Sonar performance modelling for clustered water masses

Jonas Halse Rygh[1], Karl Thomas Hjelmervik[2], and Trond Jenserud[1]

[1]Norwegian Defence Research Establishment, Nedre vei 8, 3183 Horten, Norway
[2]University of South-Eastern Norway, Raveien 215, 3184 Borre, Norway

Contact author: J. H Rygh, Norwegian Defence Research Establishment, Nedre vei 8, 3183 Horten, Norway, e-mail: jonas-halse.rygh@ffi.no

Abstract: *Solid information on the present environment and its effect on acoustic wave propagation is key to conducting and planning a successful sonar operation. In recent years, the quality and avaliability of oceanographic model data has increased significantly. However, the vastness of information contained in these models leads to high complexity and possibly very long computation times. We propose a method for organizing the ocean masses in geographically divided areas with similar sound speed profiles by using empirical orthogonal functions and clustering. The expected sonar performance and its uncertainty are estimated by using representative sound speed profiles for each area and compared with an exhaustive run for the entire model area. The method is demonstrated and explored using ocean model data sets from Eggakanten outside Lofoten in Northern Norway.*

Keywords: *Sonar performance modelling, oceanography, clustering, empirical orthogonal functions*

1. INTRODUCTION

The increasing amount and quality of relevant environmental data provides a potential for improved modelling of sonar performance, which may benefit both planning and execution of sonar operations. In particular, recent dynamic ocean models provide high resolution forecasts of the water-mass structure, from which detailed sound speed data can be obtained.

Traditionally, sonar performance modelling in an area have been based on a single sound speed profile from either measured or historical data. In areas of complex oceanography containing frontal zones, this is not sufficient for characterizing the propagation conditions. High resolution sound speed data obtained from ocean models allows us to compute a comprehensive picture of the variability in an area, but may provide too much detail to be useful for a sonar operator; this approach is also computationally intensive.

A better approach, to ease interpretation and reduce computations, may be to divide an area into regions of similar oceanographic properties. We suggest a method where the sound speed profiles of an area are first grouped into regions of similar characteristics, and the sonar performance is then computed from representative profiles within each region. A combination of the method of Empirical Orthogonal Functions (EOF) [1] and clustering techniques [2] are used for categorizing the water masses [3, 4, 5, 6].

An important questions is how well the clustered sound speed profiles can describe the sonar performance in an area; to asses this question a comparison with a full computation using all profiles is done. This approach is demonstrated for an area around Eggakanten outside the Lofoten Islands in Northern Norway, using sound speed data from a high resolution ocean model.

2. METHOD

The proposed method consists of two steps: First, the area is divided into regions of similar sound speed profiles using clustering. Clustering is based on EOF analysis; profiles with similar EOF coefficients are sorted into clusters, and average- and representative sound speed profiles are found for each cluster. Second, sonar performance is computed for each cluster, from a number of representative sound speed profiles.

A difficulty arises when using EOF analysis on sound speed data in variable bathymetry, since EOF analysis requires all profiles to have the same length. A solution to this problem is discussed below.

Clustering EOF analysis allows information compression while retaining as much of the variance in the data set as possible. Here it is used to represent the variability of the vertical structure of the sound speed profiles.

Let the data matrix $\mathbf{G}$ contain a set of sound speed profiles. In our case the sound speed profiles are estimated from the temperature and salinity profiles obtained from the Norkyst800 ocean model. We have

$$\mathbf{G} = \begin{pmatrix} \vec{c}_1 & \vec{c}_2 & \dots & \vec{c}_N \end{pmatrix} \tag{1}$$

