Acta Oceanol. Sin., 2017, Vol. 36, No. 11, P. 27–37
DOI: 10.1007/s13131-017-1136-6
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

Tracking the evolution processes and behaviors of mesoscale eddies in the South China Sea: a global nearest neighbor filter approach

YI Jiawei[1, 2], DU Yunyan[1, 2*], WANG Dongxiao[3], ZHOU Chenghu[1, 2]

[1] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Science and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] State Key Laboratory of Tropical Oceanography (LTO), South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou 510301, China

Received 27 July 2016; accepted 29 November 2016

Abstract

The eddy tracking approach is developed using the global nearest neighbor filter (GNNF) to investigate the evolution processes and behaviors of mesoscale eddies in the South China Sea (SCS). Combining the Kalman filter and optimal data association technologies, the GNNF algorithm is able to reduce pairing errors to 0.2% in tracking synthetic eddy tracks, outperforming other existing methods. A total of 4 913 eddy tracks that last more than a week are obtained by the GNNF during 1993–2012. The analysis of a growth and a decay based on 3 445 simple eddy tracks show that eddy radius, amplitude, and vorticity smoothly increase during the first half of lifetime and decline during the second half following a parabola opening downwards. The genesis of eddies mainly clusters northwest and southwest of Luzon Island whereas the dissipations concentrate the Xisha Islands where the underwater bay traps and terminates eddies. West of the Luzon Strait, northwest of Luzon Island, and southeast of Vietnam are regions where eddy splits and mergers are frequently observed. Short disappearances mainly distribute in the first two regions. Moreover, eddy splits generally result in a decrease of the radius and the amplitude whereas eddy mergers induce growing up. Eddy intensity and vorticity, on the contrary, are strengthened in the eddy splits and diminished in mergers.

Key words: mesoscale ocean eddies, eddy tracking, eddy split, eddy merger, South China Sea

Citation: Yi Jiawei, Du Yunyan, Wang Dongxiao, Zhou Chenghu. 2017. Tracking the evolution processes and behaviors of mesoscale eddies in the South China Sea: a global nearest neighbor filter approach. Acta Oceanologica Sinica, 36(11): 27–37, doi: 10.1007/s13131-017-1136-6

1 Introduction

The South China Sea (SCS) is the largest marginal sea with an area of approximately 3.5×10^6 km^2 in the Northwest Pacific. Mesoscale eddies are highly active in the SCS (Liu et al., 2008; Su, 2004). Using satellite altimeter data, Yuan et al. (2007) identified an anticyclonic eddy (AE), aka. Luzon warm eddy (LWE), originates northwest of Luzon Island nearly every summer and then migrates across the northeastern basin and propagates southwestward along the northern continental slope. Chow et al. (2008) documented the evolution characteristics of a cyclonic eddy (CE) that yearly occurred south or southwest of the Dongsha Islands, aka. Dongsha cyclonic eddy. Chen et al. (2010) observed the eddy dipole off eastern Vietnam usually occurred in the summer to fall displays a clear interannual variability and significantly impacts the thermohaline structure in the local upper water. Besides, there are comprehensive surveys of mesoscale eddy activities in the SCS from long-term altimeter observations (Chen et al., 2011; Wang et al., 2003; Xiu et al., 2010).

Automatic identification and tracking algorithms are essential for studying eddies from altimeter data set. Yi et al. (2014) categorized the existing eddy detection algorithms into three classes: (1) algorithms based on a physical parameter (Morrow et al., 2004; Okubo, 1970; Weiss, 1991); (2) algorithms based on a rotary flow pattern (Ari Sadarjoen and Post, 2000; Chaigneau et al., 2008; Nencioli et al., 2010); and (3) algorithms based on the anomaly of a sea surface height (Chelton et al., 2011). Besides, there are novel algorithms proposed in recent years, e.g., Lagrangian detection scheme (Dong et al., 2011), a statistically-driven algorithm (Williams et al., 2011), and a hybrid detection scheme (Yi et al., 2014).

Despite the richness of eddy detection schemes, there are only a handful of tracking algorithms available for ocean eddies. We categorized them into two types. The first type of algorithms searches for eddy successors based on the distance between the candidates at the next time step and the present eddy, which we refer to as a distance-based search (DBS) approach. The distance could be measured between just two points located in Euclidean space, e.g., the methods used by Isern-Fontanet et al. (2006) and Chelton et al. (2011), or between two attribute vectors of eddies in multidimensional space, e.g., the similarity measure used by Penven et al. (2005) and Chaigneau et al. (2008). The second type of algorithms assumes the eddy successor should overlap with the present eddy state, which we refer to as an overlap-based search (OBS) approaches. Henson and Thomas (2008) de-

Foundation item: The National Natural Science Foundation of China under contract Nos 41421001 and 41371378.
*Corresponding author, E-mail: duyy@lreis.ac.cn

veloped this algorithm to trace AEs in the Gulf of Alaska. Xiu et al. (2010) employed the algorithm to study mesoscale eddies in the SCS.

Although these tracking algorithms have been used to survey eddies in local or global seas, there are unresolved issues that may diminish the tracking accuracy and yield unreliable results. First, there is a lack of strategy to determine a global optimal assignment when multiple eddies share common successor candidates. While the DBS algorithm only considers the local nearest neighbor for each eddy, the OBS algorithm initiates a new track for each overlapped eddy in such complex cases (Henson and Thomas, 2008). Second, seeking only the best candidate as successor, current tracking algorithms are incapable of capturing potential eddy splits and mergers during evolution process.

To overcome such disadvantages, this study developed an auto-tracking algorithm using the global nearest neighbor filter (GNNF) and investigated the evolution processes and behaviors of mesoscale eddies in the SCS. The remainder of paper is organized as follows: Section 2 introduces eddy detection data and specifications of the GNNF algorithm; Section 3 first evaluates the tracking performances of the GNNF using synthetic eddy tracks, and examines the user-specified parameters of the method, then discusses the growth and decay in eddy evolution processes based on the tracking results obtained by the GNNF, and finally presents the spatial patterns and property changes during the evolution behaviors, e.g., eddy splits or mergers; and Section 4 concludes the whole study.

2 Methodology

2.1 *Eddy detection data*

This study adopted the hybrid detection (HD) (Yi et al., 2014, 2015) scheme to identify mesoscale eddies in the South China Sea from the daily sea level anomaly (SLA) maps during 1993–2012. The HD scheme combined the Okubo-Weiss criterion and the SLA geometry criterion to achieve a better detection accuracy. A comprehensive description of the HD algorithm is available in our previous studies (Yi et al., 2014, 2015) hence is not repeated in this study. The altimeter data sets used in this study are delay-time version of the gridded (1/4)°×(1/4)° SLA maps provided by the Copernicus Marine and Environment Monitoring Service (CMEMS).

The eddy detection data obtained by the HD are time series data that include eddy centers, represented by SLA local maxima or minima, eddy boundaries, delineated by closed SLA contours, and multi eddy structures that enclose several interacting eddies. Several eddy attributes were also recorded, including radius, amplitude, eddy intensity (*EI*), and vorticity. For each eddy state, the area (S) is delimited by the eddy boundary and the radius (r) is defined by the radius of a circle with equal area, i.e., $r=\sqrt{S/\pi}$. The amplitude (a) is defined by the absolute SLA difference between the eddy center and the eddy boundary contour, i.e., $a=|a_{\mathrm{sl,\,c}}-a_{\mathrm{sl,\,b}}|$, where $a_{\mathrm{sl,\,c}}$ is the SLA of the eddy center, and $a_{\mathrm{sl,\,b}}$ is the SLA of the eddy boundary. The eddy intensity (*EI*) describing the energy density of an eddy (Chaigneau et al., 2008) is defined by $EI=\overline{E_{\mathrm{ke}}}/S$, where the overbar denotes the average eddy kinematic energy over the eddy area and E_{ke} denotes the eddy kinematic energy defined by $\left(U^2+V^2\right)/2$, where U and V are the geostrophic velocity components derived from the SLA. The vorticity is defined by $\omega=\partial V/\partial x-\partial U/\partial y$, and in the Northern Hemisphere, vorticities of the AEs and the CEs are negative and positive values, respectively. To facilitate comparison, we use the absolute value to describe the vorticity magnitude of the AEs or the CEs.

