Classification of Deep Sea Sound Speed Profiles in the Northwest Pacific Based on Cluster Analysis

Yu Zhang, Hua Wang, Xiangjun Yu, Qinghong LI
Dalian Naval Academy, Dalian Liaoning, China

Abstract. Sound speed profile in ocean has significant seasonal and sea area distribution characteristics. Different sound speed profiles have different effects on underwater acoustic propagation, which makes the detection performance of sonar have significant differences. In order to study the distribution characteristics of sound speed profile in the Northwest Pacific, using WOA13 data set, calculate the number of types by application hierarchical clustering method, using cluster fuzzy c-means method classify sound speed profile, calculate the distribution area of various types in different seasons and typical sound speed profile structure. It is concluded that there are 6 types of sound speed profiles in Northwest pacific: Subarctic category, subarctic transitional category, subtropical transitional category, subtropical category, tropical transitional category and equatorial tropical category. The distribution area of these categories is generally zonal distribution. Except for the tropical transitional category and equatorial tropical category, the other categories have significant seasonal changes.

1. Introduction
One of the important factors affecting underwater acoustic propagation is sound speed profile. The sound speed profile varies with season, time and sea area, and the detection distance of different sound speed profiles of the same sonar has significant differences. Therefore, it is necessary to analyze the sound speed profile, especially to study the classification and distribution characteristics of sound speed profile types. Marine Atlas of Bohai Sea, Yellow Sea and East China Sea (1992)\[1\] provided the type distribution of sound speed profile in the Bohai Sea, Yellow Sea and East China Sea according to experience, but it was limited to the shallow sea shelf sea area to the north of the East China Sea. Xiejun et al. (2009)[2] used hierarchical cluster, sequential cluster analysis and SOFM neural network cluster arithmetic to divide the offshore waters of China into 3 categories and 13 types of sound speed profile. Zhang xu et al. (2010)[3] used fuzzy c-means clustering to divided the coastal waters of china into 3 categories and 11 types sound speed profile. Wang et al.(2019)[4] used the same method to study the classification of deep-sea sound speed profiles in the Indian Ocean, at present, there are few studies on the classification of deep-sea sound speed profiles in the Northwest Pacific. Mandelberg (2000)[5] used hierarchical clustering method to classify sound speed profiles in the North Atlantic and Northeast Pacific, but did not involve the classification of the types in the Northwest Pacific Ocean. The Northwest Pacific is the sea area that we need to focus on, the distribution of sound speed profiles in Northeast Pacific is necessary to studied.

In this paper, Using WOA13 data sets, hierarchical clustering method is applied to classify the number of types in the Northwest Pacific, and the fuzzy c-means clustering method is used to give the distribution areas of different types in different seasons and the typical sound speed profile structure of each type. Thus provides guidance for the use of sonar.
2. Data and methodology

2.1 Data
WOA13 (World Ocean Atlas 2013) data set published by the National Oceanographic Data Center (NODC) of the United States were chosen. The original data of this data set comes from the sea temperature and salinity profiles of historical stations, MBT, CTD, DBT and XBT in the world from 1955 to 2012. These data are controlled by range test, gradient test, statistical test and static stability test methods. Through objective analysis by Leitus, the horizontal grid is 0.25°×0.25° (102 layers in the vertical standard layer in the depth range of 0-5500 m). In this paper, the range of the Northwest Pacific is 0°~65°N, 120°~180°E. The seasonal average temperature and salinity data of the region were selected, and each season included 62400 sets of temperature and salinity profiles respectively. According to data of sea temperature, salinity and depth, the empirical formula proposed by Mackenzie (1981) is used to calculate the sound speed.

2.2 Sound speed profiles classification method
The existence of deep-sea acoustic channel makes most of the sea areas have deep-sea convergence area, which is a typical feature of deep-sea acoustic propagation. Whether the convergence area can appear or not is mainly determined by the sea depth, sound speed profile and its seasonal variation. When there is a critical depth, the sound wave will reverse and return to the sea surface due to the refraction effect of the sea, so as to avoid contact with the sea floor, so that the energy loss is small and make acoustic long-distance propagation. According to the statistical analysis of the critical depth of the Northwest Pacific, there is no critical depth in summer when the water depth is less than 2000m. Therefore, the deep sea in this paper refers to the water depth over 2000m.