The sound speed vectors $\vec{c}_n$ have length N_d, where d refers to the maximum model depth. However, each sound speed vector may have a different number of NaN numbers depending on bathymetry. The EOFs are eigenvectors that may be found by solving the following eigenvalue problem

$$\mathbf{X}\vec{u} = \lambda\vec{u} \tag{2}$$

where $\mathbf{X}$ is the covariance matrix of a *subset* of the data matrix $\mathbf{G}$. The subset matrix consists only of sound speed profiles with a local bottom depth in a neighbourhood of q. The depth q is typically the most frequent depth. The dimensionality of $\mathbf{X}$ is $N_q \times N_q$, where N_q is the number of depth samples. Consequently $\mathbf{X}$ will have N_q eigenvectors and eigenvalues. It is common to only use $K < N_q$ eigenvectors, as most of the variance in the data are represented by the first K eigenvectors, sorted by the magnitude of the eigenvalues.

EOF coefficients representing all sound speed profiles with different bottom depths are then derived from

$$\kappa_{jk} = \vec{c}_j \cdot \vec{u}_k, \quad j \in \{1, .., N\}, k \in \{1, ..K\} \tag{3}$$

where the NaN part of the sound speed vectors $\vec{c}_j$ are removed, resulting in different vector lengths. Without any modification this dot product is ill-defined when the dimensions of $\vec{c}_j$ and $\vec{u}_k$ are unequal. However, as proposed in [7], the coefficients representing sound speed profiles with local bottom depths z different than q can be estimated using the following algorithm

- if $z < q$, clip the eigenvectors to match the length of the shallow profile
- if $z > q$, clip the deeper sound speed profile to match the length of the eigenvectors

The EOF coefficients are then clustered into M clusters using k-means clustering [2] which is a fast clustering algorithm that assumes a Gaussian distribution of each coefficient. One can then get the geographical distribution of the clusters by doing an inverse transformation.

Sonar performance modelling The expected sonar performance for the cluster locations is assessed using area-equivalent detection range (AER) as a performance measure. This quantity is based on the estimated probability of detection for a given sonar depth, and is given by the following expression:

$$r_{\text{AER}}(z) = \sqrt{2 \int_0^R P_D(r', z) r' dr'} \tag{4}$$

This corresponds to the radius of a circle with an area equal to the sonar coverage area. $P_D(r, z)$ is the modelled probability of detection for a target located at range, r, and depth z. The target depth dependence of $r_{\text{AER}}(z)$ may be removed if an *a priori* distribution of the target depth, $f_z(z)$, is known. The depth independent version then becomes:

$$r_{\text{AER}} = \int_0^{z_b} f_z(z) r_{\text{AER}}(z) dz, \tag{5}$$

where z_b is the present bottom depth.

Sonar performance is computed by the acoustic ray trace model Lybin [8], using a set of representative sound speed profiles from each cluster.

An important question is how well the sonar performance of a cluster is captured by using its average sound speed profile for the acoustic modelling. A way of quantifying this is to compare the averaged profile AER with the average and spread of AER calculations for a subset of sound speed profiles in a cluster.

3. OCEANOGRAPHIC DATA

Vestfjorden is an area of complex oceanography dominated by two major currents; the Norwegian Atlantic Current (NAC) and the Norwegian Coastal Current (NCC). The currents are affected by strong tides and variable bathymetry, forming fronts and eddies [9].

The NCC is a low-salinity current. It follows the coast and forms a wedge above the NAC. The NCC splits at the entrance of the fjord; the inner branch contributes to the general circulation in the fjord together with fresh water runoff. In general, there is inflowing water along the east side of the fjord and outflowing water along the Lofoten Islands. During winter and spring the NCC is colder than the NAC; in early summer it becomes warmer.

Sound speed data was obtained from the Norkyst800 Ocean model [10]. Modelled sea surface temperature and currents for the Vestfjord area are shown in Fig. 1. Significant variability, with intricate current fields and fronts are observed, due to the interaction of the different water masses and the highly variable bathymetry.

The data set is gridded uniformly with a resolution of 800 m in the horizontal and with nonuniform gridding in depth, with highest resolution closest to the surface.

