

Abstract

Spatial variation is important for studying the mesoscale eddies in the South China Sea (SCS). To investigate such spatial variations, this study made a clustering analysis on eddies' distribution using the *K*-means approach. Results showed that clustering tendency of anticyclonic eddies (AEs) and cyclonic eddies (CEs) were weak but not random, and the number of clusters were proved greater than four. Finer clustering results showed 10 regions where AEs densely populated and 6 regions for CEs in the SCS. Previous studies confirmed these partitions and possible generation mechanisms were related. Comparisons between AEs and CEs revealed that patterns of AE are relatively more aggregated than those of CE, and specific distinctions were summarized: (1) to the southwest of Luzon Island, AEs and CEs are generated spatially apart; AEs are likely located north of 14° N and closer to shore, while CEs are to the south and further offshore; (2) the Central SCS and Nansha Trough are mostly dominated by AEs; (3) along 112° E, clusters of AEs and CEs are located sequentially apart, and the pair off Vietnam represents the dipole eddies; (4) to the southwest of Dongsha Islands, AEs are concentrated to the east of CEs. Overlaps of AEs and CEs in the northeastern and Southern SCS were further examined considering seasonal variations. The north-eastern overlap represented near-concentric distributions while the southern one was a mixed effect of seasonal variations, complex circulations and topography influences.

1 Introduction

The South China Sea (SCS) is the largest semi-enclosed marginal sea in Northwest Pacific Ocean. It has a large NE–SW oriented abyssal basin with greatest depth of 5567 m. There is an approximate 3.5×10^6 km² of total area with mean depth of 1212 m. The SCS is well linked to the adjacent seas through several straits or channels, connecting to the Pacific Ocean through the Luzon Strait, to the Sulu Sea through the

OSD

9, 3451–3483, 2012

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

2 Data and method

2.1 Eddy data

The sea surface height (SSH) data used for eddy detection are produced by the operational 1/32° global Naval Research Laboratory's (NRL) Layered Ocean Model (NLOM) nowcast/forecast system (http://www7320.nrlssc.navy.mil/global_nlom32/skill.html). The model assimilates SSH from 3 satellite altimeters (ENVISAT, JASON-1 interleaved and JASON-2). Altimeter track data over a 3-day window are assimilated each day.

The reliability of the model performance has been validated by Du et al. (2011). Comparison between anomaly of SSH (SSHA) derived by averaging NLOM SSH data over 2003–2009 and sea level anomalies (SLA) observations provided by Archiving Validation and Interpretation of Satellite Data in Oceanography (AVISO) showed that deterministic eddies founded in SLA were well reflected by the SSHA. Comparison of current speed simulated by NLOM with cruise measurements also confirms the data validation. More details can be referred to (Du et al., 2011).

The eddy identification algorithm adopted in this paper is an adaptation of SSH-based automated procedure proposed by Chelton et al. (2011). Brief descriptions can be found in Appendix A and Wang et al. (2012). The 1/32° NLOM numerical output started from 1 November 2003. There are some null data from 2003 to 2004. So, we used the output SSH field from 2005 to 2011 for eddy identification and tracking. 371 eddies (159 AEs and 212 CEs) were identified and the map of study area and birth locations of these eddies are showed in Fig. 1. To be noted that only the centre of each eddy's birth state, which denotes where it generates, is taken into clustering analysis.

2.2 Clustering analysis

The goal of clustering analysis is to group data into meaningful subgroups (clusters) (Jain and Dubes, 1988; Kaufman and Rousseeuw, 1990). The more the similarity

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the assumption that good performance will be achieved if the data are clustered by right cluster number, optimal values of evaluation parameters like SSE or SC usually hint the answers. In other words, the knees of the SSE curve or the peaks of the SC curve shall indicate the possible right cluster number.

3 Results and discussion

First, the Hopkins statistic results are presented to answer the question (a): if eddies' distribution has a clustering tendency or not. Second, correct cluster number is discussed separately for cyclonic eddy, anticyclonic eddy and the whole, since we are not sure about the answer to the question (b). The third part gives an intensive discussion about how AE and CE are different in spatial variations.

3.1 Hopkins statistic

It is suggested that the amount of sample data (P) be less than 10 % of the whole when computing the Hopkins statistic (Ripley, 1977). To facilitate observation, we made a series of experiments with increasing the sample data from 1 to 10 % of the total amount. 100 trials are carried out in each calculation of Hopkins value (H), and the mean values are plotted in Fig. 2. Results of random pattern are presented for comparison.

H value of random patterns fluctuates around 0.5, and becomes stable at 0.5 with accretion of the sample amount as expectation. Comparatively, H values of AEs, CEs and all eddies are clearly above 0.5 axis. The H mean for AEs, CEs and all eddies at the largest sample amount are 0.563, 0.560 and 0.547 respectively. But Banerjee and Dave (2004) suggested that the H value of a clustered data set should be greater than 0.7. In our opinion, high H values indicate obvious aggregated characteristics, while these lower H values exceeding 0.5 truthfully reflect the weak tendency of eddies' distribution and difficulty for clustering. So, the clustering analysis should be carried out very carefully and objectively to derive a reliable result from plenty of experiments.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



To definitely prove that eddies' distribution is not random, we applied the t -test to H values calculated at the largest sample amount. The results are showed in Table 1. The values in the last column show that no matter AEs or CEs or the whole, the assumption that the mean H equals to that of random is denied significantly at the 0.05 level. So, we conclude that eddies' distribution in the SCS shows a weak clustering tendency.

3.2 Correct number of clusters

To determine the correct number of clusters, we apply K -means by specifying K from 1 to 20 and replicate 1000 times on every K . Figure 3 shows the plots of the median of SSE and silhouette coefficient against variable K .

As mentioned in Sect. 2.2, the knees of the SSE curve or the peaks of the SC curve hint the possible right cluster number. But as we can see in Fig. 3, it is not easy to decide which is the optimal value for there are more than one local peaks and curve knees. For example, 4, 5, 6 and 7 can all be viewed as the knees of the SSE curve of all eddies (Fig. 3a) while 2, 4, 7 and 9 are reasonable peaks of the SC curve of all eddies (Fig. 3d). It should be known that the spatial pattern of eddy generation may contain a hierarchical clustering structure, and these multiple optima are direct indicators. If the partition goes too coarse, valuable spatial variations will be obscured. If it goes too fine spatial variations would be wrongly tore apart and cause difficulties for interpretation. A practical solution to finding the correct number, not too coarse and not too fine, is to take into consideration both the knees of SSE curve and peaks of SC curve. The lower the SSE, the better the partition. But we do not go to extremes since the curve of SSE is approximating zero axis and will equal 0 when cluster number equals the number of data points. The local peaks of SC curve is the higher the better. So we select as correct the number of clusters for which the SSE is small and the SC is a local maximum and as large as possible.

