC 29 Sep 2010, and the velocity field corresponds to the 850 mb pressure surface. Notice different regions comparing the two panels and also the higher level of FTLE coefficient in panel (b). Different qualitative changes at the location of marked points on panel (a) are discussed in Fig. 4. In (b), the states are labeled by their postal code for brevity.

velocity, calculate the Cauchy-Green tensor in the usual way, and then calculate the FTLE via (10). We notice different features comparing the two panels of Fig. 2. For example, the sharp ridges of the deterministic case (panel (a)) over VA, NC and some part of SC become a broad high value region in panel (b). Meanwhile, one can observe that sharp and strong deterministic ridges over PA and MD are washed out in panel (b) which suggests it is not necessary to have a broad and high value cloud of stochastic FTLE wherever deterministic high value ridges exist. This observation is important since it shows different behaviors when two similar cases—parallel and strong ridges of FTLE field—are considered. We can also see that the curly ridge over south west NC does not exist in panel (b). Note that the graininess of Fig. 2(b) is due to the inherent and discontinuous stochastic behavior.

To investigate the common patterns of the stochastic FTLE fields we use the expected value of those fields. Fig. 3(a) is the expected value of the stochastic FTLE fields for 1000 particular realizations obtained by the Monte Carlo method. As is common, the expected value smooths out the graininess of individual realizations. Fig. 3(a) shows an important point which is difficult to observe in Fig. 2(b). In this panel we observe a narrow cloud of high value SFTLE over GA at the same place where we have a strong deterministic ridge (in Fig. 2(a)). This observation shows that it is possible to have sharp deterministic FTLE ridges which remain as narrow high value regions of SFTLE; but not always.

Fig. 3(b) is the standard deviation of the stochastic FTLE fields. We observe that the value of standard deviation is much smaller than the expected values with a maximum order of 6%. This observation means that the stochastic realizations are not too "separated" from each other. This is an important remark since the numerical behavior of typical stochastic differential equations depends on their coefficients and by changing the coefficient one can theoretically get significantly different distributions, e.g., in Ornstein–Uhlenbeck processes [53]. Based on results corresponding to different days, we observe that the standard deviations in atmospheric SFTLEs are typically small comparing to their expected value.

To have a better understanding about different regional changes, we investigate the sensitivity of the probability distribution of forward and backward trajectories with respect to the reference point, i.e., the sensitivity of $A_\alpha(\mathbf{x_0}, t_0, T)$ with respect to $\mathbf{x_0}$. We observe that source (or destination) areas corresponding to different target (or starting) locations are qualitatively different. In those cases that the source (or destination) areas are small and close to the deterministic solutions, the changes in the *local* FTLE field are small, but in cases that the source (or destination) areas are large and/or the centroid of the source (or destination) areas are far from the deterministic solutions, local changes of FTLE fields are noticeable.

Fig. 4 shows three cases of the probabilistic 50%-source area (each sub-figure represents 10^6 stochastic trajectories obtained from the Monte Carlo method) corresponding to qualitative changes observed in Figs. 2(a) and 3(a). These cases correspond to different $\mathbf{x_0}$ with the same t_0. In the first case, Fig. 4(a), the source area is shown where $\mathbf{x_0}$ is located at Virginia Tech Kentland Farm ($37°11'$N, $-80°35'$W) which is considered as the reference point (0,0). Fig. 4(b) shows the probabilistic source area when $\mathbf{x_0}$ is placed at ($37°10'30''$N, $-78°55'48''$W) equal to 150 km eastward relative to the reference point, and the last case, Fig. 4(c), refers to the source area when $\mathbf{x_0}$ is at ($33°40'48''$N, $-82°$W) which corresponds to 131 km west and 396 km south respectively. For all three cases the backward integration time is $T = -24$ h. These three points and their corresponding probabilistic source areas illustrate the qualitative changes compared to the deterministic FTLE field when the unresolved turbulent velocity is added. In the first case the source area is small and compact. We observe the same qualitative behavior for the other nearby local points, i.e., points inside a circle of radius of order $\sim d$. So, we expect very small change in the local FTLE field around the point $(0, 0)$. The second case, Fig. 4(b), which is the corresponding probable source area for a point between two sharp ridges, where a band of high value SFTLE is formed, is large and stretched. And finally, the third source area refers to a region where the width of the sharp FTLE ridge is increased but not as much as the second case
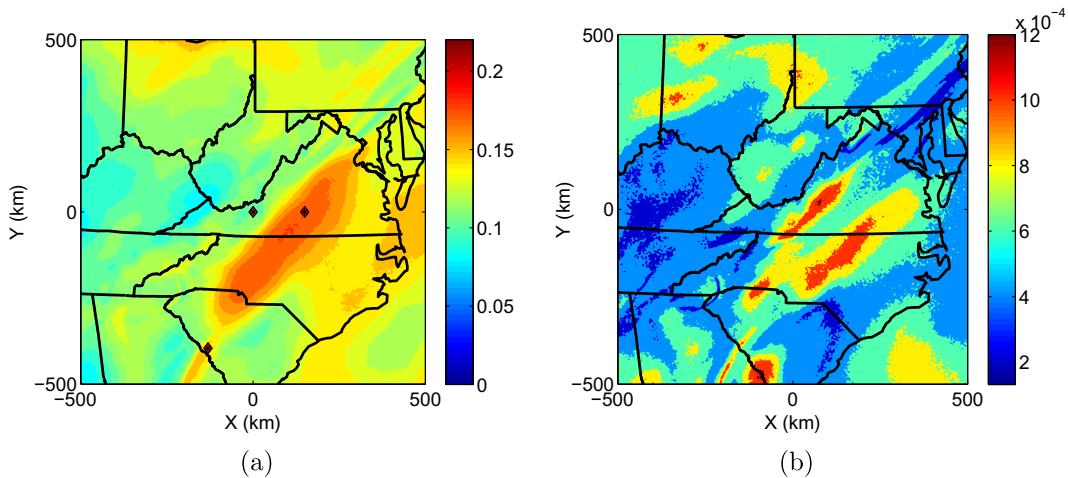

Fig. 3. 1000 stochastic realizations are used for calculation of (a) expected value of the stochastic backward FTLE fields (b) standard deviation of the stochastic backward FTLE fields; Integration time $T = -24$ h, $t_0 = 12:00$ UTC 29 Sep 2010. Different qualitative changes at the location of marked points on panel (a) are discussed in Fig. 4.
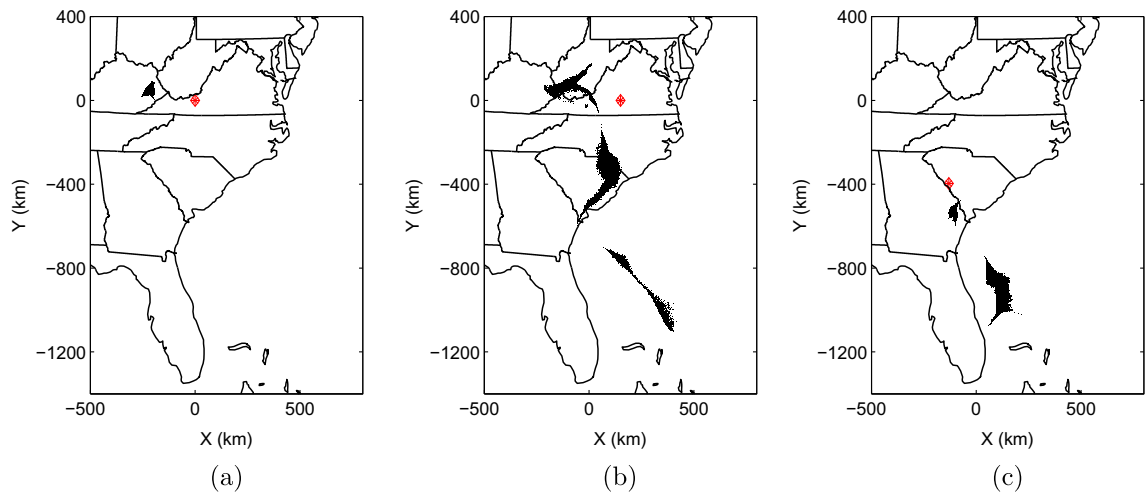
Fig. 4. Investigating the qualitative changes in a stochastic FTLE field due to the distribution of the probabilistic source areas, (a) probable source area for the reference point $(37°11'N, -80°35'W)$ corresponding to (0,0) (b) probable source area for $(37°10'30''N, -78°55'48''W)$ or 150 km east to of the reference point (c) probable source area for $(33°40'48''N, -82°W)$ equal to 131 km west and 396 km south respectively; for all the cases $\mathbf{x_0}$ is shown by a red diamond, $A_{50\%}$, $T = -24$ h, $t_0 = 12:00$ UTC 29 Sep 2010 and box size $d = 4$ km. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(referring to the ridges over GA in Figs. 2(a) and 3(a)). In Fig. 4(c) the source area is not as large as the second one but larger then the first one (red diamonds show $\mathbf{x_0}$ for each scenario and the collections of black points show the probabilistic source areas).