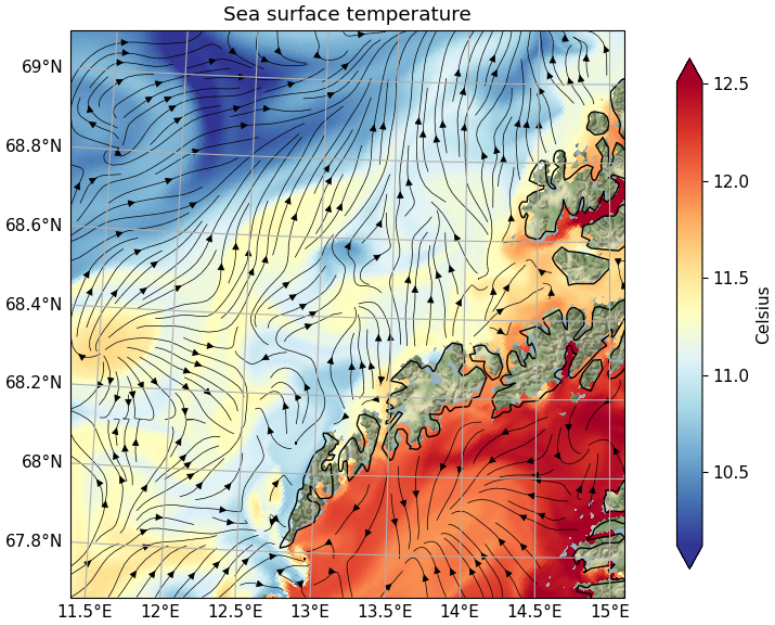

Figure 1: Modelled sea surface temperature and currents at Vestfjorden and Eggakanten, 1st of September, 2022

4. RESULTS

The coefficients for the 5 most dominant EOFs were extracted for all sound speed profiles in the model area using (3), with an explained variance of ≈ 99 %. They were then clustered using k-means clustering for an increasing amount of clusters until the coefficients of the SSPs in the resulting clusters showed a satisfactory Gaussian statistical behaviour. The Gaussian behaviour of the clusters was evaluated using the Kolmogorov-Smirnov (KS) test following the steps of [4]. The lower order EOFs (u_0, u_1, and u_2) exhibit a stronger Gaussian behaviour for a lower number of clusters, than the higher orders. For the following analysis 16 clusters are used, for which the KS-test yields approximately 40% for the more important, lower order EOFs.

Figs. 2 shows the resulting geographical distribution of the clusters and their corresponding representative sound speed profiles. We clearly see evidence of the mixing and churning of disparate ocean masses, which is expected for these waters due to the strong eddies [11] and the mixing due to the interaction between the water from the Atlantic ocean and the coastal

