Acta Oceanol. Sin., 2019, Vol. 38, No. 5, P. 134–141
https://doi.org/10.1007/s13131-019-1447-x
http://www.hyxb.org.cn
E-mail: hyxbe@263.net

A new automatic oceanic mesoscale eddy detection method using satellite altimeter data based on density clustering

Jitao Li[1], Yongquan Liang[1], Jie Zhang[2], Jungang Yang[2]*, Pingjian Song[2], Wei Cui[2, 3]

[1] College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China

[2] First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China

[3] Physical Oceanography Lab, Qingdao Collaborative Innovation Center of Marine Science and Technology, Ocean University of China, Qingdao 266100, China

Received 6 May 2018; accepted 9 July 2018

Abstract

The mesoscale eddy is a typical mesoscale oceanic phenomenon that transfers ocean energy. The detection and extraction of mesoscale eddies is an important aspect of physical oceanography, and automatic mesoscale eddy detection algorithms are the most fundamental tools for detecting and analyzing mesoscale eddies. The main data used in mesoscale eddy detection are sea level anomaly (SLA) data merged by multi-satellite altimeters' data. These data objectively describe the state of the sea surface height. The mesoscale eddy can be represented by a local equivalent region surrounded by an SLA closed contour, and the detection process requires the extraction of a stable closed contour structure from SLA maps. In consideration of the characteristics of mesoscale eddy detection based on SLA data, this paper proposes a new automatic mesoscale eddy detection algorithm based on clustering. The mesoscale eddy structure can be extracted by separating and filtering SLA data sets to separate a mesoscale eddy region and non-eddy region and then establishing relationships among eddy regions and mapping them on SLA maps. This paper overcomes the problem of the sensitivity of parameter setting that affects the traditional detection algorithm and does not require a sensitivity test. The proposed algorithm is thus more adaptable. An eddy discrimination mechanism is added to the algorithm to ensure the stability of the detected eddy structure and to improve the detection accuracy. On this basis, the paper selects the Northwest Pacific Ocean and the South China Sea to carry out a mesoscale eddy detection experiment. Experimental results show that the proposed algorithm is more efficient than the traditional algorithm and the results of the algorithm remain stable. The proposed algorithm detects not only stable single-core eddies but also stable multi-core eddy structures.

Key words: mesoscale eddy, density clustering, shape discrimination, outermost closed contour

Citation: Li Jitao, Liang Yongquan, Zhang Jie, Yang Jungang, Song Pingjian, Cui Wei. 2019. A new automatic oceanic mesoscale eddy detection method using satellite altimeter data based on density clustering. Acta Oceanologica Sinica, 38(5): 134–141, doi: 10.1007/s13131-019-1447-x

1 Introduction

The mesoscale eddy is an important form of seawater transport having a long-term closed state and a spatial scale ranging from dozens to hundreds of kilometers and a time scale ranging from a few days to a few hundred days. The mesoscale eddy is responsible for sea power transfer and mass transfer and affects oceanic elements, such as oceanic temperature and salinity distributions. There are many factors of the formation of mesoscale eddies. In general, ocean flow is affected by many factors, such as wind, topography and ambient field flow, which forms mesoscale eddies. The study of mesoscale eddies is important to marine military activities and fisheries.

The study of mesoscale eddies in the ocean can be divided into the detection, observation and tracking of mesoscale eddies. Data of the global sea level can be obtained from satellite altimeter data, and sea level anomaly (SLA) results can then be used for the detection of mesoscale eddies. In the detection of mesoscale eddies, existing automatic detection methods based on an SLA map can be divided into four categories: methods using physical parameters, methods employing wavelet analysis, methods using streamline geometry features and methods using sea surface height anomaly characteristics.

The classic physical mesoscale eddy detection algorithm is the Okubo-Weiss (OW) algorithm (Okubo et al., 1970), which realizes the automatic detection of mesoscale eddies (Morrow et al., 2004). This method is widely used in the mesoscale eddy analysis of oceans, such as the Tasman Sea (Waugh et al., 2006), the Gulf of Alaska (Henson and Thomas, 2008) and the South China Sea (Nan et al., 2011). There are two deficiencies of the OW algorithm. Firstly, the parameter calculation may enlarge error in the SLA and results in a misjudgment. Secondly, for the boundary of mesoscale eddies, the algorithm cannot provide a precise

Foundation item: The National Key R&D Program of China under contract No. 2016YFC1401800; the National Natural Science Foundation of China under contract No. 41576176; the National Programme on Global Change and Air-Sea Interaction under contract Nos GASI-02-PAC-YGST$_2$-04, GASI-02-IND-YGST$_2$-04 and GASI-02-SCS-YGST$_2$-04.
*Corresponding author, E-mail: yangjg@fio.org.cn

closed contour but a scope.

A mesoscale eddy detection method based on wavelet analysis (Doglioli et al., 2007) regards vorticity as a semaphore radar signal. By processing this vorticity obtained from SLA data, the obtained signal curve is decomposed and mapped onto a vorticity space to extract mesoscale eddies. This method makes full use of the rapidity of the wavelet transform and the signal sensitivity and is able to handle massive data. However, the limit of the dimensions of data sets leads to the low identification accuracy of eddies and the detection ability is poor for mesoscale eddies with an irregular closed contour shape of the SLA value.