Before considering other features of SFTLE fields, we consider Fig. 4(b) as an example with separated source areas to see the effect of LCSs on shaping those areas. Fig. 5(a) shows the expected stochastic forward FTLE field for 12:00 UTC 28 Sep 2010, obtained from 500 sample members. When we *back* track the probable source points for $\mathbf{x_0}$ (red diamond) on Fig. 4(b), we expect that they settle down on *parts* of a stochastic repelling LCS feature since the integration direction is backward (repelling in forward time is equal to attracting in backward time). This attraction is noticeable when Fig. 4(b) is compared with Fig. 5(a). We notice that almost all the probable source areas coincide with a *repelling* feature (which is attracting in backward time). Similarly, the separation of islands are governed by the stochastic attracting LCS features (acting as repellers in backward time). Fig. 5(b) shows that isolated and bifurcated islands are indeed the results of backward "cutting" by attracting LCS features in the vicinity of locations that the attracting and repelling structures intersect. Two of these regions are denoted by arrows in Fig. 5(b).

Panels of Fig. 6 illustrates how the probabilistic bifurcated source areas are influenced by the corresponding contemporaneous attracting LCSs during 24 h of integration (three panels correspond to 24, 12 and 6 h of backward integration). In this figure we show the position of all 10^6 virtual particles. To keep the figure clear we only show the deterministic LCSs, as we know that these features usually act as the backbone of the stochastic LCS features. The cutting effects of LCS features manifests when a smaller percentage of particles is shown, for example if we show 50% of particles, then the result would be Fig. 4(b) on which the bifurcated and isolated islands of probable source areas due to the LCSs are distinct. In the language of Bayesian statistical inference, we can say the LCSs play a role in structuring the prior probability distribution of the source location.

To study other probable features in stochastic FTLE fields we consider Fig. 7(a) which is the backward FTLE field for 12:00 UTC 8 Apr 2012. We can see a very strong, sharp and lengthy ridge over KS, MO and IL. The analog of this feature exists in Fig. 7(b), which is the expected value (the mean) of SFTLE fields for 1000 realizations.

We can see that the tail of the long ridge over KS remains fairly sharp and the head becomes a broad region over MO. The most interesting feature of these two panels is over the eastern coast of the United States. In Fig. 7(a) there is no major and extended high value ridge over the eastern part but in Fig. 7(b) we notice a vast high value area covering all the eastern part of the figure; starting from NC continuing to the northern part of NY. This large area of the SFTLE corresponds to several small features of elevated FTLE in Fig. 7(a). The applied LPDM represents horizontal diffusion based on velocity gradients and therefore the minor elevated FTLE region becomes a spread out region of high SFTLE. We can conclude that if we consider realistic conditions by adding unresolved turbulent motion, it is possible to have regions with high SFTLE values where the deterministic FTLE values are low. In fact these regions are zones of strong attraction or repulsion which can be important for describing the distribution of passive tracers. If we neglect the unresolved turbulence, then we may underestimate the importance of these regions.

We summarize our observations into four categories (note that each of these cases corresponds to different regions of a FTLE field):
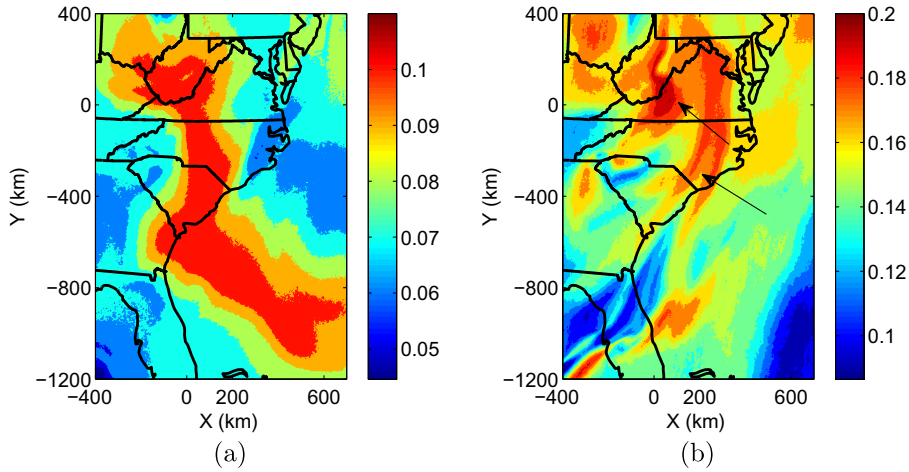
Fig. 5. Role of stochastic FTLE fields in shaping the probable source area of the virtual particles, (a) expected forward SFTLE field. High SFTLE regions act as the source location of particles (b) expected backward SFTLE field. High SFTLE regions acts as cutting edges, thus density of the virtual particles is low near them. Both cases are obtained by ensemble averaging 500 SFTLE fields, integration time $T = 24$ h, time shown is $t_0 = 12$:00 UTC 28 Sep 2010.
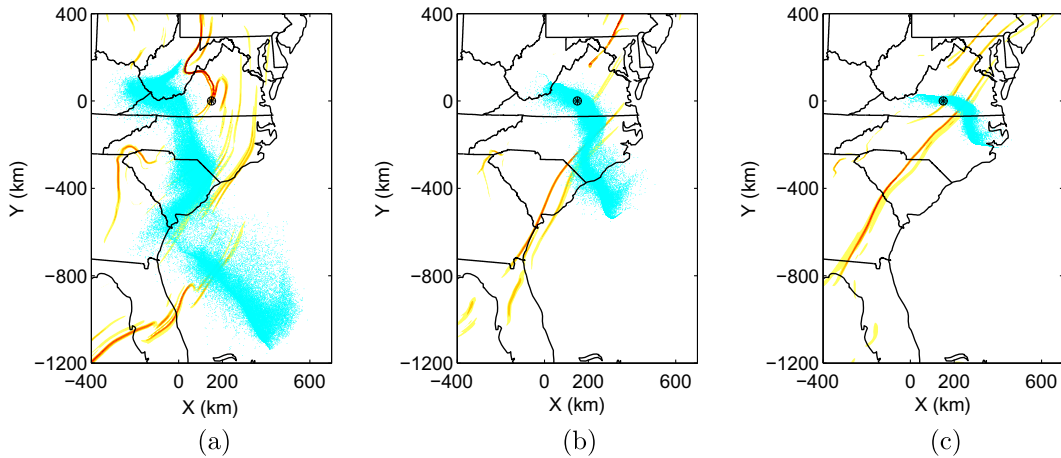

Fig. 6. Probabilistic source area ($A_{99\%}$) at (a) $T = -24$ h, (b) $T = -12$ h, (c) $T = -6$ h and the role of attracting LCS features in forming the bifurcated probabilistic areas; initialization time $t_0 = 12$:00 UTC 29 Sep 2010, reference point $\mathbf{x_0} = (37°10'30''N, -78°55'48''W)$.