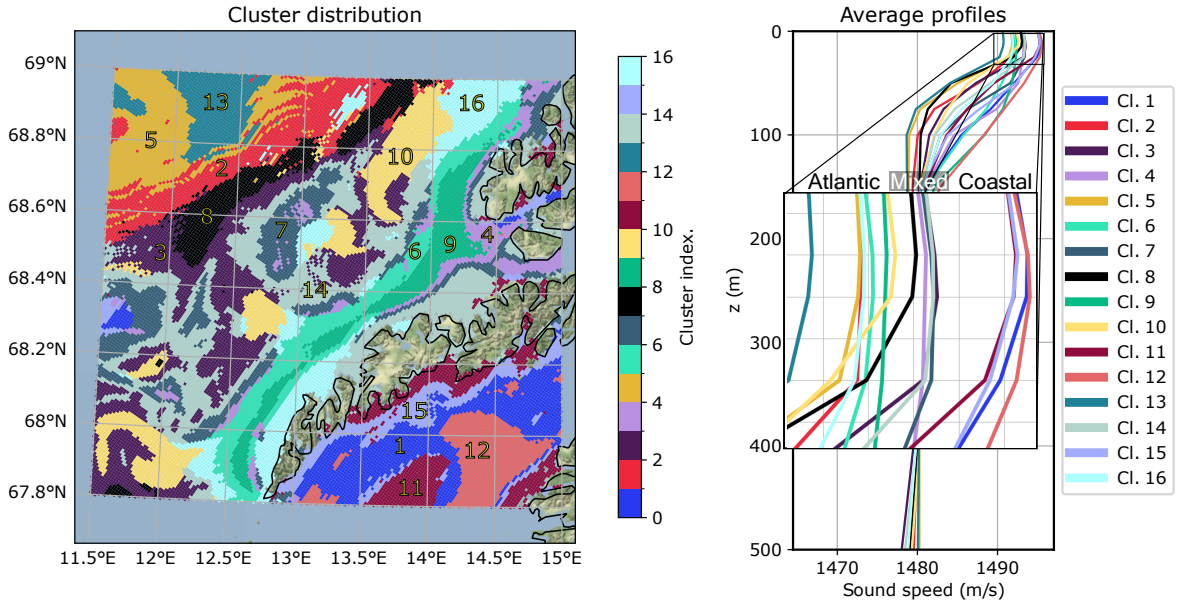

Figure 2: Geographical distribution of sixteen clusters for Vestfjorden and Eggakanten, 1st of September, 2022 (Left). The average sound speed profiles for each cluster, with the zoomed-in frame showing the profiles for the first 30 m (Right).

current [12].

The main variations are at shallow depths (<150 m). The higher sound speeds are found closest to the shore due to the relatively warm NCC and are contrasted by the relatively cold water from the NAC. The different clusters mainly represent different degrees of intermixing of the two water masses, but is also influenced by the topography. Clusters 1, 11, 12, and 15 contain sound speed profiles from the confinement of a fjord, from the NCC. For lower number of clusters, these would typically be grouped into a single cluster. A similar behaviour also applies for cluster 4, 6, 9, 14, and 16, all of which corresponds with the NCC. Clusters 2, 5 and 13 corresponds to the NAC, and the remanining clusters corresponds to the intermixing of NAC and NCC.

Depth-independent AER (5) is estimated for the following cases:

- AER for each profile, denoted by $\mathbf{A}$. Exhaustive computation using model depth and sound speed profile for each grid point. Approximately 25000 model runs.
- AER for subsamples, denoted by $\hat{\mathbf{A}}$. Calculate AER for a subsample of the profiles in the cluster for each model depth, for each cluster. Return mean AER and standard deviation. Approximately 1000 model runs.
- AER from averaged profiles, denoted by $\bar{\mathbf{A}}$. Calculate AER for the average profile for each model depth, for each cluster. Approximately 250 model runs.

In all three cases a flat, homogeneous bottom is used in the acoustic raytrace model, Lybin [13]. This assumpion is questionable in the littorals, but the inclusion of range-dependent bathymetry and sediment types, would push the computational cost past the limit for this study. For the second case, both the mean and variance AER is estimated using a Monte Carlo approach [14].

The results for each case are shown in Fig. 3. The results from case 2 and 3 compares well visually. Case 3, that only requires 250 runs, appears to offer the same fidelity as case 2, which requires four times as many runs. However, the disadvantage of case 3 is that it does not include