Based on this criterion, the correct cluster number of all eddies is four. $K = 2$ is discarded for the SSE is the largest. The SC value of $K = 4$ is the second maximum and the SSE value is relatively small. The correct number of AEs and CEs are also

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



four for the same reason. It is interesting that the cluster number equals the number of geographic zones in Wang et al.'s (2003). Comparisons of them are detailed in Sect. 3.3.1.

Besides, alternative choices are carefully made to gain finer clustering results which may unveil more details of the spatial variations. For all eddies, $K = 7$ is the better choice. It is the third maximum of SC curve and the SSE value is low. Although $K = 9$ is the fourth local maximum of SC curve, it is discarded for the separation might be too fine. Similarly, $K = 10$ is retained for AEs and $K = 16$ is discarded, $K = 6$ is retained for CEs and $K = 10$ is discarded.

3.3 Clustering results

As a matter of fact, K -means can only guarantee a local minimum by optimizing the objective function to produce relative better and acceptable clustering results. Global optimization achieved by exhaustive searches of all possible choices is unrealistic in applications. So, in this study, the clustering procedure is executed 100 times with K random initial centroids each time. And the objective function we adopted is defined as follows:

$$O(i) = SSE'(i) - SSB'(i) - SC'(i) \quad (1)$$

where i denotes the i th time execution, and SSE' , SSB' and SC' denote the Gaussian normalized value of SSE, SSB and SC respectively. The result that minimizes the objective function is selected for the following analysis and discussion.

3.3.1 Four-cluster result

Here, the four-cluster result for all eddies (Fig. 4) derived from K -means clustering is compared with the four geographic partition of SCS (Fig. 5). In Fig. 4, the four clusters are labeled by E1 (west of Luzon Strait), E2 (around the Xisha Islands), E3 (southwest of Luzon Island), and E4 (southern SCS). E3 and E4 are much overlapped with Z3 and

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A2, and A3 together represent the Northern SCS, A4, A5, and A6 represent the Central SCS, and A7, A8, A9, and A10 represent the Southern SCS.

First, in Northern SCS, A1 represents those AEs generated near the Luzon Strait. This group is situated in the most complex area of the SCS. The Luzon Strait is the deepest channel for water exchange between the SCS and Pacific Ocean. The Kuroshio frequently intrudes into the SCS and influences its circulation, producing AEs southeast of the Dongsha Islands (Li et al., 1998) and southwest of Taiwan Island (Wang and Chen, 1987). Part of the eddies generated northwest of Luzon Island is still improperly included in A1 due to the weakness of *K*-means in handling data of different density. Concentrations there are supported by (Cheng et al., 2005; Wang et al., 2003). The Ekman transport driven by summer wind stresses is a possible cause for these AEs.

A2 represents those AEs located southwest of the Dongsha Islands. This distribution characteristic is shown in many statistical studies (Wang et al., 2003; Cheng et al., 2005; Lin et al., 2007; Hwang and Chen, 2000). Model results showed that eddies may be caused by the wind stress curl or the Kuroshio intrusions. (Cai et al., 2002a, b)

As for A3, Chai et al. (2001) reported three anticyclonic eddies to the east of Hainan Island and frontal instability and topography influences were the causes. Other studies observed eddies moving around. Zhong et al. (1990) found many AEs in this area based on hydrological survey near the Northern continental shelf of SCS from 1975–1984. Li et al. (2011) also identified AEs in this area from Lagrangian drifter data from 1979 to present.

Second, in the Central SCS, A4 represents the dense cluster southwest of Luzon Island, between 14° N and 16° N. Wang et al. (2008) noted the orographic wind jets associated with the winter monsoon wind and the mountainous topography along the eastern boundary of SCS can spin up AEs and CEs. A5 represents AEs generated in the SCS center. However, while these AEs were also confirmed in Wang et al.'s (2003) statistical study, generation mechanisms about them were few. A6 represents the concentration off the Vietnam coast. By principal component analysis of SLA and wind

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

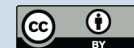
Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

stress curl, Shaw et al. (1999) found that the generation of this dipole structure was related to an eastward jet following the line of zero curl. Wang (2004, 2006) suggested vorticity transports from the nonlinear effect of the western boundary currents were crucial for the generation of the dipole structure.

5 Third, in the Southern SCS, four clusters are produced. A7 well captures the AEs along the Nansha Trough. A8 groups the eddies generated around the Nansha Islands. A9 and A10 are so adjacent off the Vietnam coast south of 12° N that they may be viewed together representing the dense population in southwest of SCS. In previous studies, Cai et al. (2002a) suggested that interaction between strong barotropic shelf currents and the local topography can generate anticyclonic eddies over the deep
10 trough (A7) in winter using a coupled single-layer and two-layer hybrid model. Fang et al. (1997) pointed out that seasonal monsoon was a important driving force for the southern gyres and eddies. Guan and Yuan (2006) analyzed eddies generated in southwest of SCS (A9 and A10) using SSHA from TOPEX/ERS and found the seasonal variations were strong, especially between July and December.

Figure 7 illustrates the six clusters of CEs. C1 represents CEs generated near the Luzon Strait. Compared with A1, CE locations are so scattered that no distinguishable patterns are recognized between northwest of Luzon Island and southwest of Taiwan Island. Many studies show that CEs northwest of Luzon Island have strong seasonal variations (Fang et al., 1998; Nitani, 1970; Qu et al., 2000), which explains the dispersed distribution of CEs. Qu et al. (2000) suggested the intense positive wind stress curl offshore northwest of Luzon Island was likely mechanism. But Liu and Su (1992) found the positive vorticity advected westward from the Kuroshio front may produce cyclonic eddies as well.

