Turkish Journal of Earth Sciences (Turkish J. Earth Sci.), Vol. 20, 2011, pp. 343–358. Copyright ©TÜBİTAK
doi:10.3906/yer-0907-11 First published online 23 August 2010

Lagrangian Prediction and Correlation Analysis with Eulerian Data

MİNE ÇAĞLAR[1], TAYLAN BİLAL[1,2] & LEONID I. PITERBARG[2]

[1] Department of Mathematics, Koç University, Sarıyer, TR–34450 İstanbul, Turkey
(E-mail: mcaglar@ku.edu.tr)
[2] Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA

Received 29 July 2009; revised typescripts receipt 26 January 2010, 08 April 2010 & 19 August 2010; accepted 23 August2010

Abstract: A velocity field obtained from the ocean surface by high-frequency radar is used to test Lagrangian prediction algorithms designed to evaluate the position of a particle given its initial position and observations of several other simultaneously released particles. The problem is motivated by oceanographic applications such as search and rescue operations and spreading pollutants, especially in coastal regions. The prediction skill is essentially determined by temporal and spatial covariances of the underlying velocity field. For this reason correlation analysis of both Lagrangian and Eulerian velocities was carried out. Space covariance functions and spectra of the velocity field are also presented to better illustrate statistical environments for the predictability studies. The results show that the regression prediction algorithm performs quite well on scales comparable with and higher than the velocity correlation scales.

Key Words: turbulent flows, stochastic flows, Lagrangian prediction, eddy, correlation, spectrum, Euler velocity field

Euler Verilerle Lagrange Yörüngelerin Tahmini ve İlintilerin İncelenmesi

Özet: Okyanus yüzeyinden yüksek çözünürlükte radarla elde edilen hız alanı verileri, başlangıç noktası ve aynı anda salıverilen başka parçacıkların gözlemleri verildiğinde bir parçacığın konumunu bulmak için tasarlanmış olan Lagrange tahmin algoritmalarını incelemek için kullanılmıştır. Bu problem, özellikle kıyıda arama ve kurtarma çalışmaları, kirli atıkların saçınımı gibi uygulama alanlarından doğmuştur. Tahmin başarısını özünde hız alanının zamansal ve uzaysal kovaryansları belirler. Bu nedenle, hem Euler hem de Lagrange hız alanının ilintileri incelenmiştir. Tahmin edilebilirlik çalışmaları için var olan istatistiksel ortamı belirlemek üzere, uzay kovaryans fonksiyonları ve hız alanı spektrumu da bulunmuştur. Sonuçlar, regresyon tahmin algoritmasının hız alanı ilinti ölçekleri ve daha üstü ölçeklerde oldukça iyi başarıma sahip olduğunu göstermektedir.

Anahtar Sözcükler: türbülanslı akışlar, stokastik akışlar, Lagrange yörünge tahmini, döngü, ilinti, spektrum, Euler hız alanı

Introduction

The ocean turbulence is mostly related to chaotic motion of coherent eddies of different size and intensity filling up the upper layers. Recently, availability of high-frequency (HF) radar has permitted the measurement of eddies with high space and time resolution. Çağlar *et al.* (2006) have estimated Eulerian characteristics of the eddy turbulence from real data, based on a stochastic velocity field that represents coherent structures. Further analysis of the data is important to provide new perspectives on advanced ocean models.

In this paper, we study Lagrangian prediction based on HF radar data for Eulerian velocity. The prediction of particle trajectories in the ocean is of practical importance for problems such as searching for objects lost at sea, designing oceanic observing systems, and studying the spread of pollutants and fish larvae. We also investigate the correlation structure of the velocity field and the trajectories. Temporal covariance analysis, via Lagrangian and Eulerian approaches, is followed by spatial covariance analysis and spectral analysis.

The application of Lagrangian prediction to search and rescue operations relies on predictor data obtained from several drifters/buoys released simultaneously at different, but known, positions on the ocean surface. The problem is to predict the trajectory of an unobservable float at any time given its initial position and the trajectories of the predictor floats. In the presence of Eulerian velocity field, Lagrangian trajectories are first computed, and then a linear regression based prediction algorithm is implemented using the computed data.

The velocity correlations are closely related to the predictability problem. Higher correlations, or dependence, imply a stronger functional relationship between the trajectories, which improves the prediction. Therefore, Lagrangian and Eulerian correlations were also studied in the data. The previous work on stochastic flows for upper ocean turbulence in particular, Lagrangian prediction and eddy parameter estimation were reviewed in Piterbarg & Çağlar (2008). A Çinlar stochastic velocity model has been used in this study to parameterize the sub-mesoscale eddies detected in the data. The presence of submesoscale eddies at the coast have a direct impact on Lagrangian prediction. Motivated by such eddies, a Çinlar stochastic velocity field model represents the flow through randomization of the eddy features. This includes random arrival of eddies, randomization of their centres, amplitudes and radii, and their exponential decay with a constant parameter. The flow is incompressible and isotropic by construction.

Monin *et al.* (1971) give a classical account of correlation analysis of Lagrangian and Eulerian velocity fields. Recently, Lagrangian velocity correlations were considered in Mordant *et al.* (2002) who approached intermittency in turbulence from a dynamical point of view. Cressman *et al.* (2004) investigated turbulent fluid motion at the surface, but in an experimental setting where the flow is compressible. Mordant *et al.* (2004) described an original acoustic method to track the motion of tracer particles in turbulent flows and resolve Lagrangian velocity across the inertial range turbulence. More recently, Lagrangian velocity correlations and timescales were studied numerically using direct numerical simulation and a large-eddy simulation coupled with a subgrid Lagrangian stochastic model in Wei *et al.* (2006).

In the rest of the paper the available data and the applicability of both the data collection and the analysis to the coastal areas in Turkey and its vicinity are described first. Secondly, the computation of Lagrangian trajectories from Eulerian data is discussed. Then, Lagrangian prediction is performed with the linear regression algorithm. In the following section, the temporal correlation results are given for both Eulerian and Lagrangian velocity. For the data, spatial covariance functions and energy spectral density are computed. Finally, the conclusions are outlined.

HF Radar Data and Potential Study Areas

The data upon which our analysis is based and the applicability of this work to the Turkish coast and its vicinity are described as follows.

HF Radar Data

In this paper, Lagrangian prediction methods are applied, based on HF radar data for Eulerian velocity. The high-resolution radar data of surface velocity were obtained by satellite observation technology in the region between the Florida Current and the coast (Shay *et al.* 2000). These snapshots are sequenced by a constant time lag of 15 minutes and cover 28 days in total. At each snapshot, there are 91x91 velocity values, each representing a grid with 125m space interval, a total area of 11.25km by 11.25km. The velocity vector at a grid point with coordinates (x,y) at time t is denoted by

$$(U(x,y,t),\ V(x,y,t))$$

where U and V are zonal and meridianal components respectively.

Coasts of Turkey and the Surrounding Areas

The methods in the present work are demonstrated by the available data from the Florida coast. The new radar technology for collecting Eulerian data and the accompanying analysis are also applicable to the coastal areas in the Black Sea and the Mediterranean. More generally, this work contributes to efforts

to build a European capacity in ocean observing systems and their analysis. The need for more data collection and analysis in Europe was emphasized by several papers in Dahlin *et al.* (2003).

As for Lagrangian studies in the Black Sea, most observations are from autonomous drifting platforms for data collection called drifters, equipped with satellite communication devices. Most recently, Tolstosheev *et al.* (2008) presented the results of the Black Sea drifter monitoring in 2002–2006 within a number of international programs and projects. Long-term data were obtained about the circulation of the surface currents in particular. Similarly, Ivanov *et al.* (2007) revealed wind induced oscillator dynamics and single gyre structures during 2002–2003. The statistical description of the Black Sea near-surface circulation is given in Poulain *et al.* (2005) using the earlier drifter observations of 1999–2003.

The availability of HF radar technology makes high resolution Eulerian observations also possible in the Black Sea, especially useful in coastal areas for predictions such as the spread of pollutants. Likewise, Maderich (1999) simulated the transport of radionuclides in the chain system of the Mediterranean seas by incorporating submodels of the Black Sea, Azov Sea, Marmara Sea, Western and Eastern Mediterranean.