The winding angle (WA) method is a representative detection algorithm based on geometrical features (Ari Sadarjoen and Post, 2000) and has been improved and applied to the detection of mesoscale eddies (Chaigneau et al., 2008). The default structure of the WA algorithm is a single-core structure. The maximum range of the SLA closed contour is found as the boundary around an initial core point. This algorithm is simple and easy to understand and is the most popular algorithm for the detection of mesoscale eddies. However, the algorithm has deficiencies in that it is unable to find multi-core structures, has high algorithm complexity and requires a sensitivity test for the setting of parameters. Another method based on the geometric streamline feature is the vector geometry (VG) method (Nencioli et al., 2010), which uses the geometrical characteristics of the velocity vector to judge eddy cores. The velocity of the flow field is a minimum around an eddy core, and the tangential velocity of the flow field increases with the distance from the core. At the eddy border, the tangential velocity of the flow field begins to abate. The VG method uses the maximum peripheral speed at the boundary. However, the situation is complex when judging the mesoscale eddy boundary and eddy core, and two additional parameters are needed to limit the detection results while the parameter setting requires a sensitivity test.

The SSH-based method is an example of a method of automatic mesoscale eddy detection without a threshold value (Chelton et al., 2011) that aims to avoid the complicated problems caused by the threshold setting. The SLA extreme-value point is used as the eddy core, the opposite polarity eddy boundary is taken as the starting point, the closed contour line is searched for in steps of 1 cm from the outer boundary toward the eddy core, and the outermost SLA closed contour is found as the boundary of the mesoscale eddy. Compared with traditional algorithms, the SSH-based method can detect multi-core structures but does not judge the shape of the eddy, and there are many unstable multicore structures in the detection results. It is necessary to initialize the eddy area in the process of searching for the multicore structure. The sensitivity of the parameter setting is high and the adaptability is poor.

The hybrid detection (HD) algorithm is a hybrid algorithm combining traditional OW and SSH-based methods (Yi et al., 2014). The combined application of the OW algorithm and SSH-based method to initial cores reduces the effect of noise on the OW algorithm and avoids the sensitivity of threshold setting during initialization. This algorithm improves the accuracy of mesoscale eddy boundary detection by reducing the incremental step size and eliminating mesoscale eddies with irregular boundaries to guarantee the stability of the selected-scale eddies. The main problem that the HD algorithm solves is that it finds multicore structures that are more stable, but the parameter setting in the eddy selection process requires a sensitivity test and the algorithm steps are complicated. Another algorithm distinguishes color attributes of an SSH image to extract mesoscale eddies (Zhang et al., 2014). The algorithm separates a region having a defined pixel density from a region of low pixel density to detect the boundaries of mesoscale eddies. The same method has been used in handling SST images (Zhang et al., 2015). In the case of an SSH-image-based method, the algorithm has too many initialization parameters and the parameter setting requires a sensitivity test, resulting in low adaptability of the algorithm. A parallel method has been added to the SSH-based algorithm to improve the operation efficiency of the algorithm (Liu et al., 2016). The computational efficiency is greatly improved by dividing the global data set into several sub-regions.

The traditional mesoscale eddy detection algorithm has deficiencies, such as the sensitivity of the parameter setting, the problem of defining the initial threshold value for the eddy core, and the stability problem of the mesoscale eddy result. To solve the above problems, the present paper proposes the density-based clustering mesoscale eddy detection algorithm (DC). Density clustering algorithms have been used for data analysis in machine learning, and stable mesoscale eddies can be filtered by establishing potential relations among the data. The advantages of the proposed algorithm are that it has fewer parameters, the parameters have higher stability, it is easier to set parameters, and parameters are initialized mostly according to the attributes of the data set. For the same attribute data sets, the parameters are the same and the method has good adaptability. The proposed algorithm does not use the core points of the mesoscale eddies as the initialization condition and thus avoids the sensitivity problem of the initialization threshold of the initial core points. For the stability problem of finding mesoscale eddies, the algorithm guarantees the stability of the results by shape discrimination, and irregular mesoscale eddy structures in the process of convergence are filtered out.

2 Data and method

2.1 *Altimetry data*

The satellite altimeter product used in this paper is fusion data of SLA recorded by multiple satellites and distributed by AVISO (http://www.aviso.oceanobs.com). The SLA data are merged altimeter data recorded by TOPEX/Poseidon, Jason-1, Jason-2, ERS-1, ERS-2 and Envisat satellites with spatial resolution of 0.25° × 0.25°. Daily altimeter data are the integration of data from at least two satellites. As a result, the data of different satellites can be calibrated, and the precision and consistency of the fusion products can be ensured at the same time.

2.2 *Eddy detection methodology*

In general, a mesoscale eddy detection algorithm aims to find the potential internal relation within SLA data and to find the outermost closed contours through interpolation and thus identify eddies. In SLA data there are plentiful regional maximum and minimum values, which are regarded as the cores of eddies. There are several closed contours of SLA data, which may contain one or more extreme points with the same polarity. Such a closed region can be seen as a mesoscale eddy, and the data points contained in the closed region can be regarded as data points having the same attributes with potential relations.

