Visualization for Ocean General Circulation Model via Multi-dimensional Transfer Function and Multivariate Analysis

Daisuke Matsuoka [1], Fumiaki Araki [1], Shinichiro Kida [1], Hideharu Sasaki [1], and Bunmei Taguchi [1]

[1] Earth Simulator Center, Japan Agency for Marine-Earth Science and Technology, Japan

1 Introduction

With advances in supercomputing technology, ultra-high resolution large-scale simulation study has been the mainstream in the field of ocean science. The dataset from such simulation is not only large in size, but also contains a lot of useful information. However, as the geophysical fluid simulation, it is difficult to understand such information hiding in the dataset due to the complexity of ocean dynamics and the number of physical variables. In order to obtain scientific knowledge from large-scale simulation data, it is important to effectively extract and to efficiently express the characteristic feature [1, 2].

The aim of this study is how to solve the above problem. In this study, a multi-dimensional transfer function to emphasis the feature such as ocean currents and vortices from OFES (Ocean general circulation model for the Earth Simulator) [3, 4] dataset is proposed. Furthermore, a multivariate analysis to extract such features is developed. This paper describes the experimental results of the applying the multi-dimensional transfer function and multivariate analysis to the visualization of Oceanic Global Circulation Model data.

2 Feature Emphasise Visualization via Multi-dimensional Transfer Function

Visualization of simulation results is effective way to understand a complex phenomenon and configuration intuitively. However, representation using a conventional transfer function mapping from one scalar value to one color and opacity is not good enough to understand its fine structure. For example, Fig.1.(a) shows the visualization result of Sea Surface Temperature (SST) distribution calculated by the OFES with horizontal resolution of 1/10 degree via the traditional transfer function of a rainbow color map. The color red indicates high value, blue is low as shown color bar in the bottom right of the figure. Fig.1.(b) shows the magnitude of sea surface velocity. The counter color in the red and black area denotes high and slow speed, respectively. In such a general color map, it is suitable for seeing the distribution of the value. However, it is not easy to represent the difference of flow speeds in the same temperature region.

In this study, a multivariate color map which not only assigns Hue (H) to one value but also assigns Saturation (S) and Brightness (V) to other variables is proposed in order to representation not only one value but also two or more values. In Fig.1.(c), rainbow colors represent the SST and brightness represents the flow speed (the bright color indicates fast flow, and the dark color indicates slow flow). The temperature and the flow speed can be recognized simultaneously by assigning two variables to H and B. The flow speed range from 0.0 m/s to 1.0 m/s corresponds to the range of V from 0.5 to 1.0. Fig.1. (c) effectively emphasis the heat transfer by the ocean currents and vortices than Fig.1. (a) and (b). For instance, the Kuroshio, North Pacific Current, Agulhas current and Antarctic Circumpolar Current and its heat transfer can be seen in this figure.