The Northwest Pacific is vast, the surface waters mass include tropical water, subtropical water and subarctic water, and there are many currents in Northwest Pacific. The main currents include the North Equatorial Current, the Equatorial Countercurrent, the Kuril Current, Oyashio current, Kuroshio current, North Pacific warm current, etc. The physical properties of seawater vary greatly. Cluster analysis can accurately divide the sound speed profiles category and the area of each category. Typical sound speed profiles of each category can also be obtained. Here, the hierarchical clustering method is used to calculate categories number of sound speed profiles in Northwest Pacific, and then the fuzzy c-means clustering is used to classify the distribution area and the sound speed profile central values of every category in Northwest Pacific.

The cluster method classifies the samples according to their similarity. The hierarchical clustering method is the most commonly used in oceanography, which is a polymerization process. Let the set of sound speed profiles in a certain season be \( X = \{x_1, x_2, \ldots, x_n\} \), The number of samples is \( n \), \( x_i \) represents the vector of sound speed values in the i-th profile, make \( n \) samples into one class, marked as 1, 2, ..., \( n \), at this time, the mean of each category is the sample value itself, and the sum of the squares of the intra-class deviation is zero. Then, the sum of the squares of the dispersion between the classes are calculated. The sum of the squares of the dispersion represents the magnitude of the two categories of similarity. According to the degree of similarity, the classes with high similarity are merged first, and then the classes with low similarity are merged. Finally, when the sum of squared deviation reaches 1000, the classification ends. Thus, the sound speed profiles categories are divided into 6 categories in Northwest Pacific.

Hierarchical clustering method is straightforward, but when the sample size is large, the computer is taking up too much memory. When analyzing sound speed profiles with 0.25°×0.25° grid in Northwest Pacific, the value of \( n \) will be large and the processing speed will be very slow. In addition, the category of sound speed profiles between any two areas is not eidetic but ambiguous. In order to make the regional division result more consistent with the actual situation, fuzzy c-means clustering is adopted for classification.

Let the set of sound speed profiles in a certain season be \( X = \{x_1, x_2, \ldots, x_n\} \), The number of
samples is \( n \), \( x_i \) represents the vector of sound speed values in the \( i \)-th profile. In the depth range of 0~2000m, there are 67 layers. Establishing the criterion function based on the Euclidean distance.

\[
J(D,V) = \sum_{j=1}^{n} \sum_{i=1}^{n} (d_{ij})^2 \| x_j - V_i \|^2
\]  

(1)

\( D = (d_{ij})_{p \times n} \), \( d_{ij} \) is the membership function of the \( j \)-th sample to the \( i \)-th class, need to meet

\[
d_{ij} \in [0,1], \quad \sum_{j=1}^{n} d_{ij} = 1, \quad \sum_{j=1}^{n} d_{ij} > 0 \;
\]

\( V_i = \{v_1, v_2, \cdots, v_p\} \), \( v_i \) is the cluster center of class \( i \); \( c \) is the fuzzy degree control constant, here, \( c = 2 \), \( P \) is the number of categories. After \( P \) is determined by the hierarchical clustering method, an initial partition matrix \( D_0 \) is given arbitrarily for the samples and the clustering center is initialized, and then \( d_{ij} \) and \( v_i \) iterative calculation is carried out. The expression is,

\[
d_{ij} = \left( \frac{1}{c} \right)^c \left[ \frac{d_{ij}}{d_{ij}^2} \right]^{2}
\]  

(2)

\[
v_i = \frac{\sum_{j=1}^{n} (d_{ij})^c x_j}{\sum_{j=1}^{n} (d_{ij})^c}
\]  

(3)

After \( l \) iterations, if the criterion function meets

\[
J(D^{l-1}, V^{l-1}) - J(D^l, V^l) < \epsilon, \quad (\text{here } \epsilon = 1 \times 10^{-5})
\]

the iteration is stopped, the clustering center of each category and the membership value of each sample to each category can be obtained. Then, according to the maximum principle of membership degree, defuzzification is carried out, and fuzzy classification is transformed into deterministic classification. The regional distribution and the central values of the sound speed profiles in different sea are obtained.