25 C2 represents CEs generated southwest of Dongsha Islands, where CEs were frequently identified by previous studies (Su et al., 1999) and the mechanism is basically the same with A2. CEs in the C3 are mainly generated north of 14° N, and form the northern part of the dipole structure off the Vietnam coast. CEs in C4 concentrate between 12° N and 14° N to the southwest of Luzon Island, which also has been reported

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



in Wang et al. (2003) and Cheng et al. (2005). The mechanism is the same with A4 and in-depth discussion can be referred to (Wang et al., 2008). C5 and C6 represent CEs formed in the cyclonic gyres of the Southern SCS, south of 12° N. Corresponding mechanism studies can be referred to (Guan and Yuan, 2006; Fang et al., 1997).

3.3.3 Comparisons between AEs and CEs

Figure 8 combines the centroids of AEs and CEs to show how their spatial patterns differ. The spatial dimension of each cluster is represented by a circle of which the radius is defined by the mean distance from these eddies' generation locations to the centroid.

First, more clusters and smaller spatial domains of AEs indicate patterns of AE are relatively more aggregated than those of CE. Second, differences between them can be summarized as follows: (1) To the southwest of Luzon Island, AEs and CEs are basically generated apart; AEs are likely to the north of 14° N and closer to the shore, whereas CEs are to the south and farther from shore. (2) The Central SCS and the Nansha Trough are dominated by AEs. (3) Clusters A3, C3 and A6 are located sequentially apart, along 112° E, and C3–A6 represent the dipole eddies off Vietnam. (4) Southwest of Dongsha Islands, AEs are concentrated to the east of CEs.

Beside these distinctions, AEs and CEs are overlapped in Northeastern and Southern SCS. Do these overlaps simply explain the similarities there, or are there any hidden differences behind the overlaps? To answer these questions, we reexamined the eddies with regard to seasonal monsoons. The winter monsoon is dominated by northeasterly wind while the summer monsoon is dominated by southwesterly wind (Hellerman and Rosenstein, 1983). So we remapped these eddies in northeastern (Fig. 9) and Southern SCS (Fig. 10) according to the two monsoon seasons.

In Northeastern SCS, eddies generated during different monsoon seasons show similar distribution patterns. Specifically, the center of Northeastern SCS is dominated by AEs which are surrounded by CEs generated around Xisha, northwest of Luzon and southwest of Taiwan Island. The similarities indicate that those anticyclonic rings

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

shedding from the loop currents caused by the Kuroshio intrusions may occur during any season of the year, which are consistent with the study of Yuan et al. (2006), and that the CEs nearby could be induced by the instability of the horizontal shear stress of velocity around the loop currents. On the other hand, during northeasterly monsoon some CEs exist in the center area of Northeastern SCS where few CEs are identified during southwesterly monsoon. Such difference could be due to different patterns of the Kuroshio intrusions. Yuan et al. (2006) pointed out that the anticyclonic intrusions of the Kuroshio are more frequent in summer than in winter, but in winter the intrusion path would turn cyclonic toward southwest at the center water of Northeastern SCS. As a result, the turns may probably be recognized as CEs which weakened the dominance of AEs in this area in winter. In summary, the spatial patterns of AEs and CEs in Northeastern SCS are mainly controlled by the Kuroshio intrusions. The impact seasonal variations make on the pattern is not strong. As for the overlap of the centroids, it is precisely caused by the concentric distribution of AEs and CEs.

In the Southern SCS, AEs are more in southwesterly monsoon while CEs are more in northeasterly monsoon. The spatial patterns are quite different in the two monsoon seasons. During southwesterly monsoon, AEs are concentrated to the southeast of Vietnam between 110° E and 112° E, while CEs are mainly located nearby between 112° E and 114° E. During northeasterly monsoon, CEs almost control the whole area, AEs are scattered. Distributions of AEs and CEs are not entirely reversed with the monsoon change. Hence, overlaps in Southern SCS cannot be explained only by seasonal variations; complex circulations and topography there can be other important causes as well.

4 Conclusions

In this paper, we carried out a clustering analysis using *K*-means approach on 371 eddies to investigate the spatial variations among their birth locations. We carefully examined the clustering tendency of eddies' distribution by Hopkins statistic. Results

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Finally, the importance of spatial variations in surveying eddies motivates us to investigate patterns as well as distinctions of eddies' distribution. Different generation mechanisms are expected to be reflected by proper geographic partitions, which will also facilitate interpreting eddies' dynamic characteristics in statistical studies. But, as we know, eddy is a dynamic phenomenon with an evolution process. We've only made a necessary effort toward in-depth analysis of eddy evolution. Further studies are required for better understanding the dynamic motions and mechanisms of mesoscale eddies in the SCS.

Appendix A

Eddy identification procedure

The identification procedure applied in this paper is an adaption of that used in Chelton et al. (2011). Eddy's boundary is delineated by contours of SSH and the center is represented by local extreme. First, to define the boundary, contours are constructed at 1-cm interval for instantaneous SSH field. 1-cm interval yields a good compromise between minimum resolvable eddy amplitude and well defined and compact eddy interiors (Chelton et al., 2011). Then, those closed contours that satisfy the following criteria are kept for next step:

(1) The shape is similar to circle or ellipse. A shape parameter, SI , is introduced to measure the similarity. It is defined as follows:

$$SI = \frac{S}{\pi \cdot \left(\frac{D}{2}\right)^2}, \quad (\text{A1})$$

where S denotes the area of the contour, D denotes the maximum distance between any pairs of points on the contour. Contours whose $SI > 0.4$ are retained.

(2) Diameter of the eddy, which is defined by the diameter of a circle with equivalent area, is limited from 45 to 500 km.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



(3) The distance between any pair of points along the contour must be less than a specified maximum. The maximum-distance threshold is 400 km for latitudes above 25°, and is increased linearly to 1200 km at the equator (Chelton et al., 2011).

Next, among these selected contours, spatial containment relationship is calculated by program interfaces provided by the commercial ArcGIS Engine Software Develop Kit (http://www.esri.com/software/arcgis/arcgisengine). Those contours that are not contained by any others are defined as the boundary of eddy, and the local extremes within them are identified as centers. Eddies with amplitude less than 2 cm and surviving no more than 2 weeks are excluded from analysis. The amplitude is defined as the absolute difference between SSH value of the center and the boundary.

Appendix B

Clustering tendency

The term “clustering tendency” refers to the problem of deciding whether data exhibit a predisposition to cluster into natural groups, without identifying these groups (Jain and Dubes, 1988). It is recommended that random and regularly spaced patterns not be submitted to clustering algorithms (Jain and Dubes, 1988). Since K -means and other clustering algorithms dutifully cluster data into groups no matter whether they are naturally clustered or purely random, it is essential to assess clustering tendency to prevent meaningless results. Hopkins statistic, which is simple and suitable for two-dimensional data, is used for testing whether eddies' distribution shows an aggregative tendency, or they are just randomly distributed.