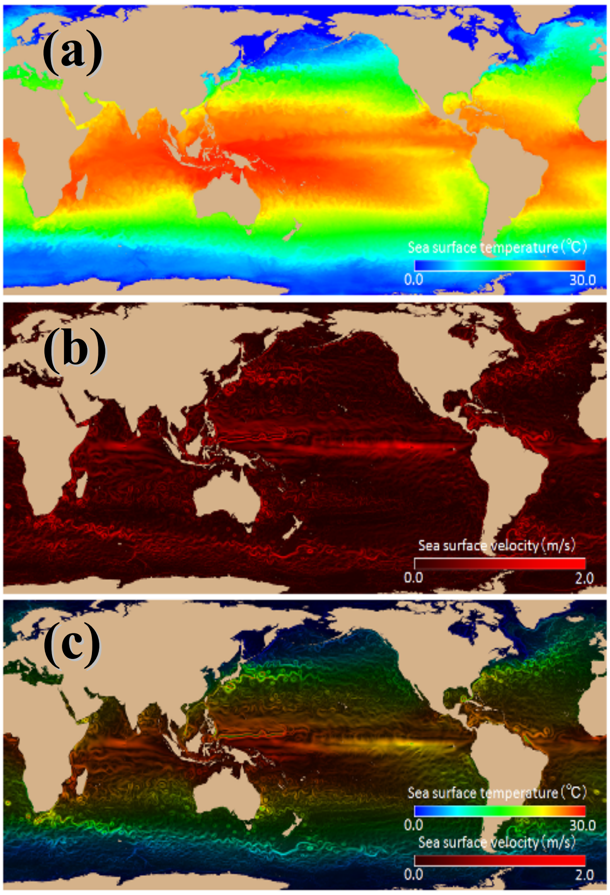

Fig. 1. The distribution of (a) Sea Surface Temperature (SST), (b) magnitude velocity and (c) mixed image of (a) and (b) via multi-dimensional transfer function.

3 Feature Extraction and Visualization via Cluster Analysis

In order to accuracy set a transfer function using three or more variables, it is necessary to extract characteristic features using semi-automatic objective analysis. In this section, cluster analysis which is one of multivariate analysis techniques is improved and applied to extract ocean currents.

3.1 Application of Density-based Cluster Analysis

Cluster analysis is the task of assigning a set of objects into groups called cluster so that the objects in the same cluster are more similar to each other than to those in other clusters. Several cluster analysis approaches have been proposed. They may be divided into following 3 types: Hierarchical method, Non-hierarchical method (k-means) and Density-based method [5].

The point about the feature of ocean currents is that two elements (data) in a same current are not always similar but continuity. Hence, we choose density-based method [6] which uses the connectivity and density in variable space. In this research, we originally improved this cluster analysis method and applied it to extraction of ocean currents. In the our method, it is necessary to set the following parameters, a position of initial seed, density in the variable space, range for calculating density, range for exploration and range for clustering. Initial seeds and clustering results is shown in Fig.2. The red and blue area indicates the Kuroshio and Oyashio, respectively. Both results can be seen that the clustering is roughly correct.

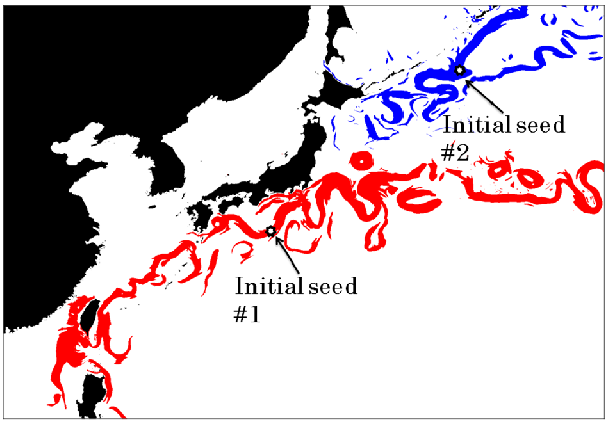

Fig. 2. Clustering result of Oyashio and Kuroshio using density based method.

3.2 Application of Non-Hierarchical and Hierarchical Hybrid Cluster Analysis

Although the density-based method can extract ocean currents, its computational cost is too high. In this study, we adopted a non-hierarchical cluster method in which the computational complexity and the memory size are not restricted.

Non-hierarchical algorithms [7] divide a set of elements into *k* partitions or groups with each partition representing a cluster. *k* initial cluster centroids are selected manually using domain knowledge. In clustering analysis, one of the important steps is to select a distance measure, which can determine how the similarity of two elements is calculated. The distance between two elements (*i* and *j*) in 4-dimensional Euclidean space is defined as:

$$D^2_{ij} = C_T(T_i - T_j)^2 + C_V(V_i - V_j)^2 + C_X(X_i - X_j)^2 + C_Y(Y_i - Y_j)^2$$

where, T, V, X and Y is temperature, flow speed, x-position, y-position, respectively, and C_T, C_V, C_X, and C_Y are constant number to make dimensionless quantity of similarity.

The resulting image of non-hierarchical cluster analysis and 16 centroids are shown in Fig.2. (a). In order to visualize the result, each cluster is indicated by a different color. Not only the Kuroshio (cluster #1 and #2), the Kuroshio Extension region (cluster #3 and #4) and the Oyashio (cluster #14 and #15) but also the mixed layer in the Kuroshio/Oyashio Extension region and the vortices in the Sea of Japan are clustering almost correctly.

Hierarchical algorithm is a method of cluster analysis which seeks a hierarchy of clusters or elements. Next, hierarchical method is adopted to extract the hierarchical structure of the small clusters obtained by the non-hierarchical method. The hierarchical structures using the hierarchical method is represented as a tree diagram called dendrogram as shown in Fig.3. (b). We can recognize from Fig.3. (b) that Kuroshio and Oyashio is composed from cluster #1 - #4 and cluster #14 and #15, respectively.

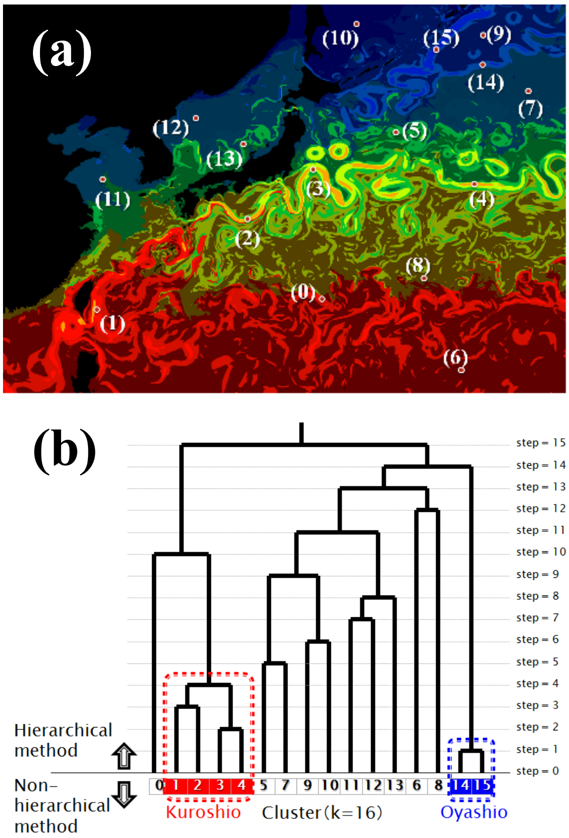

Fig. 3. Clustering results of Oyashio and Kuroshio using (a) non-hierarchical method and (b) hierarchical method.

4 Conclusion

This paper described a new multi-dimensional transfer function that can emphasis the ocean currents and multivariate analysis methods that can extract the ocean currents. These results lead to the conclusion that multi-dimensional transfer function and multivariate analysis are effective method in order to visualize and understand high resolution simulation data. In the visualization for next generation simulation, it is necessary to efficiently extract useful information hiding into the datasets and to effectively express such information.

Acknowledgements

This work was supported by KAKENHI (24650045) for Grant-in-Aid for Challenging Exploratory Research.

References

[1] C. D. Correa and K.-L.- Ma, Visibility-Driven Transfer Functions, Proc. of Pacific Viz '09, pp. 177-184, 2009

[2] N. Ohno and D. Matsuoka, Visualization of Global Geophysical Fluid Simulation, Nagare, Vol. 30, No. 5, pp. 409-414, 2011

[3] Y. Masumoto et al., A Fifty-Year Eddy-Resolving Simulation of the World Ocean –Preliminary Outcomes of OFES (OGCM for the Earth Simulator)-, Journal of the Earth Simulator, Vol. 1, pp. 35-56, 2004

[4] H. Sasaki et al, An eddy-resolving hindcast simulation of the quasi-global ocean from 1950 to 2003 on the Earth Simulator, High Resolution Numerical Modeling of the Atmosphere and Ocean, pp. 157-186, 2008

[5] L. Kaufman and P. J. Rousseuw, Finding Groups in Data: an Introduction to Cluster Analysis, John Wley & Sons, 1990

[6] A. K. Jain and R. C. Dubes, Algorithms for Clustering Data, Prentice hall, 1988

[7] M. Ester et al, A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise, Proc. of 2nd Int. Conf. on Knowledge Discovery and Data Mining, pp.226-231, 1996