We demonstrate the analysis of Eulerian data for Lagrangian prediction, as Lagrangian trajectories can be efficiently computed numerically from such data. Therefore, much of the previous analysis based on drifter data can be replicated with HF radar observations. For example, Lipphardt *et al.* (2000) applied a spectral method that was first applied to drifter and model data from the Black Sea (Eremeev *et al.* 1992), using HF radar data and model velocities in Monterey Bay. Similarly, the approach of the present paper is applicable to various coastal areas, in particular those of Turkey.

Lagrangian Trajectories from Eulerian Data

In this section, we describe our method for obtaining Lagrangian trajectories from Eulerian velocity data by interpolating its values both in space and time. Then, the linear regression method is demonstrated as a proper approach for predicting unobserved trajectories from the observed ones.

Interpolation Method

The path (X_t, Y_t) of a particle starting from the point (x,y) at time 0, is found as the solution of the flow equations

$$\frac{dX_t}{dt} = \mathrm{U}(X_t, Y_t, t) \qquad X_0 = x \tag{1}$$

$$\frac{dY_t}{dt} = V(X_t, Y_t, t) \qquad Y_0 U = y \tag{2}$$

Since the velocity values are available on a grid and only for every 15 mins, the data are interpolated as required in the numerical solution procedure.

Equations (1) and (2) are solved by Runge-Kutta fourth-order method given by (Gerald & Wheatley 2004):

$$k_{1,x} = hU(x_n, y_n, t_n)$$

$$k_{2,x} = hU(x_n + k_{1,x}/2, y_n + k_{1,y}/2, t_n + h/2)$$

$$k_{3,x} = hU(x_n + k_{2,x}/2, y_n + k_{2,y}/2, t_n + h/2)$$

$$k_{4,x} = hU(x_n + k_{3,x}, y_n + k_{3,y}, t_n + h)$$

$$x_{n+1} = x_n + (k_{1,x} + 2k_{2,x} + 2k_{3,x} + k_{4,x})/6$$

$$k_{1,y} = hV(x_n, y_n, t_n)$$

$$k_{2,y} = hV(x_n + k_{1,x}/2, y_n + k_{1,y}/2, t_n + h/2)$$

$$k_{3,y} = hV(x_n + k_{2,x}/2, y_n + k_{2,y}/2, t_n + h/2)$$

$$k_{4,y} = hV(x_n + k_{3,x}, y_n + k_{3,y}, t_n + h)$$

$$y_{n+1} = y_n + (k_{1,y} + 2k_{2,y} + 2k_{3,y} + k_{4,y})/6$$

As required by these steps, the velocity values are not only needed at the last position and time, but also at intermediate values of the grid points and intermediate times even if the time step is chosen equal to the time resolution. We first interpolate in space. The grid points and the intermediate values are illustrated in Figure 1 in a 10x10 grid as an example. The point in space to be interpolated is marked by a square.

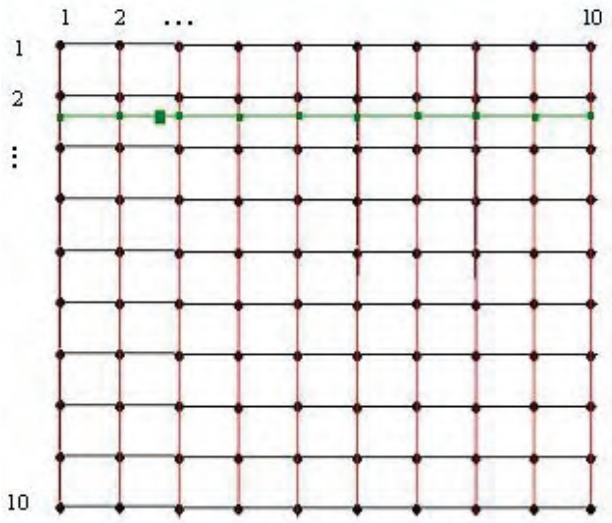

Figure 1. An example grid for velocity measurements and an intermediate position (marked with a square).

First, the velocity values are interpolated at the intersection points of the grid with the horizontal line that passes through the marked point. This is accomplished by passing cubic splines from the given data on the vertical grid lines, separately for each such intersection point. Passing cubic splines a final time using the interpolated values at the intersection points, we interpolate the velocity at the marked point. This value is obtained for several snapshots in order to interpolate in time as well. Since the time resolution is 15 mins, the snapshots for a complete day or even less yield a sufficiently large sample for interpolation in time at the market spatial point. The interpolation steps are performed for intermediate values in space and time as required for the Runge-Kutta method.

Trajectories

Initially, the time step was taken to be the time resolution 15 mins. Then, it was decreased until the computed trajectory converged within an error tolerance. In order, h= 0.5, 0.25, 0.125 time units were tried and the distance between two trajectories was found to be

$$D_{0.5-0.25} = \text{Max}\{2.2328, 2.4349\} = 2.4349 \text{ units}$$

$$D_{0.25-0.125} = \text{Max}\{0.5616, 0.2017\} = 0.5616 \text{ units}$$

where the unit is one grid spacing, namely 125 m, and the distance between the trajectories is taken to be the maximum distance in longitude and latitude directions. In view of the real dimensions of the sea and respective computational errors, we decided that h= 0.25, in which case the error is 0.5616x125 m, approximately 70 m. As shown in Figure 2, the visually closer paths are for the smaller values of h. The starting coordinates are (30,75) and the particle traverses the observation area vertically approaching its boundary in 1 hr.

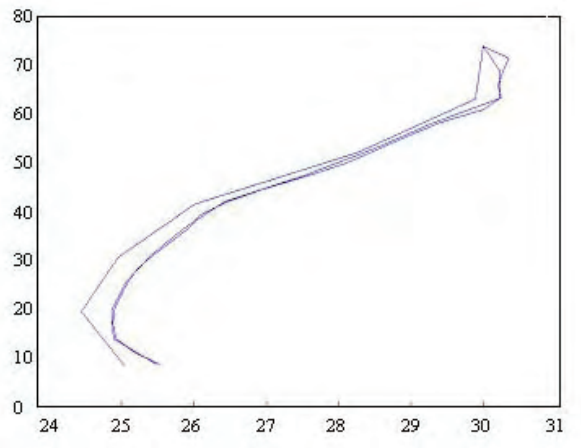

Figure 2. Particle trajectories computed with the time steps h= 0.50, 0.25, 0.125 for a total of 1 hr, from the Eulerian velocity field. Here, h denotes the fraction of the time unit, namely 15 mins. The two trajectories closer to each other correspond to h= 0.25 and h= 0.125.

Comparison with the Çinlar Model

The Çinlar stochastic velocity model represents eddy-rich flows by a sum of random number of eddies obtained by random scattering, amplification and dilation parameters. Thus, the velocity field is given by

$$\sum_{i=1}^{N} e^{-c(t-s_i)} a_i \, \mathrm{v}\left(\frac{\mathbf{r}-\mathbf{z_i}}{b}\right)$$

where **r**= (*x*,*y*), s_i are moments of eddy birth forming a Poisson process in time, hence *N* denotes the number of arrivals up to time *t*, $\mathbf{z}_i$ are eddy centres, a_i are amplitudes, b_i are radii of eddies, and as non-random parameters $c > 0$ is a decay rate and **v** is a deterministic velocity field with a compact support.

In Figure 3, a trajectory with the Çinlar model is obtained with the estimated parameters from the same Eulerian velocity data (Çağlar *et al.* 2006). Our experimentation with such trajectories has shown that it takes longer for a model particle to traverse the same distance than a simulated particle on the Eulerian data as in Figure 2. This confirms the discrepancy between the model and data about eddy decay. The average magnitude of eddies estimated from data decay linearly, whereas the model contains exponential decay to form a Markovian velocity field (Çağlar *et al.* 2006). Although the variances agree well, the model has more eddies on a given snapshot with the estimated parameters than the average number of eddies estimated from data. The observed eddies have larger average intensity to compensate for that number and yield equal variances. Therefore, in Figure 3, the particle moves from eddy to eddy and gets dispersed slowly rather than being scattered by a few strong eddies as in Figure 2. This discrepancy is aimed to be removed by modification of the model according to real eddy decay dynamics in future work.