The Hopkins statistic H is defined as follows:

$$H = \frac{\sum_{i=1}^p u_i}{\sum_{i=1}^p u_i + \sum_{i=1}^p w_i} \quad (\text{B1})$$

There are p points randomly generated in the data domain, and p sample points randomly selected from the original data. u_i is the nearest neighbor distance of the i th randomly generated point in the original data set, and w_i is the nearest neighbor distance of the i th sample point in the original data set. The H varies from 0 to 1, and it tends to approximate 0.5 when the sums of u_i and w_i are nearly identical, indicating that the data are similar to random points. Values near 1 and 0 indicate obvious clustering and regularly distributed patterns, respectively.

Appendix C

Goodness of clustering

Cohesion and separation are two measures that evaluate how well the clustering fits the data. In Euclidean space, the SSE can be used to measure the cohesion of both individual clusters (individual SSE) and overall clustering (total SSE). It is defined as follows:

$$\text{SSE} = \sum_{i=1}^K \sum_{x \in C_i} \text{dist}(c_i, x)^2 \quad (\text{C1})$$

where dist is the standard Euclidean distance, K is the cluster number, and C_i denotes the i th cluster and c_i the centroid of C_i .

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



SSB, the sum of the squared distances of a cluster centroid to the overall centroid of all data points, is a traditional measure of separation. The relationship between SSE and SSB is one-falling-and-the-other-rising (Tan et al., 2005). The SSB is defined as follows:

$$5 \quad \text{SSB} = \sum_{i=1}^K m_i \text{dist}(c_i, c)^2 \quad (\text{C2})$$

where c denotes the centroid of all data points, and m_i denotes the number of objects in the i th cluster. In evaluation, the lower SSE is better while the higher SSB is better (Tan et al., 2005).

10 Another popular method that can be used to measure goodness of clustering is called the silhouette coefficient, which combines cohesion and separation in evaluating cluster validity. The silhouette coefficient is defined as follows:

$$15 \quad s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (\text{C3})$$

where $a(i)$ denotes the average distance from the i th point to all other points within the same cluster; the average distance from the i th point to all points in any cluster not containing the points is calculated, and $b(i)$ is the minimum value with respect to all clusters.

As defined, the silhouette coefficient ranges from -1 to 1 . A well-clustered point usually has a small $a(i)$ but a relatively large $b(i)$, and the coefficient is close to 1 . On the contrary, a negative value is generated only when $a(i) > b(i)$, indicating that the point is probably wrongly clustered. In contrast to SSE and SSB, the silhouette coefficient is calculated at every data point. Thus, to measure the goodness of individual or overall clustering, we typically use the silhouette mean of the cluster, or of the whole.

25 *Acknowledgement.* This is a contribution to Research Project 41071250 of the National Science Foundation of China and 088RA500KA of the Innovation Projects of the State Key Laboratory of Resource and Environment Information System, Chinese Academy of Sciences. We

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



thank the US Naval Research Laboratory for providing the SSH dataset and appreciate the comments from reviews and readers.