1. Sharp ridges of the deterministic FTLE field deform into bands of high value expected SFTLE.
2. Sharp ridges of the deterministic FTLE field remain fairly sharp in stochastic cases.
3. Ridges of the FTLE field are washed out and disappear when unresolved stochastic motion is considered.
4. Regions with low to moderate and segmented FTLE values in deterministic field can become regions of high expected SFTLE value regions.

If we consider the deterministic solutions, we only obtain co-dimension one LCS features that are strongest in attraction or repulsion of nearby particles. But if we consider the stochasticity of the flow field, we may find regions of high attraction or repulsion strength that may not be observed by deterministic solutions. By using the archive data we can detect the backward SFTLE and their associated attracting LCSs where the particles settle down. Also by considering the stochasticity of the flow field we declare the boundaries between coherent sets as *fuzzy* instead of sharp lines. Since unresolved turbulence is similar to a diffusion process, in cases that we observe considerable diffusion such as atmospheric transports of aerosols, we could expect to see qualitative changes to the FTLE field analogous to our results for SFTLE fields.

3. Uncertainty analysis

When atmospheric forecast velocity data is used to generate the forecast FTLE fields, it is important to consider the effect of unresolved turbulence as was done in the previous section, but also the error of the results with respect to the archive-based (observation-based) FTLE fields. In a previous paper [11] we observed that even if one uses the best resolution operational forecast data for nearly real-time forecasting, one would face inevitable errors in forecasting the resultant determin-
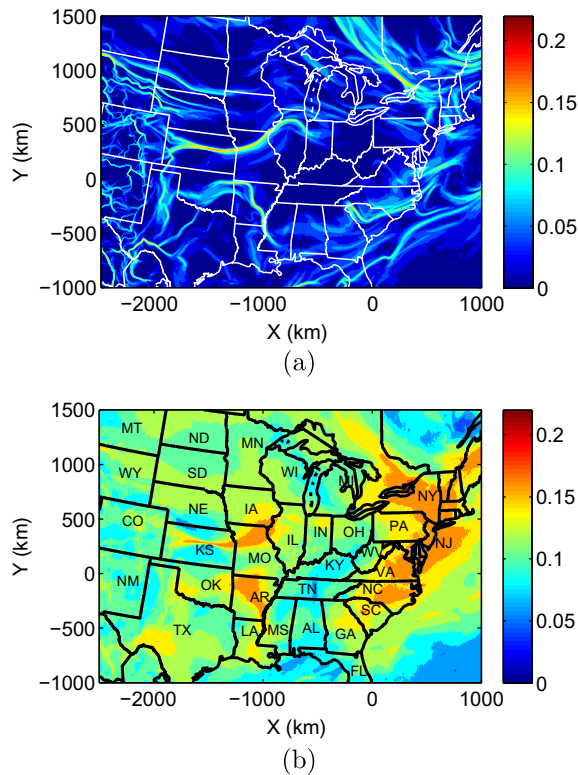
Fig. 7. (a) Deterministic backward FTLE field, (b) expected value, backward SFTLE field. Integration time $T = -24$ h, $t_0 = 12$:00 UTC 8 Apr 2012. Note the high value of SFTLE over the east coast of US where the deterministic FTLE field does not capture it. This region could be an attracting area for passive tracers. Also note the varying width of the high value SFTLE region over KS, MO and IL.

istic FTLE-LCSs fields. This fact is due to the cumulative errors in the integration process for calculating the tracer particle flow map. One might compare Eulerian forecast and archive velocity fields and find that the level of error is small and the forecasts are skillful [54], however, in the process of time integration, small errors accumulate and the forecast flow map can diverge dramatically from the flow map obtained from archive/reanalysis data (note, archive/reanalysis data is taken as the best representation of the true state of the system). The quality of the velocity forecast field is essential. However it is well known that improving the quality of forecasts is limited by the atmosphere's inherent chaotic behavior [55]. Regarding forecast FTLE fields and the associated LCSs, one has the option to either depend on the best resolution deterministic velocity fields or attempt to quantify the uncertainty of the forecast results.

The current practice for the study of uncertainty of forecast results is to use an ensemble forecasting approach. In this approach the outputs of the data assimilation process are perturbed and fed to the geophysical models which are assumed to represent the dynamics of the system. Different perturbation techniques are used in this regard, e.g., NCEP uses the breeding method while ECMWF implements the singular vector approach [20,56]. The output members of the ensemble forecasting approach are suitable for different purposes, e.g., improving the quality of forecasting by ensemble averaging, providing indications of the reliability of the forecast or providing a quantitative basis for probabilistic forecasting [20,56].

In this study we use the ensemble members of the forecast velocity field to generate ensemble members of the FTLE field. By comparing these fields we have a measure of the uncertainty of the forecast fields over different geographical regions.

3.1. Velocity ensemble data set

To perform this case study we use the NCEP-GFS ensemble forecast data in which the ensemble forecasts and the breeding cycles are initialized at 00:00 UTC each day and the control run is initialized from NCEP/NCAR reanalysis data.[3] The time resolution of GFS data is 12 h and the spatial resolution is 2.5°.

When the breeding approach is used in ensemble forecasting, a random initial perturbation vector (the "random seed") with a small fixed initial norm (ϵ) is added and subtracted to the control input data, generating a perturbed initial condition pair, see Fig. 8. The random seed is the same size as the degree of freedom of the system and contains all existing variables of the model (e.g., velocities, temperature, solar radiation, and so on). This perturbation vector is added at the beginning of the procedure, i.e., added at an initial time t, then the model is integrated for the both control and the perturbed data sets during

---

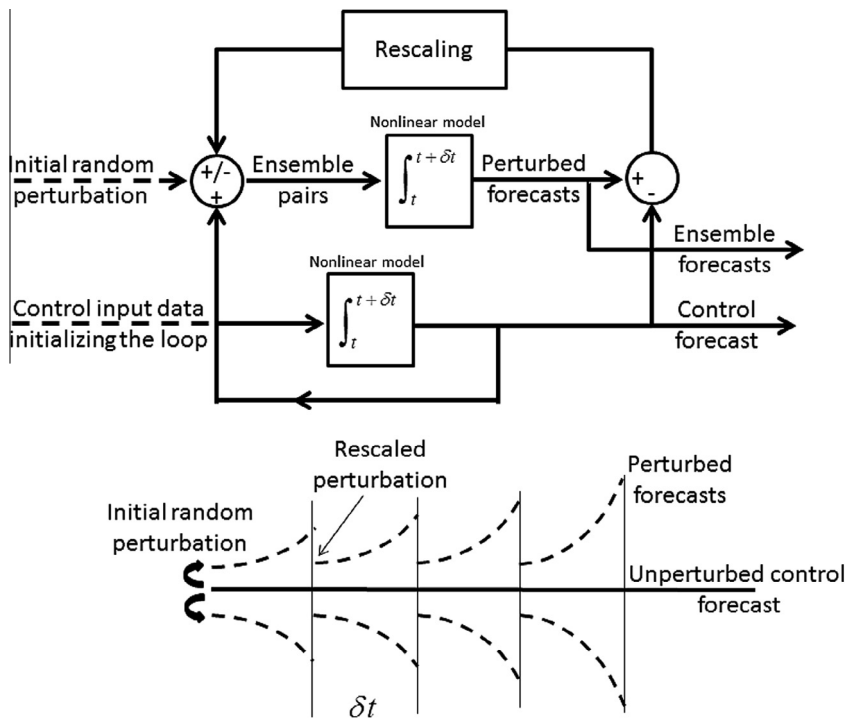[3] http://www.esrl.noaa.gov/psd/forecasts/reforecast/data.html

