A pattern recognition approach to identify biological clusters acquired by acoustic multi-beam in Kongsfjorden

Giovanni Giacalone[a], Marco Barra[b], Angelo Bonanno[a], Gualtiero Basilone[a], Ignazio Fontana[a], Monica Calabrò[a], Simona Genovese[a], Rosalia Ferreri[a], Giuseppa Buscaino[a], Salvatore Mazzola[a], Riko Noormets[d], Christopher Nuth[e], Giosuè Lo Bosco[c,1], Riccardo Rizzo[f,1] and Salvatore Aronica[a,*,1]

[a]IAS-CNR, National Research Council, Rome, Italy

[b]ISMAR-CNR, National Research Council, Rome, Italy.

[c]DMI, University of Palermo, Palermo, Italy

[d]UNIS,Department of Arctic GeologyThe University Centre in Svalbard, Longyearbyen, Norway

[e]Department of Geosciences University of Oslo, Oslo, Norway

[f]ICAR-CNR - National Research Council, Rome, Italy

ARTICLE INFO

ABSTRACT

The Svalbard archipelago is one of the most intensively studied marine regions in the Artic area; here the composition and distribution of marine assemblages are changing under the effect of global change, and marine communities are monitored in order to understand the long-term effects on marine biodiversity. In the present work, acoustic data collected in the Kongsfjorden (Svalbard) using multi-beam technology was analyzed to develop a methodology for identifying and classifying 3D acoustic patterns related to fish aggregations (schools). For each identified aggregations, different features were taken into account thus allowing to develop a multi-variate classification procedure considering morphological, energetic and bathymetric characteristics associated to such aggregations. The results obtained from clustering suggest that from a mathematical point of view three distinct groups could be identified. The proposed approach, that allows to discriminate the acoustic patterns identified in the water column, seems promising for improving the monitoring programs of the marine resources, also in view of the ongoing climate changes.

1. Introduction

Acoustic methods, employing single-beam echo-sounder technology in the last decades represented the most effective tool in monitoring the abundance and spatial distribution of marine organisms inhabiting the water column [57]. Recently, the multi-beam technology, historically adopted in morpho-bathymetric surveys, was adapted to identify targets in the water column, thus allowing fishery scientists to obtain a more detailed view of fish aggregations [35],

*Corresponding author

✉ giovanni.giacalone@cnr.it (G. Giacalone); marco.barra@cnr.it (M. Barra); angelo.bonanno@cnr.it (A. Bonanno); gualtiero.basilone@cnr.it (G. Basilone); ignazio-fontana@cnr.it (I. Fontana); monica.calabro@ias.cnr.it (M. Calabrò); simona.genovese@cnr.it (S. Genovese); rosalia.ferreri@cnr.it (R. Ferreri); giuseppa.buscaino@cnr.it (G. Buscaino); salvatore.mazzola@cnr.it (S. Mazzola); riko.noormets@unis.no (R. Noormets); christopher.nuth@geo.uio.no (C. Nuth); giosue.lobosco@unipa.it (G. Lo Bosco); riccardo.rizzo@cnr.it (R. Rizzo); salvatore.aronica@cnr.it (S. Aronica)

👤 (G. Giacalone); (M. Barra); (A. Bonanno); (G. Basilone); (I. Fontana); (M. Calabrò); (S. Genovese); (R. Ferreri); (G. Buscaino); (S. Mazzola); (R. Noormets); (C. Nuth); (G. Lo Bosco); (R. Rizzo); (S. Aronica)