It is worth noting that any other detection scheme could also be used to prepare data for the GNNF tracking method, only if the identifications meet the following requirements: (1) every eddy state identified has a unique identity, *ID*; (2) the specific date (d_s) when an eddy exists is recorded; (3) the specific geography location (X, Y) of an identified eddy is recorded; and (4) these measurements (or records) are organized as a time series at a regular time interval.

The first three requirements define the instantaneous state of an eddy, which can be formalized by [*ID*, d_s, X, Y]. The fourth record is to facilitate the time iteration process of tracking. The identification results obtained from the daily or weekly altimeter maps by the Eulerian methods satisfy all the requirements. However, those derived from the Lagrangian drifting trajectories are at an irregular time interval, and therefore should be interpolated at a constant time interval before the GNNF method could be used to track the eddies.

2.2 *Global nearest neighbor filter*

The GNNF algorithm combined the Kalman filter and optimal data association technologies to recursively recover the evolution processes of ocean eddies in two major steps: (1) use a tracker to estimate the next center location of an eddy; and (2) determine the correct successor, if any, from the candidates at the next time step based on the predicted location.

2.2.1 *Kalman filter tracker*

The first step is fulfilled by implementing eddy trackers using the Kalman filter technologies, which has wide applications in radar tracking (Blackman, 1986), video tracking (Forsyth and Ponce, 2002), and robotics (Thrun, 2000). By using the Kalman filter, the GNNF assumes that the process of an eddy is a hidden Markov chain (Bar-Shalom et al., 2009), and could be modeled by the following linear stochastic equation:

$$\vec{X}_{t+1}=\boldsymbol{A}\vec{X}_t+w\ , \tag{1}$$

where $\vec{X}_t$ is the state vector of an eddy defined by $[x_t, u_t, y_t, v_t]^{\mathrm{T}}$, in which x_t and y_t denote the coordinates of the center location, and u_t and v_t denote the zonal and meridional components of the velocity at the center; $\boldsymbol{A}$ is the state transition matrix specifying how the state vector at time step t is transited to the next state at time step t+1, which could be defined by

$$\boldsymbol{A}=\begin{bmatrix}1 & \mathrm{d}t & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & \mathrm{d}t\\ 0 & 0 & 0 & 1\end{bmatrix}, \tag{2}$$

where dt denotes the time interval, i.e., 1 d; and w is a time invariant random variable representing the white Gaussian process noise, $p(w)$~$N(0, \boldsymbol{Q})$. Assuming the process is a discrete time Wiener process (Bar-Shalom et al., 2001), the covariance matrix $\boldsymbol{Q}$ could be represented in terms of the variance σ_p^2 as follows:

$$\boldsymbol{Q}=\begin{bmatrix}\mathrm{d}t^4/4 & \mathrm{d}t^3/2 & 0 & 0\\ \mathrm{d}t^3/2 & \mathrm{d}t^2 & 0 & 0\\ 0 & 0 & \mathrm{d}t^4/4 & \mathrm{d}t^3/2\\ 0 & 0 & \mathrm{d}t^3/2 & \mathrm{d}t^2\end{bmatrix}\sigma_p^2. \tag{3}$$

The variance σ_p^2 is a user-specified value that reflects how much we can trust our process model in estimating the states. If the variance is set too small, the state estimate considers the prediction from the process model is more accurate than the noisy measurement. On the other hand, if it is set too large, the estimate favors more on accurate measurement than the prediction.

Meanwhile, the GNNF algorithm assumes the measurement be modeled by

$$\vec{Z}_t = \boldsymbol{H}\vec{X}_t + v, \tag{4}$$

where Z_t denotes the measurement vector, $[x_t, y_t]^{\mathrm{T}}$; the matrix $\boldsymbol{H}$ specifying how the eddy measurement is related to the state vector could be defined as

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \tag{5}$$

and v is a time invariant random variable representing the white Gaussian measurement noise, $p(v)$~$N(0, \boldsymbol{R})$. Assuming x and y to be independent white noise with the same variance σ_m^2, the covariance matrix $\boldsymbol{R}$ could be represented by

$$\boldsymbol{R} = \begin{bmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_m^2 \end{bmatrix}. \tag{6}$$

Given the spatial resolution of the merged SLA product is (1/4)°×(1/4)°, the variance σ_m is assumed to be no less than (1/8)°, which is approximately 14 km at the equator. To balance the weights of the measurement and model prediction in estimating the actual eddy state, the variance σ_p^2 is set equal to the variance σ_m^2 in this study.

With these models established, the Kalman filter is able to track eddies through a recursive process of prediction and correction. To be specific, the Kalman filter predicts the next state by computing the posterior $P(X_{t+1}|Z_t)$ based on the total probability theorem:

$$P(X_{t+1}|Z_t) = \int P(X_{t+1}|X_t)\, P(X_t|Z_t)\, \mathrm{d}X_t, \tag{7}$$

where $P(X_{t+1}|X_t)$ is the state transition probability obtained by Eq. (1) and the $P(X_t|Z_t)$ is the estimated posterior probability of the previous state.

When the measurement of the next state is confirmed, which is fulfilled in the second step of the GNNF, the Kalman filter corrects the state estimate by computing the posterior $P(X_{t+1}|Z_{t+1})$ according to the following Bayes' rule:

$$P(X_{t+1}|Z_{t+1}) = \eta P(Z_{t+1}|X_{t+1})\, P(X_{t+1}|Z_t), \tag{8}$$

where η is a normalizer that converts the numerator to a probability, $P(Z_{t+1}|X_{t+1})$ is the measurement probability obtained by Eq. (4), and the posterior $P(X_{t+1}|Z_t)$ is the state estimate before incorporating the most recent measurement.

The final practical formulas shown in Fig. 1 are derived by solving Eqs (7) and (8). Under the assumptions of the Kalman filter, the posteriors $P(X_{t+1}|Z_t)$ and $\mathrm{P}(X_{t+1}|Z_{t+1})$ are both Gaussians (Bar-Shalom et al., 2001) that can be represented by $N(\bar{X}, \bar{\Sigma})$ and $N(\hat{X}, \hat{\Sigma})$, respectively. Thrun et al. (2005) and Welch and Bishop (2001) presented more details about the mathematical derivation process of these formulas.

2.2.2 *Optimal data association*

The second step of the GNNF is accomplished by an optimal data association technology that defines a validation gate to refine the candidates of a correct successor and then makes the best choice by an assignment optimization algorithm. Specifically, the validation gate is placed around the predicted location of an eddy to search the association candidates inside the gate and dismiss the measurements outside that are unlikely originating from the state. Using the Mahalanobis distance, the gate is defined as