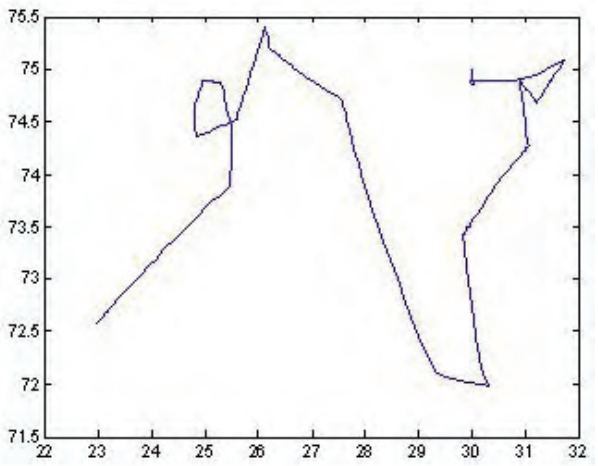

Figure 3. A particle trajectory simulated for 1 hr. from Çinlar velocity field model with parameters estimated from the Eulerian velocity field. Note that the particle path is less dispersed than that of Figure 2.

Prediction by Linear Regression

In this section, the linear regression method for Lagrangian prediction is summarized and implemented. The results are compared with those obtained by the centre of mass method (CM).

Linear Regression Method

An important application area of the Lagrangian approach is the prediction of the position of a lost item when observations of other close floating objects are available. Rigorously the problem is formulated as follows: given several particle paths, to predict an unobserved trajectory starting from a known position. The given trajectories are denoted by $\vec{r}_i$, $i = 1,\ldots,M$; in particular $\vec{r}_i(t)$ corresponds to the position vector of the i^{th} particle at time instant *t*. Suppose the unobserved path is $\vec{r}_M$. As the trajectories are random, the predictor that minimizes the mean square error is given by

$$\hat{\vec{r}}_M(T) = \mathrm{E}[\vec{r}_M(T) \mid \vec{r}_1(t),\vec{r}_2(t),\ldots,\vec{r}_{M-1}(t), 0 \le t \le T] \quad (3)$$

where E denotes the expectation operator. In other words, the predictor is the conditional expectation of the unobserved position given the observed trajectories. The error is defined as the difference between the true but observed value of $\vec{r}_M(T)$ and its predictor $\hat{\vec{r}}_M(T)$ in (3). In the linear regression method of prediction, the position at each instant is assumed to be a linear function of the initial position (Piterbarg & Özgökmen 2002) as

$$\vec{r}_i(t) = A(t)\vec{r}(0) + b(t) + y_i(t)$$

where $y_i(t)$ is the error and the functions $A(t)$ and $b(t)$ are to be estimated by the least squares method. The estimated values of *A* and *b* are found in terms of $\vec{r}_1(t),\ldots,\vec{r}_{M-1}(t)$ as

$$\hat{A}(t) = S(t)S(0)^{-1}$$
$$\hat{b}(t) = \vec{r}_c(t) - \hat{A}(t)\vec{r}_c(0)$$

where

$$\vec{r}_c(t) = \frac{1}{M-1}\sum_{i=1}^{M-1}\vec{r}_i(t)$$

$$S(t) = \sum_{i=1}^{M-1} (\vec{r}_i(t) - \vec{r}_c(t))(\vec{r}_i(0) - \vec{r}_c(0))^T$$

are the centre of mass and the dispersion matrix of *the observed* particles, respectively.

The linear regression method assumes that the unobserved path depends on the positions of the predicting trajectories. The prediction skill depends on the predictor (observed particle) density. In particular, when the numbers of predictors near the predicted (and unobservable particle) goes to infinity, the error tends toward zero. Another important detail is the initial positions of the predictors. A frequently used assumption is that the predict and is initially located close to the centroid of the polygon formed by the predictors. Such an initialization justifies the CM method which takes the predicted trajectory to be the centroid. Next, the results of the linear regression method are compared with the results of the method of centre of gravity.

Results for Lagrangian Prediction

Five predictors are initially placed on the corners of a pentagon. The particle to be predicted is positioned close to its centre. The trajectories of the particles used for prediction are first approximated as above and are assumed to be known. The known trajectories, as well as the trajectory predicted with the linear regression method, are shown in Figure 4. In this figure, the predictand is close to, but not exactly at the centroid.

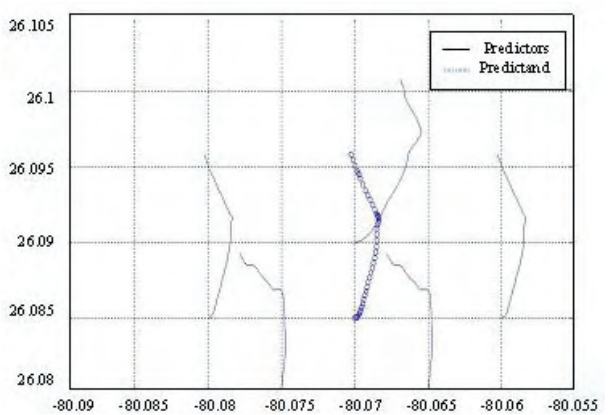

Figure 4. Known trajectories and the predicted trajectory with initial coordinates (–80.07, 26.085).

The trajectories predicted from the linear regression and CM methods are compared in Figure 5 with the true trajectory approximated from the Eulerian velocity field. Due to its nature, the CM algorithm starts with the centre of mass which is also taken as the initialization and is different from the actual starting point of the unknown trajectory.

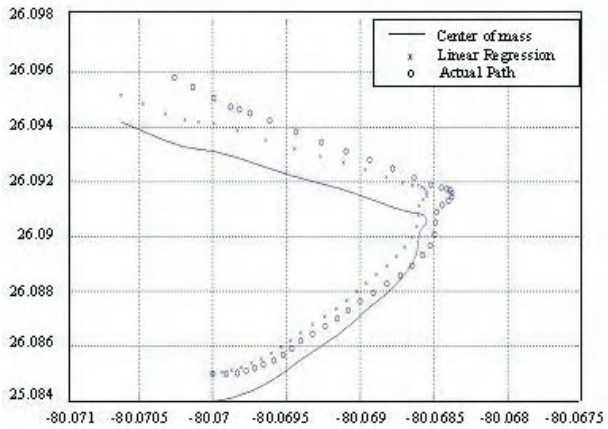

Figure 5. Predicted trajectories by two methods and the actual path computed from the velocity measurements.

The error is plotted against time in Figure 6, which shows that the linear regression does not exceed an error of 0.1 km. According to this result, in a sufficiently short time, the lost particle can be found within a circle of radius 100 m of the predicted trajectory. The error of 70 m that occurred at the calculation stage of approximate trajectories can be added to this margin of error.

If the predicted particle is initially placed exactly at the centroid, the error of the CM method is found to be lower, and comparable to that of the linear regression method. In general, we conclude that the linear regression method performs better as this type of initialization is not guaranteed in real applications. Also, this is a model independent prediction algorithm like the CM approach. In Piterbarg & Özgökmen (2002), the performance of the linear regression algorithm was compared with a Kalman filter type algorithm which makes use of flow statistics. It has also been found that regression algorithm performs better in view of simulations and real float data.

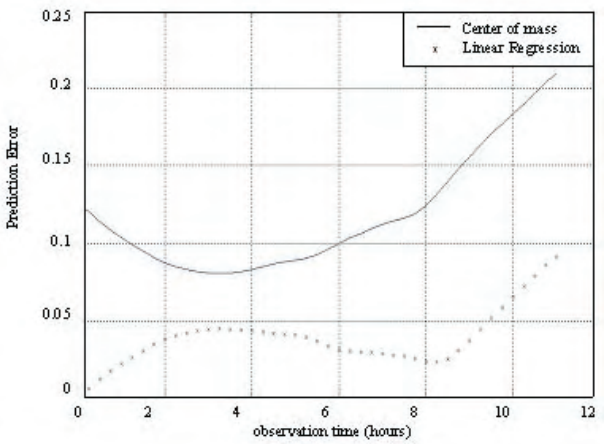

Figure 6. The errors of center of mass and linear regression methods for prediction of the true particle path.

Temporal Correlation Analysis and Results

In this section, the variance calculations will be performed for the spread of the particle trajectories and the correlation time scales will be found. The stochastic velocity model and the flow have already been analyzed by means of correlation analysis in Çağlar (2000, 2003). Therefore, the covariance analysis of the present work can be used to match the parameters of the model with data also from a Lagrangian perspective. In contrast, our earlier work (Çağlar *et al.* 2006) included parameter estimation only from Eulerian data.