References

- Banerjee, A. and Dave, R. N.: Validating clusters using the hopkins statistic, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2004), Budapest (Hungary), 2004, 149–153, 2004.
- Brown, M. B. and Forsythe, A. B.: Robust tests for the equality of variances, *J. Am. Stat. Assoc.*, 69, 364–367, doi:10.1080/01621459.1974.10482955, 1974.
- Cai, S., Su, J., Gan, Z., and Liu, Q.: The numerical study of the South China Sea upper circulation characteristics and its dynamic mechanism, in winter, *Cont. Shelf Res.*, 22, 2247–2264, doi:10.1016/s0278-4343(02)00073-0, 2002a.
- Cai, S., Su, J., Gan, Z., and Liu, Q.: The numerical study of the dynamic mechanism of the South China Sea upper circulation in summer, *Acta Oceanol. Sin.*, 24, 1–7, 2002b (in Chinese with English abstract).
- Chai, F., Xue, H., and Shi, M.: Hydrographic characteristics and seasonal variation of three anticyclonic eddies on the northern continental shelf of the South China Sea, in: *Oceanography in China*, 13, edited by: Xue, H. et al., Ocean Press, Beijing, 24, 105–116, 2001 (in Chinese with English abstract).
- Chelton, D. B., Schlax, M. G., and Samelson, R. M.: Global observations of nonlinear mesoscale eddies, *Prog. Oceanogr.*, 91, 167–216, doi:10.1016/j.pocean.2011.01.002, 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, doi:10.1029/2010jc006716, 2011.
- Cheng, X., Qi, Y. Q., and Wang, W. Q.: Seasonal and interannual variabilities of mesoscale eddies in South China Sea, *J. Trop. Oceanogr.*, 24, 51–59, 2005.
- Du, Y., Fan, X., He, Z., Su, F., Zhou, C., Mao, H., and Wang, D.: Extraction of spatial-temporal rules from mesoscale eddies in the South China Sea based on rough set theory, *Ocean Sci.*, 7, 835–849, doi:10.5194/os-7-835-2011, 2011.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Fang, G., Fang, W., Fang, Y., and Wang, K.: A survey of studies on the South China Sea upper circulation, *Acta Oceanogr. Taiwan.*, 37, 1–16, 1998.
- Fang, G., Chen, H., Wei, Z., Wang, Y., Wang, X., and Li, C.: Trends and interannual variability of the South China Sea surface winds, surface height, and surface temperature in the recent decade, *J. Geophys. Res.*, 111, C11S16, doi:10.1029/2005jc003276, 2006.
- Fang, W., Guo, X., and Huang, Y. : Observation and study on the circulation in the Southern South China Sea, *Chinese Sci. Bull.*, 42, 2261–2271, 1997.
- Fang, W., Fang, G., Shi, P., Huang, Q., and Xie, Q.: Seasonal structures of upper layer circulation in the Southern South China Sea from in situ observations, *J. Geophys. Res.*, 107, 3202, doi:10.1029/2002jc001343, 2002.
- Guan, B.: Warm eddy in the open sea east of Hainan Island, *J. Oceanogr. Huanghai Bohai Seas*, 15, 1–7, 1997 (in Chinese with English abstract).
- Guan, B. and Yuan, Y. C.: Overview of studies on some eddies in the China Seas and their adjacent seas, I. The South China Sea and the region east of Taiwan, *Acta Oceanol. Sin.*, 28, 1–16, 2006.
- Hellerman, S. and Rosenstein, M.: Normal monthly wind stress over the world ocean with error estimates, *J. Phy. Oceanogr.*, 13, 1093–1104, doi:10.1175/1520-0485(1983)013<1093:nmwsot>2.0.co;2, 1983.
- Huang, Q., Wang, W. Z., Li, Y. X., Li, Z. W., and Mao, M.: General situations of the current and eddy in the South China Sea, *Adv. Earth Sci.*, 7, 1–9, 1992.
- Hwang, C. and Chen, S.-A.: Circulations and eddies over the South China Sea derived from TOPEX/Poseidon altimetry, *J. Geophys. Res.*, 105, 23943–23965, doi:10.1029/2000jc900092, 2000.
- Jain, A. K. and Dubes, R. C.: *Algorithms for Clustering Data*, Prentice-Hall, Inc., 320 pp., 1988.
- Kaufman, L. and Rousseeuw, P. J.: *Finding Groups in Data: an Introduction to Cluster Analysis*, John Wiley, 1990.
- Li, J. X., Zhang, R., and Jin, B.: Eddy characteristics in the Northern South China Sea as inferred from Lagrangian drifter data, *Ocean Sci. Discuss.*, 8, 1575–1599, doi:10.5194/osd-8-1575-2011, 2011.
- Li, L., Nowlin Jr., W. D., and Jilan, S.: Anticyclonic rings from the Kuroshio in the South China Sea, *Deep-Sea Res. Pt. I*, 45, 1469–1482, doi:10.1016/s0967-0637(98)00026-0, 1998.
- Li, Y., Cai, W., Li, L., and Xu, D.: Seasonal and interannual variabilities of mesoscale eddies in Northeastern South China Sea, *J. Trop. Oceanogr.*, 22, 61–70, 2003.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Lin, P., Wang, F., Chen, Y., and Tang, X.: Temporal and spatial variation characteristics on eddies in the South China Sea, I. Statistical analyses, *Acta Oceanol. Sin.*, 29, 14–22, 2007.
- Liu, X. and Su, J.: A reduced gravity model of the circulation in the South China Sea, *Oceanologia Et Limnologia Sinica*, 2, 167–174, 1992.
- 5 Macqueen, J.: Some methods for classification and analysis of multivariate observations, in: *Proc. 5th Berkeley Symp. Mathematical Statist. Probability*, 281–297, 1967.
- Nitani, H.: Oceanographic conditions in the sea east of Philippines and Luzon Strait Summer of 1965 and 1966, *The Kuroshio – a Symposium on Japan Current*, Honolulu, Hawaii, 213–232, 1970.
- 10 Qu, T.: Upper-layer circulation in the South China Sea, *J. Phys. Oceanogr.*, 30, 1450–1460, doi:10.1175/1520-0485(2000)030<1450:ulcits>2.0.co;2, 2000.
- Ripley, B. D.: Modelling spatial patterns, *J. Roy. Stat. Soc. B Met.*, 39, 172–212, 1977.
- Shaw, P.-T., Chao, S.-Y., and Fu, L.-L.: Sea surface height variations in the South China Sea from satellite altimetry, *Oceanol. Acta*, 22, 1–17, doi:10.1016/s0399-1784(99)80028-0, 1999.
- 15 Su, J.: Overview of the South China Sea circulation and its influence on the coastal physical oceanography outside the Pearl River Estuary, *Cont. Shelf Res.*, 24, 1745–1760, doi:10.1016/j.csr.2004.06.005, 2004.
- Su, J., Xu, J., and Cai, S.: Gyres and eddies in the South China Sea, in: *Onset and Evolution of the South China Sea Monsoon and its Interaction With the Ocean*, edited by: Ding, Y., and Li, C., China meteorological Press, Beijing, 272–279, 1999.
- 20 Tan, P. N., Steinbach, M., and Kumar, V.: *Introduction to Data Mining*, Pearson Addison Wesley, 2006.
- Wang, G.: Discussions on the movement of mesoscale eddies in the South China Sea, Doctor, Ocean University of China, Qingdao, 2004.
- 25 Wang, G., Su, J., and Chu, P. C.: Mesoscale eddies in the South China Sea observed with altimeter data, *Geophys. Res. Lett.*, 30, 2121, doi:10.1029/2003gl018532, 2003.
- 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, doi:10.1029/2005jc003314, 2006.
- 30 Wang, G., Chen, D., and Su, J.: Winter Eddy Genesis in the Eastern South China Sea Due to Orographic Wind Jets, 3, *American Meteorological Society*, Boston, 7 pp., 2008.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[⏪](#)
[⏩](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)


- Wang, L., Koblinsky, C. J., and Howden, S.: Mesoscale variability in the South China Sea from the TOPEX/Poseidon altimetry data, *Deep-Sea Res. Pt. I*, 47, 681–708, doi:10.1016/s0967-0637(99)00068-0, 2000.
- 5 Wang, X., Du, Y., Fan, X., and Yi, J.: Research on automatic identification method of ocean mesoscale eddies, *J. Trop. Oceanogr.*, in review, 2011.
- Wang, Z. and Chen, Q. : The warm eddy in Northern South China Sea (1). *J. Oceanogr. Taiwan Uni.*, 18, 92–103, 1987.
- Wu, C.-R. and Chiang, T.-L.: Mesoscale eddies in the Northern South China Sea, *Deep Sea Res. Pt. II*, 54, 1575–1588, doi:10.1016/j.dsr2.2007.05.008, 2007.
- 10 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, doi:10.1029/2009jc005657, 2010.
- Yang, K., Shi, P., Wang, D., You, X., and Li, R.: Numerical study about the mesoscale multi-eddy system in the Northern South China Sea in winter, *Acta Oceanol. Sin.*, 22, 27–34, 2000 (in Chinese with English abstract).
- 15 Yuan, D. and Li, R.: Dynamics of eddy-induced Kuroshio variability in Luzon Strait, *J. Trop. Oceanogr.*, 27, 1–9, 2008.
- Yuan, D., Han, W., and Hu, D.: Surface Kuroshio path in the Luzon Strait area derived from satellite remote sensing data, *J. Geophys. Res.*, 111, C11007, doi:10.1029/2005jc003412, 2006.
- 20 Yuan, D., Han, W., and Hu, D.: Anti-cyclonic eddies northwest of Luzon in summer-fall observed by satellite altimeters, *Geophys. Res. Lett.*, 34, L13610, doi:10.1029/2007gl029401, 2007.
- Zhong, H.: Structures of the density circulation, in: Report of Decadal Hydrographic Series Survey of the Shelf and Adjacent Waters of the Northern South China Sea, edited by: Ma, Y. et al., Ocean Press, Beijing, 215–239, 1990 (in Chinese).

