

Spatiotemporal variability of vertical structures of temperature and salinity in the mid-latitude northwest Pacific Ocean based on unsupervised clustering

*Fumika Sambe¹, Toshio Suga^{1,2}

1. Graduate School of Science, Tohoku University, 2. JAMSTEC

In the mid-latitude northwest Pacific Ocean, the transport, stirring and mixing of subtropical and subarctic water occur, resulting in multiple ocean domains characterized by a variety of unique oceanographic structures. Many studies were conducted on the characteristics and distribution of these ocean structures, and the boundaries of domains and characteristic water masses were defined. Most of subsequent studies on the ocean structure and water mass in this ocean region have been based on the existing definitions of the structure and water mass of interest based on spatiotemporally limited data. Even while a large amount of vertical profile data has become available evenly in space and time with the expansion of the Argo observing network, the basic research framework has not changed significantly. However, with the huge amount of data accumulated by Argo with a small spatiotemporal bias, it may be possible to objectively recapture the ocean structure by data-driven scientific methods, that is, machine learning. The Profile Classification Model (PCM) developed by Maze et al. (2017) is a machine learning method for vertical profile data in the ocean. The PCM classifies profiles by unsupervised clustering using Gaussian Mixture Model, and one of its main features is that it can perform analysis using a labeling metric that quantifies the certainty of the classification. Previous studies have shown that clustering with PCM represents the vertical structure and spatial distribution of each class representing a particular ocean structure and its main distribution area. In this study, PCM is applied for the first time to the temperature and salinity profiles by Argo floats in the mid-latitude northwest Pacific Ocean, and the results are compared with previous knowledge of ocean structure. In addition, we tried to understand the spatiotemporal variability of oceanographic structures from the results of PCM by grouping the results based on the Kuroshio Extension index (KE index), which is a measure of a stable and an unstable dynamic state of the KE, and discussing the characteristics of each group.

In this study, PCM clustering was performed on 57476 temperature and salinity profiles located between 30°N and 60°N and between 140°E and 180°E. Based on statistical criteria and the initial value dependence, the number of classes was set to five. The results showed that classes 1, 2, 3, 4, and 5 were distributed to form five regions: the subarctic region, the mixed water region, the subtropical subarctic intermediate region, the northwestern part of the subtropical gyre, and the downstream of the KE region. Each class has different oceanographic features. The relationship between the distribution area and the characteristics of each class corresponded to the relationship between the ocean structure and the area described in previous studies, which mainly focused on specific water masses and vertical structures. In addition, we investigated how the distribution of each class changed depending on the KE index. We found that the shape of the distribution area of class 4 and the degree to which classes other than class 5 coexist in the downstream of the KE region changed. This change in the distribution area was consistent with the results expected from the north-south shift of the KE position due to the dynamic state of the KE. The profiles with low labeling metrics were found to be concentrated near the boundaries of the main distribution areas of the classes. This may reflect “mixing” between different classes at the domain boundary. It was also confirmed that the proportion of profiles with low labeling metric to total profiles had a significant negative correlation with KE index, suggesting that the more unstable the KE dynamic state was, the more the interclass “mixing” became active.

Keywords: Mid-latitude Northwest Pacific Ocean, Kuroshio Extension, Argo float, Unsupervised Clustering, Data-driven Science, Oceanographic Structure