The covariance function between processes *A* and *B*, is defined as:

$$R_{AB}(\tau) = E\left[\{A(t) - \mu_A\}.\{B(t+\tau) - \mu_B\}\right] \quad t,\tau \in R$$

The correlation function is defined as:

$$\rho_{AB}(\tau) = \frac{R_{AB}(\tau)}{R_{AB}(0)} \tag{4}$$

If *A* and *B* are different, the covariance (correlation) function is called the cross-covariance (cross-correlation) function, and when they are equal, it is called the autocovariance (autocorrelation) function (Bendat & Piersol 1993). Here, *A* and *B* are components of the Eulerian velocity field, i.e. they take values of *U* (zonal) and *V* (meridianal), and are not necessarily different.

There are two different approaches to determine how the flow is correlated in time; 'Lagrangian covariance' and 'Eulerian covariance'. Eulerian covariance corresponds to the covariance of the velocity data over time, whereas Lagrangian covariance relates to the particle followed in time and is found from the velocity data at the particle's position. In this paper, we only compute the autocovariance functions, and not the cross-covariance functions.

As indicated in Piterbarg & Özgökmen (2002) the error of the linear regression prediction algorithm is mostly determined by two parameters, the Lagrangian correlation time (Lagrangian velocity scale) and the velocity field space correlation radius. Here we focus on investigating the former since estimating the latter is problematic, given limited observations. The Eulerian correlation time is also briefly discussed, since it is related to the Lagrangian correlation time although an explicit functional relation is hard to find.

Lagrangian Autocovariance

The Lagrangian velocity autocovariance functions for a moving particle are defined as follows:

$$R_U^L(\tau) = E[u(t)u(t+\tau)]$$

$$R_V^L(\tau) = E[v(t)v(t+\tau)]$$

where *u* and *v* indicate the horizontal and vertical component of the velocity vector at the point where the particle resides at time *t*, respectively. We then introduce the following estimators for the autocovariance functions:

$$\hat{R}_U^L(\tau) = \frac{1}{T-\tau}\sum_{t=0}^{T-\tau} u(t)u(t+\tau)$$

$$\hat{R}_V^L(\tau) = \frac{1}{T-\tau}\sum_{t=0}^{T-\tau} v(t)v(t+\tau)$$

where *T* is the last time value before the particle leaves the grid, the time unit corresponds to 15 minutes for *t* which takes positive integer values, and $\tau = 0, 1, 2, \ldots$, must be less than *T*.

To calculate Lagrangian autocovariance, we need a particle's trajectory in the grid. For this purpose,

we choose 4 particles with respective initial positions (30,75), (35,70), (45,70), (50,75), and track them until they leave the grid. We obtain estimates of the Lagrangian autocovariance functions by averaging the functions due to these 4 particles. The estimates are plotted in Figure 7. We have used only four particles because obtaining Lagrangian velocity fields requires extensive computation time, and also the more the particles the earlier at least one particle leaves the grid in a short time. In Figure 7, the curves are smooth, indicating that the averaging over only 4 particles is sufficient. Note that the autocovariance function vanishes at about 60 time units, which is equivalent to 15 hrs.

From this estimate, autocorrelation functions $\hat{\rho}_U^L$ and $\hat{\rho}_V^L$ are easily determined by dividing the corresponding covariance function by the variance $\hat{R}_U^L(0)$ or $\hat{R}_V^L(0)$. Autocorrelation functions will be displayed in the sequel where correlation times are calculated.

Eulerian Autocovariance

Unlike the previous case, Eulerian autocovariance calculation is not related to whether the velocity field forces the particle to leave the grid or not. Eulerian covariance function depends on the coordinates of the data point, and indicates how the velocity is correlated throughout time at that particular point.

Eulerian autocovariance functions at point *(x,y)* are defined as follows

$$R_U^E(x,y,\tau) = E\left[U(x,y,t).U(x,y,t+\tau)\right]$$

$$R_V^E(x,y,\tau) = E\left[V(x,y,t).V(x,y,t+\tau)\right]$$

These expected values are estimated as

$$\hat{R}_U^E(\tau) = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}U(x,y,t)U(x,y,t+\tau)$$

$$\hat{R}_V^E(\tau) = \frac{1}{T-\tau}\sum_{t=1}^{T-\tau}\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}V(x,y,t)V(x,y,t+\tau)$$

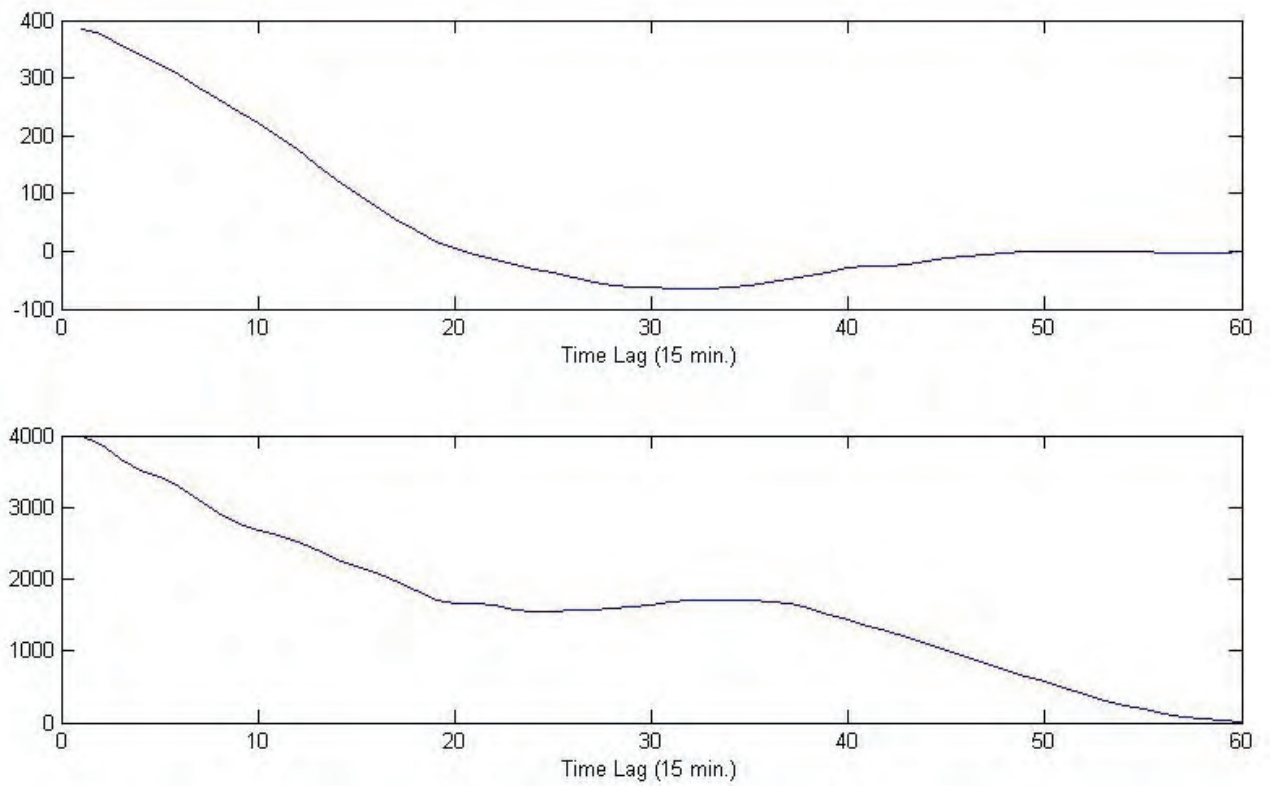

Figure 7. Top– Lagrangian Autocovariance for *u* with 4 particles; Bottom– Lagrangian Autocovariance for *v* with 4 particles.

where T is the latest time for which there is an observation.

We compute two Eulerian autocovariance functions, one for the first 14 day period, and the other for the last 14 day period, where the velocity field is stationary. Additionally, the estimations are carried on a 10-by-10 subgrid, which yields a total of 100 data points to be averaged. Estimated covariance functions are given in Figure 8. Eulerian autocorrelation functions $\hat{\rho}_U^E$ and $\hat{\rho}_V^E$ are found by using Equation (4) as before.