Fig. 8. Schematic of the breeding cycle. Control run is integrated without any perturbation. Bred vectors are initialized by adding random perturbation (random seed) to the control case. Each ensemble member is integrated (solving the systems's model with this initial condition) for a specified length of time (e.g., 6 h), and then each field is compared with the control field. The difference between control case and each ensemble member is scaled down to have the same norm as the unperturbed one. This rescaled field is added/subtracted to the control case to generate the new perturbed initial conditions, and the integration cycle is repeated until the end of forecast time. Dashed line in the lower panel show the symbolic solution of two members.

a pre-determined length of time δt, i.e., from t to $t + \delta t$. Results from the control set integration are subtracted from the perturbed one. For the next step the difference is scaled down to the initial perturbation norm, ϵ, and this new perturbation vector is added to the new state of the system calculated from the control set to produce the new perturbed initial condition. These steps are repeated for all the ensemble members until the pre-specified maximum integration time [20]. This approach is conceptually similar to the process of finding the leading order Lyapunov vector/exponent in systems of ordinary differential equations. Moreover the perturbations depend on the dynamical responses of the underlying system and it is known that for atmospheric ensemble forecasting the bred vectors do not converge to a *single* leading bred vector [57].

At NCEP the same procedure is implemented for ensemble forecasting proposes with a slight modification of the numerical procedure. In NCEP approach which is called self-breeding method, the difference between positive and negative perturbed forecasts is divided by 2 and then scaled down to the initial perturbation norm. The scaled difference is added and subtracted from the control set, generating initial condition for the new pair of forecasts. For NCEP-GFS ensemble forecast data, 7 pairs are calculated, so we have 14 ensemble members plus one control case [56,58,59].

3.2. Ensemble FTLE field results

To demonstrate the applications of ensemble FTLE fields, we use the 850 mb pressure surface data for the velocity field. The target time for the FTLE field calculation is 12:00 UTC 29 Sep 2010 the same as the first case shown in Fig. 2(a) and the backward integration time is $T = -24$ h. According to this target time, the forecast chain was initialized at 00:00 UTC 28 Sep 2010 and extended to 00:00 UTC 30 Sep 2010. Fig. 9 shows all 15 members of the ensemble FTLE fields calculated from ensemble members of the GFS velocity field. Fig. 9(a) is the FTLE field from the control velocity field and the other 14 members are the results from 7 pairs of bred modes. From this collection of FTLE fields we recognize the main patterns of the forecast result and the associated candidate LCSs. Also we notice the differences in ensemble members, e.g., the FTLE-LCS results from Fig. 9(j) and (k) are very different, but they correspond to a single pair of bred modes (i.e., the control velocity field plus an addition and subtraction of the same small perturbation velocity field).

Fig. 10(a) and (b) represents the average and standard deviation (as a measure of spread of FTLE value over each point) of the represented ensemble FTLE field members in Fig. 9. This figure shows that over some regions of the FL panhandle one can expect a high value FTLE ridge but over some other regions such as north–west SC and especially south–east Canada the resultant FTLE-LCS are uncertain (the latter is not due to the edge effect since the data domain is larger than what is
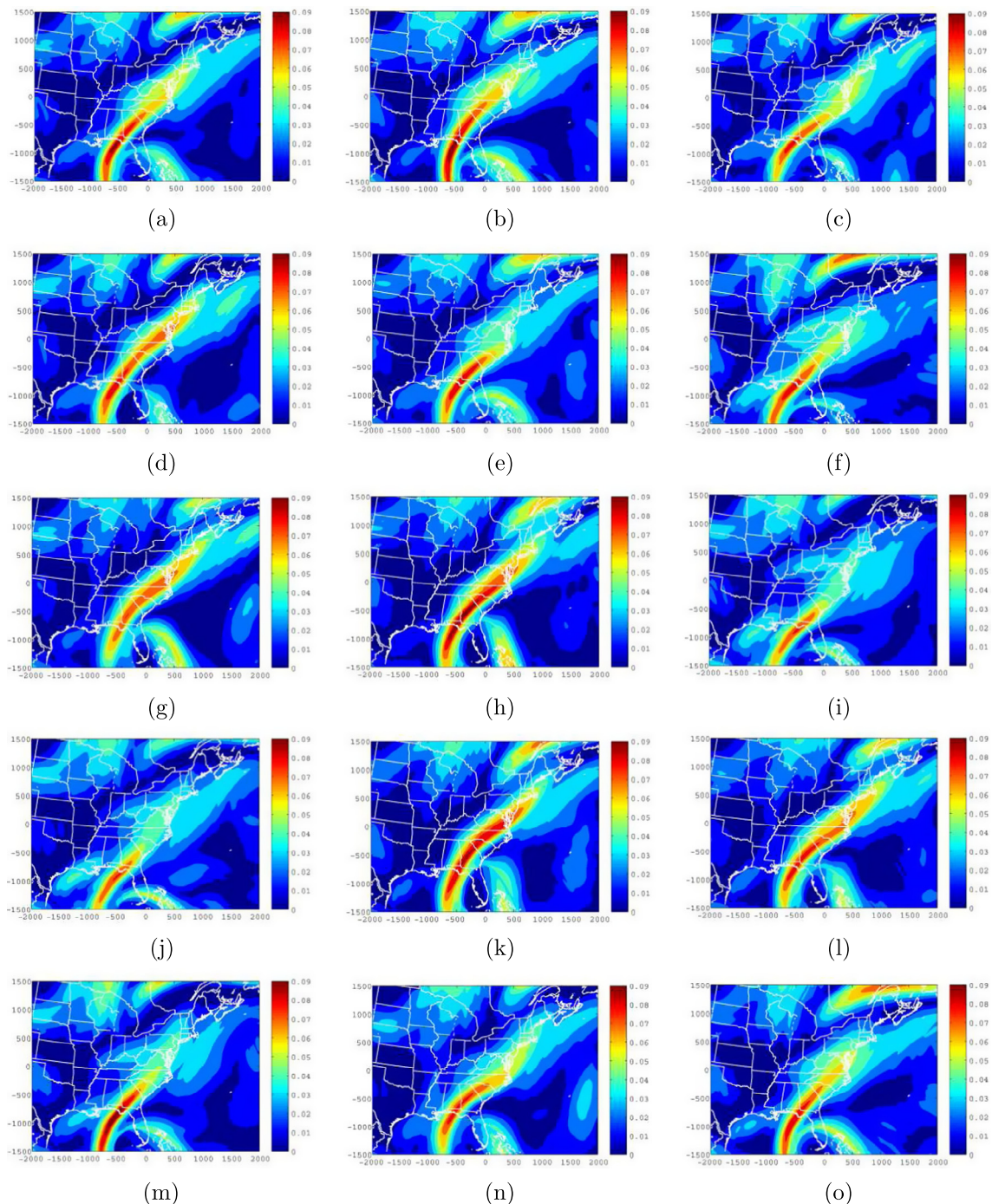
Fig. 9. 15 Ensemble members of forecast FTLE field for 12:00 UTC 29 Sep 2010, (a) FTLE from the control velocity field (b . . .o) 14 other ensemble members of forecast FTLE filed obtained from perturbed velocity fields, integration time is −24 h for all the cases.