In the field of machine learning, a clustering algorithm is used to discover potential relations within a data set. The present paper therefore includes a clustering algorithm in the detection process and transforms the algorithm into a process of discovering and recognizing a data set having the same attributes. The present paper uses a density-based clustering method for the clustering process. The advantage of the density-based clustering method lies in the method's better adaptability to the irregularity of clustering patterns and better adaptability to the number of clustering classes. For the SLA-based data set used in the present paper, data characteristics and data density are fixed values. Disadvantages of the traditional density-based clustering method, such as the high sensitivity of the parameter setting, are overcome in this case. For different data sets with the same properties, the same *MinPts* (minimum points) and ε values are chosen, which shows the strong adaptability and intelligence of the algorithm and the reduction of human intervention.

The process of detection by the proposed mesoscale eddy algorithm is divided into five steps. (1) Data are segmented to separate the non-eddy region from the eddy region. (2) Clustering is conducted to recognize mesoscale eddies and establish intraregional relations. The preprocessed data set is divided into multiple sub-data sets and each sub-data set represents a potential mesoscale eddy. (3) Data are classified and filtered to obtain sub-data sets that meet preconditions; data that are too large or too small ought to be filtered out. (4) Shape discrimination is conducted to judge the shape of a mesoscale eddy and to ensure the stability of the selected mesoscale eddy. (5) The outermost contour is determined to find the closed outermost contours and the result is output. To avoid repeating operations in the subsequent iteration process, the data point is deleted from the original data set for the sub-data set obtained a closed contour.

The remaining data sets that are not empty enter the next iteration process, the filter value is increased according to a preset increment, and the above five steps are re-executed until the required data set is empty.

In each iteration, the quantity of data required to manage becomes smaller. The algorithm converges when the data set is empty. Results are integrated to obtain all the outermost closed contours. Results are then output and the detection process finishes. A flowchart of the algorithm is shown in Fig. 1. The whole algorithm goes through several iterations, with only the data segment value changing in the iteration and the other parameters remaining unchanged.

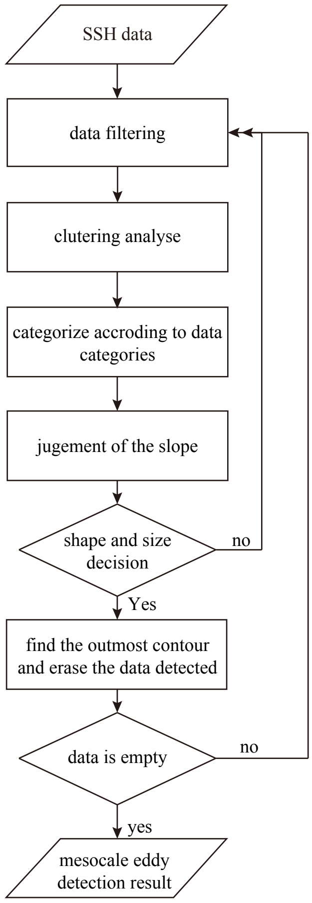

Fig. 1. Flowchart of the DC method.

2.2.1 *Data segmentation: preprocessing*

To distinguish data sets by different attributes and to separate an eddy region from the surrounding environment, the data sets are divided into two classes according to the algorithm objective. In the algorithm, the data points contained in the anticyclonic eddy range are the data points whose SLA values are greater than zero while the data points contained in the cyclonic eddy range are the data points whose SLA values are smaller than zero in most cases. The algorithm initially divides all data points into two categories, those whose SLA values are greater than zero and those whose SLA values are smaller than zero. In the process of detecting cyclonic eddies, the present paper clusters the data sets with SLA values smaller than zero, then uses a data increment of −1 cm to partition the data sets and separates an eddy region from the non-eddy region. In the process of determining the anticyclonic eddies, the present paper considers the data points whose SLA values are greater than zero, and the steps are opposite those followed in the process of recognizing a cyclonic eddy.

2.2.2 *Cluster analysis: association management*

The present paper uses the density-based clustering method to analyze the partitioned data set. The advantage of the density-based clustering method is that it can identify the class characteristics of irregular shapes, and there is no need to specify the core points in advance. In addition, it can automatically filter out noisy points in the clustering process and thus avoid their effect on the data clustering result (He et al., 2014). The algorithm complexity of the density-based clustering method is based on the number of selected data sets (n) and the complexity of the algorithm $O(n^2)$. Clustering is the most complex part of the algorithm. The complexity of the algorithm can therefore be defined as $O(n^2)$, which is similar to that of the SSH-based algorithm (Chelton et al., 2011).

The notion of density-based clustering is defined as follows.

(1) The distance measure between points in the data set (DB) is defined as

$$\mathrm{Dist}(p,q)=\sqrt{(x_p-x_q)^2+(y_p-y_q)^2},\ p,q\in\mathrm{DB}. \tag{1}$$

(2) For a point $p \in$ DB, if a given object p such that within the radius range of ε has at least *MinPts* samples, it is said that the point in the radius range belongs to the data set ε-nbhd (p).

(3) Given p, $q \in$ DB, if the distance between points q and p is smaller than the distance ε and belongs to data set ε-nbhd (p), then point q is directly density-reachable from point p.

(4) For points p and q, if the sample can be directly arrived at q from p by means of samples x_1, x_2, ..., x_t, it is called density-reachable. From the asymmetry of density reachability, it is concluded that density-reachability cannot satisfy symmetry.