Lagrangian and Eulerian Correlation Times

The correlation time τ can be estimated using the autocorrelation function $\hat{\rho}$. It is called Lagrangian correlation time τ_L, if it is derived from the Lagrangian autocorrelation function; and Eulerian correlation time τ_E if it is derived from the Eulerian autocorrelation function. There are three approaches to estimate τ.

- **Method 1:** Calculating the area under the graph of $\hat{\rho}$ between $(0, \infty)$.
- **Method 2:** Calculating the area under the curve of $\hat{\rho}$ between 0 and the first real value where $\hat{\rho}$ becomes zero.
- **Method 3:** approximating $\hat{\rho}'(0)$.

Lagrangian and Eulerian autocorrelation functions are given in Figures 9 and 10, respectively. As a result, we see that the autocorrelation in vertical, or, in other words the vertical component is larger than the horizontal one. Note that the Gulf Stream is in this direction. Although the mean flow has been eliminated from the data, the variance remains. In Figure 9, Lagrangian autocorrelation diminishes at about 20 time units, equivalent to 5 hours in the horizontal direction, and at 60 time

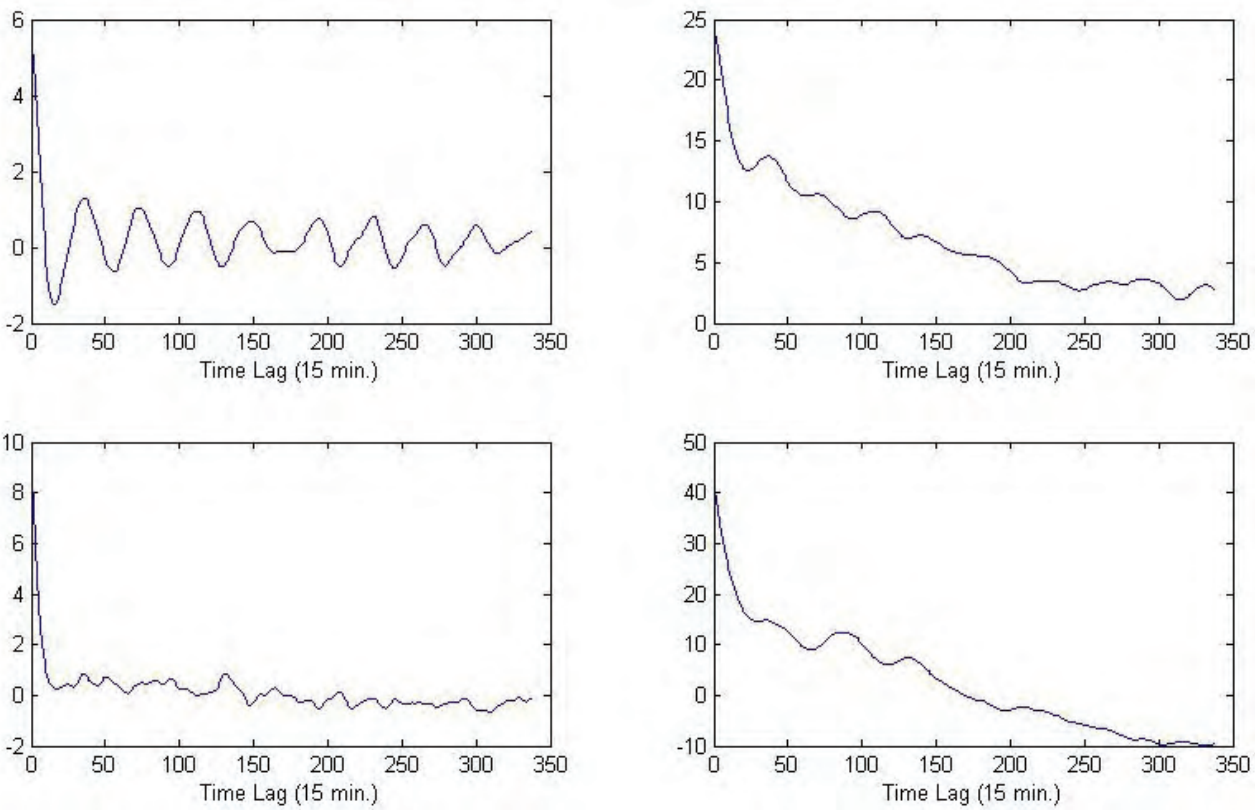

Figure 8. Top Left– Eulerian Autocovariance for *U*, First Period; Top Right– Eulerian Autocovariance for *V*, First Period; Bottom Left– Eulerian Autocovariance for *U*, Second Period; Bottom Right– Eulerian Autocovariance for *V*, Second Period.

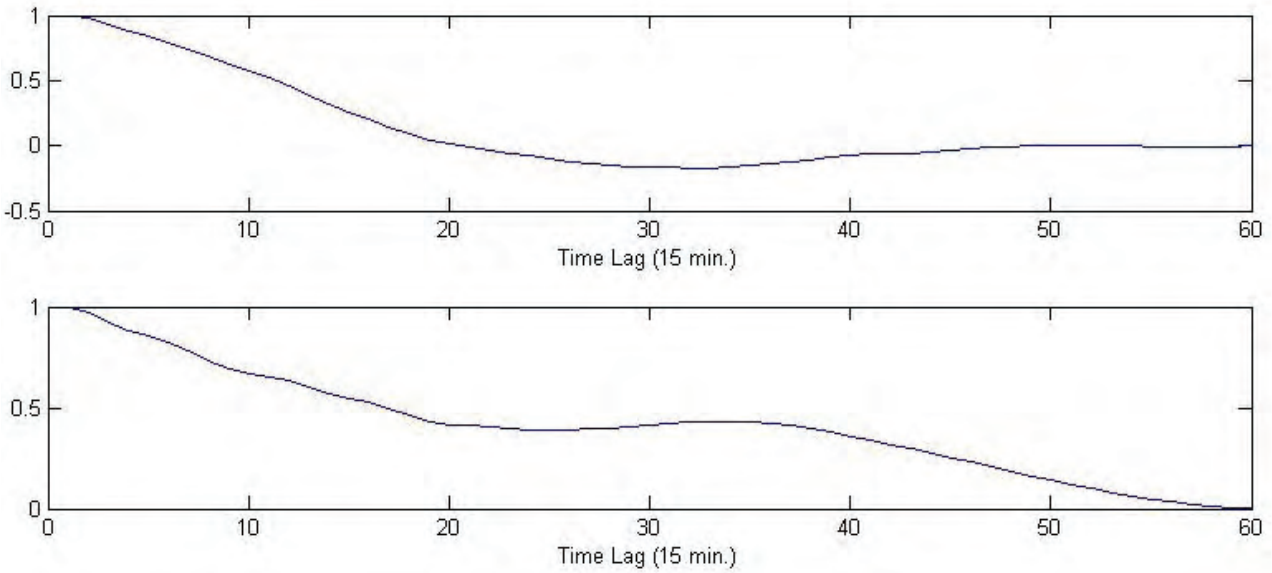

Figure 9. Top– Lagrangian autocorrelation for *u* with 4 particles; Bottom– Lagrangian autocorrelation for *v* with 4 particles.