considered for the FTLE field and we apply the linear extension technique). Panel (c) of Fig. 10 is the backward FTLE field result (integration time $T = -24$ h) from NCEP/NCAR Reanalysis 1 velocity data with spatiotemporal resolution 2.5° and 6 h.[4] By using panels (a) and (c) of this figure, we can compare the results from forecast ensemble mean with the pure reanalysis based FTLE field. For example we observe a bifurcation in the LCS structures over the east coast which the mean of the forecast ensemble does not show. In addition there is a weaker but long and wide feature over MN, SD and NE in the reanalysis-based
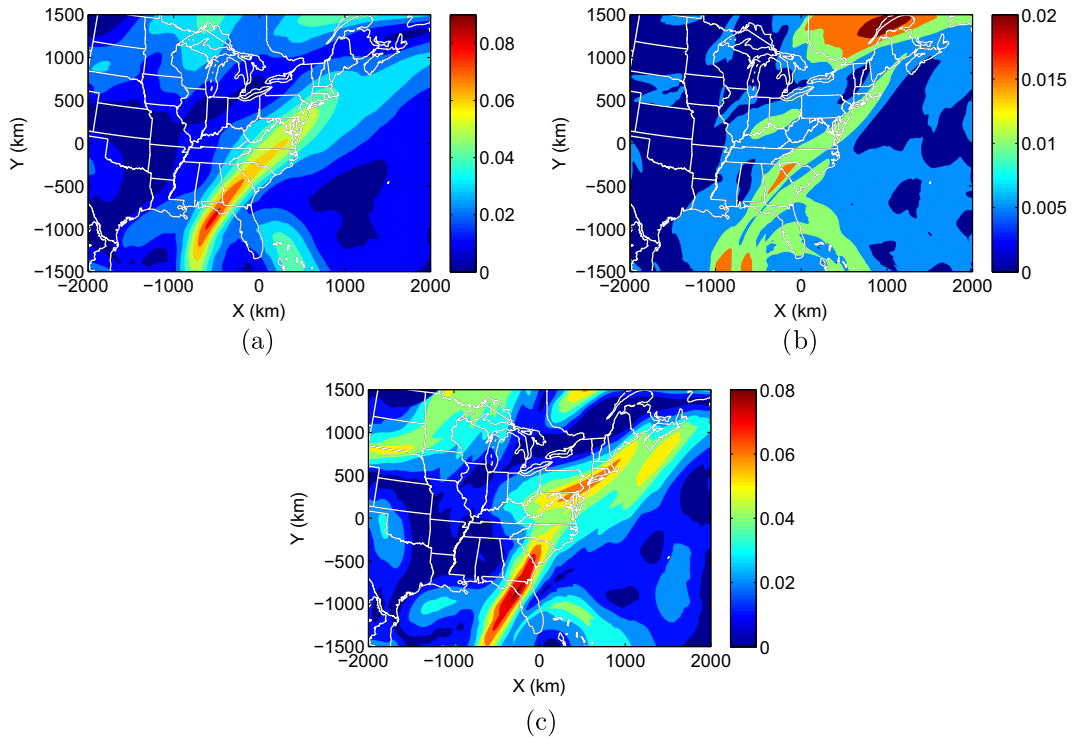
---

[4] http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html

Fig. 10. (a) Mean of the forecast ensemble members (b) standard deviation of the forecast ensemble members (c) FTLE from GFS reanalysis data; corresponding time and integration time are 12:00 UTC 29 Sep 2010 and −24 h respectively.
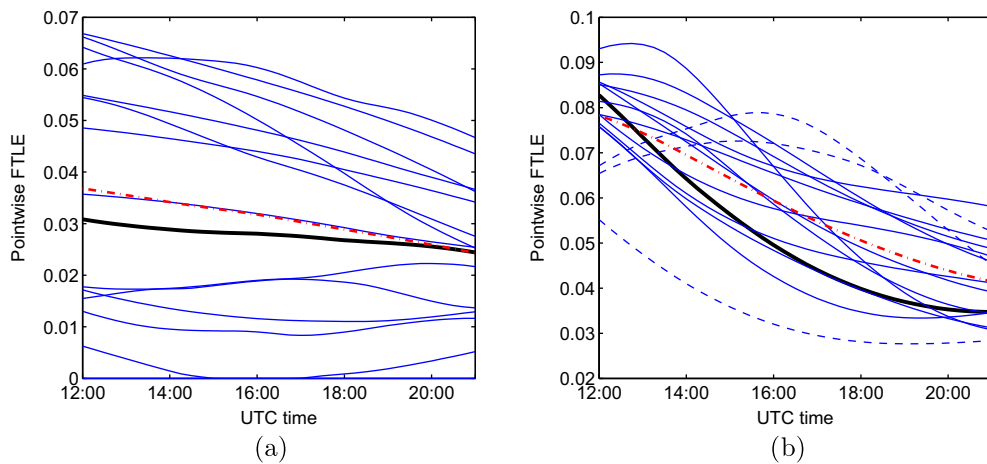

Fig. 11. Ensemble of pointwise FTLEs vs. time. Black line represents the control case, dashed red line shows the average of ensembles and the other 14 members are from each FTLE calculation, (a) highly uncertain forecast, FTLE measurement at $(48°22'N, -68°10'48''W)$ (b) fairly consistent forecast, FTLE measurement at $(30°27'N, -85°3'W)$. These two points correspond to $(1000, 1400)$ and $(-450, -750)$ km with respect to the reference point.

FTLE field which is more faint in the mean ensemble forecast. One should note that the overall similarity between Figs. 2(a) and 10(a) in common windows is due to the fact that both FTLE fields correspond to the same time, meanwhile the differences are mainly due to the spatiotemporal resolution of the velocity fields and the fact that the former is calculated from archive data and latter is generated from *pure* forecast data initialized by reanalysis data set at 00:00 UTC 28 Sep 2010.

To study the uncertainty of the FTLE fields we use the pointwise FTLE value vs. time at a fixed location. The local maxima of this time series could be regarded as representative of candidate hyperbolic LCS features that might pass over a specific
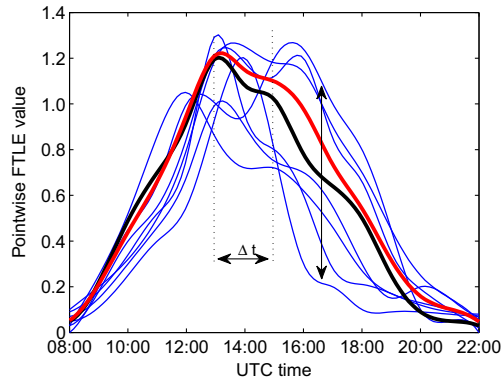
Fig. 12. Schematic of ensembles pointwise forecasts FTLE value at a specified location. Black: control case, red: mean of the ensembles, blue: other ensemble members. The number of maxima in the time interval Δt is assumed proportional to the probability of passage of a LCS over the specified location. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
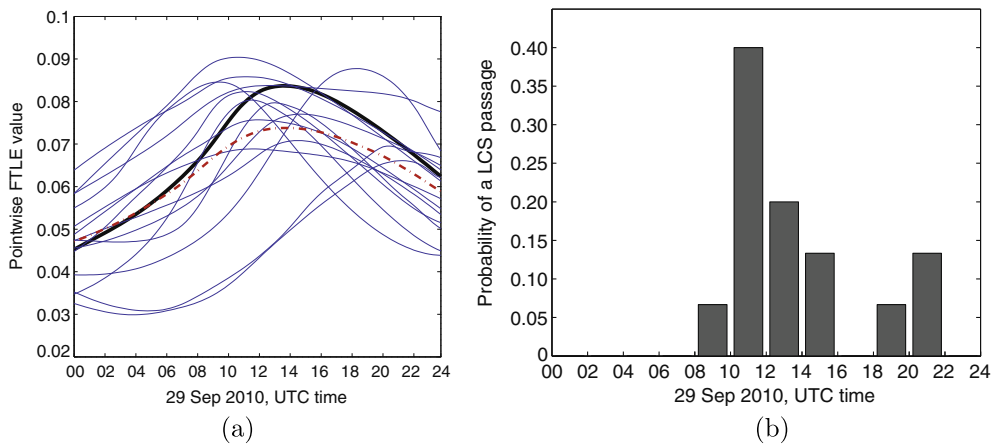