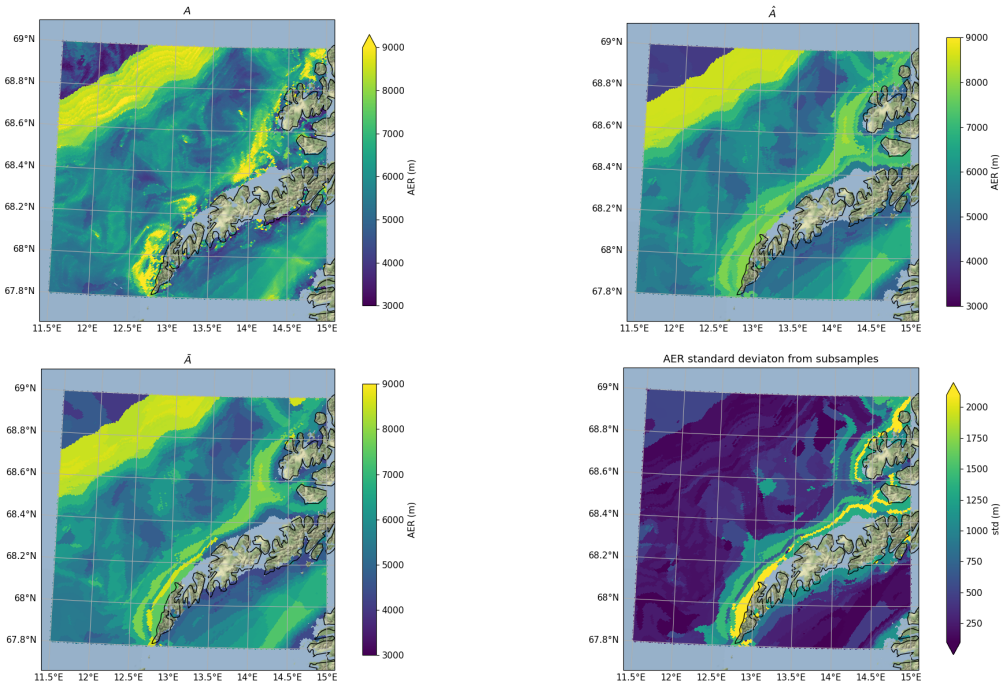

Figure 3: Estimated AER for case 1 (top left), case 2 (top right), and case 3 (bottom left), in addition to the estimated standard deviation for case 2 (bottom right)

an estimate of the standard deviation. The baseline, case 1 clearly offers more detail than the remaining two, particularly close to the coast. However, this region is expected to have large errors in the range estimates due to the approximation of a flat and homogeneous bottom.

To qualitatively assess the results of case 2 and 3, we define the criteria $|\mathbf{A} - \mathbf{A}_i| \leq n\boldsymbol{\sigma}$. The standard deviation matrix $\boldsymbol{\sigma}$ is evaluated at each grid point using the results of case 2, and ranges from 250 m to 2000 m. Within one standard deviation, $n = 1$, we get that both cases fulfills the criteria for approximately 65 % of grid points. For two standard deviations, $n = 2$, the proportions of points fulfilling the criteria are 90 % and 89 % for case 2 and 3 respectively.

This may be an acceptable error considering the significant decrease in computation time. Also, in this scenario it shows that the average profiles from each cluster performs almost as well as using a subset from each cluster. In addition, computing $|\bar{\mathbf{A}} - \hat{\mathbf{A}}| \leq \boldsymbol{\sigma}$, we find that 93 % of AER values calculated from the averaged profiles are within one standard deviation of the subsamples average AER calculations. This indicates that the averaged profiles are mostly representative for the clusters.

5. SUMMARY

The intent of the proposed method is to characterize the present operational environment by a relatively low number of oceanographic regions, sixteen in this case. This allows for a quick computation of the estimated sonar performance for the whole region. The intention is to provide decision makers with a quick overview of the sonar performance in a given area.

The proposed method was tested on ocean model data from an area close to Lofoten in Northern Norway, an area known for its significant oceanographic variability and large eddy currents. The clustering technique applied on the oceanographic data clearly demonstrated this variability, resulting in 16 different classes of watermasses, each representing different mixing

of the NAC and NCC.

The sonar performance was estimated using the fast acoustic raytracer, Lybin, and compared for three different cases; i) an exhaustive run for all sound speed profiles in the area, ii) for a selected few sound speed profiles for each cluster (included also a variance estimate), and iii) for a representative sound speed profile for each cluster. The quickly estimated cases ii) and iii) (order of minutes), resulted in sonar performance estimates within one standard deviation of the slowly estimated case i) for 65% of the investigated locations.

