A simple predictive model for the eddy propagation trajectory in the northern South China Sea

Jiaxun Li[1,2], Guihua Wang[*1], Huijie Xue[3,4], and Huizan Wang[5]

[1]Department of Atmospheric and Oceanic Sciences, Institute of Atmospheric Science, Fudan University, Shanghai, China

[2]Naval Institute of Hydrographic Surveying and Charting, Tianjin, China

[3]State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China

[4]School of Marine Sciences, University of Maine, Orono, Maine, USA

[5]Institute of Meteorology and Oceanography, National University of Defense Technology, Nanjing, China

Corresponding author: Guihua Wang, Department of Atmospheric and Oceanic Sciences, Institute of Atmospheric Science, Fudan University, Shanghai, China. (wghocean@yahoo.com)

Abstract A novel predictive model is built for eddy propagation trajectory using the multiple linear regression method. This simple model has related various oceanic parameters to eddy propagation position changes in the northern South China Sea (~~SCS~~NSCS). These oceanic parameters mainly represent the effects of ~~planetary~~ β and mean flow advection on the eddy propagation. The performance of the proposed model is examined in the ~~SCS~~NSCS based on twenty years of satellite altimeter data, and demonstrates its significant forecast skills over a 4-week forecast window comparing to the traditional persistence method. It is also found that the model forecast accuracy is sensitive to eddy polarity and forecast season.

1. Introduction

Mesoscale eddies are coherent rotating structures that are ubiquitous over most of the world's oceans (Chelton et al., 2007). They play an important role in the transport of momentum, heat, mass and chemical and biological tracers, thereby become critical for issues such as general circulation, water mass distribution, ocean biology and climate change (Wang et al., 2012; Dong et al., 2014; Zhang et al., 2014; Ma et al., 2016; Li et al., 2017). Therefore, forecasting the eddy propagation positions accurately is not only important scientifically but also important practically for problems such as ocean observing systems designing, fishing planning, and underwater acoustic detecting.

Traditionally, ocean dynamical models were used as the tool of predicting the evolution of ocean eddies (Robinson et al., 1984). Since mesoscale eddies are often associated with strong nonlinear processes and their dynamical mechanisms are quite different, the operational forecast of eddies has been a big challenge to ocean numerical model. Much progress has been made in recent years in eddy-resolving ocean prediction. With the data assimilation and the increasing of model resolution, the model increases forecast skill. Daily forecast errors of eddy center positions in the northwestern Arabian Sea and Gulf of Oman are 44-68 km in 1/12^{o} global HYCOM model, and reach to 22.5-37 km in 1/32^{o} NLOM model (Hurlburt et al., 2008). The forecast skill and predictability of dynamical models can only be increased by better assimilation schemes (initialization), sufficient data (especially the subsurface), and improving resolution (physics and computing) (Rienecker et al., 1987; Oey et al., 2005). These restrictions preclude the all-pervading operational use of dynamical models when these initial data and computing power are not feasible due to some reasons. ~~To make a useful forecast, accurately updated boundary data, satellite and in situ observation data for assimilation must be available, and a fair degree of computer power is needed (Rienecker et al., 1987; Oey et al., 2005).~~ These restrictions preclude the all-pervading operational use of dynamical models when these initial data and

computer power are not feasible due to some reasons.

In this paper, we used a simple statistical method to predict the eddy positions 1-4 weeks in advance using only the past positions of the eddy and its surrounding fields. Our "test block" of ocean is the northern South China Sea (N~~SCS~~SCS). It is a semi-enclosed sea under the dramatic influence of the East Asian Monsoon and Kuroshio intrusion (Liu and Xie, 1999; Shaw, 1991). Due to the variable external forcing and complex topography, mesoscale eddies show obvious geographic distributions and various characteristics (Wang et al., 2003; Xiu et al., 2010), but the common character is the overall westward tendency of eddy trajectories no matter of the eddy polarity (Fig. 1). We will first analyze the pattern and dynamics of the common westward movement of eddies in the ~~SCS~~NSCS, then choose the potential predictors and develop a simple predictive model of eddy trajectories, and finally discuss the impact of eddy polarity and season on the model forecast accuracy.

2. Data and Methods

2.1 Data

The sea level anomalies (SLA) are from the Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO, ftp://ftp.aviso.oceanobs.com/) (Ducet et al., 2000). The product merges the measurements of TOPEX/Poseidon, European Remote Sensing Satellite (ERS-1/2), Geosat Follow-on, Jason-1/2, and Envisat, and spans the period from October 14, 1992 to August 7, 2013. Its temporal resolution is weekly, and its spatial resolution is 0.25^{o} latitude by 0.25^{o} longitude. To estimate the large-scale geostrophic currents, we use the absolute dynamic topography (ADT), which consists of the SLAs and a mean dynamic topography (MDT). The method for calculating the MDT was introduced by Rio and Hernandez (2004), and the data is also distributed by AVISO.

The monthly climatology of observed ocean temperature and salinity from U.S. Navy

Generalized Digital Environment Model (GDEM-Version 3.0) is used to calculate the phase speed of nondispersive baroclinic Rossby waves in the ~~SCS~~NSCS. It has a horizontal resolution of 0.25° latitude by 0.25° longitude, and 78 standard depths from 0 to 6600 m with the vertical resolution varying from 2 m at the surface to 200 m below 1600 m (Canes, 2009).

The ~~SCS~~NSCS eddy trajectory data is derived from the 3rd release of the global eddy dataset (http://cioss.coas.oregonstate.edu/eddies/). The eddy positions within their trajectories are recorded at 7-day time intervals. A detailed description of the eddy trajectory dataset can be found in Chelton et al. (2011). To forecast the eddy trajectory 1-4 weeks in advance using the last position of the eddy, only eddies with a lifetime of 5 weeks or longer are retained in this study.

2.2 The Maximum Cross-Correlation Method

The maximum cross-correlation (MCC) method is a space-time lagged technique, which can estimate the surface motions from time-sequential remote sensing images. It has been successfully used to track clouds from geosynchronous satellite data (Leese et al., 1971), to compute sea-ice motion (Ninnis et al., 1986) and advective surface velocities (Emery et al., 1986) from sequential infrared satellite images, and to determine the propagation velocities of ocean eddies from satellite altimeter data (Fu, 2006~~; Zhuang et al., 2010~~, 2009). The MCC method used in this study is the same as that of Fu et al. (2006, 2009), which is a little different with that of Emery et al. (1986). In the method of Emery et al., the correlations of the image in the subwindow with all the neighboring ones in the whole window at the next time are computed, and the speed and direction of the maximum correlation can be estimated. While in the method of Fu et al., the correlations of the SLA at a given location with all the neighboring SLA at various time lags are computed, and the speed and direction of the maximum correlation can be estimated. The reason of their difference may be due to the low time-space resolution of SLA comparing with other infrared satellite images.

带格式的：字体颜色：自动设置

The MCC method mainly consists of two procedures (Fu, 2009): first, the cross-correlations of the SLA time series (h) with others within a certain range box are computed for some time lags (ΔT) in multiples of 7 days (time resolution of SLA data) at each grid node location (x, y) as:

$$C_{x,y}(\Delta x, \Delta y, \Delta T) = \overline{h(x, y, t)h(x+\Delta x, y+\Delta y, t+\Delta T)} \quad (1)$$

where Δx and Δy are the spatial lags and the over bar means time averaging. Second, the position of the maximum correlation at each time lag (ΔT) is identified and a speed can be derived from the time lag and the distance of this position from the origin. Then an average speed vector (u, v) weighted by the correlation coefficients is calculated from the estimates at various time lags as:

$$(u, v) = \frac{\sum_i (\Delta x_i / \Delta T_i, \Delta y_i / \Delta T_i) C_i}{\sum_i C_i} \quad (2)$$

where C_i is the maximum correlation at ΔT_i, and Δx_i, Δy_i are the distances between the position of maximum correlation and the origin. The average velocities are then assigned to the eddy movement velocities at the given grid point.

To focus on the global mesoscale eddy, the time lags were limited to less than 70 days and the dimension of the window was less than 400 km (Fu, 2009). While in the NSCS, the time lags should be limited to less than 42 days, since many correlation coefficients are below the 95% confidence level at larger time lags (Zhuang et al., 2010). Besides, Chen et al. (2011) found that eddies propagate with 5.0-9.0 cm/s in the NSCS. Thus the search radius can be generally limited as 300 km (9.0 cm/s*42 days≈300 km) to reduce incidence of spurious MCC vectors. Since the mean flow and associated eddy propagation in the SCS have seasonal variability, we divided the weekly SLA data from 1992 to 2013 into four groups according to four seasons (winter: December-February, spring: March-May, summer: June-August, autumn: Septermber-November). Then the seasonal climatological eddy propagation velocities can be estimated from the same seasonal group at intervals of 1 week using the MCC method.~~To focus on the mesoscale in the SCS, the time lags are limited to less than 42 days, and the dimension of the search box is generally less than 300 km. To reduce~~

~~incidental spurious MCC vectors, the maximum speed is set to 30 cm/s, since the phase speeds of baroclinic Rossby waves in the SCS are mostly lower than this threshold (Cai et al., 2008).~~

2.3 The Multiple Linear Regression ~~Method~~Model

To develop a simple statistical predictive model for relating various oceanic parameters to eddy propagation position changes, the multiple linear regression method is used for developing such statistical forecast model. The multiple linear regression is a linear approach to modeling the relationship between the response and explanatory variables. This classical method has many practical uses in oceanography and meteorology, such as the prediction of Arctic sea ice extent (Zhang, 2015), the estimation of subsurface salinity profile (Bao et al, 2019), the estimation of anthropogenic CO_2 accumulation in the Southern Ocean (Matear and McNeil, 2003), the forecast of typhoon track (Aberson and Sampson, 2003) and intensity (Demaria and Kaplan, 1994), Maddan-Julian Oscillation forecast (Seo, 2008), and ENSO prediction (Dominiak and Terray, 2005).

In this study, the predictands (dependent variables) are the zonal and meridional displacements at each forecast time from the initial position (Table 1). The choice of the predictors based on physical analysis will be shown in detail in Section 3. Since the variables used for the regression involve different scales and units, it is inappropriate to use them directly, as it may cause the fitting to deviate from the physical constraints. Thus all the variables are normalized with their anomalies divided by their corresponding standard deviations before the regressing. After that, the normalized predicted zonal (meridional) displacement DX (DY) can be estimated using a multiple linear regression model:

$$DX_j = \sum_{i=1}^{n} a_{i,j} P_{i,} \quad j = 1, 4 \qquad (3)$$

带格式的：行距：1.5 倍行距

带格式的：字体颜色：自动设置

域代码已更改

$$DY_j = \sum_{i=1}^{n} b_{i,j} P_{i,} \quad j = 1, 4 \tag{4}$$

where the subscript *j* refers to the forecast interval (1-4 weeks), the subscript *i* refers to the serial number of normalized predictors (*P*), n represents the number of selected predictors; *a* and *b* donate the regression coefficients of predictors onto *DX* and *DY*, respectively.

There are a total of 8 regression equations, i.e., both the meridional and zonal directions for the weeks of 1-4. We can separate the whole eddy trajectories into two sets: one for regressing and the other for forecasting. At week-1, we used 1981 (76%) eddy trajectory segments (the distances between the eddy positions at 7-day time interval) of 283 eddy trajectories during 1992-2008 for regressing, and 623 (24%) eddy trajectory segments of 81 eddy trajectories during 2009-2013 for forecasting. The other forecast experiments for 2, 3, and 4 weeks maintain the same periods for regressing and forecasting. To evaluate the overall forecast ability of the model, the mean forecast error is defined as the averaged distance (D) between the predicted eddy positions and the satellite observed eddy positions following great circle distance (Ali et al., 2007):

$$D = R \bullet \arccos[\sin Y_o \sin Y_F + \cos Y_o \cos Y_F \cos(X_o - X_F)] \tag{5}$$

where R is the earth radius, X_o (X_F) and Y_o (Y_F) represent the observed (predicted) longitude and latitude in degrees, respectively.

带格式的：字体颜色：自动设置

带格式的：行距：1.5 倍行距

带格式的：降低量 6 磅

带格式的：字体：四号

带格式的：段落间距段前：0.2 行

3. ~~Results~~Dynamics of Eddy Propagation in the NSCS and Choice of Predictors

~~3.1 Dynamics of Eddy Propagation in the SCS~~

3.1 Pattern and Dynamical Analysis of Eddy Propagation in the NSCS

One of the most important steps in the development of a regression model is the choice of independent variables (predictors). In choosing the potential predictors, the candidates should have a physical link (direct or indirect) with the eddy propagation. To investigate the dynamical factors associated with eddy propagation in the NSCS, the pattern of eddy propagation speeds should be estimated firstly.

Instead of a Lagrangian description of the movement of individual eddies as reported in the previous studies (e.g., Wang et al., 2003; Chen et al., 2011), the space-time lagged MCC method provides an Eulerian description of the pattern of eddy propagation speeds (Fu, 2009). As shown in Fig. 2a and 2d, the MCC method has mapped the propagation speeds of ~~the~~ eddies in the ~~SCS~~NSCS for the winter and summer season, respectively. The propagation of ~~the~~ eddies is generally westward in the ocean interior and southward in the western boundary with the typical speed of 4-10 cm/s. The propagation direction of ~~the~~ eddies generated southwest of Taiwan is southwestward along the 200-2000 m isobaths, indicating the steering effects of the ocean's bathymetry. There are two distinct differences between the winter season and the summer season: one is that the eddy propagation speed in winter is relatively larger than that in summer; and the other is that the influence of the western boundary current can be clearly seen near 16°N-18°N along the Vietnam coast in winter, creating an organized band of southward eddy propagation pattern, while this cannot

带格式的: 行距: 单倍行距

带格式的: 标题 1, 段落间距段前: 0.2 行, 段后: 0.2 行, 行距: 单倍行距

带格式的: 字体: Times New Roman, 四号

带格式的: 标题 2, 段落间距段前: 5 磅, 段后: 5 磅, 行距: 多倍行距 1.73 字行

带格式的: 字体: Times New Roman, 四号

be found in summer. The different patterns of the eddy propagation speed in winter and summer have revealed several details of the mean flow in the ~~SCS~~SCS: the large-scale circulation under the influence of northeasterly winter monsoon is stronger than that in the southwesterly summer monsoon, and the robust western boundary current in winter becomes relatively weak and unorganized in summer.

Eddies also have their own westward drift under the planetary β effect in the absence of any mean flow (Nof, 1981, Cushiman-Roisin, 1994). Their propagation speed is approximately the phase speed of the first baroclinic Rossby waves with preferences for small poleward and equatorward deflection of cyclonic and anticyclonic eddies in the global ocean, respectively (Chelton et al., 2007). Theoretically, the phase speed of the first baroclinic Rossby wave is $C_{R1} = -\beta R_1^2$, where the first baroclinic Rossby radius of deformation R_1 is estimated using the climatological GDEM temperature and salinity data. Figure 2b (2e) shows the theoretical phase speed of nondispersive baroclinic Rossby waves calculated from GDEM winter (summer) climatological temperature and salinity data. The direction of the phase speed is due west and the magnitude increases from about 2 cm/s in the north latitude to 12 cm/s in the south latitude. It should be noted that the difference between the winter and summer distributions of the phase speed of the first baroclinic Rossby wave is relatively small. The underlying reason is that the variation of seasonal stratification in the upper layer has little effect on the seasonal distribution of the first baroclinic Rossby deformation radius (Chelton et al, 1998, Cai et al., 2008).