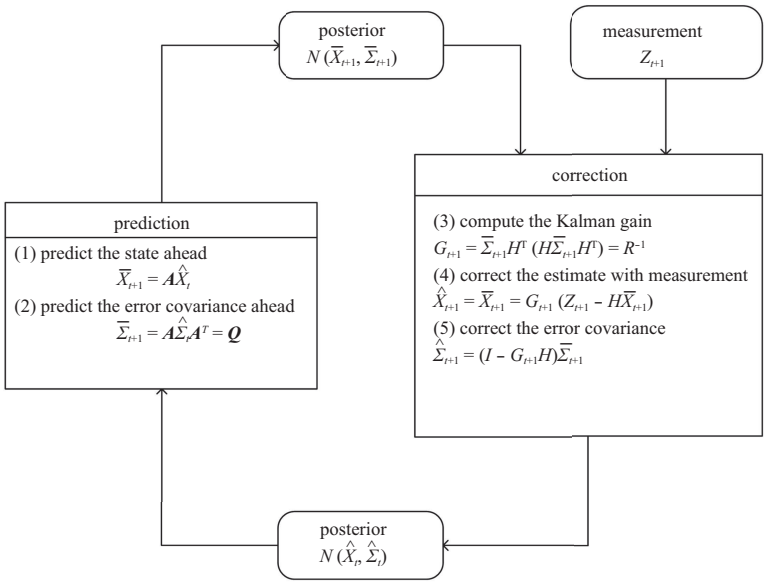

Fig. 1. The flow chart of the tracking procedure based on the Kalman filter. The posteriors $P(X_{t+1}|Z_t)$ and $P(X_{t+1}|Z_{t+1})$ are represented by Gaussians $N(\bar{X}, \bar{\Sigma})$ and $N(\hat{X}, \hat{\Sigma})$, respectively.

$$V(\gamma)=\left\{Z:\left(Z-\bar{X}\right)^{\mathrm{T}}\overline{\sum}\left(Z-\bar{X}\right)\leqslant\gamma\right\}, \tag{9}$$

where γ is the gate threshold, $\bar{X}$ and $\overline{\sum}$ denote the mean and covariance of the predicted state, and Z represents the measurement vector. In two-dimensional space, the area of the validation gate is an ellipsoid with its boundary delimited by the gate threshold γ. As the measurement and the estimate are both Gaussian, the Mahalanobis distance is a χ^2 distribution with a freedom of k degrees (Bar-Shalom and Fortmann, 1988) and the freedom k is the number of the dimensions of the state vector. The threshold γ is usually obtained from the inverse χ^2 cumulative distribution at a significance level of α. In this study, we set α equal to 0.05, indicating that the probability of the true measurements being dismissed by the validation gate is less than 5%.

To track a single eddy, we generally choose the nearest candidate as the successor. But tracking multiple eddies simultaneously is more complicated due to the assignment conflicts in which multiple tracks share the same nearest measurement in the overlap of their validation gates (Fig. 2). To find the best matches of measurements to tracks in such cases is a typical linear assignment problem that can be formally described as follows:

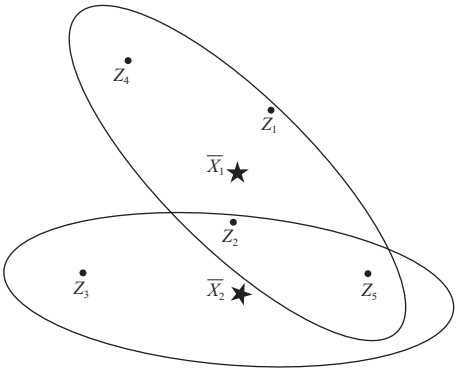

Fig. 2. A schematic diagram showing an assignment conflict. The asterisk and dot symbols indicate the posterior state estimates and the measurements, respectively. The gate region is delineated by the ellipses.

Given that there are m measurements in a gate of n tracks, an $n\times m$ cost matrix is defined by $\boldsymbol{C}=[d_{ij}]$, where d_{ij} denotes the Mahalanobis distance between the ith track and the jth measurement. An assignment is represented by a binary $n\times m$ matrix $B=[x_{ij}]$, where $x_{ij}=1$ if and only if the ith track is associated with the jth measurement. The one-to-one assignment relationship between the measurement and the track is defined by

$$\sum_{j=1}^{m}x_{ij}\leqslant 1, \tag{10}$$

and

$$\sum_{i=1}^{n}x_{ij}\leqslant 1. \tag{11}$$

Subsequently, the goal of data association is to seek the optimal assignment that minimizes the total cost defined by $\sum_{i=1}^{n}\sum_{j=1}^{m}d_{ij}x_{ij}$. The Kuhn-Munkres algorithm (Kuhn, 2010; Munkres, 1957) provides a classic solution to the linear assignment problem. A Java implementation of the algorithm is available via Github by Stern (https://github.com/KevinStern/software-and-algorithms/blob/master/src/main/java/blogspot/software_and_algorithms/stern_library/optimization/HungarianAlgorithm.java), which we have incorporated in the GNNF eddy tracking algorithm.

2.2.3 *Post-tracking procedure*

The GNNF method distinguishes itself from the DBS and the OBS methods in two major aspects. First, in searching the successor of an eddy, other algorithms compare the candidates at the next time step with the center location of the present eddy, but the GNNF compares with the predicted center location estimated by the Kalman filter assuming the eddy to be mobile other than stationary. Second, while few algorithms addressed the assignment conflicts that may occur when tracking multiple eddies, the GNNF adopts an optimal data association strategy to obtain the best matches. Despite these advantages, the GNNF also displays limitations in detecting false connections, short disappearance, splits, and mergers of ocean eddies.

So, we developed the post-tracking procedure to scrutinize the tracking result based on our prior knowledge and observation of eddies. The propagation speed of the ocean eddies in the SCS is generally in the order of 10 km/d (Chen et al., 2010), but the radius of a mesoscale eddy is at least 45 km, which suggests, with few exceptions, an eddy should have a substantial domain overlap with the state identified the next day on the SLA map. Based on this, we define the link between two consecutive states of an eddy is a false connection if there is no overlap in the eddy domain. The false connections in the tracking result are removed.

Eddies may occasionally disappear for a few days and then reappear for various reasons, e.g., sampling errors or measurement noise of the SLA, failures of the eddy identification, or simply the result of eddy-eddy interactions. Extending the search of eddies a few steps further is a straightforward strategy that many studies have adopted to recover such short breaks of eddies. However, a threshold of the extended searching time should be appropriately determined, because if the threshold is too small, eddy tracks tend to be falsely divided into pieces by the short breaks, and if it is too large, different tracks may be incorrectly patched together. In this study, we use 3 d as the maximal extensible time to search for the reappearing eddies. The searching area (or validation gate) is still centered at the predicted location given by the Kalman filter except it would expand with the extended searching time due to the increased estimation uncertainty. A reappearing eddy is confirmed if it overlaps with the predicted center location. The experiment in Section 3.1.2 shows that the error rate is less than 5%.

There were occurrences of the eddy splits and mergers in the SCS (Chu et al., 2014; Nan et al., 2011), but they are not identifiable if algorithms only consider one-to-one assignment in associating eddy measurements with the right tracks. So, in the post-tracking procedure, we further used the domain overlap of eddies as a practical and reasonable criterion to detect the eddy splits and mergers concealed in the result. Specifically, a split of an eddy is recorded if its area is overlapped by multiple eddies identified at the next time step and a merger is recorded if its area is overlapped by multiple eddies identified at the previous time step. We define an eddy as a “complex” track if its life cycle con-

tains branches due to splits or/and mergers, otherwise it has a "simple" track.

3 Results and discussion

3.1 *Method validation and parameter selection*

The tracking performances of the GNNF are compared against two major algorithms used in previous studies, i.e., the DBS and the OBS. In this study, the DBS method concatenates the closest successor candidate within a 150 km searching radius of the target eddy center whereas the OBS method identifies the first overlapping candidate eddy at the next time step as the successor. As the DBS and the OBS are incapable of detecting the eddy splits or mergers, we did not include the post-tracking procedure in this comparison. Performances of these methods were evaluated using synthetic eddy tracks. The appropriateness of the parameter selection in the GNNF method was also examined.

3.1.1 *Validation with synthetic eddy tracks*

The synthetic eddy tracks are generated based on a brief survey of the tracks that were identically recovered by the three methods during 1993–1994. There are totally 8 871 anticyclonic and 8 845 cyclonic eddy states in the survey, which corresponds to averagely 12 AEs or CEs identified per day. As shown in Fig. 3, the mean life time, radius, and propagation speed of these common eddies are about 7 weeks, 70 km, and 10 km/d (0.12 m/s), respectively, and the standard deviations of propagation speed and radius over the life process of eddies are about 15 km/d (0.17 m/s) and 20 km, respectively. Although the propagation speeds of these eddies are consistent with the westward phase speed of the first baroclinic Rossby wave which is about 0.05–0.35 m/s (Cai et al., 2008), the eddies may not be free Rossby waves due to different geographical variabilities of the speed (Chen et al., 2011).