It is important to note that the proposed method does not take into account topographic range-dependency, as this would significantly increase the computational cost. Errors introduced by this approximation are expected to be particularly large in shallow regions, such as close to the coast. The method should therefore be considered used for the open ocean, rather in the confinement of a fjord or in the shallows.

ACKNOWLEDGEMENTS

We would like to thank Mark Prior for coming up with an excellent measure for a simple presentation of the sonar performance for a given environment. The area-equivalent range provides an efficient means for comparing the sonar performance across different environments or sonar settings.

REFERENCES

[1] R. W. PREISENDORFER and C. D. MOBLEY, "Principal component analysis in meteorology and oceanography," *Principal component analysis in meteorology and oceanography*, vol. 17, pp. XVIII–425 p, 1988. Place: Amsterdam Publisher: Elsevier.

[2] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, Sept. 2007. Google-Books-ID: 1aAOdzK3FegC.

[3] J. K. Jensen, K. T. Hjelmervik, and P. Ostenstad, "Finding Acoustically Stable Areas Through Empirical Orthogonal Function (EOF) Classification," *IEEE Journal of Oceanic Engineering*, vol. 37, pp. 103–111, Jan. 2012. Conference Name: IEEE Journal of Oceanic Engineering.

[4] K. T. Hjelmervik and K. Hjelmervik, "Estimating temperature and salinity profiles using empirical orthogonal functions and clustering on historical measurements," *Ocean Dynamics*, vol. 63, pp. 809–821, July 2013.

[5] G. Maze, H. Mercier, R. Fablet, P. Tandeo, M. Lopez Radcenco, P. Lenca, C. Feucher, and C. Le Goff, "Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean," *Progress in Oceanography*, vol. 151, pp. 275–292, Feb. 2017.

[6] F. Sambe and T. Suga, "Unsupervised Clustering of Argo Temperature and Salinity Profiles in the Mid-Latitude Northwest Pacific Ocean and Revealed Influence of the Kuroshio Extension Variability on the Vertical Structure Distribution," *Journal of*

Geophysical Research: Oceans, vol. 127, no. 3, p. e2021JC018138, 2022. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2021JC018138.

[7] K. B. Hjelmervik and K. T. Hjelmervik, "Detection of Oceanographic Fronts on Variable Water Depths Using Empirical Orthogonal Functions," *IEEE Journal of Oceanic Engineering*, vol. 45, pp. 915–926, July 2020. Conference Name: IEEE Journal of Oceanic Engineering.

[8] E. Dombestein and T. Jenserud, "Improving Underwater Surveillance: LYBIN Sonar performance Prediction," in *MAST Americas 2010*, (Washington DC, USA), 2010.

[9] G. Mitchelson-Jacob and S. Sundby, "Eddies of vestfjorden, norway," *Continental Shelf Research*, vol. 21, no. 16, pp. 1901–1918, 2001.

[10] "Oceanographic data from the norwegian meteorological institute." `https://thredds.met.no/thredds/catalog.html`. Accessed: 2023-03-06.

[11] D. L. Volkov, A. A. Kubryakov, and R. Lumpkin, "Formation and variability of the lofoten basin vortex in a high-resolution ocean model," *Deep Sea Research Part I: Oceanographic Research Papers*, vol. 105, pp. 142–157, 2015.

[12] R. Sætre, *The Norwegian coastal current: oceanography and climate*. Akademika Pub, 2007.

[13] E. Dombestein, S. Mjølsnes, and F. Hermansen, "Visualization of sonar performance within environmental information," in *2013 MTS/IEEE OCEANS-Bergen*, pp. 1–6, IEEE, 2013.

[14] E. M. Bøhler, P. Østenstad, and K. T. Hjelmervik, "A Monte Carlo approach for capturing the uncertainty in sonar performance modelling," p. 8.