Fig. 13. (a) Ensembles of pointwise forecasts FTLE field value at $(30°27'N, -85°3'W)$, integration time: -24 h, red line: mean curve, and black line: control case (b) probability distribution of a LCS passage during 2 h windows. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

location [11]. The uncertainty of the pointwise results represents the uncertainty of presence and timing of strong LCS features passing over a location of interest. As an example, Fig. 11(a) shows the pointwise FTLE value at $(48°22'N, -68°10'48''W)$ equal to $(1000, 1400)$ km with respect to the reference point (Virginia Tech Kentland Farm) for the previous forecast interrogation window of 12:00 to 21:00 UTC 29 Sep 2010 where the first frames of the FTLE field for the 15 ensemble members are shown in Fig. 9. We observe that the results from ensemble members are significantly different during the interrogation time window. This figure shows high uncertainty of the forecast FTLE value over that location when different forecast velocity fields are used to calculate the FTLE-LCS features. Fig. 11(b) shows another case of pointwise FTLE vs. time at $(30°27'N, -85°3'W)$ equal to $(-430, -750)$ km with respect to the same reference point and during the same time interval. In this figure we observe that (with the exception of three members shown by dashed line) they all follow a similar pattern and their differences are small, meaning that forecasts at that location are fairly consistent during the interrogation time interval.

In addition to uncertainty quantification, this observation could lead to the concept of probabilistic forecasting of the LCS features. To obtain reliable forecast of LCS features, we will assume that one needs to have enough ensemble members that show maxima in a common time interval. Referring to Fig. 12 as a schematic of forecast ensembles, we define the probability of the passage of an LCS feature over a region during a defined time interval $\mathcal{T} = (t - \frac{\Delta t}{2}, t + \frac{\Delta t}{2})$ as

$$P_{\mathcal{T}} = n_p / N_{ens}, \tag{12}$$

where n_p is the number of maxima beyond a pre-determined threshold (for more information about that threshold refer to [11]) in the time interval $\mathcal{T}$ and N_{ens} is the total number of ensemble members, providing that each does not have more than one peak in that time interval. For example, referring to Fig. 12, N_{ens} is 8 during the time interval $\mathcal{T} = 13:00-15:00$ UTC, and

Table 1
Four randomly selected weeks, one in each season of 2012. Days that major LCS features observed are reported. Each entity represents the probability of passing LCS features over the reference point (Virginia Tech Kentland farm) during the specified time interval. These values are calculated based on 15 ensemble members of GFS forecast data set. "*" represent the passage time of LCS features over the reference point which is calculated by using the reanalysis NARR data set. In some days we observe more than one LCS over the reference point. Note that in some cases the summation of numerators in each column is less than 15 which means that some of the ensemble members do not have any peak during 24 h of the interrogation window.

| UTC	April				June			September				November			
	22	23	25	27	4	6	7	16	17	20	22	5	7	8	9
00:00–02:00	0	$\frac{10^*}{15}$	0	$\frac{1}{15}$	0	$\frac{1}{15}$	0	0	0	0	0	0	$\frac{1}{15}$	0	0
02:00–04:00	0	$\frac{2}{15}$	0	$\frac{7^*}{15}$	0	$\frac{1}{15}$	0	$\frac{2}{15}$	0	0	0	0	$\frac{6^*}{15}$	0	0
04:00–06:00	0	$\frac{3}{15}$	$\frac{7}{15}$	$\frac{4}{15}$	0	$\frac{9^*}{15}$	0	$\frac{11^*}{15}$	0	0	0	0	$\frac{8}{15}$	0	0
06:00–08:00	$\frac{1}{15}$	0	$\frac{6^*}{15}$	$\frac{3}{15}$	0	$\frac{2}{15}$	0	0^*	$\frac{1}{15}$	0	0	$\frac{1}{15}$	0	0	0
08:00–10:00	$\frac{2}{15}$	0	$\frac{2}{15}$	0	0	$\frac{2}{15}$	0	$\frac{2}{15}$	$\frac{14^*}{15}$	0	0	$\frac{1}{15}$	0	0	0
10:00–12:00	$\frac{8}{15}$	0	0	0	0	0	0	0	0	0	0	$\frac{1}{15}$	0	0	0
12:00–14:00	$\frac{3^*}{15}$	0	0	0	0	0	0	0	0	0	0	$\frac{10^*}{15}$	0	0	0
14:00–16:00	0^*	0	0	0	$\frac{1}{15}$	0	$\frac{2}{15}$	0	0	0	0	$\frac{2}{15}$	0	0	$\frac{2}{15}$
16:00–18:00	$\frac{1}{15}$	0	0	0	$\frac{12^*}{15}$	0	0^*	0	0	0	$\frac{1}{15}$	0	0	$\frac{1^*}{15}$	$\frac{6}{15}$
18:00–20:00	0	0	0	0	$\frac{2}{15}$	0	$\frac{6}{15}$	0	0	$\frac{1}{15}$	$\frac{2}{15}$	0	0	0	$\frac{3^*}{15}$
20:00–22:00	0	0	0	0	0	0	$\frac{7}{15}$	0	0	$\frac{6^*}{15}$	$\frac{11^*}{15}$	0	0	$\frac{11^*}{15}$	$\frac{2}{15}$
22:00–24:00	0	0	0	0	0	0	0	0	0	$\frac{7}{15}$	$\frac{1}{15}$	0	0	$\frac{1}{15}$	$\frac{2}{15}$
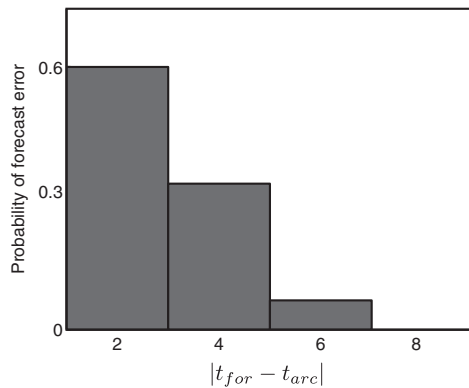
Fig. 14. Summary of Table 1. A histogram of the absolute value of the difference between forecast (t_{for}) and archive (t_{arc}) LCS passage times. For days with more than one LCS occurrence, we consider the one closest to the forecast interval result with highest probability.

the number of maxima above the hypothetical threshold 1.1 is $n_p = 5$. Thus the probability of observing an LCS feature would be $n_p/N_{ens} = 5/8$.