In the experiment, eddies were randomly generated in an area of 2 000 km×2 000 km which is approximately the same area of the SCS, with equal lifetime of 49 d, initial radius ranges from 35 to 250 km, and initial zonal and meridional velocities range from –10 to 10 km/d. Variations of velocities and radius during the evolution process of an eddy are described by specific standard deviations of normal distributions symmetric about its initial values. There are three variables in the experiment, i.e., the number (n) of eddy tracks, the standard deviation (σ_v) of velocities, and the standard deviation (σ_r) of radius, and we examined them one at a time by adjusting one variable and controlling the others. Each experiment was repeated 1 000 times to assess the average performance of methods.

As shown in Fig. 4, the mean false transition rate (FTR), which is the ratio of incorrect transitions to the total transitions, escalates linearly and exponentially with the increase of the track number and the standard deviation of velocity but displays only flat fluctuations with the increase of the standard deviation of radius, and the GNNF method achieved the lowest FTR among the three methods in every experiment. In particular, when the synthetic tracks were generated using the experiment variables (n=12, σ_v =15 km/d, and σ_r =20 km) to best approximate the survey results, the mean FTR of the GNNF, the DBS, and the OBS are about 0.2%, 0.4%, and 0.5%, confirming the outperformance of GNNF in tracking ocean eddies over the other two methods.

3.1.2 *Parameter selection in eddy tracking*

There are three user-specified parameters in the tracking algorithm, i.e., the measurement noise σ_m=14 km, the process noise σ_p=14 km, and the maximal extensible time of 3 d for

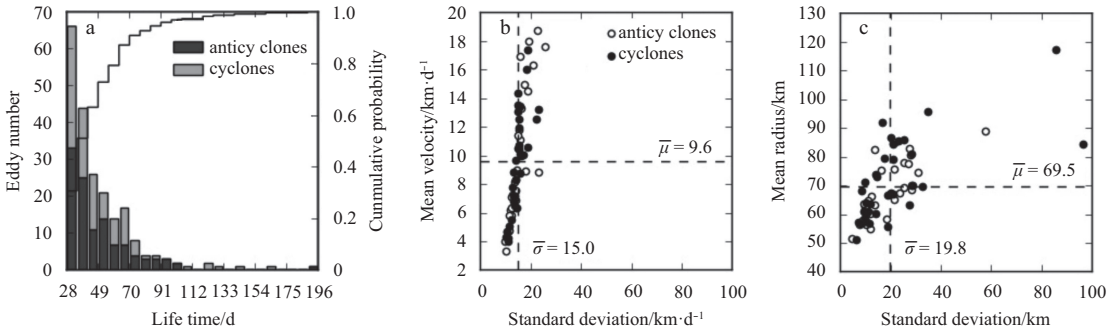

Fig. 3. A histogram graph showing the distribution of eddies with different life-spans (a), scatter plots showing the velocity (b) and radius (c) of the common eddy tracks that were identified by all the three methods.

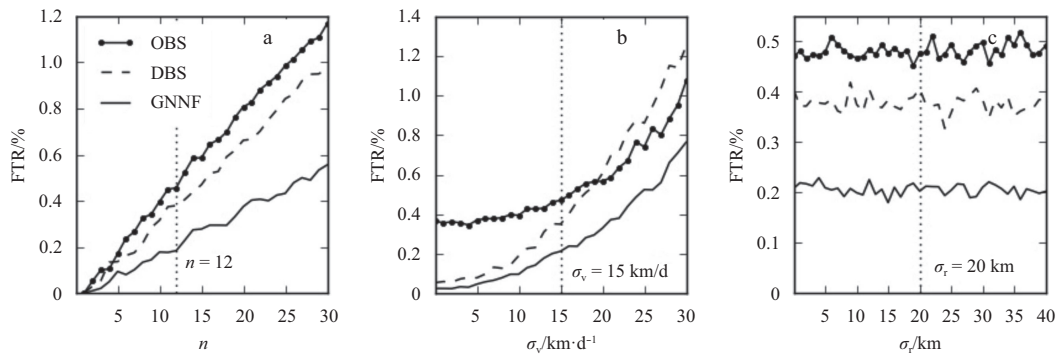

Fig. 4. Impact of track number (a), variations in velocity (b) and radius (c) on tracking performance of the three methods.

searching reappearing eddies. First, to examine the effect of the process noise and the measurement noise on the performance of the GNNF, we again used the synthetic eddy tracks generated with n=12, σ_v =15 km/d, and σ_r =20 km and conducted a series experiments with different combination values of the process noise and the measurement noise. As shown in Fig. 5, the FTR declines to less than 0.4% and generally fluctuates around 0.2% when the process noise and the measurement noise are both greater than 5 km, which suggests the noise settings in this study are acceptable for tracking ocean eddies.

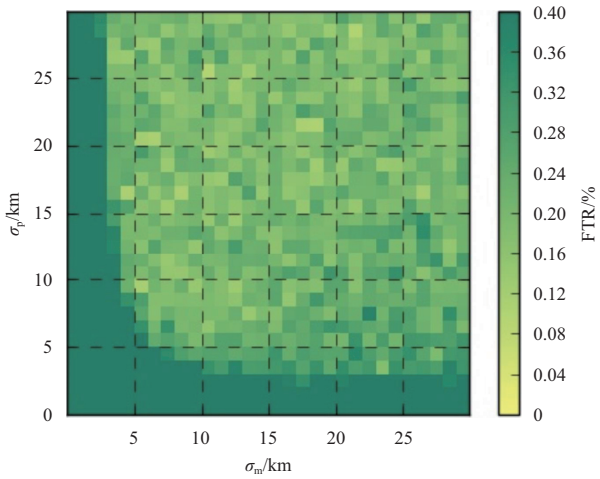

Fig. 5. False transition rate of GNNF with different value combinations of process and measurement noise.

Second, to find an appropriate threshold for the maximal extensible time, we used the GNNF algorithm (not including the post-track procedure) to determine a following eddy state for every ended track regardless how long the disappearance endured. The following eddy should overlap with the eddy center location projected by the GNNF. As we believe that a real reappearing eddy should always be the beginning state of another track, assigning a following eddy state that has histories to an ended track is regarded as a spurious disappearing event. A survey of such assignment errors is shown in Fig. 6. The error rate escalates quickly with the increase of disappearance time period. Particularly, the error rate is less than 5% when the disappearance period is shorter than 3 d, but soon climbs to nearly 20% after then. So, a maximal period of 3 d is selected in this study as the extensible time threshold to connect the tracks separated by short disappearances.

3.2 *Evolution processes*

3.2.1 *Growth and decay*

Using the GNNF algorithm, we recovered 2 394 AE and 2 519 CE tracks that survived more than a week in the SCS over 20 a. Simple eddy tracks, which account for 70% of the total number, were selected to investigate the growth and decay of eddy life-cycles. The age of each eddy was normalized by its life-span and the eddy attributes including the radius, amplitude, EI, and vorticity were standardized for each track, in order to compare eddies of different life-spans and observe the relative changes of these attributes during the life course. Besides, the life cycle was divided into ten ages at the interval of 0.1 and the attributes of different eddies were averaged at each age.