(5) If both points p and q are density-reachable, then they can be called density-connected, which satisfies symmetry.

The density-based clustering method for a data set C is as follows. For any p and q with the same class label, p and q are density-connected. For any p belonging to class C and for $q \in$ DB, if p and q are density-connected, then q can be added to class C. If there is a point that cannot be classified into any class, it is called a noisy point. In addition, if a point p has a class label and the point set does not satisfy ε-nbhd (p) from point p, then point p is classified as a border point. Therefore, after rounds of density clustering, several data sets can be obtained, and each point can be classified as a border point or noisy point.

The basic idea of the density-based clustering method is to classify high-density regions in different regions by separating the parts with high data density from those with low data density. The algorithm uses the phenomenon that the closer the mesoscale eddy data are to the eddy core, the higher the data density will be after data segmentation. In this way, a potential mesoscale eddy region with high density can be separated from its surrounding and adjacent mesoscale eddies. After density clustering, the data set can be divided into several different regions, each of which is one or several potential mesoscale eddies.

In the process of density-based clustering, the sensitivity of the parameters is tested according to the characteristics of the selected SLA data. When the search radius is 0.25 (ε = 0.25) and the minimum number of points within the radius is four (i.e., *MinPts* = 4), the separated mesoscale eddy sub-data sets can be obtained in the clustering process. When selecting different search border lengths and numbers of points, the sensitivity of parameter correlation may be too high, and the initialization definition can be complicated, which may lead to an undivided difference between different sub-data sets.

2.2.3 *Data classification and selection: data set filtering*

In clustering analysis, each data point can be assigned a class label or classified as noise, and the data set can be divided into several classes according to the class label. After eliminating data that have been identified as a noisy point, the results of each class label can be extracted separately to form a number of different sub-data sets. The algorithm regards each formed sub-data set as a mesoscale eddy containing all the data points, and the outermost contour should contain all the data points in this sub-data set.

Considering the observation error in the altimetry and resolution of the AVISO fusion product, the recognition conditions for sub-data sets and the standard for identifying mesoscale eddies in literature (Chelton et al., 2011), the specific identification conditions are as follows. (1) SLA values of all grid points in the cyclonic eddy (anticyclonic eddy) are smaller (greater) than that of boundary SLA value. (2) The number of grid cells of the eddy is not less than 5 and not more than 400. (3) The amplitude of the mesoscale eddy is not less than 3 cm. (4) The distance to a nearby grid point in the mesoscale eddy is less than 600 km.

2.2.4 *Shape discrimination: stability criterion*

For the obtained sub-data set of the same class of data, the attributes in the data set are the longitude, latitude and corresponding SLA values. Using the longitudes and latitudes of the sub-data set, the position and shape of a data point in a certain mesoscale eddy can be obtained. To ensure the shape and high stability of the mesoscale eddy, this paper refers to the shape criterion provided in the literature (Liu et al., 2016). The area of deviation is the area of a graph surrounded by a closed contour. The area of the circle adopted as a decision graph is the minimum circular area that covers the closed contour (Fig. 2a). The ratio of the formed graph to the standard graph is obtained. Existing algorithms consider the shape of the eddy to be approximately regular and the structure of the eddy to be stable when the graph ratio is greater than 55%.

In the proposed algorithm, the sub-data set is determined to have a stable shape because the sub-data set can represent all the data points contained in the selected mesoscale eddy. Finding a minimum covering square is more convenient than finding a minimum covering circle. A square is also a regular graph and is more suitable for a data set that is discrete. The decision graph used in this paper is therefore a square rather than a traditional circle. The formulas for finding a square are:

$$
\begin{aligned}
&length_{\rm wide} = \max\,(lon_{\rm i}) - \min\,(lon_{\rm i})\,,\\
&length_{\rm hight} = \max\,(lat_{\rm i}) - \min\,(lat_{\rm i})\,,\\
&length = \max\,(length_{\rm wide}, length_{\rm hight})\,,\\
&Num_{\rm square} = (length \times (1/resolution) + 1)^2,\\
&Error_{\rm shape} = Num_{\rm deviation}/Num_{\rm square}\,.
\end{aligned} \tag{2}
$$

It is straightforward to find the minimum covering square using the maximum and minimum values of longitude and latitude in the sub-data set (Fig. 2b). In this paper, it is assumed that the diameter of the minimum covering circle is equal to the side length of the minimum covering square.