The differences between the satellite observed propagation speed (Fig. 2a and 2d) and the propagation speed induced by the β effect (Fig. 2b and 2e) in winter and summer are shown in Fig. 2c and 2f, respectively, which may represent the propagation speed caused by the advection of mean flow. To further illustrate the advection effect of mean flow, the winter (summer) mean dynamic topography is

带格式的：行距：固定值 26 磅

superimposed on the propagation speed caused by the mean flow. As can be seen, there is a good spatial correlation (0.61 in the zonal direction and 0.52 in the meridional direction, both of which are significant at the 95% confidence level) between the cyclonic eddy propagation speed advected by the mean flow and the large scale surface cyclonic circulation in winter, both of which are centered northwest of the Luzon Island (Fig. 2c). Due to the weak cyclonic gyre in the ~~northern SCS~~NSCS, the spatial correspondence in summer is not as obvious as that in winter (Fig. 2f). Since the propagation speed induced by the β effect is westward, this tendency is reinforced by the mean flow in the north, but compensated by the mean flow in the south. Because the mean flow in the south is not so strong, it is not able to reverse eddy propagation from its westward motion induced by the β effect as in the Antarctic Circumpolar Current region (Klocker and Marshall, 2014) no matter in winter or summer.

To explore other possible causes of eddy propagation, Fig. 3a shows the annual mean eddy propagation speed. The most striking pattern is that the eddy propagation speed is accelerated markedly on the northern continental shelf of the ~~SCS~~NSCS (also can be seen in Fig. 2a and 2d), corresponding well to the region of negative maximum meridional topographic $\beta_T = \frac{f}{H}\frac{dH}{dy}$, where H is the water depth. Their correlation is -0.40, which is significant at the 95% confidence level. This relatively good correspondence suggests that besides the planetary β effect and advection of mean flow, the topographic β effect also contributes to the eddy propagation in some regions where the bathymetry gradient cannot be neglected.

3.2 Choice of Predictors

As mentioned above, the mean flow advection and the effects of β (both planetary

带格式的: 字体: Times New Roman

带格式的: 行距: 固定值 26 磅

and topographic) are closely related with the eddy propagation. These factors should be considered as the potential predictors, and the seasonal climatological eddy zonal and meridional motions (U_CLIM V_CLIM) derived from the MCC are calculated to represent the effects of β and the mean flow advection. It is noted that we have tried to decompose U_CLIM and V_CLIM into the effects of β and the mean flow advection and incorporate them into the regression model, but found no improvement of the forecast skill.

In reality, the large-scale circulation evolves during the forecast period, this synoptic effect of mean flow advection should also be taken into account. To help account for the time variation of the mean flow advection, the current zonal and meridional absolute geostrophic flows (U_ADT, V_ADT) derived from the satellite data are evaluated at the beginning of the forecast time along the eddy trajectory. Besides, the persistence factors should also be considered in the regression model, since they contain the "latest" pattern of eddy propagation under the effects of β and the mean flow advection. The chosen persistence factors are the initial eddy position (LON, LAT) and the eddy motion past 1-week (U_PAST, V_PAST). All the chosen eight predictors are listed in Table 2, and can be derived along the eddy trajectories. They can be divided into two categories: 1) P1-P6 related to climatology and persistence, i.e., "static predictors", and 2) P7-P8 related to the changing environmental conditions, i.e., "synoptic predictors".

The relative contribution of each predictor on each forecast period is illustrated by the normalized regression coefficient (Table 3). The larger the normalized regression coefficient, the greater its contribution to the individual forecast equation. Persistence factors (U_PAST, V_PAST) are initially the most important predictors, while after 2 weeks the most important predictors are the climatology factors (U_CLIM, V_CLIM). The synoptic predictors (U_ADT, V_ADT) contribute less to the forecast equations

带格式的：制表位： 5.81 字符, 左对齐

带格式的：行距：固定值 26 磅

comparing with persistence and climatology. The underlying reason may be that the week to week variations are too large so the representation of the initial U_ADT and V_ADT to the actual velocities in the 4-week window is not as good as the U_CLIM and V_CLIM.

~~3.2 Model Development~~

~~To develop a simple statistical predictive model for relating various oceanic parameters to eddy propagation position changes, the multiple linear regression is used for developing such statistical forecast models. This method has been successfully used in the forecast of tropical cyclone (TC) tracks (Neumann and Randrianarison, 1976; Aberson and Sampson, 2003), hurricane intensity (Demaria and Kaplan, 1994) and ENSO (Knaff and Landsea, 1997). In this study, the predictands (dependent variables) are the zonal and meridional displacements at each forecast time from the initial position. In choosing the potential predictors (independent variables), two factors are considered. First, the candidates should have a physical link (direct or indirect) with the eddy propagation. Second, the candidates must be available and accessible in advance. Based on these two considerations, eight potential predictors that are associated with eddy propagation are chosen (Table 1). All of these are derived along the eddy trajectories.~~

~~These eight predictors can be divided into two categories: 1) those related to climatology and persistence, i.e., "static predictors", and 2) those related to the changing environmental conditions, i.e., "synoptic predictors". The static predictors consist of the first six predictors, while the last two are the synoptic predictors. Since the initial eddy position (LON, LAT) and the eddy motion past 1-week (U_PAST, V_PAST) represent the initial conditions of the eddy, these persistence factors are crucial for the next position of the eddy. The climatological eddy zonal and meridional motions (U_CLIM V_CLIM) derived from the MCC method are chosen to take into account the effects of β and the mean flow advection, as discussed in~~

Section 3.1. In reality, the large scale circulation evolves during the forecast period, but this effect is not taken into account in the climatology and persistence factors. To help account for the time variation of the mean flow advection, the current zonal and meridional absolute geostrophic flows (U_ADT, V_ADT) derived from the satellite data are evaluated at the beginning of the forecast time along the eddy trajectory. The relative contribution of each predictor on each forecast period is illustrated by the normalized regression coefficient (Table 2). To generate the normalized coefficients, both the predictors and the predictands are normalized before they are incorporated into the regression model. The larger the normalized regression coefficient, the greater its contribution to the individual forecast equation. Persistence factors (U_PAST, V_PAST) are initially the most important predictors, while after 2 weeks the most important predictors are the climatology factors (U_CLIM, V_CLIM). The synoptic predictors (U_ADT, V_ADT) contribute less to the forecast equations comparing with persistence and climatology. The underlying reason may be that the week to week variations are too large so the representation of the initial U_ADT and V_ADT to the actual velocities in the 4-week window is not as good as the U_CLIM and V_CLIM.

There are a total of 8 regression equations, i.e., both the meridional and zonal directions for the weeks of 1-4. We can separate the data into two sets: one for regressing and the other for forecasting. At week-1, we used 1981 (76%) eddy trajectory segments (the distances between the eddy positions at 7-day time interval) of 283 eddy trajectories during 1992-2008 for regressing, and 623 (24%) eddy trajectory segments of 81 eddy trajectories during 2009-2013 for forecasting. The other forecast experiments for 2, 3, and 4 weeks maintain the same periods for regressing and forecasting. To evaluate the overall forecast ability of the model, the mean forecast error is defined as the averaged distance (D) between the predicted eddy positions and the satellite observed eddy positions following great circle distance (Ali et al., 2007):

$$D = R \cdot \arccos[\sin Y_o \sin Y_F + \cos Y_o \cos Y_F \cos(X_o - X_F)]$$

~~where R is the earth radius, X_o (X_F) and Y_o (Y_F) represent the observed (forecast) longitude and latitude in degrees, respectively.~~

4. Performance of the Multiple Regression Model

带格式的：标题 1，段落间距段前：0.2 行，段后：0.2 行，行距：单倍行距

4.1 Comparison with the persistence method

To evaluate the performance of our prediction model, the persistence method (no change of propagation speed from the initial state, Fig. 4a) and our model are used to predict the eddy trajectories during 2009-2013. Then the root-mean-square error (RMSE) and correlation coefficient between the predicted and actual longitudes (latitudes), mean distance errors of our model and persistence method over a 4-week horizon are computed.

Table ~~3~~ 4 lists the~~ number of sampling cases,~~ comparison of prediction results.~~root mean square error (RMSE) and correlation coefficient between the predicted and actual longitudes (latitudes), mean distance errors of our model and persistence method (no change of propagation speed from the initial state, Fig. 4a) over a 4-week horizon.~~ It shows that ~~the~~ our multiple linear regression model ~~developed model~~ beats the persistence method and indicates our model has some forecast skill (Table 5): the RMSE between the predicted and the actual longitudes (latitudes) throughout the 4-week horizon is ~~0.33~~32.7-~~0.89~~89.2 km (~~0.30~~29.5-~~0.73~~73.5 km) ~~degrees~~ with the correlation coefficients >0.93 (>0.95).

As an example, Fig. 5 compares the 1-2 weeks forecast performances of our model (blue) and the persistence method (green) with the observation (red). Generally, the eddy trajectory predicted 1-2 weeks in advance by our model coincides well with the observed trajectory with an overall average error of 27.6 km (week-1) and 42.5 km

(week-2), and even the convoluted pattern can be reproduced properly (Fig. 5 (right)) though the mean error is slightly larger than the smooth case. In contrast, although the persistence forecast trajectory at week-1 is relatively consistent with the observation (Fig. 5a and 5b), the persistence method cannot forecast the eddy trajectories properly when the forecast horizon increases (Fig. 5c and 5d). To further compare their differences, their forecast distance errors are normalized with the Rossby radius on each forecast grid over 4-week forecast window, respectively. The correlation between the normalized forecast distance errors of the persistence method and our model decreases from 0.67 at week-1 to 0.38 at week-4. This is consistent with the above judgement and confirms the superiority of our multiple linear regression model over the persistence method.~~This shows the superiority of our forecast model over the persistence method.~~

4.2 Sensitive Performance of Different Eddy Polarity and Season

~~**3.3 Sensitive Performance of Different Eddy Polarity and Season**~~

Previous studies have shown that anticycloinc eddies and cyclonic eddies in the ~~SCS~~NSCS have different dynamic characteristics, such as generation sites, rotation speeds and propagation trajectories, and the seasonal variability of these eddies is robust (Wang et al., 2006; Wang et al., 2008; Li et al., 2011). Two natural questions arise: 1) is there any difference on the model forecast ability between anticyclonic eddies (Fig. 1a) and cyclonic eddies (Fig. 1b)? 2) If so, is there any difference on the forecast ability for one type of eddies in winter (Fig. ~~6a~~ 7a and ~~7a~~8a) and summer (Fig. ~~6b~~ 7b and ~~7b~~8b)? This section will explore the different model performances on two types of eddies and during different seasons in the ~~SCS~~NSCS.

The period considered for regressing and predicting the anticyclonic eddy and cyclonic eddy positions is the same as that used in developing the predictive model in

Section 3.2. The mean forecast errors of anticyclonic (cyclonic) eddies from week-1 to week-4 are 36.9 km (41.1 km), 62.6 km (68.1 km), 81.0 km (88.5 km), and 102.0 km (108.2 km), respectively (Fig. ~~1e~~6). These results show that the predicted trajectory errors of anticyclonic eddies are less than those of cyclonic eddies in all forecast horizon, and the maximum error difference can reach 7.5 km at week-3. To investigate the underlying reasons of different model performances for anticyclonic eddies and cyclonic eddies, we use the persistence error ($CC' = \sqrt{AB^2 + BC^2 - 2AB \bullet BC \bullet \cos\theta}$ in Fig. 4a) at week-1 as an index to measure the difficulty of trajectory forecast. The underlying reason in physics is that CC', which includes the effects of winding angel (θ, measuring the trajectory curvature) and the eddy propagation distances in the former and latter periods (AB and BC, measuring the eddy propagation speed), is an integral characteristic of eddy trajectory. The correlation between this integrated index and eddy trajectory forecast error is relatively high with R=0.62 (Fig. 4b), which is significant at the 95% confidence level and shows its ability of measuring the inherent difficulty of trajectory forecast: the larger the index, the more difficult the trajectory forecast, thus the larger the forecast error. Because the indices (mean persistence errors) of all the anticyclonic and cyclonic eddy trajectories in the ~~SCS~~NSCS are 46.6 km and 53.0 km, respectively, it is not difficult to understand why the mean forecast error of anticyclonic eddy trajectories is smaller than that of cyclonic eddy trajectories in the ~~SCS~~NSCS. The index difference between anticyclonic and cyclonic eddy trajectories is caused by these different trajectory patterns (Fig. 1a and 1b), which could be due to the opposing meridional drifts of anticyclonic and cyclonic eddies expected from the combination of β effect and self-advection (Morrow et al., 2004).

Figure ~~6c~~ 7c (Fig. ~~7e~~8c) shows the mean forecast errors of anticyclonic (cyclonic) eddy trajectories in winter and summer over a 4-week horizon. Because the mean persistence error (42.0 km) of anticyclonic eddy trajectories in winter is smaller than that (51.9 km) in summer, as expected, the mean forecast error of anticyclonic eddy

trajectories in winter is smaller than that in summer for all cases. This is also the case for the cyclonic eddy: since the mean persistence error (54.6 km) of cyclonic eddy trajectories in winter is relatively larger than that (52.8 km) in summer, the mean forecast error of cyclonic eddy trajectories in winter is larger than that in summer. The index difference of one type of eddy trajectories between winter and summer is also caused by the different trajectory patterns. Why do the anticyclonic and cyclonic eddies follow different trajectories in winter (Fig. ~~6a~~ 7a and ~~7a~~8a) and summer (Fig. ~~6b~~ 7b and ~~7b~~7b)? One possible dynamical reason is the different interactions between the eddies and seasonal mean flows. Other underlying factors such as eddy generation mechanisms and eddy-topography interactions in different seasons may also contribute. This is beyond the scope of this study and needs further investigation using numerical models.

5. Summary and Discussion

In this study we have investigated the underlying dynamics of the eddy propagation in the ~~SCS~~NSCS and found the propagation of ~~SCS~~NSCS eddies is mainly driven by the combination of the planetary β effect and mean flow. In addition, the topographic β effect also has some contribution to the eddy propagation where the bathymetry gradient cannot be neglected, like the steep continental shelf in the ~~northern~~ ~~SCS~~NSCS (Fig. 1a).

Based on the dynamical analysis, a simple statistical predictive model for relating various oceanic parameters to eddy propagation position changes is developed using the multiple linear regression method. This model is made up of ~~two~~ eight predictands (zonal and meridional displacements over 1-4 weeks) and eight predictors (six static predictors, two synoptic predictors). The six static predictors are associated with the initial position, the zonal and meridional motions past 1-week, and the climatological eddy zonal and meridional motion. The other two synoptic predictors account for the

time variation of the mean flow advection. Results showed that this simple model has significant forecast skills over a 4-week forecast horizon comparing the traditional persistence method. Moreover, the model performance is sensitive to eddy type and forecast season: 1) the predicted trajectory errors of anticyclonic eddies are smaller than those of cyclonic eddies; 2) the predicted trajectory errors of anticyclonic eddies in winter are smaller than those in summer; while the contrary is the case for the cyclonic eddy. The predictive model performance strongly depends on the inherent difficulty of trajectory forecast.

Although the performance of the proposed predictive model is encouraging, it could be refined further. Further improvement may be possible by including the effect of eddy-eddy interactions on the eddy propagation, which is supposed to help induce the eddy trajectory curve or loop (Early et al., 2011). Another possible improvement is to use artificial neural network (ANN) in developing the forecast model. ANN has been successfully used in the predicting cyclone tracks (Ali et al., 2007) and loop current variation (Zeng et al, 2015). ANN can represent both linear and non-linear relationships learned directly from the data being modeled. It mainly contains three layers: the input layer, the hidden layer, and the output layer. To be consistent with the multiple linear regression model, both the input layer and the output layer include the same predictors and predictands as the regression model, respectively. The hidden layer consists of two layers of neural variables. Through iterations on backward propagation of the error, the neural network learns by itself to achieve an optimum weighting function and a minimum error. The forecast errors of ANN for 1-4 weeks are listed in Table 4. We can see that some improvements (0.3-4.2 km during 1-4 weeks forecast horizon) have been shown comparing with the linear regression method. Recently, Jiang et al. (2018) have found the deep learning algorithm of neural networks performs better than the simple ANN for the parameterization of

typhoon-ocean feedback in typhoon forecast models. These enhancements (both physics and algorithms) are topics warranting future research and development.~~Another possible improvement is to replace the multiple regression method with machine learning (ML) techniques. Ashkezari et al. (2016) have shown that the ML methods particularly stand out in analyzing complex systems yet not fully understood, like estimating the eddy lifetime. These enhancements are topics warranting future research and development.~~

Data availability. The SLA and MDT data can be downloaded from AVISO (ftp://ftp.aviso.oceanobs.com/), and the ~~SCS~~NSCS eddy trajectory data can be derived from the 3rd release global eddy dataset (http://cioss.coas.oregonstate.edu/eddies/).