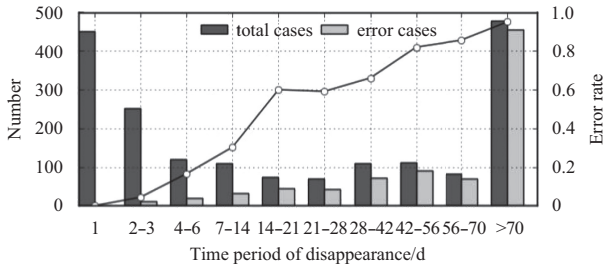

Fig. 6. Error cases in searching the reappearing eddies.

As shown by Fig. 7, the changes of the eddy radius, amplitude, and vorticity are well approximated by parabolas that climax at the age of 0.5. These attributes increase during the first half of lifetime and decline during the second half. The standardized variation ranges of the radius and the amplitude are about 0.5, greater than the range of vorticity which is about 0.3. The variation of EI, however, is approximated by a shallow parabola that declines during first stages and increases after the age of 0.5. The EI of eddies which varies between –0.1 and 0.1 throughout the life does not change as much as other attributes. There is little difference between the AEs and the CEs in the variations of these attributes except the CEs show slightly greater values during the first half of lifetime and the AEs the second half. Surveying the mesoscale eddies in the North Pacific Ocean, Liu et al. (2012) found a mature stage that eddy attributes generally stayed stable for 60% of their life-span. However, this stage was not identified in our analysis probably because the seasonal circulations, the orographic winds, and the strong currents of the SCS made it difficult for eddies to survive or travel as long and freely as the eddies in the North Pacific Ocean where the eddy movements were mainly driven by the β effect (Liu et al., 2012).

3.2.2 *Complex evolution process*

Among 1 468 complex tracks of the results, we selected an anticyclonic eddy as a case study to examine the evolution processes and behaviors recovered by the algorithm. Figure 8 schematically shows the evolution process of this particular AE from August 16 to September 18, 2007. During this short period, a merging event and a splitting event occurred northwest of Luzon on August 21 and September 9, respectively (Fig. 9). The eddy merger was also identified by Nan et al. (2011) in weekly SLA maps and *in situ* observations. Owing to the split and merger, the AE is separated into five evolving sequences, namely Eddies 1–5. Eddy 1 was interacting with Eddy 2 in a multieddy structure before the occurrence of merger. Also, Eddy 4 and Eddy 5 constituted another multieddy structure after the splitting event and were completely isolated after September 17. Eddies 1 and 4 were relatively weak but still satisfied the eddy detection criteria of the HD scheme (Yi et al., 2014). The vertical structure of Eddy 1 in geostrophic currents was well captured by Nan et al. (2011). The algorithm successively identified the temporal disappearance of Eddy 4 on September 11.

Figure 10 displays the property changes during this evolution process. As the result of the merging event, Eddy 3 has a larger radius of about 140 km but the EI and vorticity in absolute value were both lower than those of Eddies 1 and 2. The amplitude of Eddy 3 was the same as Eddy 2 at the beginning but then increased from 0.13 to 0.18 m from August 21 to September 9. After the splitting, Eddy 4 spawned from Eddy 5 at the northeast part of

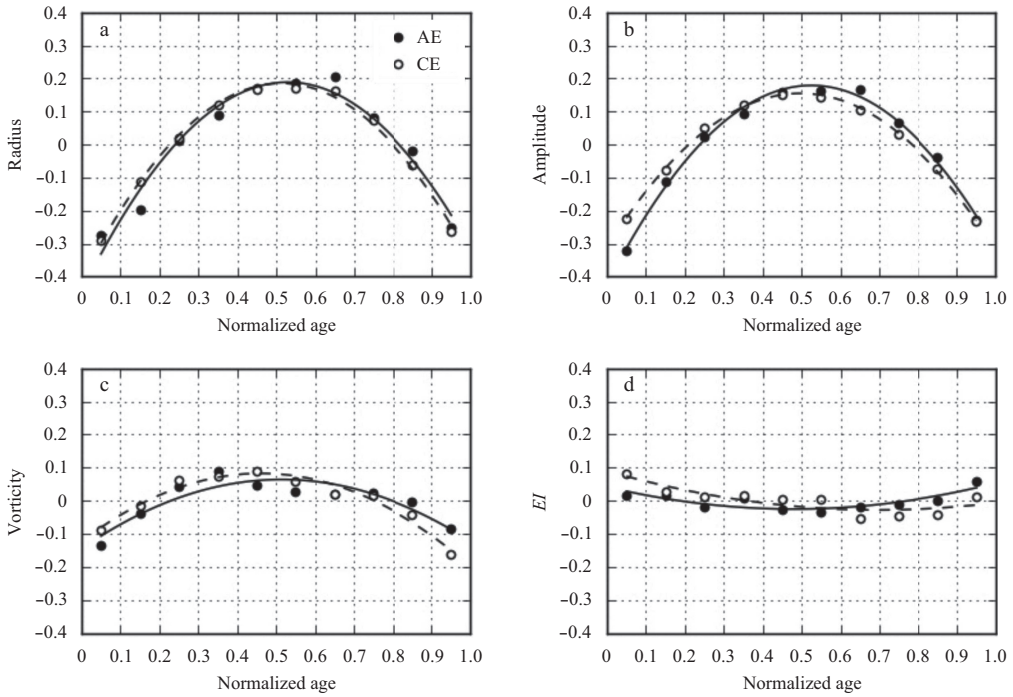

Fig. 7. Variations of standardized eddy properties as functions of normalized age.

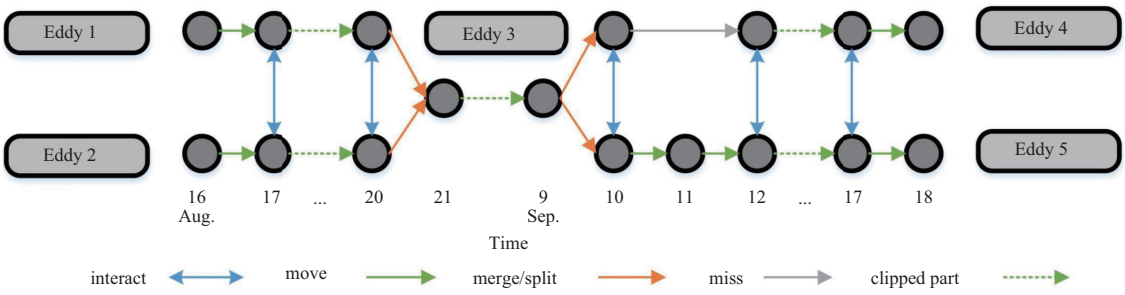

Fig. 8. A diagram showing the evolution process of the anticyclonic eddy from 2007-08-16 to 2007-09-18.

the eddy body, weak but of strong vorticity, and survived another 8 d until it dissipated. Eddy 5 absorbing most of the energy of Eddy 3, however, lasted more than 3 months until another eddy split occurred (not shown).

3.3 *Evolution behaviors*

3.3.1 *Spatial distributions*

The GNNF algorithm allows us to identify several behaviors during eddy life cycles, such as the generation, dissipation, short disappearance, split, and merger. To investigate the spatial distribution of these evolution behaviors, we calculated the kernel density of the locations where the behaviors occurred. The kernel function is based on a quadratic kernel function described in Silverman (1986), and the output is a smooth surface with the magnitude of each cell showing the density estimate.

Figure 11a displays the density difference between the eddy generation and the dissipation locations. The positive densities are mainly located northwest and southwest of Luzon, indicating the generation there is dominant over dissipation. As discussed by Wang et al. (2008), the orographic winds during the winter monsoon season are responsible for the eddy genesis in these regions. He et al. (2015) found the Kuroshio intrusion could be as important as wind stress curl in the generation of deep sea CEs. Also, eddies in these areas demonstrate different responses to El Niño-Southern Oscillation (He et al., 2016). The dissipation areas are in principle located along the west boundary of the SCS, which is the result of eddies' westward propagation driven by the beta effect. The negative densities near the Xisha Islands indicate the eddy dissipation is more frequent than generation. Studying the cyclonic eddies that yearly form south of Dongsha and travelled along the northern continental slope, Chow et al. (2008) found the underwater bay in this location trapped and terminated all the eddies they observed. Nan et al. (2011) traced the evolution of three long-lived AEs generated in the northern SCS and also found their collapses between the Xisha Islands and Hainan Island.