It is assumed that the radius of the minimum covering circle

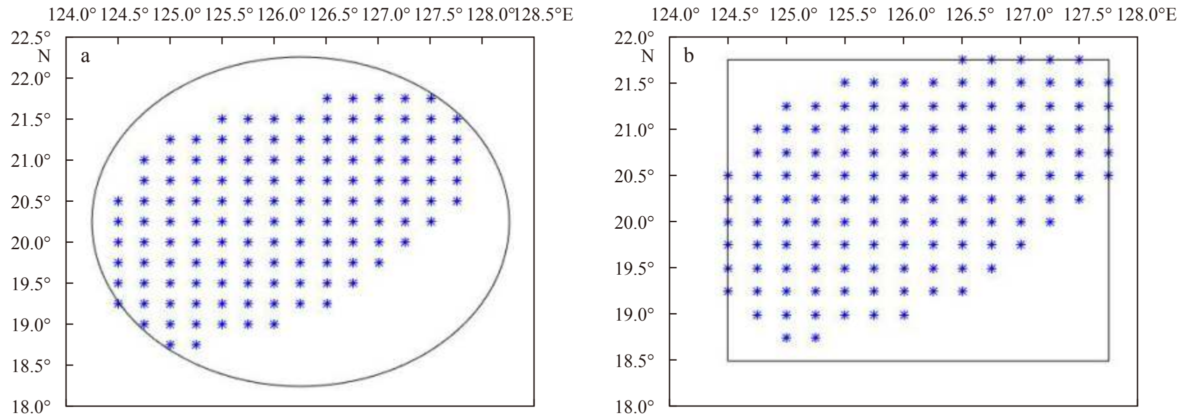

Fig. 2. Shape discrimination, with blue points representing positions of the data point set. a. A minimum covering circle, and b. a minimum covering square.

is 1, the area of the circle is 3.14, the area of the corresponding square is 4, and the coverage of the circular area is 55%, corresponding to coverage of the area of the square area of 43%. A ratio is taken of the number of data in the sub-data set belonging to the same category provided by the filtering process divided by the number of data points included in the minimum covering square. When the ratio is greater than 43%, it is believed that the enclosed graph is a regular graph. A sub-data set enters the next step when it is judged as a regular data set. Otherwise it skips the next step and waits for the next loop for graph judgment again.

A sensitivity test of the parameters shows that the results are more regular if the shape discriminant parameter set is large. However, the mesoscale eddy will be too small or it is not possible to detect some mesoscale eddies directly. At the same time, the computation time is long and there are many iterations of the algorithm. When the parameters are set small, there are many irregular eddy structures. These eddy structures should be divided into smaller mesoscale eddies. Additionally, some mesoscale eddies will be found twice because the selection range is too large in the search process, leading to the repeated identification of eddies.

2.2.5 *Outermost closed contour determination: output results*

After the shape judgment step, the algorithm moves to the step of searching for closed contours. As the closed contour detected should contain all points of the corresponding sub-data sets, it is necessary to extend the selected sub-data sets to find a closed contour over a larger range. A rectangle can be used to represent the position at which the sub data set is located (Fig. 3a). To extend the data set, this paper selects all the encircled longitudes and latitudes to add and subtract a resolution data point and obtain an extended data set. The resulting data set is a rectangular data set with a slightly larger range than the original data set; the border shown in Fig. 3a marks the expanded range of the data set. All contours can then be found using the extended data set (Fig. 3b).

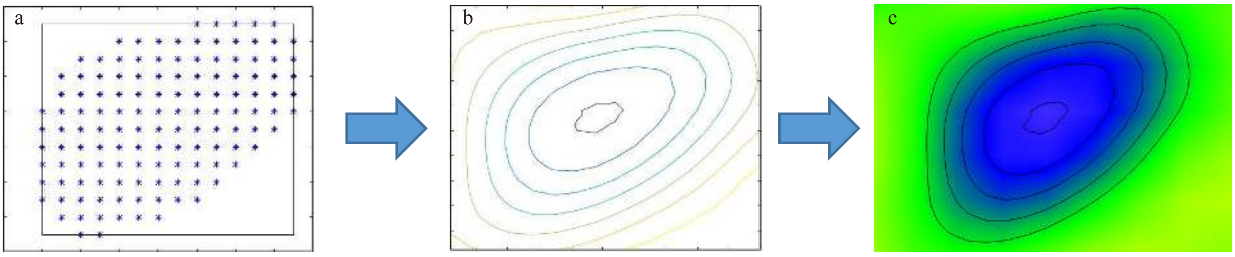

Fig. 3. Process of finding the outermost closed contour. a. Distribution of the clustering results, b. all contours obtained after the extension process, and c. all closed contours on the SLA map and the identification of eddy structure.

In this step, it is necessary to find the closed contour among all closed contours that are equal to the partition values in Section 2.2.1, and this closed contour certainly exists.

2.3 *Eddy detection*

Figure 4 shows the process of the regional segmentation and detection of an eddy. Figure 4a shows the SLA map before the detection process while Figs 4b–g shows the sub-data set after clustering analysis with a modified value of the data segmentation process. Different colors represent different clustering effects. Shape discrimination is carried out in accordance with the data classification and selection preconditions. In the case that the shape is judged to be regular, the outermost closed contour in this region is determined and the result is output; otherwise the process jumps to the next iteration.

Regions C and D are the two parts after the division of Region B (marked in Fig. 4e), while region E is a part after the division of Region D (marked in Fig. 4f), and suitable mesoscale eddies are found in Regions A, C and E. Figure 4h shows three mesoscale eddy structures as the final result.