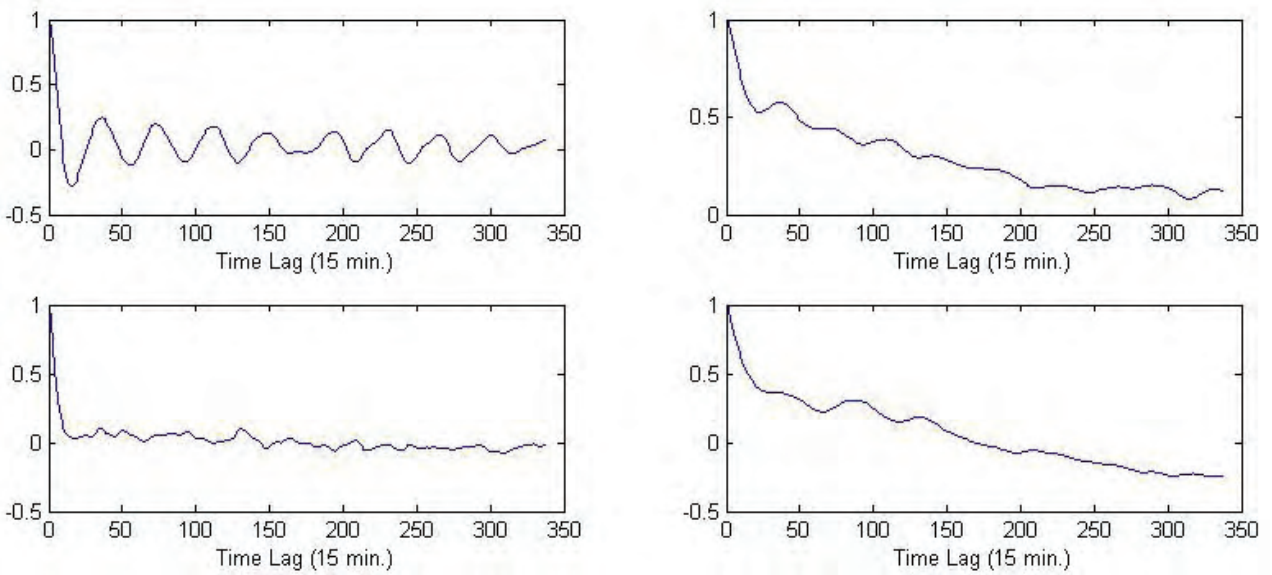

Figure 10. Top Left– Eulerian Autocorrelation for *U*, First Period; Top Right– Eulerian Autocorrelation for *V*, First Period; Bottom Left– Eulerian Autocorrelation for *U*, Second Period; Bottom Right– Eulerian Autocorrelation for *V*, Second Period.

units, equivalently 15 hours in the vertical direction. However, the behaviour of Eulerian autocorrelation is quite different, as shown in Figure 10. Eulerian autocorrelation in the first period can be compared with Lagrangian autocorrelation as it is obtained from the flow in the first period. Eulerian autocorrelation seems to decay faster in horizontal direction, while it oscillates for a longer time. On the other hand, it decays more slowly than Lagrangian autocorrelation in the vertical direction. As for comparison of the first and the second periods, Eulerian autocorrelation seems to decay more slowly in the second period in

the horizontal direction but with no oscillations. In the vertical direction, the autocorrelation becomes negative after a while indicating a slower decay to 0, possibly with further oscillations. These latter observations are consistent with the results of Çağlar *et al.* (2006) where the second period was indicated to have larger variance.

Numerical integrations for evaluating the correlation times τ_L and τ_E are accomplished using Simpson's method. The results can be found in Tables 1 and 2, which show the computed values in seconds as well as the same values in hours in parentheses. While Eulerian correlation time is close to Lagrangian correlation time for the horizontal component of the velocity, it is significantly larger than Lagrangian time for the vertical component as given in Table 2. Also, Eulerian correlation time is larger in the first period in both directions. This result can be reconciled with the discussion of Figure 10 as follows. The oscillations in the horizontal direction contribute to τ_E in the first period. In the second period, the autocorrelation is observed to be negative and any further oscillations have not been observed. This yields a lower correlation time τ_E as a result.

Table 1. Correlation Times for u or U (in seconds)

	Method 1	Method 2	Method 3
τ_L	6529 (1.8hrs)	8991 (2.5 hrs)	37030 (10.3 hrs)
τ_E (1st period)	11877 (3.3 hrs)	3816 (1.1 hrs)	12778 (3.6 hrs)
τ_E (2nd period)	4230 (1.2 hrs)	8235 (2.3 hrs)	7732 (2.2 hrs)

Table 2. Correlation Times for v or V (in seconds)

	Method 1	Method 2	Method 3
τ_L	21731 (6.0 hrs)	21731 (6.0 hrs)	32140 (8.9 hrs)
τ_E (1st period)	87406 (24.3 hrs)	87406 (24.3 hrs)	44130 (12.3 hrs)
τ_E (2nd period)	19162 (5.3 hrs)	40966 (11.4 hrs)	23635 (6.6 hrs)

Spatial Covariance Function

Spatial covariance indicates how the velocity data are correlated with respect to the distance between observation locations. In this case, our covariance and correlation functions will be two-dimensional. For each snapshot of time, we will obtain a covariance and a correlation function by averaging over time, or, more precisely, over two periods of time, one of which is the first 14 days of the 28 days data, and the other is the last 14 days.

As our observation area is finite, we need an estimator formula. Covariance functions at time t are estimated using the following expression

$$\hat{R}_{AB}(\Delta x,\Delta y,t)=\frac{1}{(M-\Delta x)(N-\Delta y)}\sum_{x=1}^{M-\Delta x}\sum_{y=1}^{N-\Delta y}A(x,y,t)B(x+\Delta x,y+\Delta y,t) \quad (5)$$

where Δx and Δy go from 1 to the number of data points of each side of the square grid, namely M= N= 91. Spatial covariance (5) is averaged over time for each period as

$$\hat{R}^1_{AB}(\Delta x,\Delta y)=\frac{1}{14m}\sum_{t=1}^{14m}\hat{R}_{AB}(\Delta x,\Delta y,t)$$

$$\hat{R}^2_{AB}(\Delta x,\Delta y)=\frac{1}{14m}\sum_{t=14m+1}^{28m}\hat{R}_{AB}(\Delta x,\Delta y,t)$$

where the unit of time is still 15 minutes and m=(4)(24)= 96 is the total number of snapshots in a given day. The components $\hat{\rho}_{AB}(\Delta x,\Delta y)$ of the so-called correlation tensor are defined as the corresponding covariance tensor component divided by the zero spatial lag covariance, i.e. variance (Mathieu & Scott 2000). Thus, we estimate the correlation functions using

$$\hat{\rho}^j_{AB}(\Delta x,\Delta y)=\frac{\hat{R}^j_{AB}(\Delta x,\Delta y)}{\hat{R}^j_{AB}(0,0)}$$

where j=1, 2.

We only give the estimated autocovariance functions in Figure 11, as correlation functions,