Fig. 13(a) shows the ensemble forecast members for 29 Sep 2010 at location $(30°27'N, -85°3'W)$. Since the spatiotemporal resolution of ensemble GFS data is coarse $(2.5°, 12\text{ h})$ the ensemble members of the pointwise FTLEs do not show sharp peaks, however this figure shows the main concept of probabilistic FTLE-LCS forecasting. Fig. 13(b) shows the corresponding probability distribution of having a candidate LCS during that time interval when $\Delta t = 2$ h. This figure shows that the probability of an LCS passage is maximum in the interval $\mathcal{T}$ between 10:00 and 12:00 UTC. We should note that to generate Fig. 13(a) we use the most updated reanalysis data as the initial condition for each forecast data frame. Thus, we use reanalysis data corresponding to 00:00 UTC 28 and 29 Sep 2010 as the initial conditions for the forecast velocity fields. Note that Figs. 11(b) and 13(a) are slightly different in their common window (after 12:00 UTC) because in the former one we just use the 00:00 UTC 28 as the initial condition while in the latter one we use the most updated initial conditions.

To test this approach for more days and to know more about statistical distribution of errors we use GFS ensemble forecast data to calculate the forecast FTLE-LCS features (the same procedure as described for Fig. 13), and then we compare these results with reanalysis based FTLE-LCSs which are obtained from the NCEP North American Regional Reanalysis (NARR).[5] Note that spatiotemporal resolution of NARR data is 0.3 degrees and 3 h. For both cases of forecasts and reanalysis results we consider a time interval of 24 h and an integration time for FTLE calculations of $T = -24$ h (backward FTLE). Results in Table 1 correspond to four weeks, one in each season of 2012. The selected weeks are randomly chosen. We report all days of each week that major LCS features have been observed over the Virginia Tech Kentland Farm.

We use a time resolution of 15 min for the FTLE-LCS calculations, so for each day we have 97 time slices. For each day all 15 members of the GFS based forecast FTLE-LCS are calculated (the total number of frames per day is 1455 = 97 slices times

---

[5] http://www.esrl.noaa.gov/psd/data/gridded/data.narr.html

the 15 members) and then the probability distribution of having a LCS over the reference point is obtained. The result of this process is similar to Fig. 13(b) for each time slice. Table 1 contains data for 15 days, so it is the abstract of comparing 21825 ensemble forecast frames (1455 frames per day times 15 days) with 1455 reanalysis based frames of FTLE-LCSs. We divide each day into 2 h intervals. The probability of LCS passage over the reference point is calculated as introduced in (12). Table 1 shows the corresponding probabilities as factors of $(\frac{1}{15})$ since the number of ensemble members is 15, so the numerator shows the number of ensemble peaks in each time interval. Note that in some cases, such as Sep 20 and Nov 8, the summation of peaks is less than 15 which means that some member of the ensemble does not have any maximum during the 24 h interrogation window or the peak is very flat. The passage time of the reanalysis (NARR) based LCSs over the reference point for each day is shown by "∗" in each column.

Results of this comparison show that in 9 out of 15 cases, the probabilistic forecast passage times are in the same time 2-hour interval as the reanalysis based results, in 5 cases they just differ one time interval (maximum error 4 h) and in 1 case the difference is equal or more than 2 time intervals. Also in 4 days we have more than one LCS passage based on reanalysis velocity field where for 3 days, the ensemble forecasting captures at least one of them. A summary of this comparison is represented in Fig. 14. For days with more than one LCS occurrences, we consider the closest one to the forecast interval result with highest probability. Table 1 and Fig. 14 help us to have quantitative measure of the distribution and reliability of the forecast results which is not available with single deterministic forecast outputs. As an example we observe that for about 60% of the cases the error of ensemble forecasting with respect to the reanalysis result is within 2 h.

4. Discussion

The aim of this study was to consider two issues related to prediction of atmospheric FTLE fields and Lagrangian coherent structures (LCSs). The first concern was the effect of unresolved turbulence on the resultant FTLE fields. By using the Monte Carlo method and adding a stochastic turbulent velocity term, we introduced the notion of stochastic FTLE (SFTLE) field. A significant difference between the deterministic solution and the centroid of the (in general, non-Gaussian) probability distributions was observed. The role of FTLE-LCSs in shaping the probability distributions was discussed.

The second concern originated from previous observations about comparisons between archive- and forecast-based FTLE fields [11]. These comparisons show that the major source of errors in forecast FTLE fields is errors in the forecast velocity field which is unavoidable due to the chaotic dynamics of atmosphere. Thus, a major concern is the level of confidence regarding the forecast results. We introduced an ensemble method for FTLE-LCSs for quantitative study of uncertainty and probabilistic forecasting, using relatively low spatiotemporal resolution GFS data. We also introduced a method for probabilistic forecasting of LCS features, particularly the passage time of an LCS over a particular geographic point, which was previously found to be associated with aeroecological events [5]. We found that one can predict the passage time with an accuracy of 2 h 60% of the time which lends confidence to the practical application of LCS passage prediction for forecasting movement of biological and chemical fronts.

Subgrid scale stochastic effects and uncertainty will be important in real applications of transport phenomena such as the monitoring of volcanic ash and large scale dust storms or aeroecology of microorganisms related to human health and agriculture (topics which can be interrelated; [60]). For the latter, given the association of FTLE fields and associated LCS features to moving fronts of microorganisms [5], there is a possibility of predicting significant changes in the population structure of airborne pathogens, or finding constraints on their dispersal range, with implications for improved management decision making by stakeholders.

References

[1] Haller G, Poje A. Finite time transport in aperiodic flows. Physica D 1998;119:352–80.
[2] Haller G, Yuan G. Lagrangian coherent structures and mixing in two-dimensional turbulence. Physica D 2000;147:352–70.
[3] Shadden SC, Lekien F, Marsden JE. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. Physica D 2005;212:271–304.
[4] Tallapragada P, Ross SD. A set oriented definition of finite-time Lyapunov exponents and coherent sets. Commun Nonlinear Sci Numer Simul 2013;18:1106–26.
[5] Tallapragada P, Ross SD, Schmale DG. Lagrangian coherent structures are associated with fluctuations in airborne microbial populations. Chaos 2011;21:033122.
[6] Tallapragada P, Identifying dynamical boundaries and phase space transport using Lagrangian coherent structures [Ph.D. thesis]. School Virginia Tech; 2010.
[7] Schmale DG, Ross SD, Fetters T, Tallapragada P, Wood-Jones A, Dingus B. Isolates of Fusarium graminearum collected 40 to 320 meters above ground level cause Fusarium head blight in wheat and produce trichothecene mycotoxins. Aerobiologia 2012;28:1–11.
[8] Schneider RW, Hollier CA, Whitam HK, Palm ME, McKemy JM, Hernandez JR, Levy L, DeVries-Paterson R. First report of soybean rust caused by Phakopsora pachyrhizi in the continental united states. Plant Dis 2005;89. 774–774.

[9] Lin B, BozorgMagham AE, Ross SD, Schmale DG. Small fluctuations in the recovery of fusaria across consecutive sampling intervals with unmanned aircraft 100 m above ground level. Aerobiologia 2013;29(1):45–54.

[10] Lin B, Movement and structure of atmospheric populations of Fusarium [Ph.D. thesis]. School Virginia Tech; 2013.

[11] BozorgMagham AE, Ross SD, Schmale DG. Real-time prediction of atmospheric lagrangian coherent structures based on uncertain forecast data: an application and error analysis. Physica D 2013;258:47–60.

[12] BozorgMagham AE, Atmospheric Lagrangian transport structures and their applications to aerobiology [Ph.D. thesis]. School Virginia Tech; 2014.

[13] Griffa A, Piterbarg L, Ozgokmen T. Predictability of Lagrangian particle trajectories: Effects of smoothing of the underlying Eulerian flow. J Marine Res 2004;62:1–35.

[14] Ozgokmen T, Griffa A, Mariano A, Piterbarg L. On the predictability of Lagrangian trajectories in the ocean. J Atmos Oceanic Technol 2000;17:366–83.