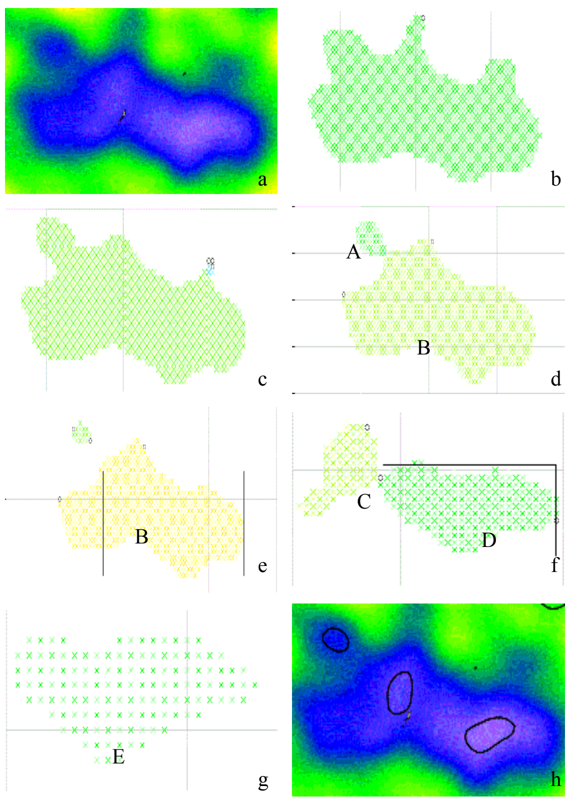

Fig. 4. Process of regional segmentation and detection.

3 Experiment on mesoscale eddy detection

Employing the proposed algorithm, we detected mesoscale eddies in the Northwest Pacific and South China Sea (0°–50°N, 100°–160°E). In this region, there are many mesoscale eddies, the numbers of cyclonic and anticyclonic eddies are relatively balanced, and there are various eddy forms. The region is thus well suited to testing the performance of the proposed algorithm.

3.1 *Comparison with the SSH-based method*

Mesoscale eddy detection algorithms have been well researched in the literature. Classic algorithms include WA and SSH-based methods. The SSH-based method is a widely used eddy detection algorithm that is able to find multi-core structures. This paper compares the SSH-based method with the DC algorithm in the mesoscale region (30°–45°N, 140°–155°E) of the Northwest Pacific.

To ensure large differences in the eddy morphology between two time nodes, we picked nodes more than 40 days apart for the comparison experiments. With three time nodes, we can better test the adaptability of the two algorithms to different data sets and thus better compare the performances of the two algorithms (Fig. 5). In general, both the DC method and traditional SSH-based algorithm detect the main scopes of mesoscale eddies. However, the SSH-based algorithm faces the problems of threshold sensitivity and results instability.

Table 1 shows that the DC algorithm in most cases detects more eddies than the SSH-based algorithm and with better accuracy, as the DC algorithm avoids the threshold sensitivity problem. Meanwhile, avoiding error due to an excessive scope of the closed contour, detection results obtained with the DC algorithm and SSH-based algorithm are compared in Fig. 6. As the traditional SSH-based method simply aims to obtain closed contours and the result would thus have irregular eddy shapes, results obtained by the DC algorithm are filtered with the shape maximum approximating a circle to guarantee the stability of the eddies detected. In comparison with the traditional SSH-based method, the SSH-based method used here regards the area between the multiple eddies also as an eddy structure while the proposed algorithm regards this region as a non-eddy area without processing. This type of eddy structure is relatively small (Fig. 7).

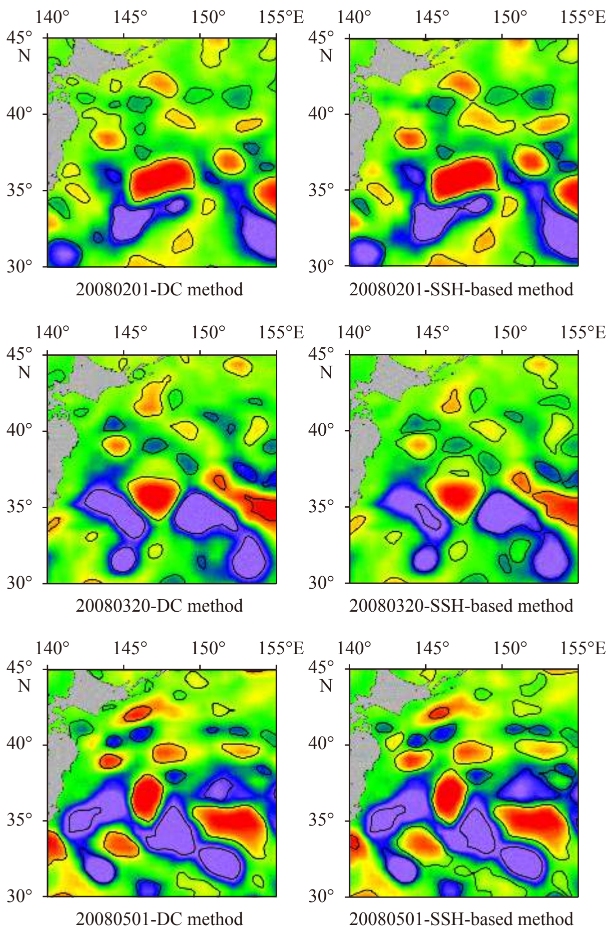

Fig. 5. Comparison of detection results obtained with the DC algorithm and SSH-based algorithm for a 2008 SLA map.

3.2 *Mesoscale eddy detection results*

Figure 8 shows the results of mesoscale eddy detection in the Northwest Pacific and the South China Sea obtained using the algorithm proposed in this paper and the traditional SSH-based method. There are many single-core structures and several multi-core structures in this region. The present paper focuses on solving the problem of how to select and maintain a stable multi-core structure or how to divide an unstable multi-core structure into several stable single-core structures.