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 1. *T*-test analysis.

	p	Descriptive statistics			Sig. of Levene's ^b for Normal Distribution	Sig. of Levene's Test ^b for Equality of Variance	Sig. of <i>T</i> -test for Equality of Means
		<i>N</i>	Mean	Standard deviation			
AEs	16	100	0.563	0.055	0.061	0.046	0.000 ^a
Random		100	0.491	0.047			
CEs	21	100	0.560	0.047	0.500	0.044	0.000 ^a
Random		100	0.488	0.038			
All	37	100	0.547	0.031	0.500	0.331	0.000
Random		100	0.508	0.034			

^a Denotes the *t*-test result is computed under the assumption of equal variance.

^b Levene's test is an inferential statistic used to assess the equality of variances (Brown and Forsythe, 1974).

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

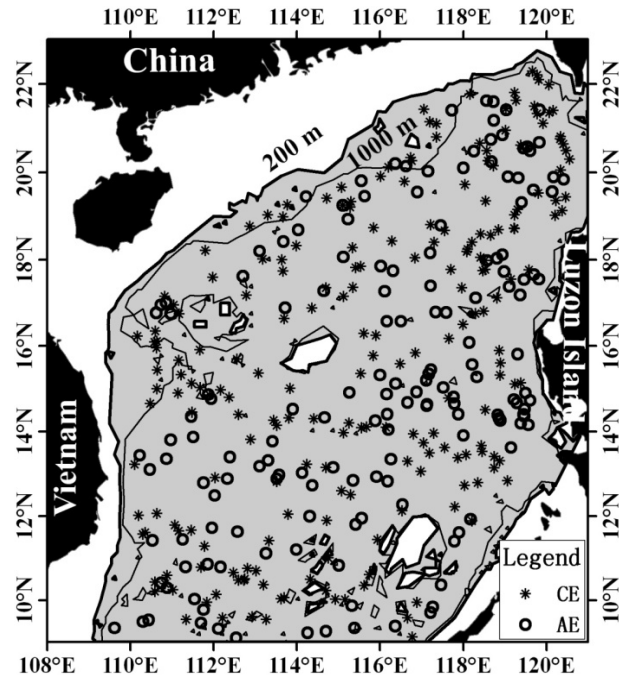


Fig. 1. Map of study area. Birth locations of cyclonic and anticyclonic eddies are symbolized by the asterisk and circle, respectively. Black solid lines indicate 200-m (bold) and 1000-m (thin) isobaths.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

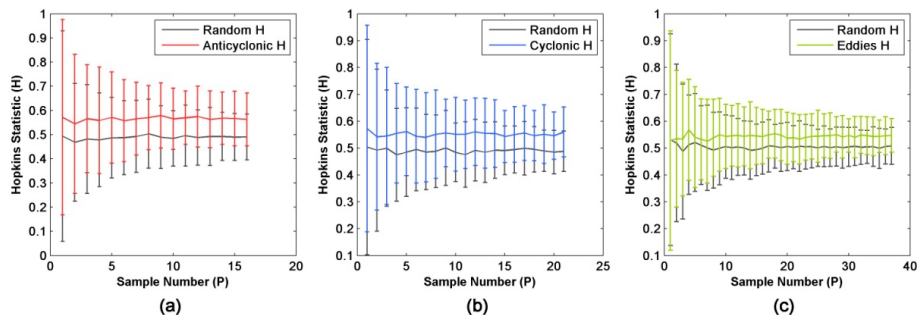


Fig. 2. Error bar plot of Hopkins statistic. Error bar is twice the standard deviation, and curve is mean of H. **(a)** Hopkins statistic of AE and random sample. **(b)** Hopkins statistics of CE and random sample. **(c)** Hopkins statistics of all eddies and random sample.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