Acknowledgements. This work is supported by the National Key Research and Development Program of China (2017YFC1404103), the National Basic Research Program of China (2013CB430301), the Natural Science Foundation of China (91428206, 41621064, 91528304), the National Programme on Global Change and Air-Sea Interaction (GASIIPOVAI-04), and the China Postdoctoral Science Foundation (2016M601493).

References

Aberson, S. D. and Sampson, C. R.: On the predictability of tropical cyclone tracks in the northwest pacific basin, Mon. Wea. Rev., 131, 1491-1497, 2003.

Ali, M. M., Kishtawal, C. M., and Jain, S.: Predicting cyclone tracks in the north Indian Ocean: An artificial neural network approach, Geophys. Res. Lett., 34, L04603, http://doi.org/10.1029/2006GL028353, 2007.

Bao, S., Zhang, R. Wang, H., Yan, H., and Yu, Y.: Salinity profile estimation in the Pacific Ocean from satellite Surface salinity observations, J. Atmos. Oceanic. Technol., 36, 53-68, 2019.

525 Cai, S., Long, X., Wu, R., and Wang, S.: Geographical and monthly variability of the first
526 baroclinic rossby radius of deformation in the south china sea, J. Mar. Syst., 74, 711-720, 2008.
527 Canes, M. R.: Description and evaluation of GDEM-V3.0, Rep.NRL/MR/7330-09-9165, Nav. Res.
528 Lab, Washington, D. C, 2009.
529 Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale
530 eddies, Prog. Oceanogr., 91, 167-216, 2011.
531 Chelton, D. B., Schlax, M. G., Samelson, R. M., and de Szoeke, R. A.: Global observations of
532 large oceanic eddies, Geophys. Res. Lett., 34, L15606, http://doi.org/10.1029/2007GL030812,
533 2007.
534 Chelton, D., DeSzoeke, R., and Schlax, M.: Geographical variability of the first baroclinic rossby
535 radius of deformation, J. Phys. Oceanogr., 28, 433-460, 1998.
536 Chen, G., Hou, Y., and Chu, X.: Mesoscale eddies in the South China Sea: Mean properties,
537 spatiotemporal variability, and impact on thermohaline structure, J. Geophys. Res., 116, C06018,
538 http://doi.org/10.1029/2010JC006716, 2011.
539 Cushiman-Roisin, B.: Introduction to Geophysical Fluid Dynamics, Prentice Hall, 320 pp, 1994.
540 Demaria, M. and Kaplan, J.: A statistical hurricane intensity prediction scheme (SHIPS) for the
541 Atlantic basin, Weather Forecast, 9, 209-220, 1994.
542 Dominiak, S., and Terray, P.: Improvement of ENSO prediction using a linear regression model
543 with a southern Indian Ocean sea surface temperature predictor, Geophys. Res. Lett., 32,
544 L18702, doi:10.1029/2005GL023153, 2005.
545 Dong, C., Mcwilliams, J. C., Liu, Y., and Chen, D.: Global heat and salt transports by eddy
546 movement, Nature Communications, 5, 3294, 2014.
547 Ducet, N., Le Traon, P. Y., and Reverdin, G.: Global high-resolution mapping of ocean circulation
548 from TOPEX/Poseidon and ERS-1 and -2, J. Geophys. Res., 105, 19477–19498,
549 http://doi.org/10.1029/2000JC900063, 2000.
550 Early, J. J., Samelson, R. M., and Chelton, D. B.: The evolution and propagation of
551 quasigeostrophic ocean eddies, J. Phys. Oceanogr., 41, 1535-1555, 2011.
552 Emery, W. J., Thomas, A. C., Collins, M. J., Crawford, W. R., and Mackas, D. L.: An objective
553 method for computing advective surface velocities from sequential infrared satellite images, J.
554 Geophys. Res., 91, 12865–12878, http://doi.org/ 10.1029/JC091iC11p12865, 1986.

带格式的：图案：清除

555 Fu, L.-L.: Pathways of eddies in the South Atlantic Ocean revealed from satellite altimeter
556 observations, Geophys. Res. Lett., 33, L14610, http://doi.org/10.1029/2006GL026245, 2006.
557 Fu, L.-L.: Pattern and speed of propagation of the global ocean eddy variability, J. Geophys.
558 Res., 114, C11017, http://doi.org/10.1029/2009JC005349, 2009.
559 Hurlburt, H., Chassignet, E., Cummings, J., Kara, A., Metzger, E., Shriver, J., Smedstad, O.,
560 Wallcraft, A., and Barron., C.: Eddy Resolving Global Ocean Prediction, Washington D. C.,
561 American Geophysical Union Geophysical Monograph Series, 353-381, 10.1029/177GM21,
562 2008.
563 Jiang, G.-Q., Xu, J., & Wei, J.: A deep learning algorithm of neural network for the
564 parameterization of typhoon-ocean feedback in typhoon forecast models, Geophys. Res. Lett.,
565 45, https://doi.org/10.1002/2018GL077004, 2018.
566 Klocker, A. and Marshall, D. P.: Advection of baroclinic eddies by depth mean flow, Geophys. Res.
567 Lett., 41, 3517–3521, http://doi.org/10.1002/2014GL060001, 2014.
568 Leese, J. A., Novak, C. S., and Clark, B. B.: An automated technique for obtaining cloud motion
569 from geosynchronous satellite data using cross correlation, J. Appl. Meteor., 10, 18-132, 1971.
570 Li, J., Zhang, R., and Jin, B.: Eddy characteristics in the northern South China Sea as inferred
571 from Lagrangian drifter data, Ocean Sci., 7, 1575-1599, 2011.
572 Li, J., Wang, G., and Zhai, X.: Observed cold filaments associated with mesoscale eddies in the
573 South China Sea, J. Geophys. Res. Oceans, 122, 762–770,
574 http://doi.org/10.1002/2016JC012353, 2017.
575 Liu, W. T. and Xie, X.: Space-based observations of the seasonal changes of South Asian
576 monsoons and oceanic response, Geophys. Res. Lett., 26, 1473-1476, 1999.
577 Ma, X. and Coauthors: Western boundary currents regulated by interaction between ocean eddies
578 and the atmosphere, Nature, 535, 533, 2016.
579 Matear, R. J., and McNeil, B. I.: Decadal accumulation of anthropogenic CO2 in the Southern
580 Ocean: A comparison of CFC-age derived estimates to multiple-linear regression estimates,
581 Global Biogeochem. Cycles, 17(4), 1113, doi:10.1029/ 2003GB002089, 2003.
582 Morrow, R., Birol, F., Griffin, D., and Sudre, J.: Divergent pathways of cyclonic and anti-cyclonic
583 ocean eddies, Geophys. Res. Lett., 31, L24311, http://doi.org/10.1029/2004GL020974, 2004.
584 Ninnis, R. M., Emery, W. J., and Collins, M. J.: Automated extraction of pack ice motion from

advanced very high resolution radiometer imagery, J. Geophys. Res., 91, 10725–10734, http://doi.org/10.1029/JC091iC09p10725, 1986.

Nof, D.: On the β-induced movement of isolated baroclinic eddies, J. Phys. Oceanogr., 11, 1662-1672, 1981.

Oey, L.-Y., Ezer, T., Forristall, G., Cooper, C., DiMarco, S., and Fan, S.: An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico, Geophys. Res. Lett., 32, L12611, http://doi.org/10.1029/2005GL023253, 2005.

Rienecker, M. M., Mooers, C. N. K., and Robinson, A. R.: Dynamical interpolation and forecast of the evolution of mesoscale features off northern California, J. Phys. Oceanogr., 17, 1189-1213, 1987.

Rio, M.-H. and Hernandez, F.: A mean dynamic topography computed over the world ocean from altimetry, in situ measurements, and a geoid model, J. Geophys. Res., 109, C12032, http://doi.org/10.1029/2003JC002226, 2004.

Robinson, A. R., Carton, J. A., Mooers, C. N. K., Walstad, L. J., Carter, E. F., Rienecker, M. M., Smith, J. A., and Leslie, W. G.: A real-time dynamical forecast of ocean synoptic/mesoscale eddies, Nature, 309, 781-783, 1984.

Seo, K. H., Wang, W., Gottschalck, J., Zhang, Q., Schemm, J.-K., Higgins, W., and Kumar, A.: Evaluation of MJO Forecast Skill from Several Statistical and Dynamical Forecast Models, J Clim., 22(9), 2372-2388, 2009.

Shaw, P. T.: Seasonal variation of the intrusion of the Philippine Sea water into the South China Sea, J. Geophys. Res., 96, 821-827, 1991.

Wang, G., Chen, D., and Su, J.: Generation and life cycle of the dipole in the South China Sea summer circulation, J. Geophys. Res., 111, C06002, http://doi.org/10.1029/2005JC003314, 2006.

Wang, G., Chen, D., and Su, J.: Winter eddy genesis in the eastern South China Sea due to orographic wind-jets, J. Phys. Oceanogr., 38, 726–732, http://doi.org/10.1175/2007JPO3868.1, 2008.

Wang, G., Li, J., Wang, C., and Yan, Y.: Interactions among the winter monsoon, ocean eddy and ocean thermal front in the South China Sea, J. Geophys. Res., 117, C08002, http://doi.org/10.1029/2012JC008007, 2012.

Wang, G., Su, J., and Chu, P. C.: Mesoscale eddies in the South China Sea observed with altimeter data, Geophys. Res. Lett., 30, 2121, http://doi.org/10.1029/2003GL018532, 21, 2003.

Xiu, P., Chai F., Shi, L., Xue, H., and Chao, Y.: A census of eddy activities in the South China Sea during 1993–2007, J. Geophys. Res., 115, C03012, http://doi.org/10.1029/2009JC005657, 2010.

Zhang, R.: Meachanims for the low frequency of summer Arctic sea ice extent, Proc. Natl. Acad. Sci., 112(15), 4570-4575, 2015.

Zhang, Z., Wang, W., and Qiu, B.: Oceanic mass transport by mesoscale eddies, Science, 345, 322, 2014.

Zeng, X., Li, Y., and He, R.: Predictability of the loop current variation and eddy shedding process in the Gulf of Mexico using an artificial neural network approach, J. Atmos. Oceanic. Technol., 32, 1098-1111, 2015.

Zhuang W., Du, Y., Wang, D., Xie, Q., and Xie, S.-P.: Pathways of mesoscale variability in the South China Sea, Chin. J. Oceanol. Limnol., 28, 1055-1067, 2010.

Figure and Table Captions

Figure 1. The trajectories of (a) anticyclonic and (b) cyclonic eddies with lifetime ≥ 5 weeks in the northern South China Sea (SCS). The solid circle represents the ending position of each trajectory. In Fig. 1a, TI: Taiwan Island, LI: Luzon Islands, VN: Vietnam. The two isobaths are for 200 m and 2000 m, respectively.

Figure 2. Winter climatology of (a) eddy propagation speed directions (vectors) and magnitudes (color, cm/s), (b) The phase speed directions (vectors) and magnitudes (color, cm/s) of the first baroclinic Rossby wave. (c) The speed difference (vectors) between (a) and (b) superimposed on the winter mean absolute dynamic topography (color, cm). (d), (e) and (f) are the same as (a), (b) and (c), respectively, but for the summer.

Figure 3. (a) Annual mean of eddy propagation speed directions (vectors) and magnitudes (color, cm/s). (b) Meridional distribution of the topographic β effect (color shading).

Figure 4. (a) Schematic of persistence method. A, B, and C are three observed eddy positions on the trajectory every 1 week interval. C' is the predictive eddy position 1 week in advance by persistence method, that is BC'=AB. Thus CC' is the persistence error at week-1. (b) Scatterplot of persistence error versus forecast error of our model at week-1 with best fit linear regression.

Figure 5. A comparison of the satellite observed trajectory (red), the predicted trajectory by our model (blue) and persistence trajectory (green) at (a) week-1, (c) week-2. (b), (d) are the same as (a) and (c), respectively, but for a recurved trajectory. The biweekly eddy positions on each trajectory are shown by the solid circles.

Figure 6. Comparison of the mean forecast errors between anticyclonic eddies (red) and cyclonic eddies (blue) over a 4-week window.

Figure 7. The trajectories of anticyclonic eddies in (a) winter and (b) summer with

lifetime≥5 weeks in the northern South China Sea. The solid circle represents the ending position of each trajectory. (c) Comparison of their mean forecast errors over a 4-week window.

Figure 8. The same as Fig. 6, but for the cyclonic eddies.

Table 1. The eight predictands used in the predictive model.

Table 2. The eight predictors used in the predictive model.

Table 3. Normalized regression coefficients $a_{i,j}$ ($b_{i,j}$) for use with the eddy zonal (meridional) motion prediction equation.

Table 4. Comparison of mean forecast distance errors (km) of the persistence, multiple linear regression (MLR), and artificial neural network (ANN) method.

Table 5. Statistics of our muliple linear regression model for different forecast time of eddy propagation positions in terms of longitudes (latitudes).

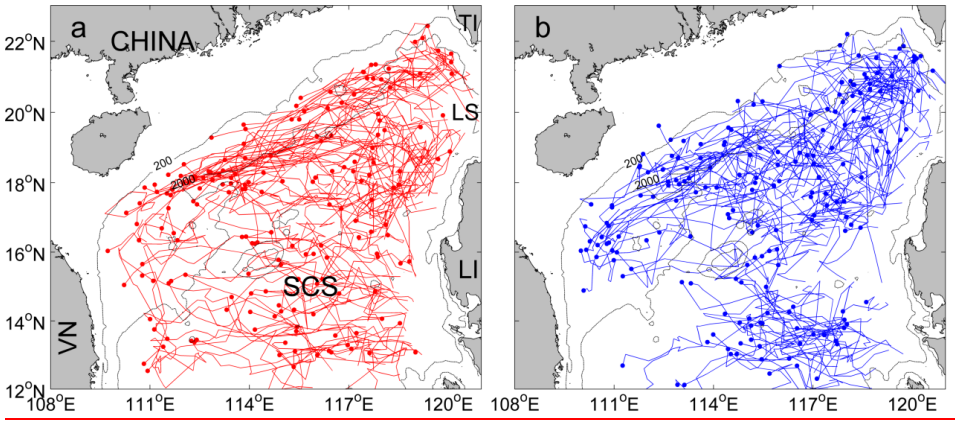


Figure 1. The trajectories of (a) anticyclonic and (b) cyclonic eddies with lifetime ≥ 5 weeks in the South China Sea (SCS). The solid circle represents the ending position of each trajectory. In Fig. 1a, TI: Taiwan Island, LI: Luzon Islands, VN: Vietnam. The two isobaths are for 200 m and 2000 m, respectively.