A multi-core structure is a mesoscale eddy with two or more eddy cores having the same polarity within the boundary. A mesoscale eddy will sometimes undergo division, merging or other processes in the course of its life. There are many multi-core structures in the Northwest Pacific region, and the question of whether to maintain or to break them down has seldom been

Table 1. Comparison of the number of eddies obtained with the DC algorithm and SSH-based algorithm, where CE refers to a cyclonic eddy and AE to an anticyclonic eddy

Date	DC method		SSH-bd method	
	CE	AE	CE	AE
20080201	20	16	14	13
20080320	19	14	16	13
20080501	19	13	19	15

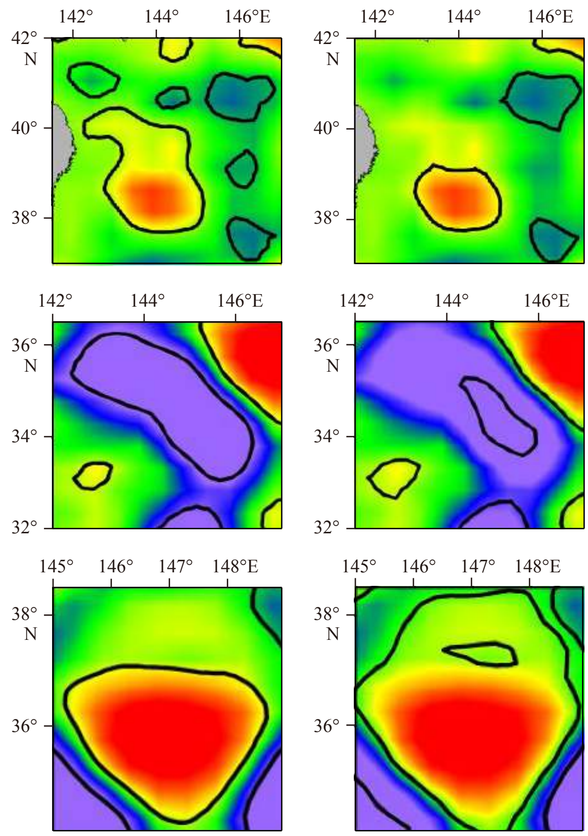

Fig. 6. Detailed comparison of detection results obtained with the DC algorithm and SSH-based algorithm. The left column presents results of the DC algorithm and the right column results of the SSH-based algorithm.

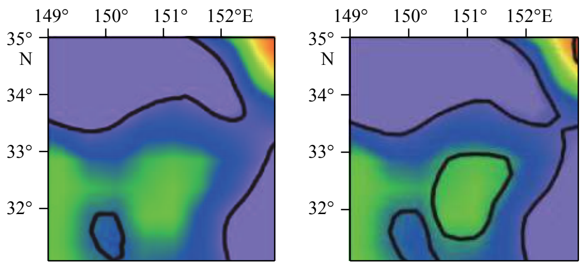

Fig. 7. Initialization difference between the DC algorithm and SSH-based algorithm. a. DC algorithm, and b. SSH-based algorithm.

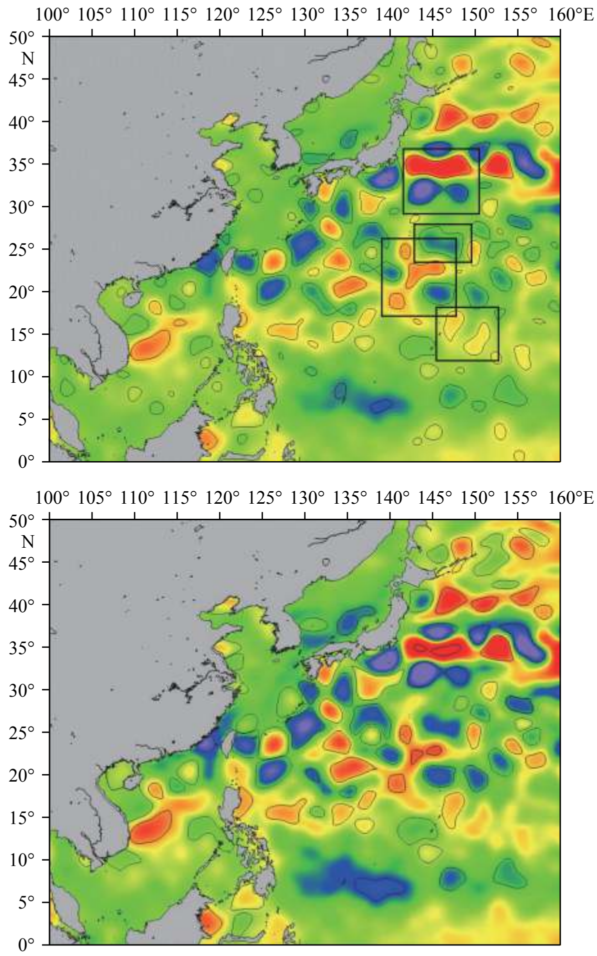

Fig. 8. Results of eddy detection for the region of the Northwest Pacific and South China Sea on May 9, 2016. Contours fill warm color indicate anticyclonic eddies while those filled with a cool color indicate cyclonic eddies. a. Results of DC algorithm, and b. results of traditional SSH-based algorithm.

mentioned in previous work. The first time that this problem was addressed was the proposal of the SSH-based method for the detection of a multi-core structure (Chelton et al., 2011), but owing to the instability of the multi-core structure, the form is more complex. In many cases, although the SSH-based algorithm can detect a closed contour, it is unable to track the movement in the process. The SSH-based method as a kind of detection algorithm of a multi-core structure. It lacks a judgment of the eddy shape. Although it overcomes the particular problem of traditional algorithms that each eddy can have only one core, it cannot automatically identify an irregular eddy structure and decompose it to a more stable structure.