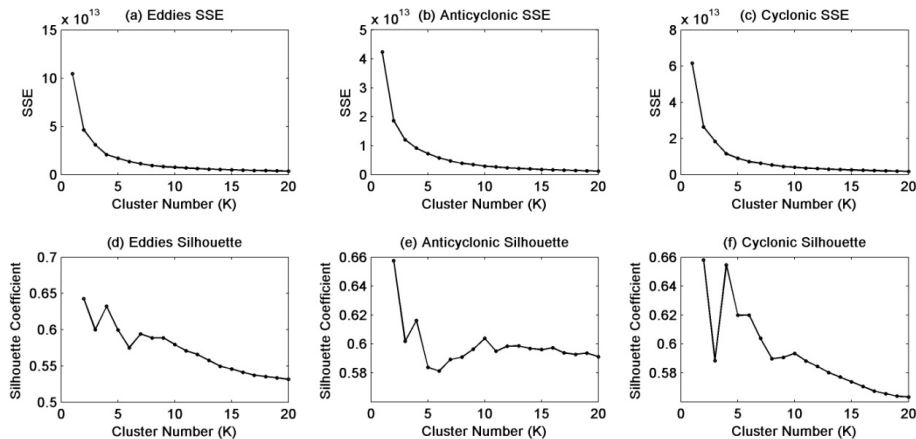
Full Screen / Esc

Printer-friendly Version

Interactive Discussion

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

**Fig. 3.** Plots of SSE and silhouette coefficient.[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

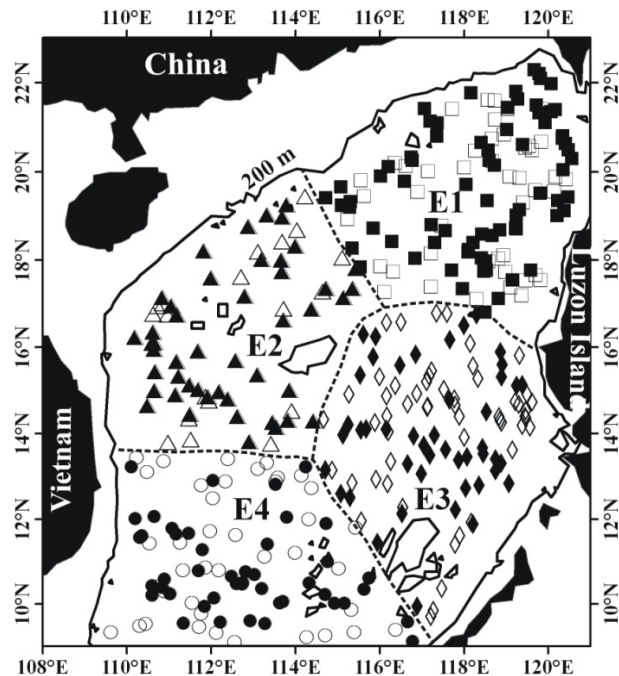


Fig. 4. Four-cluster results. The dashed lines schematically depict the boundary of the four regions. Symbols of anticyclonic eddies are unfilled while those of cyclonic eddies are filled with black. Different groups are distinguished by different geometry shapes.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

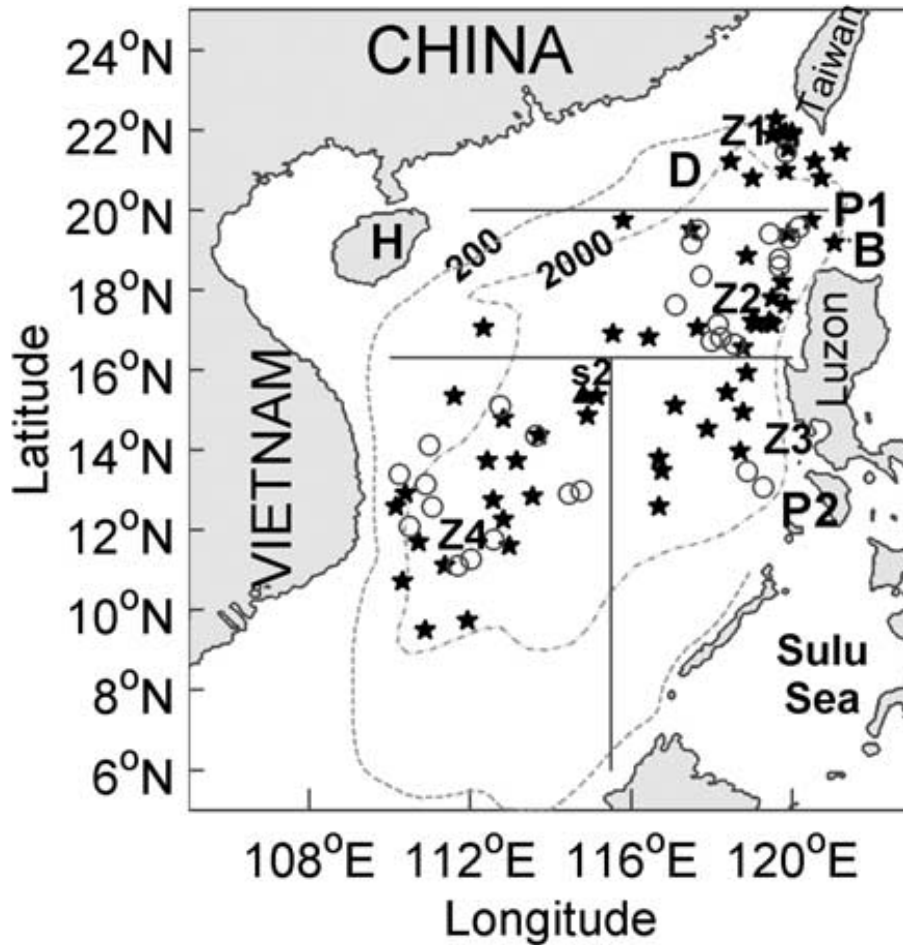


Fig. 5. Four geographic zones of SCS, adopted from Wang et al. (2003) Circles denote cyclonic eddies, and stars anticyclonic eddies.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

⏪ ⏩

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

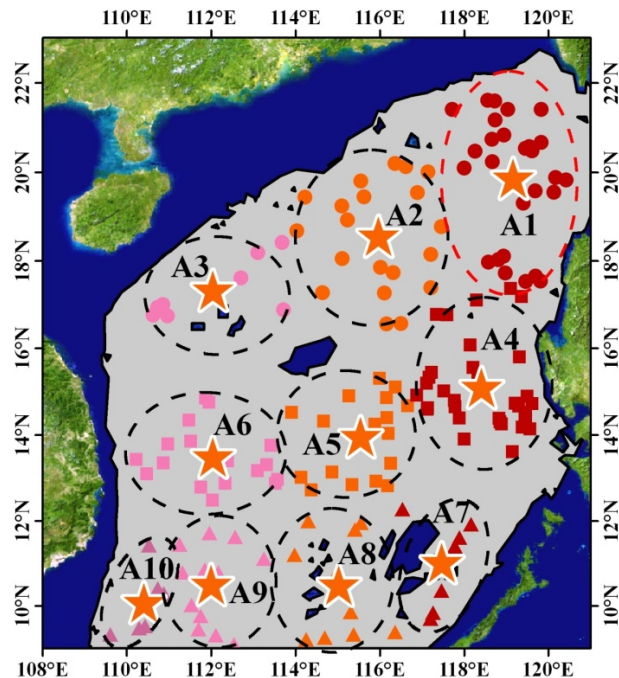


Fig. 6. Ten-cluster result for anticyclonic eddies. Orange stars are cluster centroids. Black dashed lines outline approximate area of clusters. Different groups are distinguished by different colors or symbols.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[⏴](#)[⏵](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