Figures 11b to d further display the density distribution of short disappearance, splits, and mergers. The short disappearances of eddies were concentrated west of the Luzon Strait and northwest of Luzon Island. Besides the two regions, the eddy splits and mergers also frequently occurred southeast of Vietnam. The concentrations of such eddy behaviors over these regions suggest frequent eddy-eddy interactions and evolution interruptions by marine topography (Simmons and Nof, 2000) or strong currents, such as the Kuroshio intrusions or the eastward

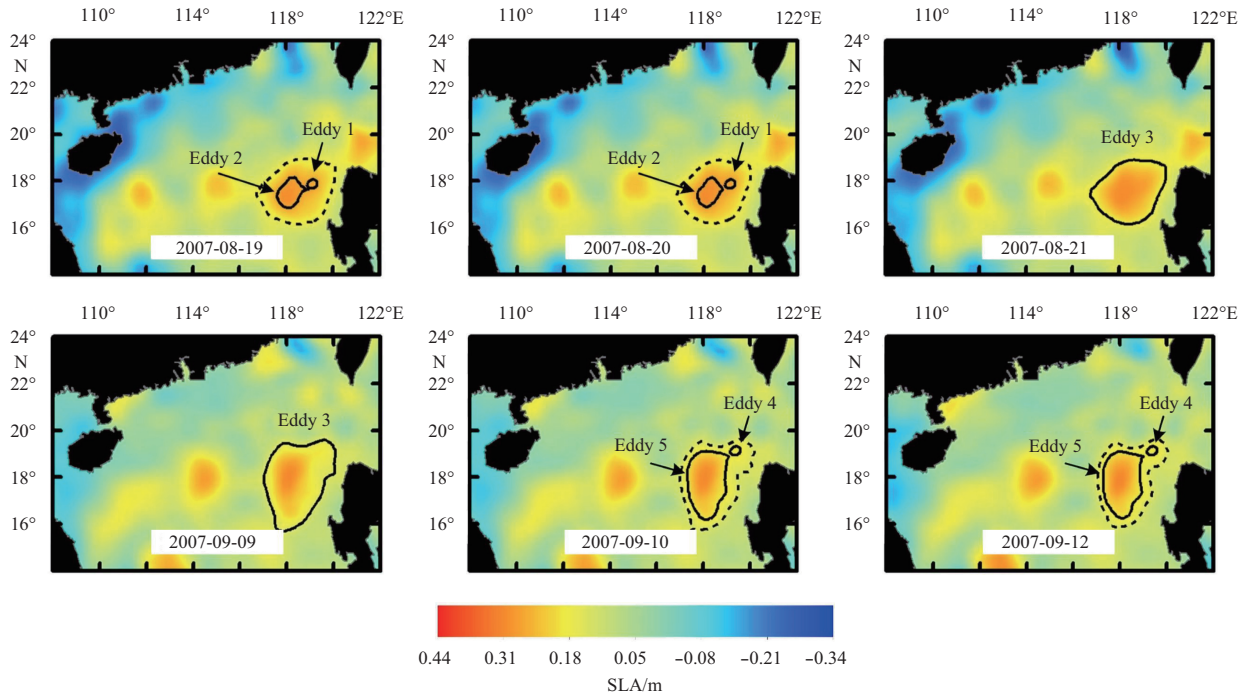

Fig. 9. Snapshots that capture the merging and splitting event.

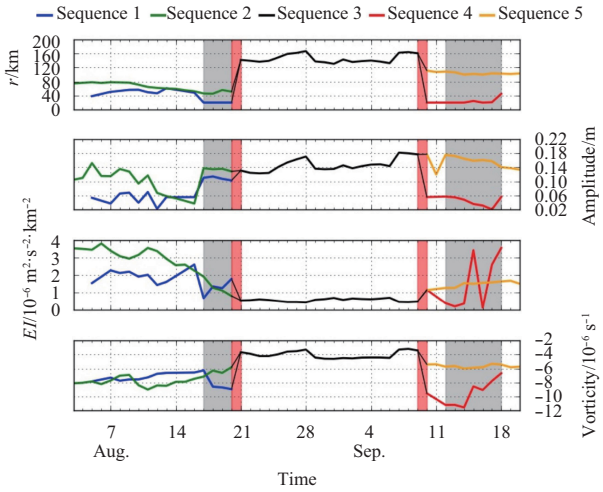

Fig. 10. Property changes during the evolution process. The grey and red shading areas denote eddy interactions and eddy split/merger, respectively.

offshore jets southeast of Vietnam. But the specific physical mechanism needs further investigations.

3.3.2 *Property changes*

To evaluate the impact of eddy split and merger on the eddy properties, e.g., radius, amplitude, *EI*, and vorticity, we calculated the difference between the standardized property values of the eddy state before and after a splitting/merging event. As summarized in Table 1, the eddy splits generally result in a decrease of the radius and the amplitude whereas the eddy mergers induce growing up. The *EI* and the vorticity, on the contrary, increase during splits and slides down during mergers. The Mann-Whitney *U* test shows that these changes are significant although there are overlaps between their probability density distributions (Fig. 12).

4 Conclusions

This study developed an automatic tracking algorithm using the global nearest neighbor filter to investigate the evolution processes and behaviors of mesoscale eddies in the SCS. The new tracking algorithm incorporated the Kalman filter to estimate the eddy center location at the next time step, and the optimal data association strategy to find the right successor. A post-tacking procedure was further implemented based on the moving characteristics of eddies to address the issues like false connections, short disappearances, the eddy mergers, and splits. The GNNF algorithm outperformed the other two commonly used algorithms, i.e., DBS and OBS, in the method evaluations, which showed the false transition rate of the GNNF was about 0.2% in tracking synthetic eddy tracks.

The GNNF algorithm recovered a total of 4 913 eddy tracks that survived more than a week in the SCS from 1993 to 2012. The analysis of the growth and decay based on 3 445 simple eddy tracks showed the eddy radius, amplitude, and vorticity smoothly increased during the first half of lifetime and declined during the second half following a parabola opening downwards, whereas the variation of EI was little throughout the life. Mature stages during which eddies maintain stable properties were not observed in the evolution processes of eddies in the SCS.

Several evolution behaviors were identified by the GNNF algorithm, including generation, dissipation, short disappearance, split, and merger. The genesis of eddies mainly clustered northwest and southwest of the Luzon Island whereas the dissipations concentrated the Xisha Islands where the underwater bay trapped and terminated eddies. West of Luzon Strait, northwest of the Luzon Island, and southeast of Vietnam are regions where the eddy splits and mergers were frequently observed. The short appearances mainly distributed in the first two regions. Moreover, the eddy splits generally resulted in a decrease of the radius and the amplitude whereas the eddy mergers induced growing up. The *EI* and the vorticity, on the contrary, were strengthened in the eddy splits and diminished in the mergers.

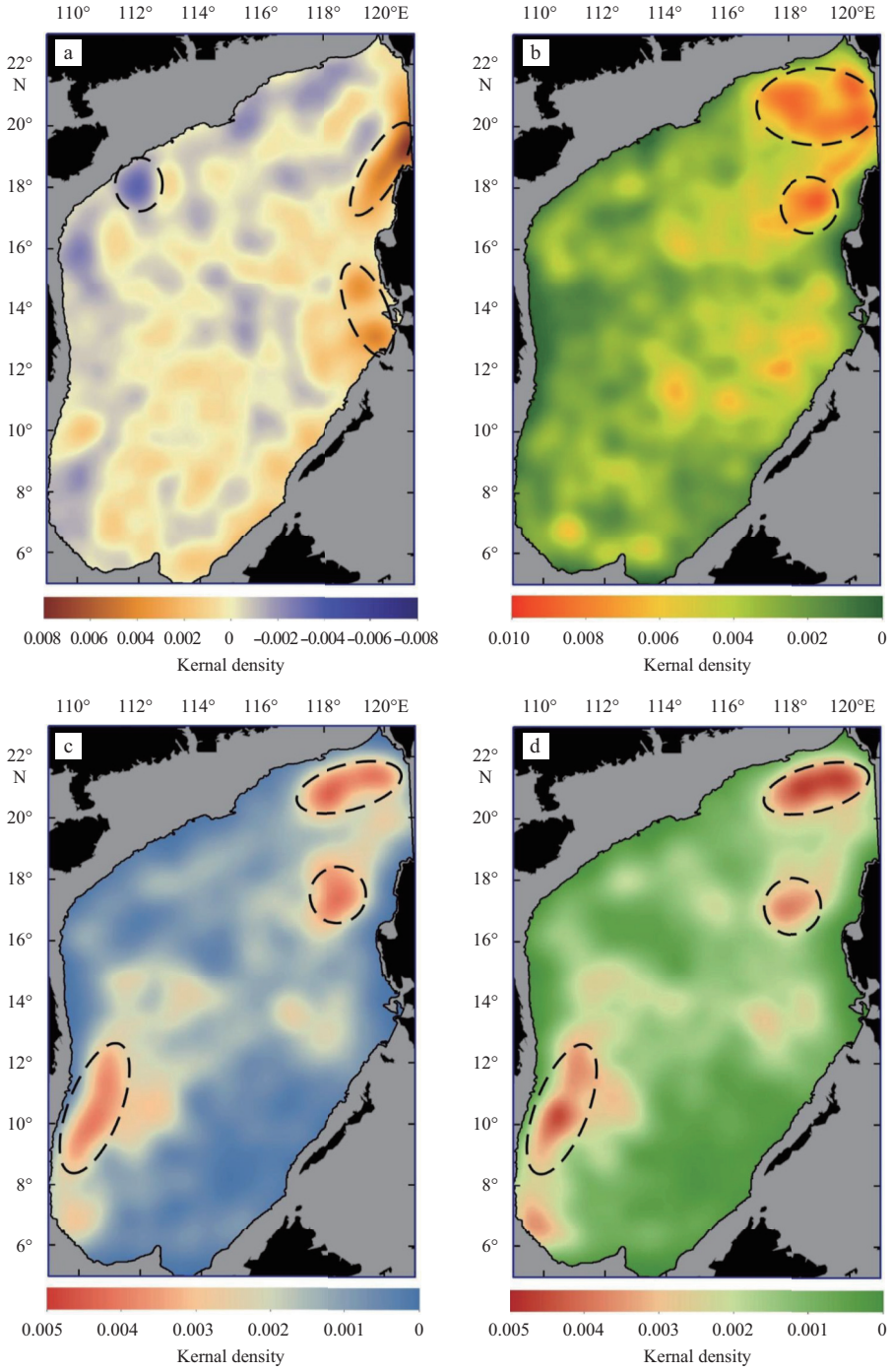

Fig. 11. Kernel density distribution of evolution behaviors. Density difference between eddy generation and dissipation (a), and density maps of short disappearances (b), eddy splits (c), and eddy mergers (d). The color shadings show the magnitudes of the kernel density estimate.

Table 1. Eddy property changes in splits and mergers

	Split				Merger				
	Mean difference			Std Dev	Mean differences			Std Dev	*P*-value[1)]
	AE	CE	Total		AE	CE	Total		
Radius[2)]	–0.54	–0.52	–0.53	1.81	0.58	0.53	0.56	1.78	0
Amplitude[2)]	–0.40	–0.39	–0.39	1.03	0.41	0.33	0.37	1.06	0
EI[2)]	0.77	0.74	0.75	1.26	–0.77	–0.75	–0.76	1.30	0
Vorticity[2)]	1.19	1.21	1.20	1.05	–1.18	–1.15	–1.16	1.09	0

Note: [1)] The *p*-value of Mann-Whitney *U* test; [2)] these attributes are normalized so that the values have no unit. Std Dev represents the standard deviation of the attribute value changes during eddy splits (or mergers).

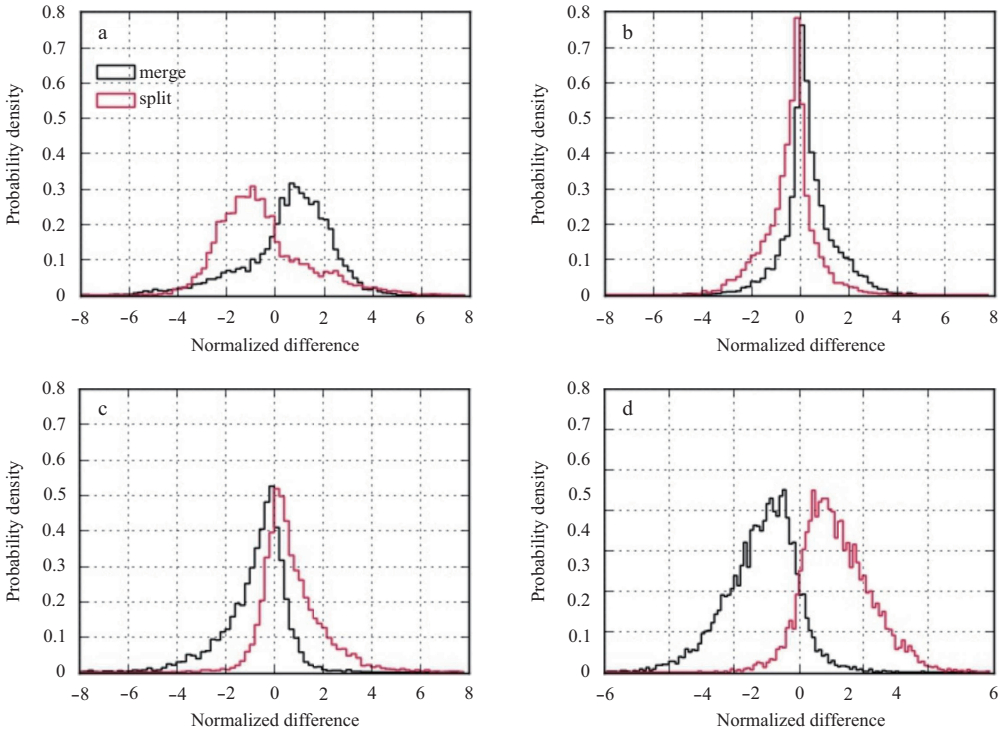

Fig. 12. Probability density functions of variations of eddy radius (a), amplitude (b), *EI* (c), and vorticity (d) during eddy splits and mergers.

Acknowledgements

The Ssalto/Duacs altimeter products were produced and distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS).

References

Ari Sadarjoen I, Post F H. 2000. Detection, quantification, and tracking of vortices using streamline geometry. Computers & Graphics, 24(3): 333–341, doi: 10.1016/S0097-8493(00)00029-7

Bar-Shalom Y, Daum F, Huang J. 2009. The probabilistic data association filter. IEEE Control Systems Magazine, 29(6): 82–100, doi: 10.1109/MCS.2009.934469

Bar-Shalom Y, Fortmann T E. 1988. Tracking and Data Association. Boston: Academic Press

Bar-Shalom Y, Li X R, Kirubarajan T. 2001. Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software. New York: John Wiley & Sons, Inc, doi: 10.1002/0471221279

Blackman S S. 1986. Multiple-target Tracking with Radar Applications. Dedham, MA: Artech House, Inc

Cai Shuqun, Long Xiaoming, Wu Renhao, et al. 2008. Geographical and monthly variability of the first baroclinic Rossby radius of deformation in the South China Sea. Journal of Marine Systems, 116(C6): C06018, doi: 10.1029/2010jc006716

Chaigneau A, Gizolme A, Grados C. 2008. Mesoscale eddies off Peru in altimeter records: identification algorithms and eddy spatio-temporal patterns. Progress in Oceanography, 79(2–4): 106–119, doi: 10.1016/j.pocean.2008.10.013

Chelton D B, Schlax M G, Samelson R M. 2011. Global observations of nonlinear mesoscale eddies. Progress in Oceanography, 91(2): 167–216, doi: 10.1016/j.pocean.2011.01.002

Chen Gengxin, Hou Yijun, Chu Xiaoqing. 2011. Mesoscale eddies in the South China Sea: mean properties, spatiotemporal variability, and impact on thermohaline structure. Journal of Geophysical Research, 116(6): C06018, doi: 10.1029/2010JC006716

Chen Gengxin, Hou Yijun, Zhang Qilong, et al. 2010. The eddy pair off western Vietnam: interannual variability and impact on thermohaline structure. Continental Shelf Research, 30(7): 715–723, doi: 10.1016/j.csr.2009.11.013

Chow C H, Hu J H, Centurioni L R, et al. 2008. Mesoscale Dongsha cyclonic eddy in the northern South China sea by drifter and satellite observations. Journal of Geophysical Research, 113(C4): C04018, doi: 10.1029/2007JC004542

Chu Xiaoqing, Xue Huijie, Qi Yiquan, et al. 2014. An exceptional anticyclonic eddy in the South China Sea in 2010. Journal of Geophysical Research, 119(2): 881–897, doi: 10.1002/2013JC009314

Dong Changming, Liu Yu, Lumpkin R, et al. 2011. A scheme to identify loops from trajectories of oceanic surface drifters: an application in the Kuroshio extension region. Journal of Atmospheric and Oceanic Technology, 28(9): 1167–1176, doi: 10.1175/JTECH-D-10-05028.1

Forsyth D A, Ponce J. 2002. Computer Vision: A Modern Approach. Englewood Cliffs, NJ: Prentice Hall

He Yinghui, Cai Shuqun, Wang Dongxiao, et al. 2015. A model study of Luzon cold eddies in the northern South China Sea. Deep Sea Research: Part I. Oceanographic Research Papers, 97: 107-123, doi: http://dx.doi.org/10.1016/j.dsr.2014.12.007

He Yinghui, Xie Jieshuo, Cai Shuqun. 2016. Interannual variability of winter eddy patterns in the western South China Sea. Geophysical Research Letters, 43(10): 5185–5193, doi: 10.1002/2016GL068842

Henson S A, Thomas A C. 2008. A census of oceanic anticyclonic eddies in the Gulf of Alaska. Deep Sea Research: Part I. Oceanographic Research Papers, 55(2): 163–176, doi: 10.1016/j.dsr.2007.11.005

Isern-Fontanet J, García-Ladona E, Font J. 2006. Vortices of the Mediterranean Sea: an altimetric perspective. Journal of Physical Oceanography, 36(1): 87–103, doi: 10.1175/JPO2826.1

Kuhn H W. 2010. The Hungarian method for the assignment problem. In: Jünger M, Liebling T M, Naddef D, et al., eds. 50 Years of Integer Programming 1958–2008: From the Early Years to the State-of-the-Art. Berlin Heidelberg: Springer, 29-47, doi: 10.1007/978-3-540-68279-0_2

Liu Yu, Dong Changming, Guan Yuping, et al. 2012. Eddy analysis in the subtropical zonal band of the North Pacific Ocean. Deep

Sea Research: Part I. Oceanographic Research Papers, 68: 54–67, doi: 10.1016/j.dsr.2012.06.001

Liu Qinyu, Kaneko A, Su Jilan. 2008. Recent progress in studies of the South China Sea circulation. Journal of Oceanography, 64(5): 753–762, doi: 10.1007/s10872-008-0063-8

Morrow R, Birol F, Griffin D, et al. 2004. Divergent pathways of cylonic and anti-cyclonic ocean eddies. Geophysical Research Letters, 31(24): L24311, doi: 10.1029/2004GL020974

Munkres J. 1957. Algorithms for the assignment and transportation problems. Journal of the Society for Industrial and Applied Mathematics, 5(1): 32–38, doi: 10.1137/0105003

Nan Feng, He Zhigang, Zhou Hui, et al. 2011. Three long-lived anticyclonic eddies in the northern South China Sea. Journal of Geophysical Research, 116(C5): C05002, doi: 10.1029/2010JC006790

Nencioli F, Dong Changming, Dickey T, et al. 2010. A vector geometry-based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the Southern California Bight. Journal of Atmospheric and Oceanic Technology, 27(3): 564–579, doi: 10.1175/2009JTECHO725.1

Okubo A. 1970. Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. Deep Sea Research and Oceanographic Abstracts, 17(3): 445–454, doi: 10.1016/0011-7471(70)90059-8

Penven P, Echevin V, Pasapera J, et al. 2005. Average circulation, seasonal cycle, and mesoscale dynamics of the Peru Current System: a modeling approach. Journal of Geophysical Research, 110(C10): C10021, doi: 10.1029/2005JC002945

Silverman B W. 1986. Density Estimation for Statistics and Data Analysis. New York: Chapman & Hall

Simmons H L, Nof D. 2000. Islands as eddy splitters. Journal of Marine Research, 58(6): 919–956, doi: 10.1357/002224000763485755

Su Jilan. 2004. Overview of the South China Sea circulation and its influence on the coastal physical oceanography outside the Pearl River Estuary. Continental Shelf Research, 24(16): 1745–1760, doi: 10.1016/j.csr.2004.06.005

Thrun S. 2000. Probabilistic algorithms in robotics. Ai Magazine, 21(4): 93–109, doi: 10.1609/aimag.v21i4.1534

Thrun S, Burgard W, Fox D. 2005. Probabilistic Robotics. Cambridge, Mass: MIT Press

Wang Guihua, Chen Dake, Su Jilan. 2008. Winter eddy genesis in the western South China Sea due to orographic wind jets. Journal of Physical Oceanography, 38(3): 726–732, doi: 10.1175/2007JPO3868.1

Wang Guihua, Su Jilan, Chu P C. 2003. Mesoscale eddies in the South China Sea observed with altimeter data. Geophysical Research Letters, 30(21): 2121, doi: 10.1029/2003GL018532

Weiss J. 1991. The dynamics of enstrophy transfer in two-dimensional hydrodynamics. Physica D: Nonlinear Phenomena, 48(2–3): 273–294, doi: 10.1016/0167-2789(91)90088-Q

Welch G, Bishop G. 2001. An Introduction to the Kalman filter. Los Angeles, CA, USA: ACM Press

Williams S, Petersen M, Bremer P T, et al. 2011. Adaptive extraction and quantification of geophysical vortices. IEEE Transactions on Visualization and Computer Graphics, 17(12): 2088–2095, doi: 10.1109/TVCG.2011.162

Xiu Peng, Chai Fei, Shi Lei, et al. 2010. A census of eddy activities in the South China Sea during 1993–2007. Journal of Geophysical Research, 115(C3): C03012, doi: 10.1029/2009JC005657

Yi J, Du Y, He Z, et al. 2014. Enhancing the accuracy of automatic eddy detection and the capability of recognizing the multi-core structures from maps of sea level anomaly. Ocean Science, 10(1): 39–48

Yi Jiawei, Du Yunyan, Zhou Chenghu, et al. 2015. Automatic identification of oceanic multieddy structures from satellite altimeter datasets. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(4): 1555–1563

Yuan Dongliang, Han Weiqing, Hu Dunxin. 2007. Anti-cyclonic eddies northwest of Luzon in summer-fall observed by satellite altimeters. Geophysical Research Letters, 34(13): L13610, doi: 10.1029/2007GL029401