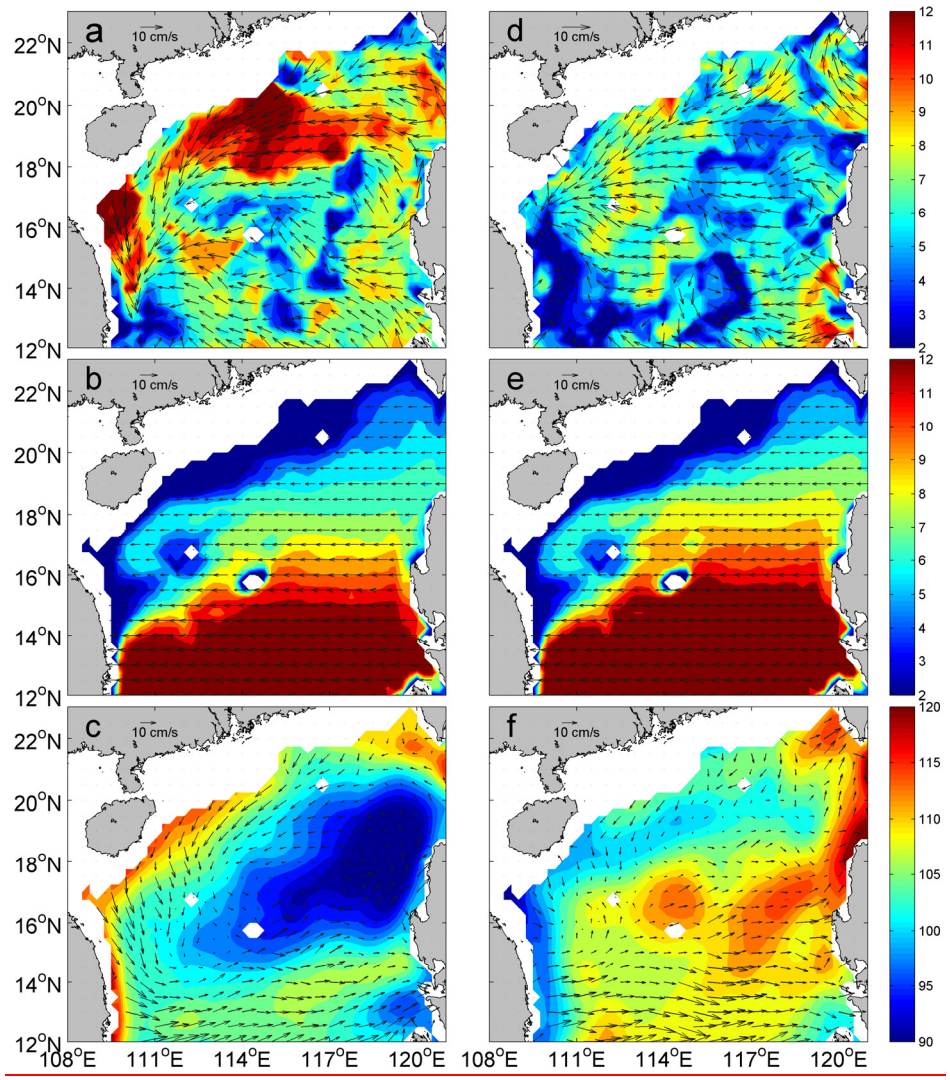


Figure 2. Winter climatology of (a) eddy propagation speed directions (vectors) and magnitudes (color, cm/s), (b) The phase speed directions (vectors) and magnitudes (color, cm/s) of the first baroclinic Rossby wave. (c) The speed difference (vectors) between (a) and (b) superimposed on the winter mean absolute dynamic topography (color, cm). (d), (e) and (f) are the same as (a), (b) and (c), respectively, but for the summer.

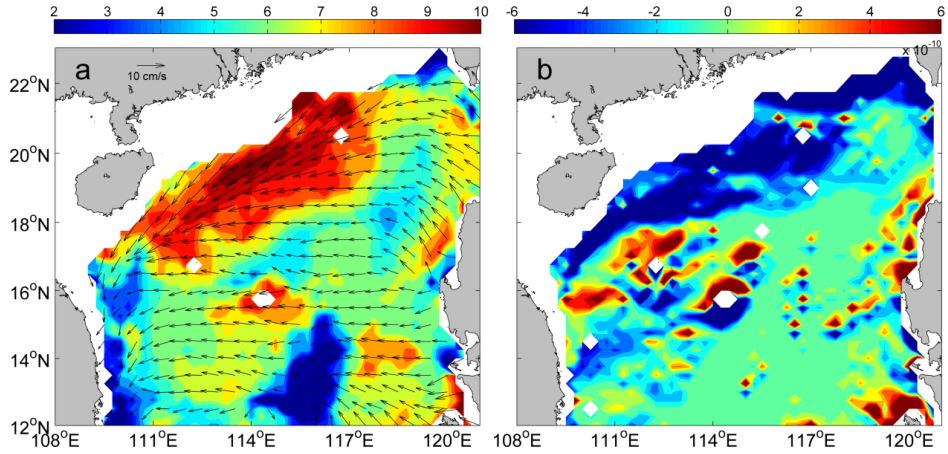
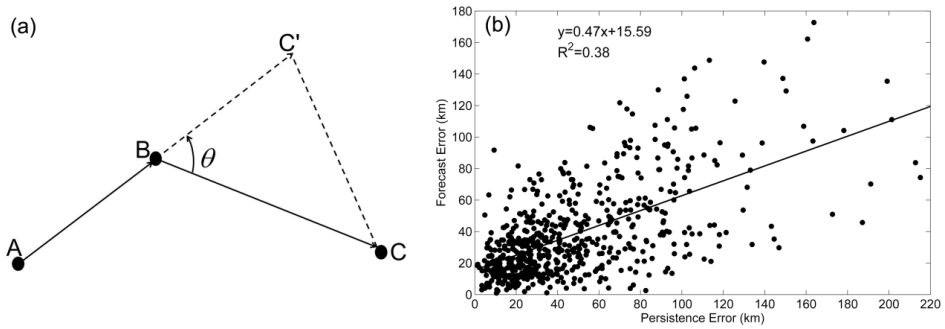


Figure 3. (a) Annual mean of eddy propagation speed directions (vectors) and magnitudes (color, cm/s). (b) Meridional distribution of the topographic β effect (color shading).



694

695

696 **Figure 4.** (a) Schematic of persistence method. A, B, and C are three observed eddy
697 positions on the trajectory every 1 week interval. C' is the predictive eddy position 1
698 week in advance by persistence method, that is BC'=AB. Thus CC' is the persistence
699 error at week-1. (b) Scatterplot of persistence error versus forecast error of our model
700 at week-1 with best fit linear regression.

701

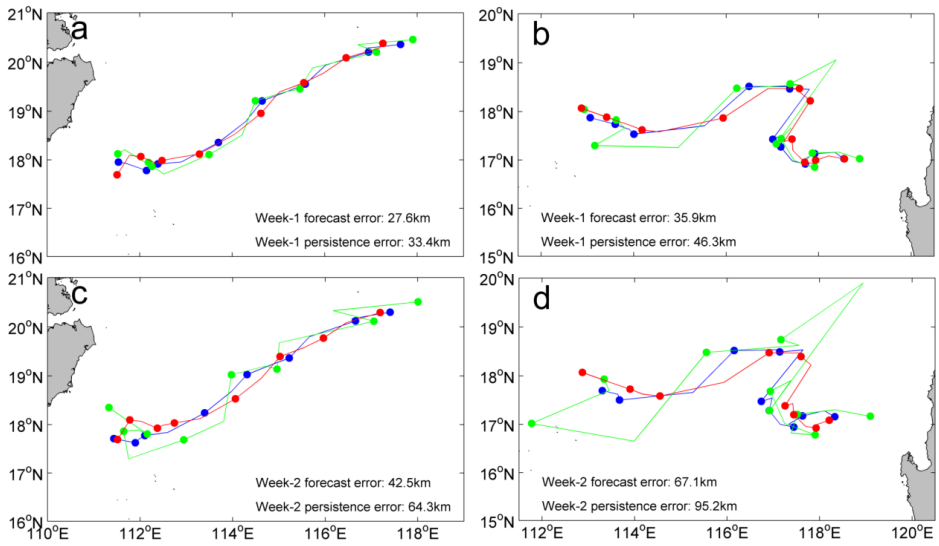


Figure 5. A comparison of the satellite observed trajectory (red), the predicted trajectory by our model (blue) and persistence trajectory (green) at (a) week-1, (c) week-2. (b), (d) are the same as (a) and (c), respectively, but for a recurved trajectory. The biweekly eddy positions on each trajectory are shown by the solid circles.

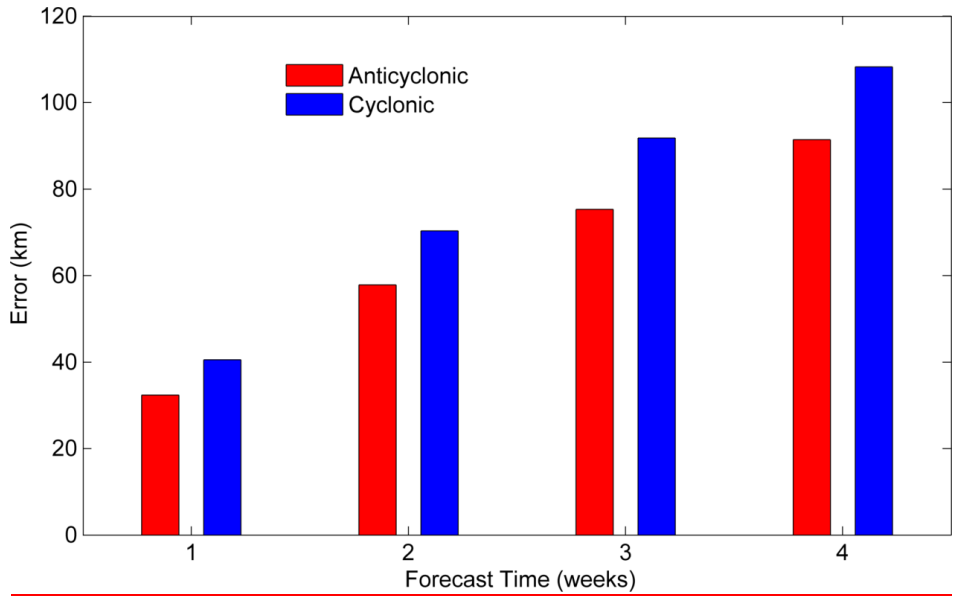


Figure 6. Comparison of the mean forecast errors between anticyclonic eddies (red) and cyclonic eddies (blue) over a 4-week window.

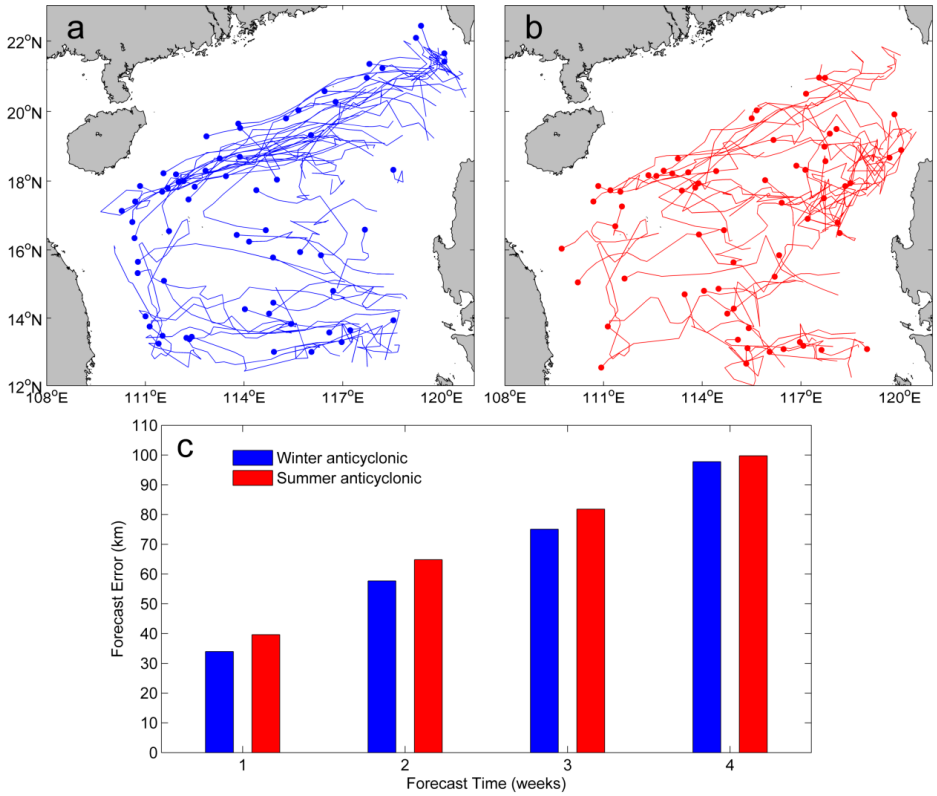
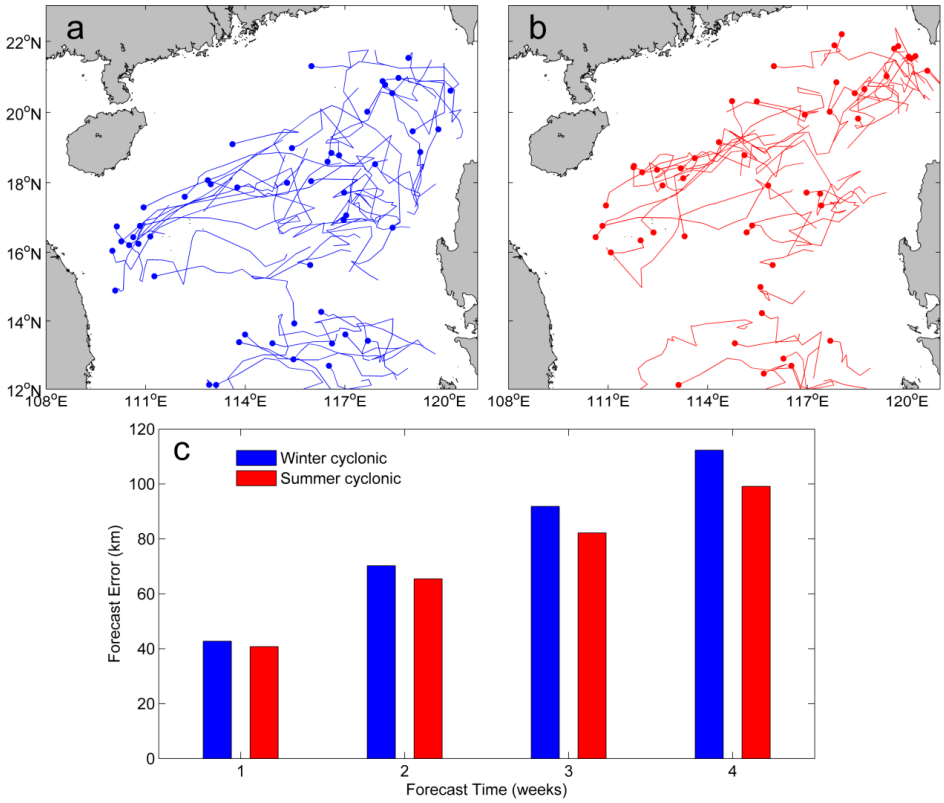


Figure 7. The trajectories of anticyclonic eddies in (a) winter and (b) summer with lifetime≥5 weeks in the South China Sea (SCS). The solid circle represents the ending position of each trajectory. (c) Comparison of their mean forecast errors over a 4-week window.



721

722

723 **Figure 8.** The same as Fig. 6, but for the cyclonic eddies.

724

725

Table 1. The Eight Predictands Used in the Predictive Model

Predictand	Symbol
1-week zonal displacement	DX_1
1-week meridional displacement	DY_1
2-week zonal displacement	DX_2
2-week meridional displacement	DY_2
3-week zonal displacement	DX_3
3-week meridional displacement	DY_3
4-week zonal displacement	DX_4
4-week meridional displacement	DY_4

730 **Table 2.** The Eight Predictors Used in the Predictive Model

731

Predictor	Symbol
Initial longitude (LON)	P_1
Initial latitude (LAT)	P_2
Eddy zonal motion past 1-week (U_PAST)	P_3
Eddy meridional motion past 1-week (V_PAST)	P_4
Climatological eddy zonal motion from MCC (U_CLIM)	P_5
Climatological eddy meridional motion from MCC (V_CLIM)	P_6
Initial zonal absolute geostrophic flow (U_ADT)	P_7
Initial meridional absolute geostrophic flow (V_ADT)	P_8

732

Table 3. Normalized Regression Coefficients $a_{i,j}$ ($b_{i,j}$) for Use with the Eddy Zonal (Meridional) Motion Prediction Equation

	j=1	j=2	j=3	j=4
i=1	-0.10 (0.03)	-0.14 (0.04)	-0.18 (0.05)	-0.24 (0.06)
i=2	0.10 (0.02)	0.13 (0.01)	0.16 (0.00)	0.18 (-0.03)
i=3	0.26 (0.00)	0.21 (0.03)	0.19 (0.07)	0.18 (0.09)
i=4	-0.02 (0.19)	-0.01 (0.10)	0.01 (0.08)	0.00 (0.08)
i=5	0.14 (0.09)	0.19 (0.13)	0.23 (0.16)	0.26 (0.16)
i=6	0.05 (0.17)	0.07 (0.23)	0.09 (0.26)	0.16 (0.27)
i=7	-0.05 (0.02)	-0.07 (0.02)	-0.07 (0.02)	-0.07 (0.03)
i=8	-0.03 (-0.07)	-0.01 (-0.08)	0.02 (-0.09)	0.04 (-0.09)

738 **Table 4.** Comparison of mean forecast distance errors (km) of the persistence, multiple linear
739 regression (MLR), and artificial neural network (ANN) method
740

Forecast weeks	Persistence	MLR	ANN
1	47.6	38.1	37.8
2	95.2	64.8	64.1
3	135.0	86.6	84.7
4	180.5	106.5	102.3

Table 5. Statistics of our Muliple Linear Regression Model for Different Forecast Time of Eddy Propagation Positions in Terms of Longitudes (Latitudes)

Forecast weeks	Total/Predicted Number of Points	RMSE, km	Correlation Coefficient	Mean Distance Error, km
1	2604/623	32.7 (29.5)	0.99 (0.99)	38.1
2	2310/549	55.1 (47.3)	0.97 (0.98)	64.8
3	2016/475	72.5 (61.4)	0.95 (0.97)	86.6
4	1722/401	89.2 (73.5)	0.93 (0.95)	106.5

Note: the total/predicted number of points refers to the eddy positions at 7-day time interval in the whole/predicted eddy trajectories during 1992-2013/2009-2013; the RMSE is the root mean square error between the predicted and the observed longitude (latitude).

Responses to Referee # 1:

Accurately forecasting the eddy propagation is a major challenge not only to consider the classes of response to atmospheric forcing but also considering the relative impact of the atmospheric forcing, updated boundary data and different ocean data types. The article of "A simple predictive model for the eddy propagation trajectory in the South China Sea" tries to build a predictive model using the multiple linear regression to predict the positions of the long-lifetime eddy tracks in the SCS.

Response: We greatly appreciate the reviewer for the time spending on providing the valuable comments. We made every effort to clarify our results and improve our manuscript according to the comments. Next our response to each comment will be labeled in blue.

Here, I disagree the reliability of this method applied to forecast the actual eddy track in the SCS because there are two main points:

1) The MCC method although more objective, still includes the assumption that displacements are translational and negate rotational and deformational motions, although Kamachi (1989) modified the MCC to include rotational effects. As they said these eight predictors: those from climatology and persistence and those "synoptic predictors" are geostrophic flows. Just like they found: "The synoptic predictors contribute less to the forecast equations comparing with persistence and climatology.", which means it mostly depends on the persistent inputs. It could be more accurate to regard this model as a diagnosing or corrected persistent forecasting.

Response: Thanks for the comment. During the past 20 years, mesoscale eddies in the South China Sea (SCS) have drawn much attention, and their statistical characteristics, generation mechanisms, and impact on the atmosphere and ocean have been widely studied (e.g., Wang et al., 2003; Chen et al., 2011; Li et al., 2017). However, studies on the forecast of eddies in the SCS are rare because of their complex dynamics and high nonlinearity. Just recently, Xu et al. (2018) used modern ocean dynamical model to predict two eddy cases in the northern SCS, found the eddy propagation paths can

774 be predicted (forecast distance errors are 81-132 km from the third to fifth week) only
775 when the eddy amplitude is larger than 8 cm. To the best of my knowledge, our work
776 is the first attempt at forecasting the eddy propagation trajectories statistically in the
777 SCS. Comparing to the dynamical method, our simple statistical method don't need
778 boundary and forcing conditions and partial differential equation discretization, thus
779 the computation is much faster than ocean models. Also our model is independent of
780 eddy amplitude, and the forecast distance error is comparable with that of the
781 dynamical model. Therefore, although our statistical model can be regarded as a
782 diagnosing or corrected persistent forecasting model, it may provide an alternative
783 and fast means for an operational forecast, which is especially useful to practical
784 applications, such as naval military operation.
785
786 2) The currently results miss the independent validation. Page 4 Line 82-84: "To
787 forecast the eddy trajectory 1-4 weeks in advance using the last position of the eddy,
788 only eddies with a lifetime of 5 weeks or longer are retained in this study". It clearly
789 shows the eddy tracks in 2009-2013 used for evaluation here have been artificially
790 filtered, and results the underestimation of the related failure events. Consequently,
791 this model cannot be regarded as a successfully predictive model and it is loss of
792 enough values to be published on OS.
793 Response: Given the accuracy of satellite altimeter product and to avoid sporadic
794 eddy events, eddy which lifetime is not shorter than 4 weeks is considered in the eddy
795 detection and tracking (e.g., Chelton et al., 2011; Chen et al., 2011; Wang et al., 2003).
796 Thus the 3rd release of the global eddy dataset used in this study discarded the eddies
797 with lifetime shorter than 4 weeks by Chelton et al. (2011). To forecast the eddy
798 trajectory 4 weeks in advance using the last position of the eddy, only eddies with a
799 lifetime of 5 weeks or longer are retained. Table R1 lists the 1-3 week forecast results
800 of the original eddy tracks with lifetime not shorter than 4 weeks and the "filtered"
801 eddy tracks with lifetime not shorter than 5 weeks, which shows the forecast results
802 are comparable and verify our predictive model is stable.
803

Table R1. Comparison of forecast distance errors (km) between the original eddy tracks with lifetime not shorter than 4 weeks and the filtered eddy tracks with lifetime not shorter than 5 weeks.

Forecast weeks	Original tracks	Filtered tracks
1	38.7	38.1
2	66.9	64.8
3	88.3	86.6

Specific Comments:

1) Before applying the MCC analysis to the images prepared, certain parameters describing the statistical method needed to be set like subwindow size, search window size as well as cross-correlation coefficient. What about their sensibilities? And finally what about the setting?

Response: Thanks for the comment. The MCC method used in this study is the same as that of Fu et al. (2006, 2009), which is a little different with that of Emery et al. (1986). In the method of Emery et al., the correlations of the image **in the subwindow** with all the neighboring ones in the whole window at the next time are computed, and the speed and direction of the maximum correlations can be estimated. While in the method of Fu et al., the correlations of the SLA **at a given location** with all the neighboring SLA at various time lags are computed, and the speed and direction of the maximum correlations can be estimated. The reason of their difference may be due to the low time-space resolution of SLA comparing with other infrared satellite images, such as AVHRR.

In the MCC method of Fu et al. (2009), the size of the time-space window for computing the correlations were determined by the time and space scales of interests. To focus on the global mesoscale eddy, the time lags were limited to less than 70 days and the dimension of the window was less than 400 km. However, the time lags should be limited to less than 42 days in the SCS, since many correlation coefficients

are below the 95% confidence level at larger time lags (Zhuang et al., 2010). Besides, Chen et al. (2011) found that eddies propagate with 5.0-9.0 cm/s in the northern SCS. Thus the search radius can be generally limited as 300 km (9.0 cm/s*42 days≈300 km) to reduce incidence of spurious MCC vectors. We add several sentences in the introduction of MCC method to clarity the parameters and their setting.

2) Here all SLA data and eddy dataset have a time resolution of 7 days. In fact, the new version based on the DT-2014 daily "two-sat merged" sea level anomaly (MSLA) fields (formerly referred to as the REF dataset) posted online by AVISO for the 22-year period January 1993–April 2015. So using the daily dataset could be more interesting, and some new knowledge can be expected.

Response: Thanks for the comment. Since the maximum westward propagation speed is about 20 cm/s in the subtropics (Chelton et al., 2011), the maximum propagation distance in one day is about 17.3 km, which is less than one grid dimension (25 km) of AVISO SLA. This may cause some uncertainties in the eddy trajectory forecasting using the daily dataset. Therefore, the weekly SLA data is still used in the eddy forecasting exercises (Oey et al., 2005; Zeng et al., 2015, and Xu et al., 2018).

3) Chen et al. (2011) also find that “Eddy propagation in the western basin to the east of Vietnam is quite random, with no uniform propagate direction”. Are there some effects or comments to that?

Response: The generation mechanism of eddy in the western basin to the east of Vietnam is complicated. Many dynamical factors may contribute, such as the wind jet, the eastward current, the flow instability and the coastal and island topography. It is also found that eddies associated with the eastward current is under the impact of the wind jet, which shows intraseasonal variability (Xie et al., 2007). We suppose these underlying factors may cause eddies in this region propagate with no inform direction.

4) The right panels of Fig. 2 showing the differences which should keep the nan areas as in (a) and (d).

Response: Corrected in the revised manuscript.

5) Page 9, Line 217: "there are a total of 8 regression equations"? Could you provide the related formal or equations to clearly distinguish the explanatory variables, the response variables, and the input regression data sources?

Response: Thanks for the suggestion. The predicted zonal (meridional) displacement *DX* (*DY*) can be estimated using a multiple linear regression approach:

$$DX_j = \sum_{i=1}^{8} a_{i,j} P_{i,} \qquad j = 1, 4$$

$$DY_j = \sum_{i=1}^{8} b_{i,j} P_{i,} \qquad j = 1, 4$$

where the subscript *j* refers to the forecast interval (1-4 weeks), the subscript *i* refers to the serial number of eight normalized predictors (*P*), *a* and *b* donate normalized regression coefficients of predictors onto *DX* and *DY*, respectively. To distinguish the input predictors, the forecasted variables, and the related regression equations clearly, we revise Table 2 and 3, add a new Table 1, and add Section 3.2 to describe these in the revised manuscript.

6) Figure 5 only shows one trajectory. Could you replaced by all trajectories in the SCS during the same time periods, which could be more objectively to explore the credibility of this method.

Response: Thanks for the suggestion. For the sake of concise layout of the paper, we only selected two cases from all the 74 forecasted results to show the comparison. To verify the credibility of this method, the forecast distance errors of all the predicted eddy trajectories over a 4-week window are shown in Figure R1.

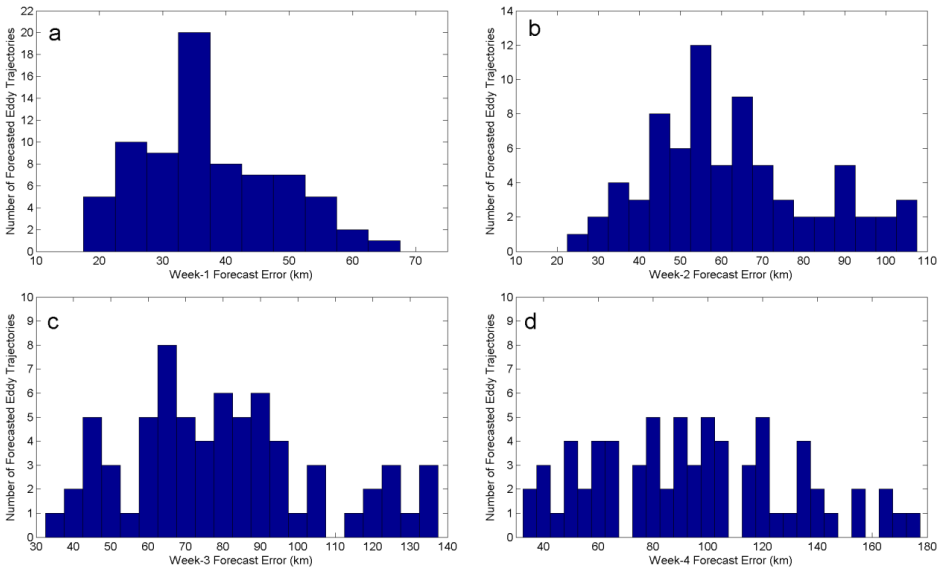


Figure R1. Histogram of the forecast distance errors of all the predicted eddy trajectories at week-1 (a), week-2 (b), week-3 (c) and week-4 (d).

7) The predictive equation should explicitly presented in text. Although the effects of planetary beta and mean flow advection are highlighted many times, the quantitative effect on the inputs or the predictive equations still are not clear.

Response: Thanks for the suggestion. (1) In the revised manuscript, the predictive regression equations have been presented in two equations of Section 2.3, and their coefficients have been shown in Table 3 in the revised version. (2) Actually, the climatological eddy zonal and meridional motions (U_CLIM V_CLIM) derived from the MCC method consist of the effects of beta and the mean flow advection. We have decompose U_CLIM and V_CLIM into these factors and incorporate them into the regression model, but found no improvement of the forecast skill. We add several sentences in Section 3.2 to clarify it.

8) Page 5, Line 99: whether the cross-correlations have been normalized by the variances of the two time series?

Response: Yes, the cross-correlations have been normalized by the variances of the

two time series.

9) Page 6, Figure 2 only show at north of 12N. Does it mean this study only investigate the eddy tracks in the northern SCS. If right, the concerned statement and title should be replaced by the northern SCS.

Response: Yes, this study only investigates the eddy tracks in the northern SCS. We have revised the statement and title using the northern SCS (NSCS).

10) Page 9 Line 198: Are the climatological eddy motions divided into 12 months or only annual mean?

Response: The climatological eddy motions are divided into four seasons (winter: 12-2, spring: 3-5, summer: 6-8, autumn: 9-11), since the mean flow and associated eddy propagation in the SCS have seasonal variability. We add several sentences in Section 2.2 to clarity it.

11) Based on the 17 years (1992-2009) of satellite altimeter data, Chen et al. (2011) identified 827 eddy (lifetime >=28 days) tracks in the SCS. However, here uses 1981 eddy trajectories during 1992-2008. Why there are so big gap between them?

Response: Thank you for the comment. 1981 is the number of eddy trajectory segments (a segment refers to the distance between two neighboring eddy center positions at 7-day interval on a single eddy trajectory), not eddy number. We add one sentence in this paragraph to explain this.

12) The eddy forecast error has been discussed by Hurlburt et al. (2008). Related to the previous evaluation, it is valuable to comment.

E. Hurlburt, Harley Chassignet, Eric A. Cummings, James Birol Kara, A Metzger, E F. Shriver, Jay Smedstad, Ole J. Wallcraft, Alan N. Barron, Charlie. (2008). EddyResolving Global Ocean Prediction. Washington DC American Geophysical Union Geophysical Monograph Series. 353-381. 10.1029/177GM21.

Table 2: Eddy center location errors in ocean prediction models compared to ocean

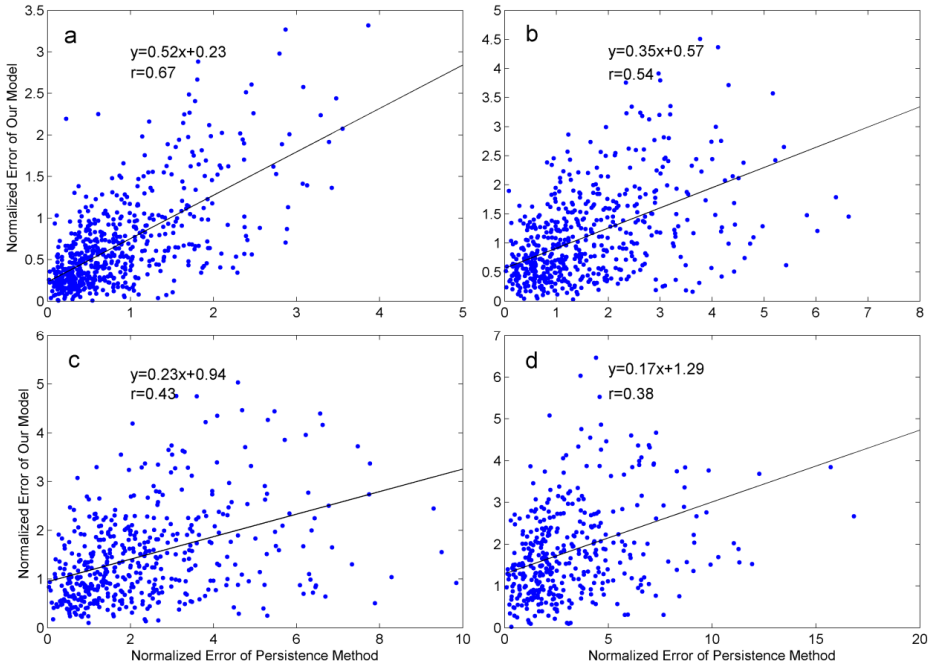
color from SeaWiFS in the northwestern Arabian Sea and Gulf of Oman

Response: Thanks for the comment. Because mesoscale eddies are often associated with strong nonlinear processes and their dynamical mechanisms are quite different, the operational forecast of eddies has been a big challenge to ocean numerical model. Much progress has been made in recent years in eddy-resolving ocean prediction. With the data assimilation and the increasing of model resolution, the model increases forecast skill. Eddy center position daily forecast errors in the northwestern Arabian Sea and Gulf of Oman is 44-68 km in 1/12^{o} global HYCOM model, and reaches to 22.5-37 km in 1/32^{o} NLOM model (Hurlburt et al., 2008). The forecast skill and predictability of dynamical models can only be increased by better assimilation schemes (initialization), sufficient data (especially the subsurface), and improving resolution (physics and computing power). We have added this reference and some sentences in the second paragraph of Section 1 in the revised manuscript.

13) In this study, the distance errors are presented by degree or km only. The relative error considering the eddy radius is more important to directly understanding the uncertainty.

Response: Thanks for the comment. Actually, we once considered the relative errors by normalizing the forecast distance errors with the Rossby radius on each forecast grid. Figure R2 shows the differences and correlation of relative errors between the persistence method and the proposed method over 4-week forecast window. Their correlation decreases from 0.67 at week-1 to 0.38 at week-4. This conclusion based on the relative errors is consistent with that of the comparison of forecast distance errors between the two methods: although the persistence forecast trajectory at week-1 is relatively consistent with the observation, the persistence method cannot forecast the eddy trajectories properly when the forecast horizon increases. Considering the forecast distance errors presented by km have been widely accepted by operational ocean eddy forecasting (e.g., Oey et al., 2005; Zeng et al., 2015) and tropical cyclone track forecasting (e.g., Aberson et al., 2003; Ali et al., 2007), the forecast distance errors by km is still used in the evaluation of forecast performance for the

961 convenience of common readers.



962

963 **Figure R2**. Scatterplot of the normalized forecast distance errors of persistence
964 method vs. the normalized forecast distance errors of out linear regression model with
965 best fit linear regression at week-1 (a), week-2 (b), week-3 (c) and week-4 (d).

966

967

968 **References:**

969 Aberson, S. D. and Sampson, C. R.: On the predictability of tropical cyclone tracks in the
970 northwest pacific basin, Mon. Wea. Rev., 131, 1491-1497, 2003.
971 Ali, M. M., Kishtawal, C. M., and Jain, S.: Predicting cyclone tracks in the north Indian Ocean:
972 An artificial neural network approach, Geophys. Res. Lett., 34, L04603,
973 http://doi.org/10.1029/2006GL028353, 2007.
974 Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale
975 eddies, Prog. Oceanogr., 91, 167-216, 2011.
976 Chen, G., Hou, Y., and Chu, X.: Mesoscale eddies in the South China Sea: Mean properties,
977 spatiotemporal variability, and impact on thermohaline structure, J. Geophys. Res., 116, C06018,

http://doi.org/10.1029/2010JC006716, 2011.

Emery, W. J., Thomas, A. C., Collins, M. J., Crawford, W. R., and Mackas, D. L.: An objective method for computing advective surface velocities from sequential infrared satellite images, J. Geophys. Res., 91, 12865–12878, http://doi.org/ 10.1029/JC091iC11p12865, 1986.

Hurlburt, H., Chassignet, E., Cummings, E., et al.: Eddy Resolving Global Ocean Prediction, Washington D. C., American Geophysical Union Geophysical Monograph Series, 353-381, 10.1029/177GM21, 2008.

Fu, L.-L.: Pathways of eddies in the South Atlantic Ocean revealed from satellite altimeter observations, Geophys. Res. Lett., 33, L14610, http://doi.org/10.1029/2006GL026245, 2006.

Fu, L.-L.: Pattern and speed of propagation of the global ocean eddy variability, J. Geophys. Res., 114, C11017, http://doi.org/10.1029/2009JC005349, 2009.

Li, J., Wang, G., and Zhai, X.: Observed cold filaments associated with mesoscale eddies in the South China Sea, J. Geophys. Res. Oceans, 122, 762–770, http://doi.org/10.1002/2016JC012353, 2017.

Oey, L.-Y., Ezer, T., Forristall, G., Cooper, C., DiMarco, S., and Fan, S.: An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico, Geophys. Res. Lett., 32, L12611, http://doi.org/10.1029/2005GL023253, 2005.

Wang, G., Su, J., and Chu, P. C.: Mesoscale eddies in the South China Sea observed with altimeter data, Geophys. Res. Lett., 30, 2121, http://doi.org/10.1029/2003GL018532, 21, 2003.

Xie, S.-P., Chang, C.-H., Xie, Q., and Wang, D.: Intraseasonal variability in the summer South China Sea: Wind jet, cold filament, and recirculaions, J. Geophys. Res., 112, C10008, doi:10.1029/2007JC004238, 2007.

Xu, D., Zhuang, W., Yan, Y.: Could the mesoscale eddies be reproduced and predicted in the northern South China Sea: case studies, Ocean Sci. Discuss., http://doi.org/10.5194/os-2018-74, 2018.

Zeng, X., Li, Y., and He, R.: Predictability of the loop current variation and eddy shedding process in the Gulf of Mexico using an artificial neural network approach, J. Atmos. Oceanic. Technol., 32, 1098-1111, 2015.

Zhuang W., Du, Y., Wang, D., Xie, Q., and Xie, S.-P.: Pathways of mesoscale variability in the South China Sea, Chin. J. Oceanol. Limnol., 28, 1055-1067, 2010.

Responses to Referee # 2:

The submitted work proposes a regression model to forecast the trajectories of eddies in the South China Sea. The method is based on using the velocity field obtained through the Maximum Cross Correlation technique applied to sea level anomalies as capturing the combination of several dynamical components (self-propagating beta effect, advection by mean flow, etc.). The basic dynamical idea of applying the MCC to altimetry to analyze trajectories of eddies was introduced several years ago (e.g. L.L. Fu JGR, 2006). The novelty here is to go a step forward to develop a linear regression model to forecast such trajectories and assuming some dynamical elements affecting eddies propagation. The authors compare their approach against forecasting using a "persistence" approach.

Response: Thanks so much for the helpful comments. We made every effort to clarify our results and improve our manuscript according to your comments. Next our response to each comment will be labeled in blue.

From my point of view, there are not great concerns on the scientific content of the paper. The authors discuss quite adequately the main assumptions leaving for future work potential refinements of their methodology. However, the major drawback when one is trying to provide a forecast method is to analyze with major detail the robustness in the choice of parameters.

Response: During the past 20 years, mesoscale eddies in the South China Sea (SCS) have drawn much attention, and their statistical characteristics, generation mechanisms, and impact on the atmosphere and ocean have been widely studied (e.g., Wang et al., 2003; Chen et al., 2011; Li et al., 2017). However, studies on the forecast of eddies in the SCS are rare because of their complex dynamics and high nonlinearity. Just recently, Xu et al. (2018) used modern ocean dynamical model to predict two eddy cases in the northern SCS, found the eddy propagation paths can be predicted only when the eddy amplitude is larger than 8 cm. To the best of my knowledge, our work is the first attempt at forecasting the eddy propagation trajectories statistically in

the SCS, and our forecasted results (forecast distance error is 86.6-106.5 km from the third to fourth week) are comparable with those of dynamical model (forecast distance error is 81-132 km from the third to fifth week). Comparing to the dynamical method, our simple statistical method don't need boundary and forcing conditions and partial differential equation discretization, thus the computation is much faster than ocean models. Also our model is independent of eddy amplitude, and the forecast distance error is comparable with that of the dynamical model. Therefore, our study may provide an alternative and fast means for an operational forecast, which is especially useful to practical applications, such as naval military operation.

There are some aspects the authors should present more carefully:

* The different choices of window search and time lag. The authors indicate the upper sizes of such values for the mesoscale in SCS (lines 107-109) but some quantitative indications on how it affects the results is necessary.

Response: Thanks for the comment. The MCC method used in this study is the same as that of Fu et al. (2006, 2009), which is a little different with that of Emery et al. (1986). In the method of Emery et al., the correlations of the image **in the subwindow** with all the neighboring ones in the whole window at the next time are computed, and the speed and direction of the maximum correlations can be estimated. While in the method of Fu et al., the correlations of the SLA **at a given location** with all the neighboring SLA at various time lags are computed, and the speed and direction of the maximum correlations can be estimated. The underlying reason of their difference may be due to the low time-space resolution of SLA comparing with other satellite images, such as AVHRR.

In the MCC method, the size of the time-space window for computing the correlations were determined by the time and space scales of interests. To focus on the global mesoscale eddy, the time lags were limited to less than 70 days and the dimension of the window was less than 400 km. However, the time lags should be limited to less than 42 days in the SCS, since many correlation coefficients are below the 95%

confidence level at larger time lags (Zhuang et al., 2010). Besides, Chen et al. (2011) found that eddies propagate with 5.0-9.0 cm/s in the northern SCS. Thus the search radius can be generally limited as 300 km (9.0 cm/s*42 days≈300 km) to reduce incidence of spurious MCC vectors. We add several sentences in the introduction of MCC method to clarity the parameters and their setting.

* The regression coefficients are computed over a limited temporal interval (1992-2008). The authors should analyze the stability of these coefficients as a function of the chosen time interval and how the results depend on it.

Response: Thanks for the comment. Two sensitive forecast experiments (Table R1) in different temporal intervals were performed to further evaluate the stability and effectiveness of our model. As we can see, when the regressed temporal interval is shorter than that of this study, the number of eddy trajectories for the regressing is expected less and the forecast errors increases 0.5-3.9 km over 1-4 week forecast window. When the regressed temporal interval is longer than that of this study, the number of eddy trajectories for the regressing is expected larger and the forecast errors decreases 0.3-3.2 km over 1-4 week forecast window. The forecast errors have small fluctuation (The stand deviation (Sd) is relatively small from 0.4-3.6 over 1-4 weeks), indicating the stability and effectiveness of the results of our model.

Table R1. Settings of two sensitive forecast experiments and our study

	Regressed Time Interval	Regressed Eddy Trajectory No.	Predicted Time Interval	Predicted Eddy Trajectory No.
Exp1	1992-2006	247	2007-2013	117
Exp2	1992-2010	321	2011-2013	43
This study	1992-2008	283	2009-2013	81

Table R2. Comparison of forecast distance errors (km) of two sensitive forecast experiments and our study

	Week-1	Week-2	Week-3	Week-4
Exp1	38.6	65.6	88.2	110.4
Exp2	37.8	63.9	84.3	103.3
This study	38.1	64.8	86.6	106.5
Sd	0.4	0.9	2.0	3.6

* The regression model introduces a climatological term U_CLIM estimated from the MCC. How this climatology is built remains unclear? Is a mean over the same regression period (weekly, monthly, seasonal, ...)? How it depends on such climatology? Such questions should be clarified in a quantitative way.

Response: Thanks for the comment. Since the mean flow and associated eddy propagation in the SCS have pronounced seasonal variability, we followed Zhuang et al. (2010) and divided the weekly SLA data from 1992 to 2013 into four groups according to four seasons (winter: 12-2, spring: 3-5, summer: 6-8, autumn: 9-11). Then the seasonal climatological propagation velocities (U_CLIM, V_CLIM) can be estimated in the same seasonal group using the MCC method of Fu (2006, 2009). We add several explanations in this Section 2.2 to clarity it.

Finally, the structure of the paper is decompensated with a central section ("Results") that mixes the methodological approach, the results and discussions. The "Data and Methods" section (section 2) include two subsections devoted to present datasets and to explain the MCC respectively, while the forecasting model is presented in detail in subsection 3.2 "Model Development" as part of the Results section (section 3). MCC is relevant for the system they propose but is just one of the elements of their methodology and is a well-known classical method in the context of the oceanography. Thus my suggestion is that the regression model should immediately follow the MCC description, both elements more coherently integrated in the "Data and methods" section and leaving the Results section to show the performance of the forecasting system.

Response: Thanks for the suggestion. We add a new Section 2.3: The Multiple Linear Regression Model in the "Data and methods" section to describe the regression model. Considering the importance of the selection of the predictors based on the dynamical analysis, we add a new Section 3: “Dynamics of Eddy Propagation in the NSCS and Choice of Predictors”, and leave Section 4: “Performance of the Multiple Regression Model” to show the performance of the forecasting system exclusively in the revised version.

Apart from these considerations I have the following list of small comments:

Section 2.2 and 3.1 (but also in the whole manuscript): In the text there is an abuse of the "eddy" word. Sometimes "eddy" is used in the context of deviation respect to a statistical mean while sometimes is used to refer to a dynamical coherent structure (an eddy). The MCC velocity field as applied to SLA maps is not only representative of the evolution of coherent eddies but also include many other structures as waves, filaments, fronts, etc. that may also evolve and propagate. Thus the velocity field are not necessarily the velocity of eddies understood as coherent structures alone. This must be mentioned and a careful use of the word "eddy" along the whole manuscript should be checked to avoid misinterpretations.

Response: Thank you for the comment. Yes, the MCC method cannot distinguish the various forms of mesoscale variability, such as filaments, fronts, and planetary waves. In the study of the pattern and velocity of global ocean eddies, Fu (2009) pointed out: when the space and time lags of the correlation analysis are chosen for the mesoscales, the MCC estimated velocities can represent the speed and direction of the propagation of ocean eddies. Chelton et al. (2011) compared the latitudinal variation of the mean eddy speed computed from the global eddy trajectories with that from the MCC method of Fu (2009), and found they are comparable well.

To focus on the global mesoscale eddy, Fu (2009) chose less than 70 days as the time lags and less than 400 km as the dimension of the window. In this work, the time lags should be limited to less than 42 days in the SCS, since many correlation coefficients

are below the 95% confidence level at larger time lags (Zhuang et al., 2010). The search radius can be generally limited as 300 km (9.0 cm/s*42 days ≈ 300 km, since 9 cm/s is the maximum eddy speed in the northern SCS (Chen et al. (2011))) .We add several sentences in the introduction of MCC method to explain this.

Section 3.1 (fig. 2b, e and lines148-165): Perhaps I'm wrong but I don't appreciate much differences between the winter and summer distributions of the phase speed of the first baroclinic Rossby wave. I may suppose that it is because at such latitudes the seasonal stratification does not change too much? A small comment may guide the general readers.

Response: Good comment. Yes, the difference between the winter and summer distributions of the phase speed of the first baroclinic Rossby wave is relatively small. The underlying reason is that the variation of seasonal stratification in the upper layer has little effect on the seasonal distribution of the first baroclinic Rossby deformation radius (Chelton et al, 1998, Cai et al., 2008). We add this comment and two references in the revised manuscript to guide the general readers.

Section 3.2 Model Description. It is needed to introduce the opportune equations representing the linear regression model with the variables involved besides of listing them in table 2.

Response: Thanks for the suggestion. The predicted zonal (meridional) displacement *DX* (*DY*) can be estimated using a multiple linear regression approach:

$$DX_j = \sum_{i=1}^{8} a_{i,j} P_{i,} \qquad j = 1, 4$$

$$DY_j = \sum_{i=1}^{8} b_{i,j} P_{i,} \qquad j = 1, 4$$

where the subscript *j* refers to the forecast interval (1-4 weeks), the subscript *i* refers to the serial number of eight normalized predictors (*P*), *a* and *b* donate normalized regression coefficients of predictors onto *DX* and *DY*, respectively. To distinguish the input predictors, the forecasted variables, and the related regression equations clearly,

To introduce the opportune equations representing the linear regression model, we revise Table 1 and 3, add a new Table 2 and add a new Section 2.3 in the revised manuscript.

Section 3.2: If I have understood, the MCC fields introduced into U_CLIM and V_CLIM are the characteristic mean from the whole altimetric period computed at intervals of 1 week? Please may you clarify it? See the remark above.