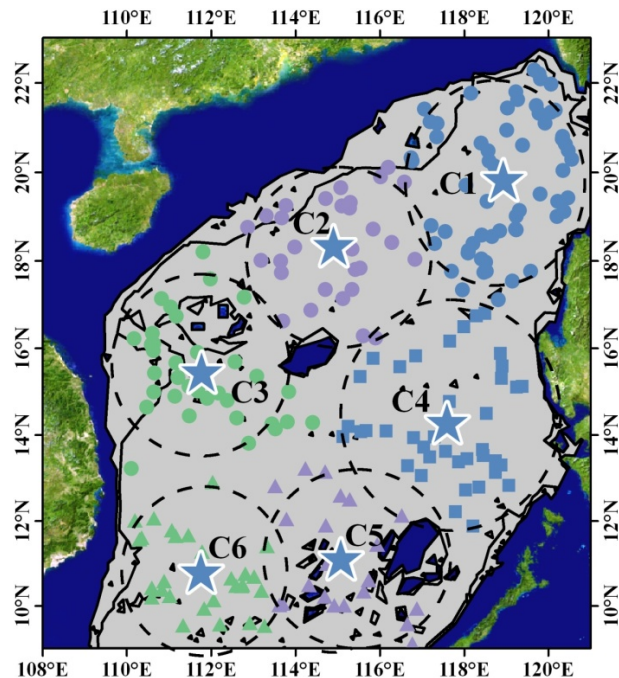


Fig. 7. Six-cluster result for cyclonic eddies. Blue stars are cluster centroids. Black dashed lines show approximate area of clusters. Different groups are distinguished by different colors or symbols.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

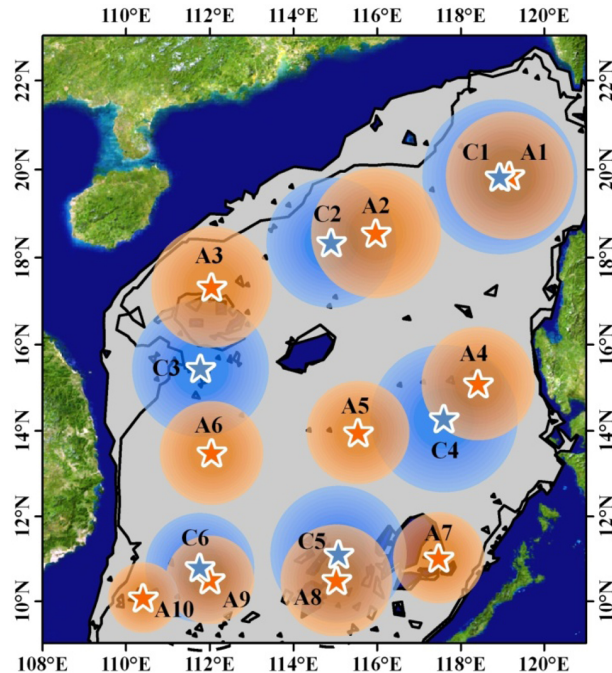


Fig. 8. Centroids and spatial domains of cyclonic and anticyclonic eddies. Centroids of AEs and CEs are shown by stars of different colors. The circles in different colors represent the spatial dimension of each cluster, and the radius is defined by the mean distance from these eddies' generation locations to corresponding centroid.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

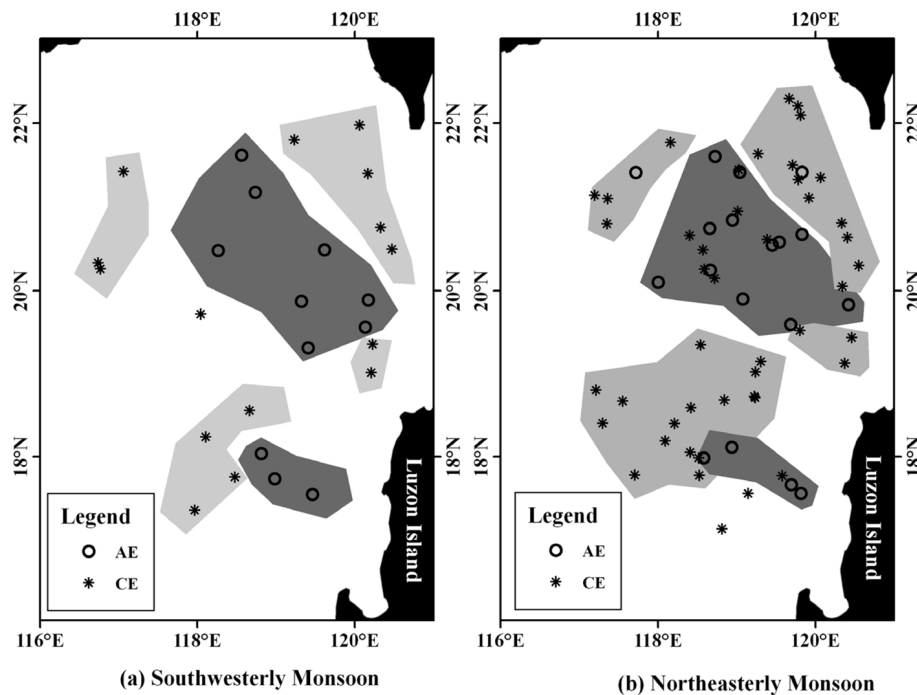


Fig. 9. Eddies of A1 and C1 generated in Northeastern SCS during different monsoon seasons. Concentration regions of AEs and CEs are generally represented by different colors.

A clustering analysis of eddies' spatial distribution in the South China Sea

J. Yi et al.

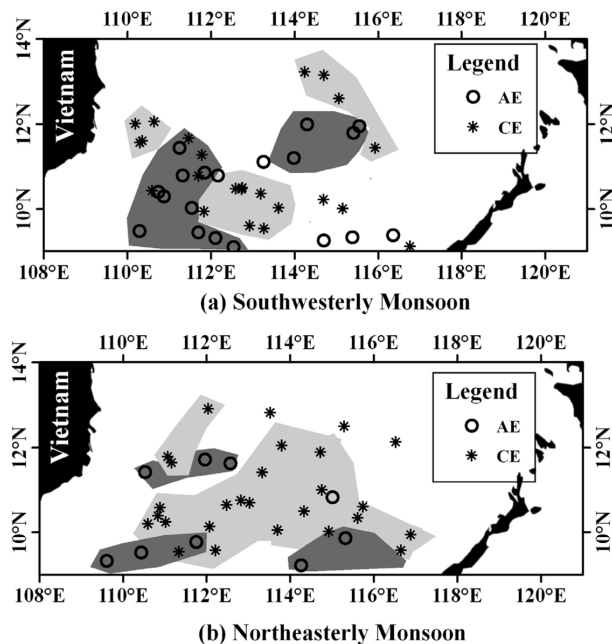


Fig. 10. Eddies of A8, A9, A10, C5 and C6 generated in Southern SCS during different monsoon seasons. Concentration regions of AEs and CEs are generally represented by different colors.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion