Time-varying Visualization of Ocean Flow Data with Clustered Heatmap

Midori Yano *
Graduate School of Humanities and Sciences, Ochanomizu University

Satsuki Kumatani†
Graduate School of Humanities and Sciences, Ochanomizu University

Yumi Inoue ‡
JAMSTEC

Takayuki Itoh§
Graduate School of Humanities and Sciences, Ochanomizu University

Abstract

Heatmap is one of the effective representation for time-varying data visualization. It may require large display spaces when an input dataset contains large number of date items or time steps. We may often want mechanisms to interactively filter non-important data items or time steps, so that we can form appropriate sizes of heatmaps and focus on important data items or time steps. This paper presents a heatmap-based time-varying data visualization technique featuring an interactive mechanism to display meaningful data times and time steps. This technique firstly calculates distances between arbitrary pairs of data items, and constructs a dendrogram consisting the data items. It then generates clusters of the data items and displays the data items belonging to the specified sizes of clusters in the heatmap, so that we can focus on groups of similar or correlated data items. It applies a similar mechanism to a set of time steps so that we can remove outlier time steps from the heatmap. Our implementation features two scatterplots which represent distribution of data items and time steps respectively, and slider widgets to interactively adjust the thresholds of the clustering process. We can intuitively understand how clusters of data items or time steps are constructed, by looking at the scatterplots while operating the sliders.

Keywords: Time-varying data visualization, Heatmap, Scatterplot, Clustering

1 Introduction

Observation of ocean flow is important for variety of academic and industrial fields, such as weather forecasting, fishery, sailing traffic navigation, and tourist trades. Experiments and computer simulation of ocean flow has a long history of research. OFES (Ocean general circulation models For the Earth Simulator) [1] is one of large-scale computer simulation of ocean models which performed a fifty-year eddy-resolving simulation of the world ocean. The model covers a near-global region extending from 75°S to 75°N with horizontal grid spacing of 0.1° and 54 vertical levels. The simulation result was visually well explored applying 2D scientific visualization techniques while assigning two of longitude, latitude, and depth to the x- and y-axes of the visualization spaces [1].

Information visualization techniques have been recently applied to variety of scientific problems. For example, multidimensional or time-varying data visualization techniques have been applied to visualization of multivariate or time-varying volume datasets. Or, relationship or categorization of various scientific simulation or visualization results have been represented by applying tree or network visualization techniques. We expect fruitful knowledge regarding ocean flow simulation will be discovered by applying such information visualization techniques.

We have developed heatmap-based time-varying data visualization techniques recently [2]. We suppose a two dimensional orthogonal coordinate system where time is assigned to the horizontal axis and data items (corresponding to the variables) are arranged along the vertical axis. The heatmap-based visualization technique divides a display space into a grid along the coordinate system, and paints the small rectangles in the grid according to the values at each time step.

In this poster, we introduce a case study of heatmap-based time-varying data visualization technique to the time-varying visualization of ocean flow around Japan calculated by the Earth Simulator. We extracted particular grid-points in the neighboring waters in Japan from the ocean flow simulation result and constructed a time-varing dataset. We visualized that the temporal changes of scalar values including salinity and temperature are clearly divided to two clusters corresponding to two major ocean currents so called “Kuroshio” and “Oyashio”.

2 Data Definition

We extracted grid-points of the simulation results corresponding to the islands and neighboring waters in Japan from the simulation result. The range of longitude of the extracted grid-pointes was [116.5°E, 200.0°E], while the range of latitude was [14.9°N, 64.9°N].

Warm and cold currents are important factors which affect to weather and fishery. Kuroshio and Tsushima currents are major warm currents around Japan, while Oyashio and Liman currents are major cold currents. Especially, Kuroshio and Oyashio are large scale currents moving in the pacific ocean. We observed how temporal change of scalar values are different between Kuroshio and Oyashio.

Figure 1 shows the extracted region and major currents around Japan. The extracted regions consists of 835x500x54 grid-points since the dataset has horizontal grid spacing of 0.1° and 54 vertical levels. Each grid-point has a three-dimensional flow vector and scalar values including salinity and temperature. We extracted these values at 60 timesteps corresponding to 60 days in the dataset.

3 Visualization of time-varying scalar values along Kuroshio and Oyashio

We applied a heatmap-based time-varying data visualization technique to the ocean flow simulation data. We sampled pre-defined number of grid-points from the dataset, and extracted the specific scalar values for a certain period. Let the sampled grid-points $G = \{g_1, g_2, ..., g_m\}$, where m is the number of sampled grid-points. Also, let the extracted scalar values of a grid-point $g_i = \{v_{i1}, v_{i2}, ..., v_{in}\}$, where n is the number of periods. In our experiment, the scalar value is one of velocity of the flow vector, salinity, or temperature.

We calculated distances between arbitrary pairs of the grid-points. Our current implementation defines the distance between a pair of grid-points as $1.0 - |Ken(g_1, g_j)|$, where $Ken(g_i, g_j)$ is

*e-mail: drsn@itolab.is.ocha.ac.jp
†e-mail: satsuki@itolab.is.ocha.ac.jp
‡e-mail: inouey@jamstec.go.jp
§e-mail: itot@computer.org

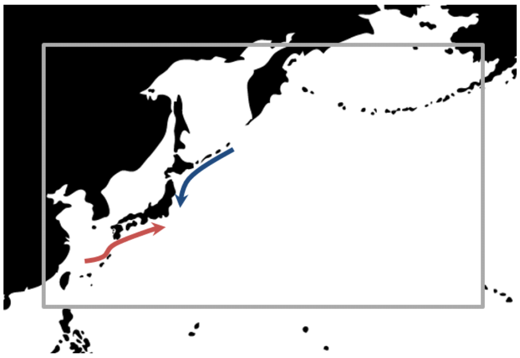

Figure 1: Kuroshio and Oyashio around Japan. Arrow lines drawn in pink and blue are Kuroshio and Oyashio respectively.

Kendall correlation coefficient between g_i and g_j. We then constructed a dendrogram of grid-points and applied a hierarchical clustering algorithm with the above-mentioned distances, to divide the grid-points into correlated groups. Finally, we displayed the set of time-varying scalar values of the grid-points as a heatmap, where timesteps are assigned to the x-axis and the grid-points are aligned along the y-axis in the order defined by the dendrogram.

Firstly, we sampled grid-points in two regions where Kuroshio and Oyashio run through respectively, as shown in Figure 2(Left). We extracted salinity at the sampled grid-points and visualized by a heatmap, as shown in Figure 2(Right). Obviously, the grid-points could be clearly divided into two clusters corresponding to Kuroshio and Oyashio. Warm colors in the heatmap depict higher salinity, where cold colors depict lower salinity. The visualization result illustrates that there are large number of small high- or low-salinity portions which frequently appear and disappear in Kuroshio. On the other hand, small number of large high- or low-salinity portions in Oyashio move constantly.

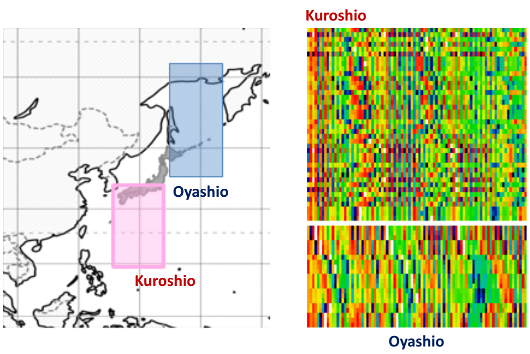

Figure 2: Visualization example with a heatmap.

Our heatmap-based visualization can represent negative temporal correlations as well as positive correlations, since it calculated distances among individuals based on absolute values of correlation coefficients. For example, we could find clear negative correlation of velocity in particular regions in two regions (a) and (b) shown in Figure 3. In the region (a) we can observe positive correlations between 1 and 4, as well as between 2 and 3. We can also observe negative correlation between these two pairs. Similarly, positive correlation between 2 and 4 is observed while they are negatively correlated with 3, in the region (b). Our visualization has a good property to discover such sets of positively and negatively correlated temporal patterns. We would like to discuss what kinds of physical mechanisms bring such correlations as a future issue.

4 Conclusion

We presented a case study of heatmap-based time-varying data visualization applied to ocean flow data simulated by the Earth Simulator. This poster showed visualization examples focusing on temporal change of scalar values including salinity and velocity around Japanese two major currents called Kuroshio and Oyashio.

This study is just in an early stage and therefore we have many future issues. Following are examples of our future directions.

Improvement of heatmap-based visualization. There are variety of distance calculation schemes and clustering algorithms implementable to heatmap-based visualization. We would like to determine which distance and algorithm are most suitable for our study. Also, we would like to apply various color maps and interaction mechanisms to improve explorative analysis of users.

Test with large datasets. The study presented in this poster intensively extracted grid-points in regions of Kuroshio and Oyashio. The problem should be much more complicated if we extract grid-points from whole the dataset. Also, we may discover more interesting phenomena if we visualize longer term of temporal change of scalar values. We would like to test the visualization with larger datasets containing wider regions and longer terms.

Linked view with 3D volume visualization. We may want to observe how clusters of scalar values geographically distribute. Or, we may also want to observe how the clusters form 3D shapes in the ocean. Since it is difficult to imagine such spatial structures just by looking at the heatmap, we would like to develop a linked view mechanism to connect the heatmap-based time-varying data visualization component with a 3D volume visualization component.

Reasoning with experts in marine science. Currently we have not discussed the relationships between visualization results and knowledge in marine science. We would like to discuss what the visualization results suggest fruitful knowledge in marine science.

User evaluation. We would like to test this case study with participants and verify how it is easy to use.

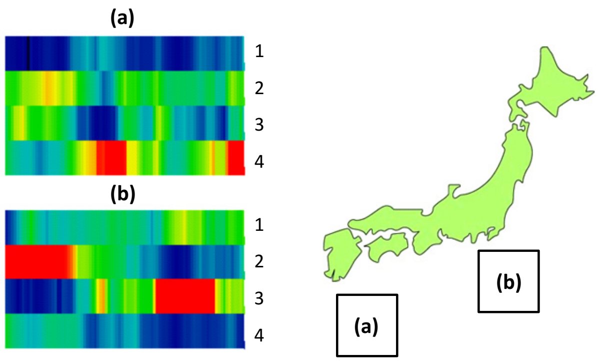

Figure 3: Visualization example with a heatmap.

References

[1] 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.

[2] H. Suematsu, S. Yagi, T. Itoh, Y. Motohashi, K. Aoki, S. Morinaga, A Heatmap-Based Time-Varying Multi-Variate Data Visualization Unifying Numeric and Categorical Variables, 18th International Conference on Information Visualisation (IV2014), pp. 84-87, 2014.