3. Categories and characteristics of sound speed profiles

The categories of sound speed profiles are divided into 6 categories in Northwest Pacific by cluster analysis, as shown in Fig. 1. Fig. 1a, b, c and d represent the classification results of spring (northern hemisphere), summer, autumn and winter respectively, these season is represented by AMJ season, JAS season, OND season and JFM season respectively. In the figure, white represents the area with water depth less than 2000m, and the color scale value represents the categories of sound speed profiles. It can be seen from the figure that each category has a zonal band distribution feature. This is mainly caused by the difference in water temperature and salinity distribution between the surface and subsurface water. The temperature and salinity distribution of each water mass is also zonal distribution. According to the distribution of surface and subsurface water masses, 6 categories are named from north to south: Subarctic category (category 1), subarctic transitional category (category 2), subtropical transitional category (category 3), subtropical category (category 4), tropical transitional category (category 5) and equatorial tropical category (category 6). The subarctic and subropical areas are larger in the western Northwest Pacific and smaller in the east; the subarctic transitional and subtropical transitional areas are the opposite; the seasonal variation of the areas of each category is not obvious.

Subarctic category (category 1): located in the subarctic area, the surface and subsurface are controlled by subarctic water[7], and the latitude is generally higher than 44°N in the Pacific Ocean, including the northern part of the Northwest Pacific, part of the Bering Sea, the Okhotsk sea and the Japan sea. The maximum surface sound speed is about 1488m / s in summer, and the minimum is
about 1456 m/s in winter, as shown in Fig. 2. Figure 2 shows the center values of various categories of sound speed profiles. Figure 2a, b, c and d show the center values of sound speed profiles in spring, summer, autumn and winter respectively. In the figure "*" indicates the depth of sound channel. In winter, the category of sound speed gradient is positive. In other seasons, there is seasonal sonicline. In summer, the seasonal sonicline is the most significant, but the depth of the sonicline is very shallow, no more than 100 m. There is no main sonicline. The overall structure is "seasonal sonicline + deep-sea channel + deep-sea positive gradient". The depth of the sound channel is about 100 m. The area has the largest coverage in winter and the smallest in summer.

Subarctic transitional category (category 2): located in the transitional zone between extratropical and subtropical zone, The latitude is generally between 40°N and 44°N. there are two fronts in this area, the Oyashio front and the subarctic front. The sound speed of surface water also has significant seasonal variation. The surface sound speed is about 10 m/s higher than that of the Subarctic category. The maximum surface sound speed is about 1510 m/s in summer and 1478 m/s in winter, as shown in Figure 2. In winter, the vertical distribution of sound speed presents the structure of "sound layer + negative gradient + deep-sea channel + deep-sea positive gradient". The depth of the sound layer is about 100 m; there is no sound layer in other seasons, the seasonal sonicline is significant in summer, but the depth is very shallow, not more than 100 m. There is no main sonicline in spring, summer and autumn, and the overall structure is "seasonal sonicline + deep-sea channel + deep-sea positive gradient", and the depth of the
sound channel is deeper, about 400 m.

Subtropical transitional category (category 3): located in the north of subtropical area, with latitude from 35°N to 40°N. This area is where the Kuroshio and Kuroshio Extension flow. The maximum surface sound speed is 1525 m/s in summer and 1500 m/s in winter, as shown in Figure 2. In winter, the structure of sound speed gradient is "sound layer + main sonicline + deep-sea channel + deep-sea positive gradient", and the depth of sound layer is about 85 m lower than that in the Subarctic transitional region; in other seasons, there is no sound layer, but there is a main sonicline. In summer, the seasonal sonicline is significant, in spring, summer and autumn, the structure is "seasonal sonicline + main sonicline + deep-sea channel + deep-sea positive gradient", and the depth of sound channel is further increased, the depth is about 700 ~ 800 m.

Subtropical category (category 4): located in the subtropical sea area, the latitude is generally between 25°N and 35°N. The surface water in this area is mainly subtropical water, and the seasonal variation of surface water is further reduced. The maximum surface sound speed is about 1538 m/s in summer, and the minimum is about 1520 m/s in winter, as shown in Figure 2. The category of sound speed gradient in winter is the same as that in subtropical transitional category. The depth of sound layer is about 85 m in winter and 35 m in autumn. There is no sound layer in other seasons. The sound speed gradient in the main sonicline is significantly stronger than that in the north. In summer, the phenomenon of double sonicline appears, and the seasonal sonicline is significant. In spring and summer, the structure is "seasonal sonicline + main sonicline + deep-sea channel + deep-sea positive gradient", and the depth of sound channel is further deepened, the depth is about 1000 m.
Tropical transition category (category 5): mainly located in the tropical area, with latitude from 14°N to 25°N. The surface layer in this area is mainly tropical water, which is similar to the equatorial tropical category. The seasonal sonicline disappears, and the seasonal variation of sound speed profile is not obvious, but there are still differences. The maximum surface sound speed is 1543 m/s in summer and 1536 m/s in winter, as shown in Figure 2. The vertical distribution of sound speed presents a structure of "sound layer + main sonicline + deep-sea channel + deep-sea positive gradient". The depth of sound layer is about 70 m in winter, 50 m in autumn and 25 m in summer. In spring, it presents a structure of "main sonicline + deep-sea channel + deep-sea positive gradient". The depth of the sound channel decreases slightly due to the rise of the main sonicline, about 950 m.

Equatorial tropical category (category 6): located in the equatorial area, the latitude is generally between 0°N and 14°N, the sound speed structure is relatively stable, the seasonal variation is small, the surface sound speed basically has no seasonal variation, the maximum sound speed is 1542.6 m/s in summer and the minimum is 1540.8 m/s in winter. The biggest difference between it and other categories of profiles is that the subsurface is occupied by equatorial water, so the main sonicline is shallow and thin, the depth is less than 500 m [3], the thickness is about 100~300 m, the surface sound layer is about 50 m and the depth of sound channel is about 1000 m.

4. Conclusion
Deep water sound speed profiles are classified by hierarchical clustering method and fuzzy c-means clustering method in Northwest Pacific. The conclusion is that the deep water sound speed profiles in the Northwest Pacific are divided into 6 categories: Subarctic category, subarctic transitional category, subtropical transitional category, subtropical category, tropical transitional category and equatorial tropical category. The results show that the distribution areas of these categories are generally zonal. Except for the tropical transition category and the equatorial tropical category, the other categories have significant seasonal changes. The most obvious feature of each category is that the depth of the sound channel increases with the decrease of latitude.

These categories of sound speed profiles have a significant impact on underwater sound propagation, and the most prominent impact is the distance of convergence zone. Except for the subarctic category, there is no convergence zone, there is convergence zone in the waters with critical depth in the Northwest Pacific. The classification of deep sea sound speed profiles in the Northwest Pacific and the research on the characteristics of convergence zone are of great significance for the use of sonar and underwater warfare.

References
[1] Ocean Atlas Editorial Board. Bohai sea, Yellow Sea, East China Sea Atlas (Hydro Section). Beijing, Ocean Press, 1992. 299-302.
[2] Xie Jun, Hu Junchuan, Da lianglong, et al. Study on automatic clustering of sound speed profile in the ocean. Acta Oceanologica Sinica, 2009, 31(2) :34-39.
[3] Zhang Xu, Zhang Yonggang, Huang Feiling, et al. Mode characteristics of sound speed profile in sea areas of China. Marine Science Bulletin, 2010, 29(1) :17-26.
[4] Wang Hua, et al. Cluster Analysis of Deep Water Sound Speed Profiles in Indian Ocean. IOP Conf. Series: Earth and Environmental Science 310 (2019) 052048.
[5] Mandelberg M D, Frizzell-Makowski L J F. Acoustic Provincing of Ocean Basins. OCEANS 2000 MTS/ IEEE Conference and Exhibition. 2000:105-108.
[6] Mackenzie, K.V. Nine-term Equation for Sound Speed in the Oceans. J Acoust Soc, 1981,70 :807-812.
[7] Li Fengqi, Su Yusong. Ocean water mass analysis. Qingdao, Qingdao Ocean University Press, 2000: 267-282.
[8] Tomczak M. Regional Oceanography: An Introduction. Oxford, Pergamon, 2005.
[9] Chen Shangji, Ma Jirui. Marine data processing analysis method and its application. Beijing, Ocean Press, 1991:181-239.