The DC algorithm proposed in this paper increases the stability of found multi-core structures through the shape judgment.

Figure 9 shows the results of multi-core structure detection. Figure 9a shows a segmentation of a larger scale eddy of irregular shape, which contains more than one local maximum or minimum point on a larger scale. Using the proposed algorithm, we get two stable single-core structures and one multi-core structure having two extremum cores. Figures 9b–d shows that multi-core structures are found and preserved. These multi-core structures have relatively regular shapes with little fluctuation of the SLA value within the contour.

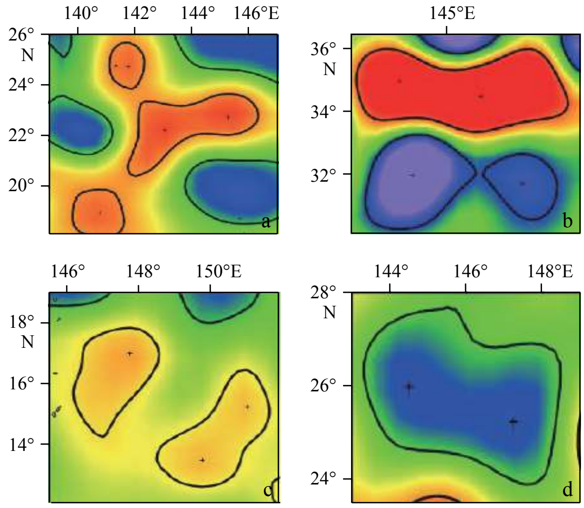

Fig. 9. Examples of eddy structure detection. "+" indicates the extremum core.

4 Conclusions

To solve the problem of automatically detecting mesoscale eddies from an SLA map, this paper proposed a combination of the density clustering algorithm used in the field of machine learning with a mesoscale eddy detection algorithm: the density-based clustering mesoscale eddy detection algorithm.

The basic theory of this algorithm and traditional methods, such as the WA and SSH-based algorithms, is the same in that we search for the outermost closed contour with which to locate mesoscale eddies. However, the proposed algorithm does not require an initialization process that searches for a local extremum to define the cores of eddies and extracts a closed contour, and rather extracts data features to detect regular mesoscale eddies step by step.

The proposed algorithm takes advantage of the strong adaptability of the density clustering algorithm to an unsupervised data set and identifies potential relations by means of density clustering for regions in which mesoscale eddies can be separated. By including eddy shape discrimination in the detection process, the boundary of potential eddies is fitted within a minimum coverage square, and the coverage ratio is determined to approximately judge whether the eddy shape is regular. The method ensures the stability of the selected eddy structure, and regular mesoscale eddy patterns are maintained to the maximum extent.

In comparison with the traditional SSH-based algorithm, the proposed algorithm requires less parameter setting, and as an automatic algorithm that eliminates manual intervention, it avoids the sensitivity of the threshold setting. The detection results of eddy structures are more stable, and the eddy shapes are more similar to regular patterns. The discovered mesoscale eddies can also be used for subsequent research work and prediction.

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

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

Doglioli A M, Blanke B, Speich S, et al. 2007. Tracking coherent structures in a regional ocean model with wavelet analysis: application to Cape Basin eddies. Journal of Geophysical Research: Oceans, 112(C5): C05043

He Yaobin, Tan Haoyu, Luo Wuman, et al. 2014. MR-DBSCAN: a scalable MapReduce-based DBSCAN algorithm for heavily skewed data. Frontiers of Computer Science, 8(1): 83–99, doi: 10.1007/s11704-013-3158-3

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

Liu Yingjie, Chen Ge, Sun Miao, et al. 2016. A parallel SLA-based algorithm for global mesoscale eddy identification. Journal of Atmospheric and Oceanic Technology, 33(12): 2743–2754, doi: 10.1175/JTECH-D-16-0033.1

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

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

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

Waugh D W, Abraham E R, Bowen M M. 2006. Spatial variations of stirring in the surface ocean: a case study of the Tasman Sea. Journal of Physical Oceanography, 36(3): 526–542, doi: 10.1175/JPO2865.1

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, doi: 10.5194/os-10-39-2014

Zhang Chunhua, Li Honglin, Liu Songtao, et al. 2015. Automatic detection of oceanic eddies in reanalyzed SST images and its application in the East China Sea. Science China Earth Sciences, 58(12): 2249–2259, doi: 10.1007/s11430-015-5101-y

Zhang Chunhua, Xi Xiaoliang, Liu Songtao, et al. 2014. A mesoscale eddy detection method of specific intensity and scale from SSH image in the South China Sea and the Northwest Pacific. Science China Earth Sciences, 57(8): 1897–1906, doi: 10.1007/s11430-014-4839-y