[15] Kahl JD, Samson PJ. Uncertainty in trajectory calculations due to low resolution meteorological data. J Clim Appl Meteorol 1986;25:1816–31.

[16] Wilson J, Sawford B. Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere. Boundary Layer Meteorol 1996;78:191–210.

[17] Stohl A. Computation, accuracy and applications of trajectories – A review and bibliography. Atmos Environ 1998;32:947–66.

[18] Palmer T. Predicting uncertainty in forecasts of weather and climate. Rep Progr Phys 2000;63:71–116.

[19] Ehrendorfer M. Predicting the uncertainty of numerical weather forecasts: a review. Meteorologische zeitschrift 1997;6:147–83.

[20] Kalnay E. Atmospheric Modelling. Cambridge: Data Assimilation and Predictability; 2003.

[21] Haller G. Lagrangian coherent structures from approximate velocity data. Phys Fluids 2002;14:1851–61.

[22] Lermusiaux PF, Chiu CS, Gawarkiewicz GG, Abbot P, Robinson AR, Miller RN, Haley PJ, Leslie WG, Majumdar SJ, Pang A, et al., Quantifying uncertainties in ocean predictions, Tech. Rep., Institution DTIC Document; 2006.

[23] Olcay AB, Pottebaum TS, Krueger PS. Sensitivity of Lagrangian coherent structure identification to flow field resolution and random errors. Chaos 2010;20:017506.

[24] Peng J, Peterson R. Attracting structures in volcanic ash transport. Atmos Environ 2012;48:230–9.

[25] Draxler R, Hess G. An overview of the HYSPLIT4 modelling system for trajectories, dispersion and deposition. Aust Meteorol Mag 1998;47:295–308.

[26] Fay B, Glaab H, Jacobsen I, Schrodin R. Evaluation of Eulerian AND Lagrangian atmospheric transport models at the Deutscher-Wetterdienst using ANATEX surface tracer data. Atmos Environ 1995;29:2485–97.

[27] Legg B, Raupach M. Markov-chain simulation of particle dispersion in inhomogeneous flows – the mean-drift velocity induced by a gradient in Eulerian velocity variance. Boundary Layer Meteorol 1982;24:3–13.

[28] ksendal B. Stochastic differential equations. Springer; 2003.

[29] Lemons DS, Langevin P. An introduction to stochastic processes in physics. JHU Press; 2002.

[30] Thomson D. Criteria for the selection of stochastic-models of particle trajectories in turbulent flows. J Fluid Mech 1987;180:529–56.

[31] Stohl A, Forster C, Frank A, Seibert P, Wotawa G. Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. Atmos Chem Phys 2005;5:2461–74.

[32] Houtekamer PL, Mitchell HL. Data assimilation using an ensemble Kalman filter technique. Mon Weather Rev 1998;126(3):796–811.

[33] Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res Oceans (1978–2012) 1994;99(C5):10143–62.

[34] Stohl A, Wotawa G, Seibert P, Kromp-Kolb H. Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. J Appl Meteorol 1995;34(10):2149–65.

[35] Danielsen EF. Trajectories: isobaric, isentropic and actual. J Atmos Sci 1961;18:479–86.

[36] Branicki M, Mancho AM, Wiggins S. A Lagrangian description of transport associated with a front–eddy interaction: Application to data from the North-Western Mediterranean Sea. Physica D 2011;240(3):282–304.

[37] Sulman MH, Huntley HS, Lipphardt Jr B, Kirwan Jr A, Leaving flatland: diagnostics for lagrangian coherent structures in three-dimensional flows, Physica D.

[38] Kaltschmitt M, Streicher W, Wiese A. Renewable energy: technology, economics and environment. Springer; 2007.

[39] Rodean HC. Stochastic Lagrangian models of turbulent diffusion. American Meteorological Society; 1987.

[40] Isard SA, Gage SH, Comtois P, Russo JM. Principles of the atmospheric pathway for invasive species applied to soybean rust. BioScience 2005;55(10):851–61.

[41] Fitt B, McCartney H, West JS. Dispersal of foliar plant pathogens: mechanisms, gradients and spatial patterns. In: The epidemiology of plant diseases. Springer; 2006. p. 159–92.

[42] Cecchi L, Malaspina TT, Albertini R, Zanca M, Ridolo E, Usberti I, Morabito M, DallAglio P, Orlandini S. The contribution of long-distance transport to the presence of Ambrosia pollen in central northern Italy. Aerobiologia 2007;23(2):145–51.

[43] Viljanen-Rollinson S, Parr E, Marroni M. Monitoring long-distance spore dispersal by wind—a review. New Zealand Plant Protect 2007;60:291–6.

[44] Andrade D, Pan Z, Dannevik W, Zidek J. Modeling soybean rust spore escape from infected canopies: model description and preliminary results. J Appl Meteorol Climatol 2009;48(4):789–803.

[45] Hernández-Ceballos M, García-Mozo H, Adame J, Domínguez-Vilches E, Bolívar J, De la Morena B, Pérez-Badía R, Galán C. Determination of potential sources of Quercus airborne pollen in Córdoba city (southern Spain) using back-trajectory analysis. Aerobiologia 2011;27(3):261–76.

[46] Hernández-Ceballos M, Adame J, Bolívar J, De la Morena B. Vertical behaviour and meteorological properties of air masses in the southwest of the Iberian Peninsula (1997–2007). Meteor Atmos Phys 2013;119(3–4):163–75.

[47] Sadyś M, Skjth C, Kennedy R. Back-trajectories show export of airborne fungal spores (Ganoderma sp.) from forests to agricultural and urban areas in England. Atmos Environ 2014;84:88–99.

[48] Csanady GT. Turbulent diffusion in the environment, vol. 3. Springer; 1973.

[49] Smagorinsky J. General circulation experiments with the primitive equations. Mon Weather Rev 1963;91:99–164.

[50] Barkai E, Metzler R, Klafter J. From continuous time random walks to the fractional Fokker–Planck equation. Phys Rev E 2000;61(1):132.

[51] Kinzelbach W. The random walk method in pollutant transport simulation. In: Groundwater flow and quality modelling. Springer; 1988. p. 227–45.

[52] Karrasch D. Comment on a variational theory of hyperbolic Lagrangian coherent structures, Physica D 240 (2011) 574–598. Physica D 2012;241:14701473.

[53] Risken H, Eberly J. The Fokker–Planck equation, methods of solution and applications. J Opt Soc Am B Opt Phys 1985;2:508.

[54] Murphy AH. Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon Weather Rev 1988;116(12):2417–24.

[55] Lorenz EN. Deterministic nonperiodic flow. J Atmos Sci 1963;20:475–80.

[56] Toth Z, Kalnay E. Ensemble forecasting at NCEP and the breeding method. Mon Weather Rev 1997;125(12):3297–319.

[57] Kalnay E, Corazza M, Cai M, Are bred vectors the same as Lyapunov vectors? In: EGS general assembly conference abstracts, vol. 27, 6820; 2002.

[58] Caplan P, Derber J, Gemmill W, Hong S-Y, Pan H-L, Parrish D. Changes to the 1995 NCEP operational medium-range forecast model analysis-forecast system. Weather Forecast 1997;12(3):581–94.

[59] Uppala SM, Kållberg P, Simmons A, Andrae U, Bechtold V, Fiorino M, Gibson J, Haseler J, Hernandez A, Kelly G, et al. The ERA-40 re-analysis. Q J R Meteorol Soc 2005;131(612):2961–3012.

[60] Kellogg CA, Griffin DW. Aerobiology and the global transport of desert dust. Trends Ecol Evol 2006;21(11):638–44. ISSN 0169-5347, http://dx.doi.org/10.1016/j.tree.2006.07.004.