Response: Yes, U_CLIM and V_CLIM are the characteristic mean estimated using the MCC. Since the mean flow and associated eddy propagation in the SCS have pronounced seasonal variability, we followed Zhuang et al. (2010) and divided the weekly SLA data from 1992 to 2013 into four groups according to four seasons (winter: 12-2, spring: 3-5, summer: 6-8, autumn: 9-11). Then the seasonal climatological propagation velocities (U_CLIM, V_CLIM) can be estimated from the same seasonal group at intervals of 1 week using the MCC method of Fu (2006, 2009). We add several explanations in this Section 2.2 to clarity it.

Section 3.2: The initial step in the forecasting procedure is to provide an initial starting point of a given eddy. How is this provided, manually upon a first visual inspection of maps or using some method to automatically identify coherent structures in SLA maps? Please precise.

Response: Thanks for the comment. We used the SCS eddy trajectory data derived from the 3rd release of the global eddy dataset. (http://cioss.coas.oregonstate.edu/eddies/). This eddy dataset is developed based on the weekly AVISO SLA data by Chelton et al. (2011), and contains several parameters of the detected eddies at 7-day time interval, such as: eddy positions, eddy radius, eddy amplitude. We have clarified this in the Section 2.1 in the revised manuscript.

Line 207: How the predictands and predictors are normalized? please explain.

Response: Thank you for the comment. Suppose X is the time series of one predictor (or predictand), it is normalized by:

$$X^* = \frac{X - \mu}{\sigma}$$

where, X^* is the normalized X, μ is the mean of X, and σ is the sample standard deviation. We add one sentence in the text to explain it.

Line 217: "There are a total of 8 regression equations...", please see the comments above.

Response: We add two Equations in Section 2.3 to show the opportune regression equations with the input predictors, the forecasted variables, and the related equations coefficients listed in Table 1, 2, and 3 in the revised manuscript.

Lines 232 and following and Table 3: The table caption and the description of the parameters listed in the table are not enough detailed. Why the RMSE is given in degrees? The use of parentheses in the table whether they mean latitudes, persistence or predicted may somehow confuse the reader. Please clarify it and try to make a more detailed description in the table caption which is extremely synthetic. An interesting way of presenting the differences between the persistence method and the proposed method could be to normalize distances with the Rossby radius on grid size in order to see if their differences are relevant or not.

Response: Thanks for the helpful comment. According to the comment, we make three changes in Table 5 in the revised version (Table 3 in the older version): (1) One note describing the parameters below the table is added below the table in the revised version; (2) The RMSE is given in km, which is consistent with mean distance error; (3) To make this table more clearly, the forecast results of the persistence method is removed from this table, and incorporated into the new Table 4 in the new version.

We follow the last suggestion and compare the differences between the persistence method and the proposed method by normalizing distances with the Rossby radius on each forecast grid over 4-week forecast window. As can be seen from Figure R1, their correlation decreases from 0.67 at week-1 to 0.38 at week-4. This further verifies the

result of the comparison of forecast distance errors between the two methods: although the persistence forecast trajectory at week-1 is relatively consistent with the observation, the persistence method cannot forecast the eddy trajectories properly when the forecast horizon increases. We add several sentences to show this point in the revised manuscript.

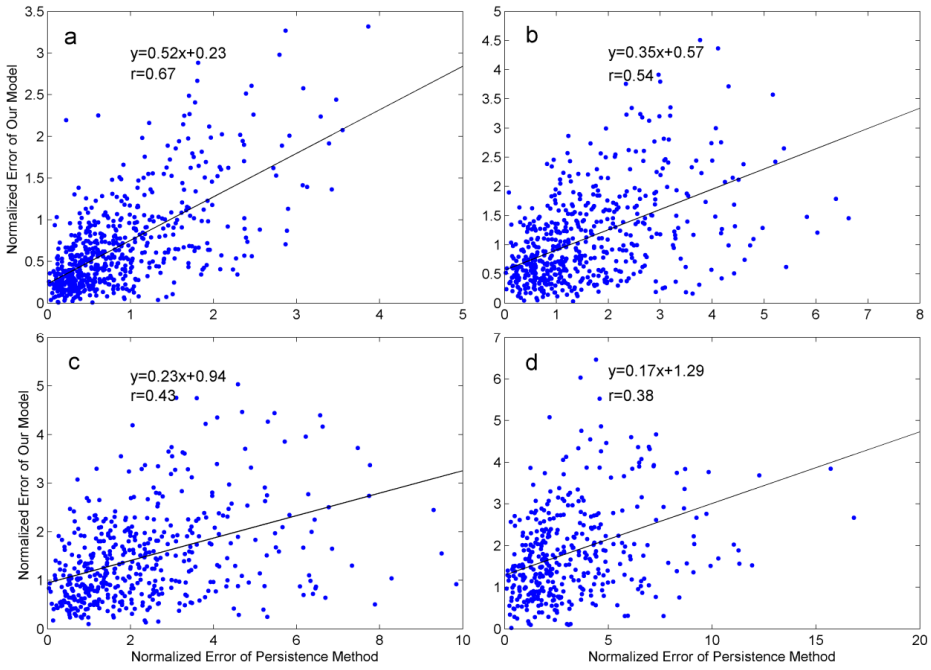


Figure R1. Scatterplot of the normalized forecast distance errors of persistence method vs. the normalized forecast distance errors of out linear regression model with best fit linear regression at week-1 (a), week-2 (b), week-3 (c) and week-4 (d).

Summary and Discussion: The authors only test the performance on seasonality and polarity but perhaps other processes as dissipation, merging or splitting which can be quite common and linked to eddy dynamical parameters as for example vorticity may affect the performance. Some comments or discussions on that should be welcome in this section but some examples on how the forecast are in such cases could also be illustrated.

Response: Thanks for the comment. Actually, in the developing process of our model,

we have tried to incorporate the eddy dynamical parameters, such as eddy radius, eddy amplitude, eddy rotational speed, and eddy vorticity into our model, but sadly, no improvements have been shown. The underlying reasons may be that the propagation of eddies are related to strong nonlinear processes, which have not been fully understood and resolved.

Another possible improvement in the model forecast skill is to use artificial neural network (ANN) in developing the forecast model. ANN has been successfully used in the predicting cyclone tracks (Ali et al., 2007) and loop current variation (Zeng et al, 2015), and the salinity profile estimation from satellite surface observations (Bao et al., 2019). ANN can represent both linear and non-linear relationships learned directly from the data being modeled. It mainly contains three layers: the input layer, the hidden layer, and the output layer. To be consistent with the multiple linear regression, the input layer also includes the same eight predictors, and the output layer includes the two predictands. The hidden layer consists of two layers of neural variables. Through iterations on backward propagation of the error, the neural network learns by itself to achieve an optimum weighting function and a minimum error. The forecast errors of ANN for 1-4 weeks are listed in Table R3. We can see that some improvements (0.3-4.2 km during 1-4 weeks forecast horizon) have been shown comparing with linear regression method. We add these sentences in the Summary and Discussion Section of the revised manuscript.

Table R3. Comparison of forecast distance errors (km) of three methods

Forecast weeks	Persistence	Linear Regression	ANN
1	47.6	38.1	37.8
2	95.2	64.8	64.1
3	135.0	86.6	84.7
4	180.5	106.5	102.3

Figures:

In fig 1 the subplot C is part of the results and I recommend to move it to the results section.

Response: Good suggestion. We move subplot C in Figure 1 to the results section as new Figure 6 in the revised version.

Fig 2 is very small in size and hard to appreciate the velocity fields.

Response: Thanks for the comment. We redraw Figure 2 with the 2*3 subplots to enlarge each panel and make the velocity fields more distinguishable.

References:

Ali, M. M., Kishtawal, C. M., and Jain, S.: Predicting cyclone tracks in the north Indian Ocean: An artificial neural network approach, Geophys. Res. Lett., 34, L04603, http://doi.org/10.1029/2006GL028353, 2007.

Bao, S., Zhang, R. Wang, H., et al.: Salinity profile estimation in the Pacific Ocean from satellite Surface salinity observations, J. Atmos. Oceanic. Technol., 36, 53-68, 2019.

Cai, S., Long, X., Wu, R., and Wang, S.: Geographical and monthly variability of the first baroclinic rossby radius of deformation in the south china sea, J. Mar. Syst., 74, 711-720, 2008.

Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale eddies, Prog. Oceanogr., 91, 167-216, 2011.

Chelton, D., DeSzoeke, R., and Schlax, M.: Geographical variability of the first baroclinic rossby radius of deformation, J. Phys. Oceanogr., 28, 433-460, 1998.

Chen, G., Hou, Y., and Chu, X.: Mesoscale eddies in the South China Sea: Mean properties, spatiotemporal variability, and impact on thermohaline structure, J. Geophys. Res., 116, C06018, http://doi.org/10.1029/2010JC006716, 2011.

Emery, W. J., Thomas, A. C., Collins, M. J., Crawford, W. R., and Mackas, D. L.: An objective method for computing advective surface velocities from sequential infrared satellite images, J. Geophys. Res., 91, 12865–12878, http://doi.org/ 10.1029/JC091iC11p12865, 1986.

Fu, L.-L.: Pathways of eddies in the South Atlantic Ocean revealed from satellite altimeter observations, Geophys. Res. Lett., 33, L14610, http://doi.org/10.1029/2006GL026245, 2006.

Fu, L.-L.: Pattern and speed of propagation of the global ocean eddy variability, J. Geophys. Res., 114, C11017, http://doi.org/10.1029/2009JC005349, 2009.

Li, J., Zhang, R., and Jin, B.: Eddy characteristics in the northern South China Sea as inferred from Lagrangian drifter data, Ocean Sci., 7, 1575-1599, 2011.

Wang, G., Su, J., and Chu, P. C.: Mesoscale eddies in the South China Sea observed with altimeter data, Geophys. Res. Lett., 30, 2121, http://doi.org/10.1029/2003GL018532, 21, 2003.

Xu, D., Zhuang, W., Yan, Y.: Could the mesoscale eddies be reproduced and predicted in the northern South China Sea: case studies, Ocean Sci. Discuss., http://doi.org/10.5194/os-2018-74, 2018.

Zeng, X., Li, Y., and He, R.: Predictability of the loop current variation and eddy shedding process in the Gulf of Mexico using an artificial neural network approach, J. Atmos. Oceanic. Technol., 32, 1098-1111, 2015.

Zhuang W., Du, Y., Wang, D., Xie, Q., and Xie, S.-P.: Pathways of mesoscale variability in the South China Sea, Chin. J. Oceanol. Limnol., 28, 1055-1067, 2010.

Responses to Referee # 3:

Accurately forecasting eddy propagation is a major challenge, requiring one not only to consider the classes of response to atmospheric forcing, but to also consider the relative impact of atmospheric forcing, updated boundary data and different ocean data types. The article of "A simple predictive model for the eddy propagation trajectory in the South China Sea" tries to build a predictive model using multiple linear regression to predict the positions of long-lifetime eddy tracks in the SCS.

Response: We greatly appreciate the reviewer for the time spending on providing the valuable comments. We made every effort to clarify our results and improve our manuscript according to the comments. Next our response to each comment will be labeled in blue.

Here, I have some concerns about the reliability and applicability of the model:

1) As presented here the MCC method, although more objective, still includes the assumption that displacements are translational. It should be acknowledged that Kamachi (1989) modified the MCC to include rotational effects. Deformational effects should also be discussed. The dependence of the predictive trajectory on the parameters used also needs to be explained clearly.

Response: Thanks for the comment. The MCC method used in this study is the same as that of Fu et al. (2006, 2009), which is a little different with that of Emery et al. (1986). In the method of Emery et al., the correlations of the image **in the subwindow** with all the neighboring ones in the whole window at the next time are computed, and the speed and direction of the maximum correlations can be estimated. While in the method of Fu et al., the correlations of the SLA **at a given location** with all the neighboring SLA at various time lags are computed, and the speed and direction of the maximum correlations can be estimated. The reason of their difference may be due to the low time-space resolution of SLA comparing with other infrared satellite images, such as AVHRR.

In the MCC method of Fu et al., the size of the time-space window for computing the correlations were determined by the time and space scales of interests. To focus on the global mesoscale eddy, the time lags were limited to less than 70 days and the dimension of the window was less than 400 km. However, the time lags should be limited to less than 42 days in the SCS, since many correlation coefficients are below the 95% confidence level at larger time lags (Zhuang et al., 2010). Besides, Chen et al. (2011) found that eddies propagate with 5.0-9.0 cm/s in the northern SCS. Thus the search radius can be generally limited as 300 km (9.0 cm/s*42 days$\approx$300 km) to reduce incidence of spurious MCC vectors. We add several sentences in the introduction of MCC method to clarity the parameters and their setting.

2) As the core novelty in this study, the regression equations need to be clearly presented. This is important for other users or readers to independently validate or to further improve the method in real conditions.

Response: Thanks for the suggestion. The predicted zonal (meridional) displacement *DX* (*DY*) can be estimated using a multiple linear regression approach:

$$DX_j = \sum_{i=1}^{8} a_{i,j} P_{i,} \qquad j = 1, 4$$

$$DY_j = \sum_{i=1}^{8} b_{i,j} P_{i,} \qquad j = 1, 4$$

where the subscript *j* refers to the forecast interval (1-4 weeks), the subscript *i* refers to the serial number of eight normalized predictors (*P*), *a* and *b* donate normalized regression coefficients of predictors onto *DX* and *DY*, respectively. We add Section 2.3 to describe the multiple linear regression method and the regression equations with the coefficients listed in Table 3.

3) In discussing the eight predictors (Lines 211-13): "The synoptic predictors contribute less to the forecast equations comparing with persistence and climatology". Does it mean that the forecast mostly depends on the persistent inputs and climatology? And then, are the U_clim and V_clim derived from the MCC method

and the history trajectories from 1992-2013? Please clarify.

Response: Thanks for the comment. (1) Yes, the forecast mostly depends on the persistence and climatology. We suppose it may be that the week to week variations are too large so the representation of the initial U_ADT and V_ADT to the actual velocities in the 4-week window is not as good as the U_CLIM and V_CLIM. (2) The climatological eddy zonal and meridional motions (U_CLIM V_CLIM) are derived from SLA data (not historical eddy trajectories) from 1992 to 2013 using the space-time lagged MCC method. We add some sentences in the first paragraph of Section 3.2 to clarify it.

4) Another point is the accounting for the Beta effect (Lines:198-200) in the predictors (U_clim, V_clim). The associated figure and illustration verified the importance of the effect, but how to modify the predictors is not clear. Please comment.

Response: Actually, the climatological eddy zonal and meridional motions (U_CLIM V_CLIM) derived from the MCC method consist of the effects of beta and the mean flow advection. We have tried to decompose U_CLIM and V_CLIM into these factors and incorporate them into the regression model, but found no improvement of the forecast skill. We add several sentences in Section 3.2 to explain it.

5) The current predictive model needs full independent validation. Page 4 Line 82- 84: "To forecast the eddy trajectory 1-4 weeks in advance using the last position of the eddy, only eddies with a lifetime of 5 weeks or longer are retained in this study". It clearly shows the eddy tracks in 2009-2013 used for evaluation here have been artificially filtered, and together with the above point 3 I think the model limit is only used for long-life eddy and the current results can be regarded as hindcast rather than prediction. So I suggest the authors consider using the current regression model to validate the new trajectories after 2013.

Response: Thanks for the comment. Given the accuracy of satellite altimeter product and to avoid sporadic eddy events, eddy which lifetime is not shorter than 4 weeks is

considered in the eddy detection and tracking (e.g., Chelton et al., 2011; Chen et al., 2011; Wang et al., 2003). Thus the 3rd release of the global eddy dataset used in this study discarded the eddies with lifetime shorter than 4 weeks by Chelton et al. (2011). To forecast the eddy trajectory 4 weeks in advance using the last position of the eddy, only eddies with a lifetime of 5 weeks or longer are retained. Table R1 lists the 1-3 week forecast results of the original eddy tracks with lifetime not shorter than 4 weeks and the filtered eddy tracks with lifetime not shorter than 5 weeks, which shows the forecast results are comparable and verify our predictive model is stable.