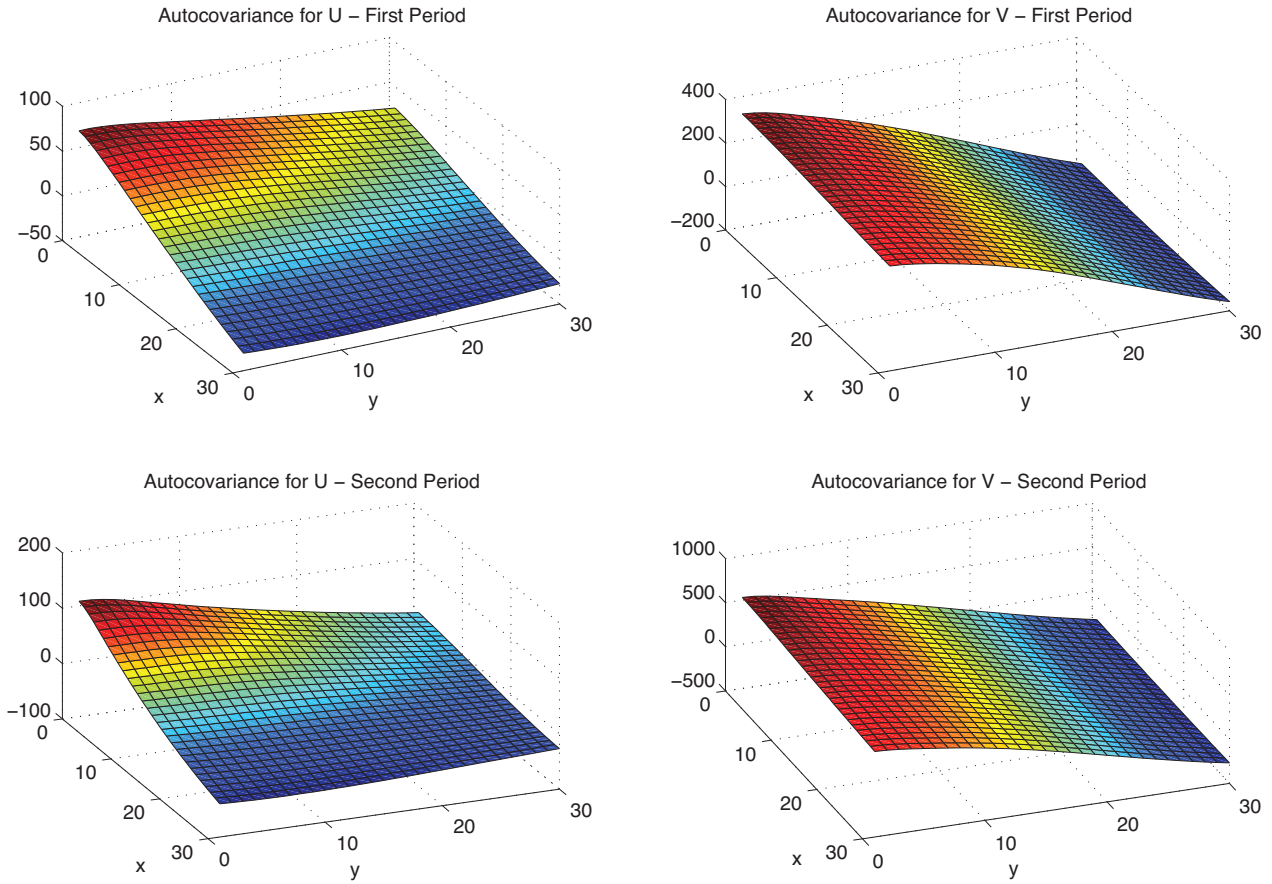

Figure 11. Spatial autocovariance functions – first and second period of observations. Velocity in the second period is more correlated.

being their scaled versions, are qualitatively the same. In computations, the space lag is changed up to 30 grids, which is about 1/3 of the total number (91). The spatial covariance is qualitatively similar in both time periods. The spatial autocovariance of the horizontal component U of the velocity decays fast in the horizontal direction at about 15 grid points, equivalently 1.9 km, but more slowly in the vertical direction, considerably decreasing at 30 grid spacing, namely 3.75 km. The autocovariance of the vertical component v decays in the vertical direction at about 20 grids, equivalently 2.5 km, while large covariance values persist in the horizontal direction even at 30 grid spacing, equivalently 3.75 km. In other words, the vertical velocity components are highly correlated along the same path. This may be an effect of the Gulf Stream in the same direction, although the mean flow has been subtracted.

Spectral Functions

In this section, the spectral functions to express Eulerian and Lagrangian velocity fields in the frequency domain are described, and the energy functions among other results for the spectra are computed.

Power Spectral Density

Spectra via Correlation Functions– Assuming that Eulerian velocity field is a stationary random process, the Power Spectral Density (PSD) can be computed by taking the Fourier transform of the autocovariance function estimate (Stanišić 1988; Bendat & Piersol 1993). The spectral functions are defined as the two dimensional Fourier transform

$$S_{AB}(k_x,k_y,t)=\frac{1}{(2\mathrm{p})^2}\iint R_{AB}(x,y,t).e^{-i(xk_x+yk_y)}dxdy$$

If *A* and *B* are different, then S_{AB} is called *cross-spectral density function*, and when they are equal, it is called *autospectral density function*, or more often *power spectral density function (PSD)* which can in brief be denoted as S_A (Bendat & Piersol 1993).

Again, an estimation method must be introduced because of the finite data. A discrete Fourier transform is applied to the obtained covariance functions, since spectral and covariance functions are Fourier transform pairs (Bendat & Piersol 1993). The resulting estimator is denoted by $\hat{S}_A$ below

$$\begin{aligned}\hat{S}_A^1(k_x,k_y)&=\frac{1}{14m}\sum_{t=1}^{14m}\hat{S}_A(k_x,k_y,t)\\ \hat{S}_A^2(k_x,k_y)&=\frac{1}{14m}\sum_{t=14m+1}^{28}\hat{S}_A(k_x,k_y,t)\end{aligned}\qquad(6)$$

where *A* stands for *u* or *v* and the estimation is performed for the two different time periods as before.

We obtain the energy function $E(k_x,k_y)$ by summing the two PSD functions as

$$\hat{E}^j(k_x,k_y)=\hat{S}_U^j(k_x,k_y)+\hat{S}_V^j(k_x,k_y)\qquad(7)$$

where *j*= 1,2 indicate the first and second 14-day periods. In isotropic turbulence, it is conventional to define the energy spectrum $\hat{E}(k)$ by integrating (7) over annular regions of (k_x, k_y) with $k = |(k_x, k_y)|$. Therefore, the resulting power spectrum depends on the wave number magnitude *k* and gives the energy spectral density versus *k* as shown in Figures 12 and 13 for the first and second observation periods, respectively. For comparison, the line $k^{-5/3}$ is also plotted on these figures. We conclude that the velocity data approximately obey the $k^{-5/3}$ spectra associated with Eulerian turbulence (Mathieu & Scott 2000, p. 242).

When the spectrums in (6) are considered separately, it is found that the meridianal component is more energetic than the zonal one during both periods while the total energy is higher in the second period. No significant extremes are found, meaning that there is no space periodicity in the velocity variability.

Spectra via Fourier Transform of the Data– PSD and the Cross-Spectral Density functions can alternatively be obtained without calculating the covariance functions as follows

$$S_{AB}(k_x,k_y,t)=E\left[\tilde{A}(k_x,k_y,t)\tilde{B}(k_x,k_y,t)\right]$$

where $\tilde{A}$ and $\tilde{B}$ are the Fourier transforms of *A* and *B*, respectively, given by

$$\tilde{A}(k_x,k_y,t)=\frac{1}{(2\mathrm{p})^2}\iint A(x,y,t).e^{-i(xk_x+yk_y)}dxdy$$

All the functions of the previous section can therefore be calculated using this point of view. The only difference in the estimation process in this technique is applying the discrete Fourier transform to the data itself and estimating $\hat{S}_{AB}$ by

$$\hat{S}_{AB}(k_x,k_y,t)=\tilde{A}(k_x,k_y,t)\,\tilde{B}(k_x,k_y,t)$$

using the functions $\tilde{A}$ and $\tilde{B}$. Due to averaging over 14 day periods, we define the two functions $\hat{S}^1$ and $\hat{S}^2$ in the same way as in Equation (6). Similarly, Equation (7) is used for finding the energy. The results are found to be very close to those in figures 12 to 13, as these are equivalent methods.

Conclusions

The high-resolution observations of surface velocity in the region between the Florida Current and the coast obtained by high-frequency radar are investigated to test Lagrangian prediction algorithms. The prediction skill is essentially determined by temporal and spatial covariances of the underlying velocity field. Higher correlations imply a stronger functional relationship between different trajectories. For this reason correlation analysis of both Lagrangian and Eulerian velocities has been carried out.

The North–South velocity components are highly correlated along the same path. This may be an effect of the Gulf Stream flowing in the same direction although the mean flow has been subtracted. While Eulerian correlation time is close to Lagrangian

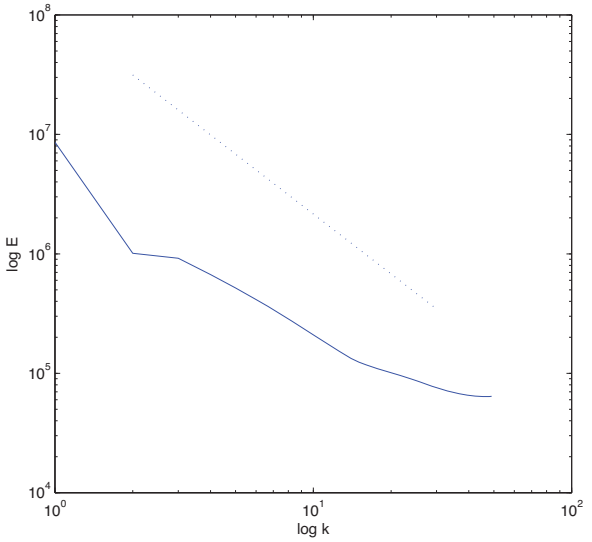

Figure 12. Energy spectra for the first period, the dashed line has slope –5/3 for comparison with the inertial range.

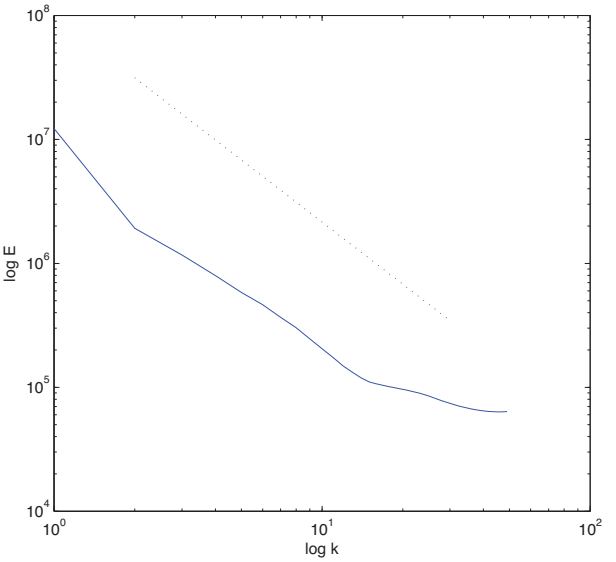

Figure 13. Energy spectra for the second period, the dashed line has slope –5/3 for comparison with the inertial range.

correlation time for the horizontal component of the velocity, it is significantly larger than Lagrangian time for the vertical component. Furthermore, Eulerian correlation time is larger in the first period of the observation horizon for both velocity components.

Space covariance functions and spectra of the velocity field have been presented to better illustrate statistical environments for the predictability studies. The meridianal component is more energetic than the zonal one during both periods while the total energy is higher in the second period. Spectrum decay with increasing wave number observed in all instances is typical for the upper ocean turbulence. No space periodicity in the velocity variability is found.

The results show that the regression prediction algorithm performs quite well on scales comparable with the velocity correlation scales and higher. Also, it is better than CM algorithm in terms of prediction error.

In future work, the covariance analysis of the present work can be used to match the parameters of the model with the data also from the Lagrangian perspective. Further improvement of the stochastic model and its integration in advanced ocean models would then be possible.

Acknowledgments

This research was supported by TÜBİTAK-NSF Project 103Y147 (Mine Çağlar, Taylan Bilal), an ONR grant N00014-09-1-0322, and a NSF Grant CMG-0530893 (Leonid Piterbarg). Mine Çağlar would like to thank former students Neşe Umut and Ferhan Türe for their studies which inspired this work.

References

Bendat, J.S. & Piersol, A.G. 1993. *Engineering Applications of Correlation and Spectral Analysis, Second Edition*. John Wiley and Sons.

Çağlar, M. 2000. Simulation of homogeneous and incompressible Çinlar flows. *Applied Mathematical Modelling* **4**, 297–314.

Çağlar, M. 2003. Dispersion of mass by two-dimensional homogeneous and incompressible Çinlar flows. *Applied Mathematical Modelling* **27**, 997–1011.

Çağlar, M., Piterbarg, L.I. & Özgökmen, T. 2006. Parameterization of submeso-scale Eddy-Rich flows using a Stochastic Velocity Model. *Journal of Atmospheric and Oceanic Technology* **23**, 1745–1758.

Cressman, J.R., Davoudi, J., Goldburg, W.I. & Schumacher. J. 2004. Eulerian and Lagrangian studies in surface flow turbulence. *New Journal of Physics* **6**, Article Number 53, DOI: 10.1088/1367-2630/6/1/053.

Dahlin, H., Flemming, N.C., Nittis, K. & Petersson, S.E. (eds) 2003. *Building the European Capacity in Operational Oceanography*. Volume 69. Proceedings 3rd EuroGOOS Conference, Elsevier Oceanography Series.

Eremeev, V.N., Ivanov, L.M. & Kirwan, A.D.Jr. 1992. Reconstruction of oceanic flow characteristics from Quasi-Lagrangian Data. 1. Approach and methamatical methods. *Journal of Geophysical Research-Oceans* **97**, 9743–9753.

Gerald, C.F. & Wheatley, P.O. 2004. *Applied Numerical Analysis*. Addison Wesley.

Ivanov, L.M., Melnichenko, O.V., Collins, C.A., Eremeev, V.N. & Motyzhev, S.V. 2007. Wind induced oscillator dynamics in the Black Sea revealed by Lagrangian drifters. *Geophysical Research Letters* **34** (13), L13609, **DOI:** 10.1029/2007GL030263.

Lipphardt, B.L., Kirwan, A.D., Grosch, C.E., Lewis, J.K. & Paduan, J.D. 2000. Blending HF radar and model velocities in Monterey Bay through normal mode analysis. *Journal of Geophysical Research-Oceans* **105**, 3425–3450.

Maderich, V.S. 1999. Reconstruction and prediction of radionuclide transport in the Mediterranean Sea chain. *Journal of Environmental Radioactivitiy* **43**, 205–219.

Mathieu, J. & Scott, J. 2000. *An Introduction to Turbulent Flow*. Cambridge University Press.

Monin, A.S., Yaglom, A.M. & Lumley, J.L. 1971. *Statistical Fluid Mechanics*. Dover.

Mordant, N., Delour, J., Le´veque, E., Arneodo, A. & Pinton, J.-F. 2002. Long time correlations in Lagrangian dynamics: a key to intermittency in turbulence. *Physical Review Letters* **89**, 254502, **DOI:** 10.1103/PhysRevLett.89.254502.

Mordant, N., Lévêque, E. & Pinton, J.-F. 2004. Experimental and numerical study of the Lagrangian dynamics of high Reynolds turbulence. *New Journal of Physics* **6**, 116, **DOI:** 10.1088/1367-2630/6/1/116.

Piterbarg, L.I. & Özgökmen, T.M. 2002. A simple prediction algorithm for the Lagrangian motion in two-dimensional turbulent flows. *SIAM Journal on Applied Mathematics* **63**, 116–148.

Piterbarg, L.I. & Çağlar, M. 2008. Research trends in Stochastic Flows Modeling the upper ocean turbulence. *In*: Alvarez, M.P. (ed), *Leading-Edge Applied Mathematical Modeling Research*, 67–110, Nova.

Poulain, P.M., Barbanti, R., Motyzhev, S. & Zatsepin, A. 2005. Statistical description of the Black Sea near-surface circulation using drifters in 1999–2003. *Deep Sea Research-Part I* **52**, 2250 –2274.

Stanišić, M.M. 1988. *The Mathematical Theory of Turbulence, Second Edition*. Springer-Verlag.

Shay, L.K., Cook, T.M., Haus, B.K., Martinez, J., Peters, H., Mariano, A.J., VanLeer, J., An, P.E., Smith, S., Soloviev, A., Weisberg, R. & Luther, M. 2000. VHF radar detects oceanic submesoscale vortex along Florida coast. *EOS Transactions* **81**, 209–213.

Tolstosheev, A.P., Lunev, E.G. & Motyzhev, V.S. 2008. Development of means and methods of drifter technology applied to the problem of the Black Sea research. *Oceanology* **48**, 138–146.

Wei , G.X., Vinkovic, I., Shao, L. & Simoëns, S. 2006. Lagrangian velocity correlations and time scales by a large-eddy simulation. *Comptes Rendus Mecanique* **334**, 298–303.

Appendix

Fourier Transform Implementation in Spectrum Calculation

When calculating spectra, before applying Fourier transform, we extended our function's domain in order to make the function even. Then, we applied the Fourier transform to the extended even function, and applied 'fftshift' function in Matlab, which is a simple swapping of the two sides of equal length of the array. Then as our Fourier-transformed function is complex, we took the magnitudes of the complex numbers. Finally, the relevant part of this extended array was cropped and plotted. The Matlab codes for the 1D and 2D cases are provided below in that order.

```
function R = evenfft(U)

[n N]  = size(U);
if N  == 1
   U = U'; N = n;
end
UU = [U(end:-1:2) U];
RR = fftshift(abs(fft(UU)));
R = RR(N:end);
return
```

```
function R = evenfft2(U)

[M N]  = size(U);
UU = [U(:,end:-1:2) U; U(end:-1:2,end:-1:2) U(end:-1:2,:)];
RR = fftshift(fft2(UU));
R = RR(M:end,N:end);
return
```

Note that in the 2D case, we did not utilize the whole 2D covariance functions, since there is not enough data for the high spatial lag values. We used a rectangular window (precisely of size 30) to filter out the unwanted effects of low sample size.