Table R1. Comparison of forecast distance errors (km) between the original eddy tracks with lifetime not shorter than 4 weeks and the filtered eddy tracks with lifetime not shorter than 5 weeks.

Forecast weeks	Original tracks	Filtered tracks
1	38.7	38.1
2	66.9	64.8
3	88.3	86.6

Other specific comments:

1) Before applying the MCC analysis to the images prepared, certain parameters describing the statistical method need to be set, like subwindow size, search window size as well as cross-correlation coefficient. Can you comment on their impacts, and their settings?

In the MCC method of Fu et al., the size of the time-space window for computing the correlations were determined by the time and space scales of interests. To focus on the global mesoscale eddy, the time lags were limited to less than 70 days and the dimension of the window was less than 400 km. However, the time lags should be limited to less than 42 days in the SCS, since many correlation coefficients are below the 95% confidence level at larger time lags (Zhuang et al., 2010). Besides, Chen et al. (2011) found that eddies propagate with 5.0-9.0 cm/s in the northern SCS. Thus the search dimension can be generally limited as 300 km (9.0 cm/s*42 days≈300 km) to

reduce incidence of spurious MCC vectors. We add several sentences in the introduction of MCC method to clarity the parameters and their setting.

2) Here all SLA data and eddy dataset have a time resolution of 7 days. In fact, the new version based on the DT-2014 daily "two-sat merged" sea level anomaly (MSLA) fields (formerly referred to as the REF dataset) posted online by AVISO for the 22-year period January 1993–April 2015. So using the daily dataset could be more interesting, and some new knowledge can be expected.

Response: Thanks for the comment. Since the maximum westward propagation speed is about 20 cm/s in the subtropics (Chelton et al., 2011), the propagation distance in one day is about 17.3 km, which is less than one grid dimension (25 km) of AVISO SLA. This may cause some uncertainties in the forecasting using the daily dataset. Therefore, the weekly SLA data is still used in the eddy forecasting exercises (Oey et al., 2005; Zeng et al., 2015, and Xu et al., 2018).

3) Chen et al. (2011) also find that “Eddy propagation in the western basin to the east of Vietnam is quite random, with no uniform propagate direction”. Do you find a southern limit to the trajectory predictive model?

Response: Good suggestion. Eddy propagation in the western basin to the east of Vietnam is quite random, which cause the eddy trajectories irregular or even convoluted. As we can see, the forecast errors in the southern area (Figure R1a) are larger than those of other regions (Figure R1b).

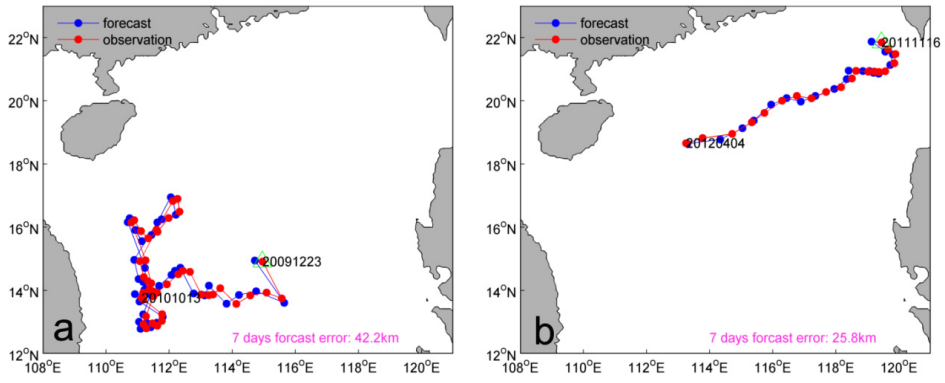


Figure R1. Eddy trajectory 1-week forecast error in (a) the western basin to the east of Vietnam and (b) the northern basin.

4) The right hand panels of Fig. 2 (c, f) showing differences should keep the NaN areas as in (a) and (d).

Response: Corrected.

5) Page 9, Line 217: "there are a total of 8 regression equations"? Could you provide the equations, to clearly distinguish the explanatory variables, response variables, and input regression data sources?

Response: Thanks for the suggestion. The predicted zonal (meridional) displacement *DX* (*DY*) can be estimated using a multiple linear regression approach:

$$DX_j = \sum_{i=1}^{8} a_{i,j} P_{i,} \qquad j = 1, 4$$

$$DY_j = \sum_{i=1}^{8} b_{i,j} P_{i,} \qquad j = 1, 4$$

where the subscript *j* refers to the forecast interval (1-4 weeks), the subscript *i* refers to the serial number of eight normalized predictors (*P*), *a* and *b* donate normalized regression coefficients of predictors onto *DX* and *DY*, respectively. To distinguish the input predictors, the forecasted variables, and the related regression equations clearly, we revise Table 2 and 3, add a new Table 1, and add Section 3.2 to describe these in the revised manuscript.

6) Figure 5 only shows one trajectory. Could you show all trajectories in the SCS during the time periods in question, in order to more thoroughly test the credibility of this method?

Response: Thanks for the suggestion. For the sake of concise layout of the paper, we only selected two cases from all the 74 forecasted results during 2008-2013 to show the comparison. To verify the credibility of this method, the forecast distance errors of all the predicted eddy trajectories over a 4-week window are shown in Figure R2.

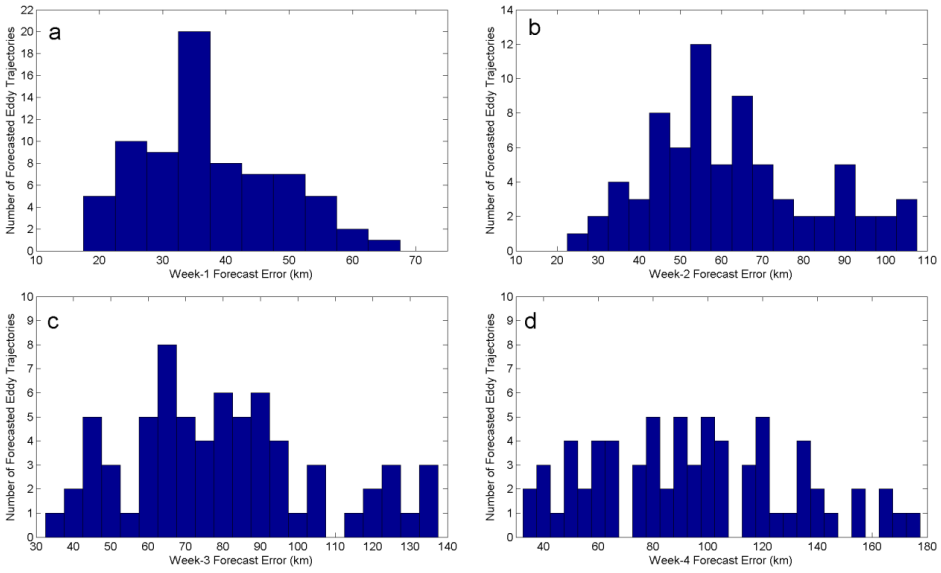


Figure R2. Histogram of the forecast distance errors of all the predicted eddy trajectories at week-1 (a), week-2 (b), week-3 (c) and week-4 (d).

7) The predictive equation should be explicitly presented in the text. Although the effects of planetary b and mean flow advection are highlighted many times, the quantitative effect on the inputs or the predictive equations still are not clear.

Response: Thanks for the suggestion. (1) In the revised manuscript, the predictive regression equations have been presented in two equations of Section 2.3, and their coefficients have been shown in Table 3 in the revised version. (2) Actually, the climatological eddy zonal and meridional motions (U_CLIM V_CLIM) derived from

the MCC method consist of the effects of beta and the mean flow advection. We have decompose U_CLIM and V_CLIM into these factors and incorporate them into the regression model, but found no improvement of the forecast skill. We add several sentences in Section 3.2 to explain it.

8) Page 5, Line 99: Have the cross-correlations been normalized by the variances of the two time series?

Response: Yes, the cross-correlations have been normalized by the variances of the two time series.

9) Page 6, Figure 2 only shows the region north of 12°N. Does it mean this study only investigates the eddy tracks in the northern SCS? If so, relevant statements, and the title, should be qualified as pertaining to the northern SCS.

Response: Yes, this study only investigates the eddy tracks in the northern SCS. We have revised the statements and title using the northern SCS (NSCS).

10) Page 9 Line 198: Are the climatological eddy motions divided into 12 months or only annual mean?

Response: The climatological eddy motions are divided into four seasons (winter: 12-2, spring: 3-5, summer: 6-8, autumn: 9-11), since the mean flow and associated eddy propagation in the SCS have seasonal variability. We add several sentences in Section 2.2 to clarity it.

11) The eddy forecast error has been discussed by Hurlburt et al. (2008). Comment upon this previous evaluation would be valuable.

E. Hurlburt, Harley & Chassignet, Eric & A. Cummings, James & Birol Kara, A & Metzger, E & F. Shriver, Jay & Smedstad, Ole & J. Wallcraft, Alan & N. Barron, Charlie. (2008). EddyResolving Global Ocean Prediction. Washington DC American Geophysical Union Geophysical Monograph Series. 353-381. 10.1029/177GM21.

Response: Thanks for the comment. Because mesoscale eddies are often associated

with strong nonlinear processes and their dynamical mechanisms are quite different, the operational forecast of eddies has been a big challenge to ocean numerical model. Much progress has been made in recent years in eddy-resolving ocean prediction. With the data assimilation and the increasing of model resolution, the model increases forecast skill. Eddy center position daily forecast errors in the northwestern Arabian Sea and Gulf of Oman is 44-68 km in 1/12^{o} global HYCOM model, and reaches to 22.5-37 km in 1/32^{o} NLOM model (Hurlburt et al., 2008). The forecast skill and predictability of dynamical models can only be increased by better assimilation schemes (initialization), sufficient data (especially the subsurface), and improving resolution (physics and computing power). We have added this reference and some comments in the Section 1 of the revised manuscript.

12) In this study, the distance errors are presented by degree or km only. The relative error, relative to the eddy radius, is more important to directly understand the uncertainty.

Response: Thanks for the comment. Actually, we once considered the relative errors by normalizing the forecast distance errors with the Rossby radius on each forecast grid. Figure R3 shows the differences and correlation of relative errors between the persistence method and the proposed method over 4-week forecast window. Their correlation decreases from 0.67 at week-1 to 0.38 at week-4. This conclusion based on the relative errors is consistent with that of the comparison of forecast distance errors between the two methods: although the persistence forecast trajectory at week-1 is relatively consistent with the observation, the persistence method cannot forecast the eddy trajectories properly when the forecast horizon increases. Considering the forecast distance errors presented by degree or km have been widely accepted by operational ocean eddy forecasting (e.g., Oey et al., 2005; Zeng et al., 2015) and tropical cyclone track forecasting (e.g., Aberson et al., 2003; Ali et al., 2007), the forecast distance errors by km is still used in the evaluation of forecast performance for the convenience of common readers.

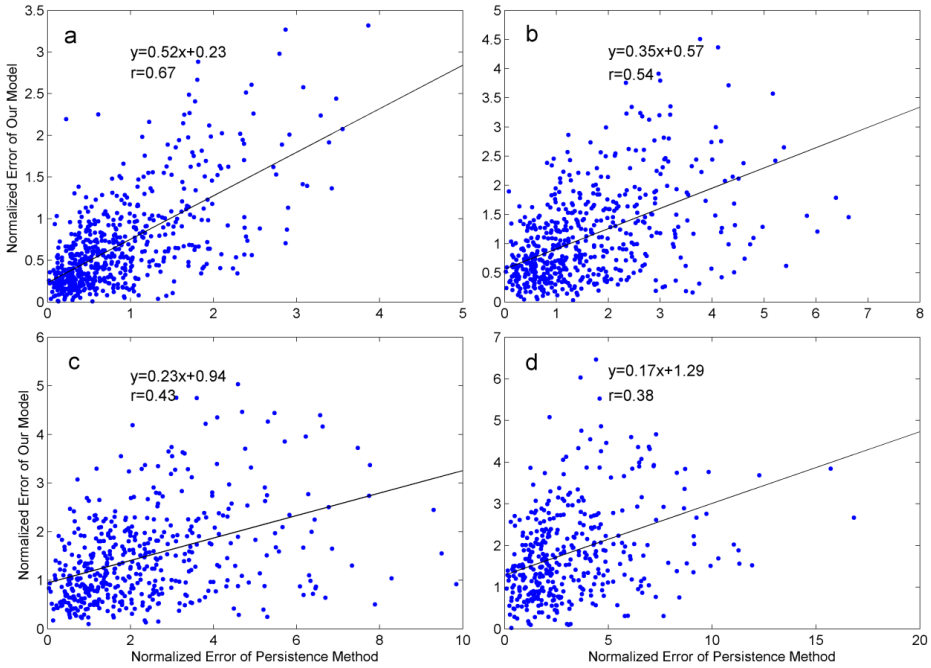


Figure R3. Scatterplot of the normalized forecast distance errors of persistence method vs. the normalized forecast distance errors of out linear regression model with best fit linear regression at week-1 (a), week-2 (b), week-3 (c) and week-4 (d).

References:

Aberson, S. D. and Sampson, C. R.: On the predictability of tropical cyclone tracks in the northwest pacific basin, Mon. Wea. Rev., 131, 1491-1497, 2003.

Ali, M. M., Kishtawal, C. M., and Jain, S.: Predicting cyclone tracks in the north Indian Ocean: An artificial neural network approach, Geophys. Res. Lett., 34, L04603, http://doi.org/10.1029/2006GL028353, 2007.

Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale eddies, Prog. Oceanogr., 91, 167-216, 2011.

Chen, G., Hou, Y., and Chu, X.: Mesoscale eddies in the South China Sea: Mean properties, spatiotemporal variability, and impact on thermohaline structure, J. Geophys. Res., 116, C06018, http://doi.org/10.1029/2010JC006716, 2011.

Emery, W. J., Thomas, A. C., Collins, M. J., Crawford, W. R., and Mackas, D. L.: An objective method for computing advective surface velocities from sequential infrared satellite images, J. Geophys. Res., 91, 12865–12878, http://doi.org/ 10.1029/JC091iC11p12865, 1986.

Hurlburt, H., Chassignet, E., Cummings, E., et al.: Eddy Resolving Global Ocean Prediction, Washington D. C., American Geophysical Union Geophysical Monograph Series, 353-381, 10.1029/177GM21, 2008.

Fu, L.-L.: Pathways of eddies in the South Atlantic Ocean revealed from satellite altimeter observations, Geophys. Res. Lett., 33, L14610, http://doi.org/10.1029/2006GL026245, 2006.

Fu, L.-L.: Pattern and speed of propagation of the global ocean eddy variability, J. Geophys. Res., 114, C11017, http://doi.org/10.1029/2009JC005349, 2009.

Oey, L.-Y., Ezer, T., Forristall, G., Cooper, C., DiMarco, S., and Fan, S.: An exercise in forecasting loop current and eddy frontal positions in the Gulf of Mexico, Geophys. Res. Lett., 32, L12611, http://doi.org/10.1029/2005GL023253, 2005.

Wang, G., Su, J., and Chu, P. C.: Mesoscale eddies in the South China Sea observed with altimeter data, Geophys. Res. Lett., 30, 2121, http://doi.org/10.1029/2003GL018532, 21, 2003.

Xu, D., Zhuang, W., Yan, Y.: Could the mesoscale eddies be reproduced and predicted in the northern South China Sea: case studies, Ocean Sci. Discuss., http://doi.org/10.5194/os-2018-74, 2018.

Zeng, X., Li, Y., and He, R.: Predictability of the loop current variation and eddy shedding process in the Gulf of Mexico using an artificial neural network approach, J. Atmos. Oceanic. Technol., 32, 1098-1111, 2015.

Zhuang W., Du, Y., Wang, D., Xie, Q., and Xie, S.-P.: Pathways of mesoscale variability in the South China Sea, Chin. J. Oceanol. Limnol., 28, 1055-1067, 2010.