ORCID(s): 0000-0001-8172-4710 (G. Giacalone); 0000-0002-8621-504X (M. Barra); 0000-0002-3868-751X (A. Bonanno); 0000-0002-5732-5055 (G. Basilone); 0000-0002-0527-9827 (I. Fontana); 0000-0002-9972-0440 (M. Calabrò); 0000-0001-9570-178X (S. Genovese); 0000-0001-7339-4942 (R. Ferreri); 0000-0002-4848-4185 (G. Buscaino); 0000-0002-4622-8129 (S. Mazzola); 0000-0002-2832-386X (R. Noormets); 0000-0002-1063-2832 (C. Nuth); 0000-0002-1602-0693 (G. Lo Bosco); 0000-0001-5007-6925 (R. Rizzo); 0000-0003-3489-1473 (S. Aronica)

[1]Equal contribution

[41]. In this framework, a key aspect to promptly develop specific sustainable management strategies is the design of efficient monitoring programs able to provide early warnings highlighting changes in marine community. Monitoring programs are even more important also under the climate-change scenario. In particular, Artic areas are experiencing the most evident climate-changes related modifications; for example, in some arctic regions, such as Greenland, in 2019 the highest ice mass loss was observed, reaching in July 2019 about 5 million tons per minute [30], [31], [56]. The marine ecosystem strongly depends on sea ice, which serves as a platform for marine life and plays a crucial role in controlling the physical and chemical processes that affect biological activity [62], [42]. The invasion of alien species in the areas surrounding the Arctic is also an expanding phenomenon negatively impacting on endemic species through complex trophic interactions. The case of Atlantic cod (*Gadus morhua*) competing with the Arctic cod (Polar cod, *Boreogadus saida*) [23] is one such example. In the present work, multi-beam data collected in the framework of the CalvingSEIS project (Glacier dynamic ice loss quantified through seismic eyes) were analyzed. Under the umbrella of CalvingSEIS project, a group of researchers belonging to different European institutions (University of Oslo, Norway; National Research Council, Italy; University in Kiel, Germany; GAMMA Remote Sensing Ag, Switzerland; NORSAR, Norway) evaluated the influence of glaciers melting dynamics on terrestrial and marine Arctic ecosystems. For the first time, a multi-beam scientific echo-sounder was used for a bathymet-

ric survey in Kongsfjorden (Norway), allowing to consistently evaluate the morphological and energetic features of fish aggregations inhabiting the fjord. The collected acoustic data was analyzed to develop a methodology for identifying and classifying 3D acoustic patterns related to fish aggregations (schools).

2. Materials and Methods

2.1. Study area

The Svalbard archipelago is located between the parallels 74° N and 81° N, is heavily influenced by Atlantic circulation [33], and extends over an area of 63,000 km^2. Here the ice cover 57% of land with a mix of land and sea glaciers and ice caps [46]. The Kongsfjorden is an area of highly scientific interest due to its biodiversity [33]. The innermost part of the fjord is characterized by one of the largest glaciers of Svalbard, called Kronebreen. Its front ends directly on the sea, making an interface from which icebergs are created in the Kongsfjord. The environmental complexity results in highly diverse ecosystem characterized by the association of Atlantic and Arctic species [7] (Fig. 1)

Figure 1: The study area of the Kongsfjorden in the context of Svalbard. Source: http://cruise-handbook.npolar.no/site-images-new/kongsfjorden/Kongsfjorden_2010_3-5-10.jpg

Due to the ice presence, most of the studies carried out in the Kongsfjorden were conducted during the summer [33], [32], [60], when this area is accessible for navigation, although different life processes also take place during polar winter [24]. In the area, besides several species of marine mammals, three species belonging to Gadidae family dominate the water column. One that has been found to have the greatest distribution is the Polar cod (*Boreogadus saida*) [33]. The other two species are the Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) [53], that have been favored by the increase in the inflow of warmer water from the Atlantic to the western coast of Svalbard. The Polar cod is a brackish water bony fish inhabiting the water column [37], [59], [14]; it is a circum-arctic species common in open and ice-covered waters of the seas of the Arctic shelf [49], [51], [15] that performs offshore-inshore migrations associated with spawning and ice movement [43]. The reproduction period is between November and March [32], [51]. The Atlantic cod, widely present in the area, is a ma-

rine bony demersal fish as it lives and feeds on or near the bottom, but can assume a semi-pelagic lifestyle in specific life stages. Atlantic cod reproduces between December and June [14], [21]. This species is widely distributed in a variety of habitats. Juveniles prefer shallow sub-littoral waters (less than 10-30 m deep) with complex habitats, areas with gravel, rocks or boulders, which provide protection from predators [10]. Adults are usually found in deeper and colder waters. During the day, they form schools that move about 30-80 m from the bottom, while during the night they disperse in the water column to feed. This is an eurythermal and euryhaline species tolerating wide temperatures and salinity ranges [18]. Finally, the haddock is a saltwater demersal fish that is generally found between 10 - 200 m [45]. Adults are most commonly found between 80 and 200 m depth; females predominate in shallow water and males in offshore areas. The temperature range in which they can be found goes from 4 to 10 °C. Among these three species, typically arctic and sub-arctic, the one characterized by a wider distribution is the Polar cod (*Boreogadus saida*) [33]. The Atlantic cod prevails in the shallow waters [7]. Hence, Atlantic species have gradually established themselves in Arctic regions with a significant impact on ecosystem structure and ecological functioning [54]. Although there appears to be an overlap in the distribution between Atlantic cod and haddock, the former, tolerating lower temperatures, typically expands further north than haddock [16], [4]. There seems to be no trace of the haddock both in the innermost part of the fjord [24] and in the proximity of the Kronebreen glacier [8].

2.2. Multi-beam data processing

Marine species characterized by aggregative behaviour and inhabiting the water column are generally monitored by using acoustic techniques [22], [6], [5]. These techniques allow, with a non-invasive methodology, to investigate large portions of the sea in a relatively short time. Even if the main aim of acoustic survey is the evaluation of the abundance and spatial distribution of target species, acoustic methods allow also to identify the morphological, energetic [25], and bathymetric characteristics of these patterns [1]. (Fig. 2). Multi-beam data were acquired in the period 23th to 26th August 2016 on board the R/V "Viking Explorer". The data were acquired using a Kongsberg-Maritime EM2040 MBES characterized by an operating frequency of 300 kHz and pole-mounted on one side of the boat. The sampling design focused on the interface with the glacial front inside the fjord and the path that goes from the base of NY Alesund up to it (Fig. 1). All the acquired data were stored in binary format on dedicated server. Subsequently, the collected files were converted in a format suitable for analysis and ad-hoc algorithms were applied to detect fish schools [25]. For each identified fish school, the morphological and energetic characteristics were extracted (Table 1). In addition, in order to take into account for the position of the schools in the water column, the school depth was also considered (Table 1).
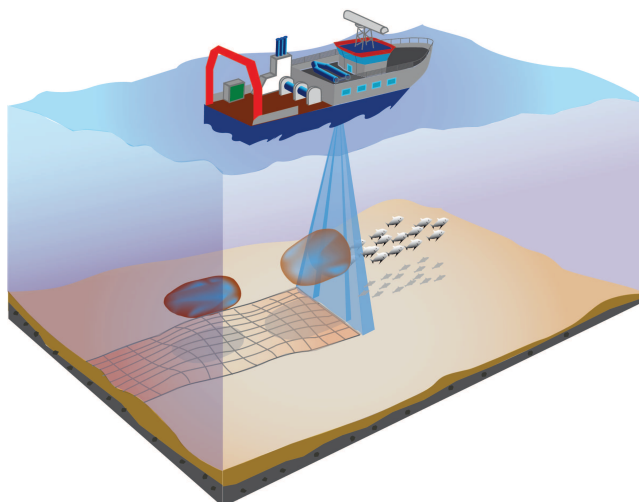
Figure 2: Diagram showing multi-beam instrument acquisition.

Figure 3: Frequency distribution of the 13 characteristic parameters.

CHARACTERISTICS OF THE EXTRACTED PATTERNS			
Energetic	Sv Mean	**Morphological**	Height
	Std Sv		Width
	Sv Max		Length
	Sv Min		Volume
Bathymetric	Depth mean		Surface area
	Depth max		Roughness
	Depth min		

Table 1
Energetic, bathymetric and morphological characteristics extracted.

Figure 4: Histograms of variables following logarithmic transformation.
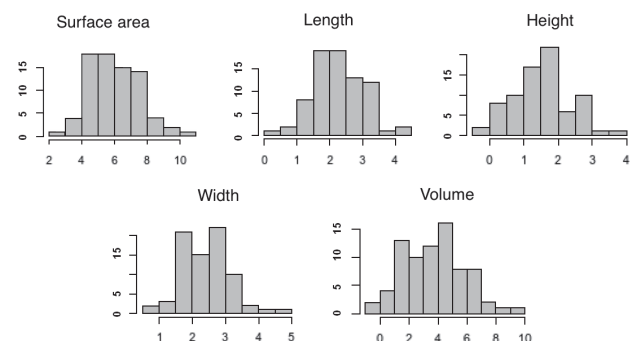
2.3. Exploratory data analysis

Clustering methods are known to be influenced by the presence of outliers, and acoustic data are often characterized by the presence of skewed distribution and extreme values. Thus, in order to avoid possible biases, the frequency distribution of considered variables and the skewness index [26] were computed. The obtained histograms evidenced in some cases the presence of skewed distribution (see Fig. 3). The combined analysis of the frequency histograms and the asymmetry index allowed to identify the most problematic variables (Table 2; the variables subjected to log transformation are highlighted in bold). To reduce the influence of the extreme values, linked to long-tailed frequency distributions, a log transformation was performed. The frequency distribution of log-transformed variables are shown in Fig. 4. Finally, in order to avoid scale-related effects during the clustering procedure, variables were standardized.

2.4. Clustering method

In this paper the k-means algorithm was used to identify different kind of patterns along the water column. The k-means algorithm requires as input the number of clusters (k) to be identified. Thus, as the number of clusters was not known a priori, a preliminary analysis was performed to evaluate the partitioning of observations among the different clusters, considering a number of clusters between 2 and 10.

For each k value, 10 iterations were performed, taking into consideration the relative average values obtained. In order to further evaluate how many clusters to take into consideration to determine an optimal solution, the internal validity indices were applied. Only the indices suitable for k-means (Table 3) were considered, using the NBClust package [11] in R statistical environment program [50]. Once the number of clusters was identified, boxplots were used to evidence for each variable the differences among the identified groups.
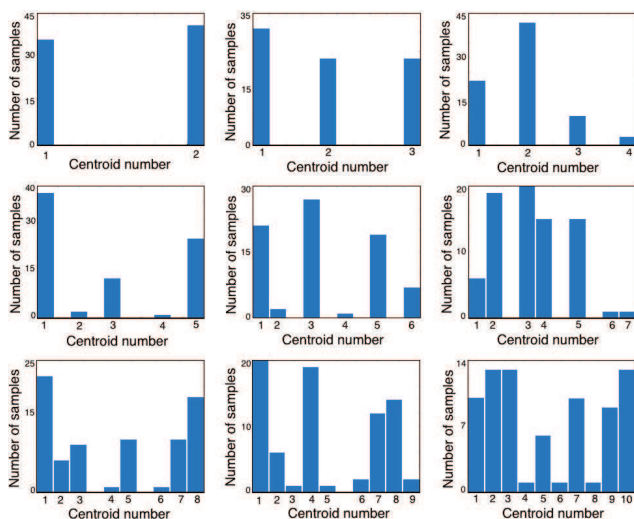
3. Results

The exploratory analysis suggested that the optimal number of clusters is 2 or 3 (Fig. 5). In particular, for both k = 2 and k = 3 the observations were almost equally distributed among the clusters. On the contrary, increasing the number of clusters, some of the clusters were consistently characterized by a low number of observations. Validation indices provided an equal probability for 2, 3 and 4 clusters, as shown in Table 3. In order to evaluate the differences among the groups, the box-plots for each of the variables obtained with k between 2 and 4 were examined. From the analysis of the box plots obtained with $k = 4$ in Fig. 6, it is not possible to highlight a clear separation between the variables in the 4 clusters. For k = 4, a pairwise similarity among the vari-

	min	Q2.5	Q25	Median	Q75	Q97.5	max	mean	Sk
Surface area	15.20	33.57	120.67	336.67	1240.61	11047.24	38150.85	1622.19	**6.17**
Length	1.59	2.53	5.76	9.37	16.22	45.92	83.85	13.28	**3.52**
Width	1.91	2.85	6.56	10.50	18.87	41.77	90.26	14.47	**3.19**
Depth min	3.43	4.17	11.15	17.15	66.32	105.03	137.20	37.27	0.90
Depth max	4.91	6.79	15.64	24.28	72.55	112.94	144.29	43.46	0.88
Height	0.81	1.01	2.68	4.55	7.09	20.08	36.44	6.20	**2.47**
Volume	0.60	1.19	8.10	39.83	180.62	2764.08	9227.65	338.57	**6.32**
Depth mean	4.07	5.48	13.45	21.60	70.03	108.34	140.67	40.39	0.88
Roughness	3.02	3.35	4.86	12.12	17.01	28.89	32.28	12.44	0.59
Sv mean	-65.91	-64.29	-41.10	-33.26	-25.68	-17.06	-3.99	-34.79	-0.52
Std Sv	-66.02	-64.53	-34.64	-26.66	-18.26	-8.44	13.60	-28.18	-0.54
Sv max	-59.27	-56.08	-23.94	-15.56	-6.14	7.68	33.44	-16.25	-0.41
Sv min	-117.12	-114.92	-97.08	-83.92	-75.33	-67.12	-65.56	-86.34	-0.46

Table 2
Basic statistical parameters used to pre-process the data. The bold rows indicates the parameters that strongly deviates from the simmetry.

Figure 5: Bar-plots obtained from a clustering with k values from 2 to 10.

Index Name	Number of clusters
"ch" [9]	2
"gap" [61]	2
"mcclain" [38]	2
"silhouette" [55]	2
"cindex" [34]	3
"db" [17]	3
"ratkowsky" [52]	3
"ball" [2]	3
"duda" [19]	4
"beale" [3]	4
"ptbiserial" [40][39]	4
"sdindex" [28]	4
"kl" [36]	7
"hartigan" [29]	8
"sdbw" [27]	10

Table 3
Indexes used to evaluate the clusters number [12], with the corresponding condition and the number of cluster obtained from the condition.

ous clusters for the morphological characteristics was identified, while clear pattern were not identified for the energetic and bathymetric characteristics. Therefore, taking into account these considerations and the results obtained by the exploratory analysis (Fig. 5), which highlighted the presence of two or three main clusters, the existence of 4 clusters may likely be excluded. In the case of k = 3, a better separation emerged (Fig. 7) even if, it can be noted that in terms of morphological variables, classes 2 and 3 show some similarities, while only a slight separation is evident for the same classes in the energetic characteristics. In the case with k = 2 (Fig. 8), there is a clear separation between the classes in all morphological variables considered and in the minimum of the Sv variable. Then, among the possible solutions obtained, the ones that better represent the dataset are the division into two or three distinct clusters. To choose between these two options, it was decided to check the labeling of the patterns for both clustering cases (k = 2 and k = 3). In case of k-means algorithm application with k =

3, it emerges that observations belonging to clusters 2 and 3 (respectively blue and light-blue in Fig. 7) are grouped into a single cluster (cluster 1 in red color in Fig. 8) when k-means is applied with k = 2, with the exception of only 4 observations out of a total of 77. The remaining cluster 1, obtained with k-means for k = 3, groups the same observations as cluster 2 obtained with k-means for k = 2; this cluster also contains the four previously excluded observations. A more in-depth view of the results allows us to state that three is the number of clusters that better represents the data set from a numerical point of view. To further investigate the relationships among the identified clusters, the spatial pattern was explored (Fig. 9). It is important to highlight that the positions (in terms of longitude and latitude) of the analyzed fish schools were not used as input variable for the clustering algorithm. From visual observation, we can detect a clear spatial separation of the three clusters. In detail, the class identified with the blue color is located in the area
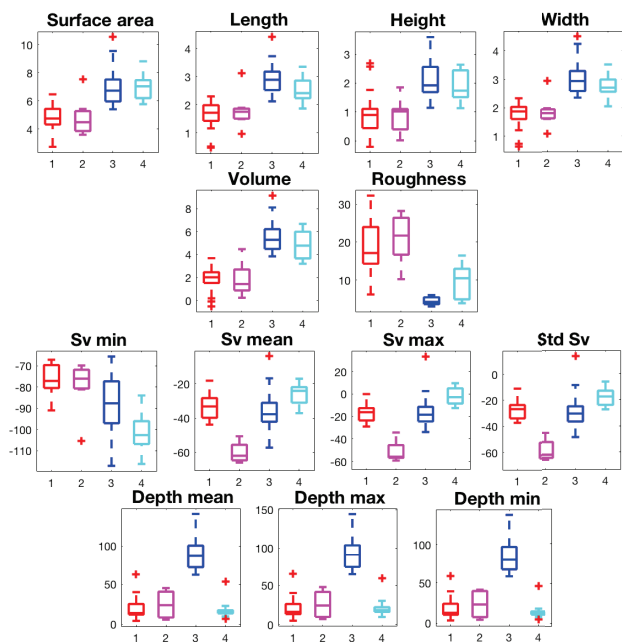
Figure 6: Box-plots obtained from a clustering by k-means equal to 4.

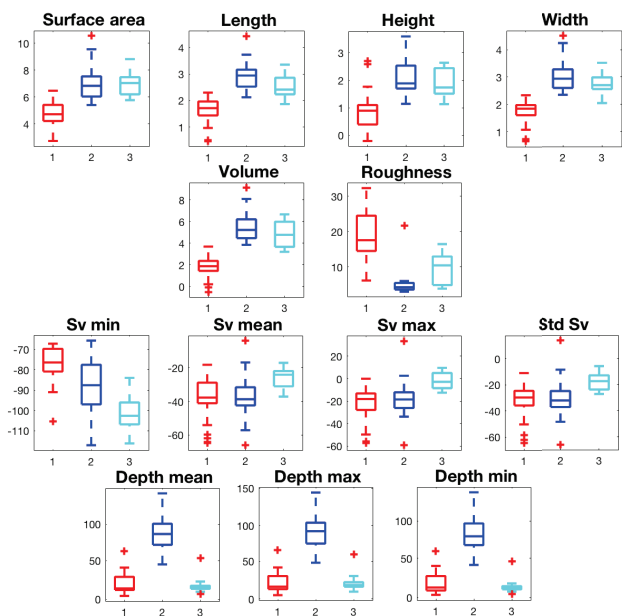**Figure 8:** Box-plot obtained from a clustering by k-means equal to 2.

Figure 7: Box-plot obtained from a clustering by k-means equal to 3.

farthest from the glacier and, considering the average depth, such fish schools are positioned at the greatest depths. On the contrary, the fish schools belonging to the class identified with the light blue color are mainly found in the central part of the study area and are characterized by a lower depth than the other clusters (about 30 m deep). Finally, the red class includes all fish schools that are positioned in the area closest to the glacier, also occupying shallow depths. However, the morphological characteristics analyzed show strong similarities between th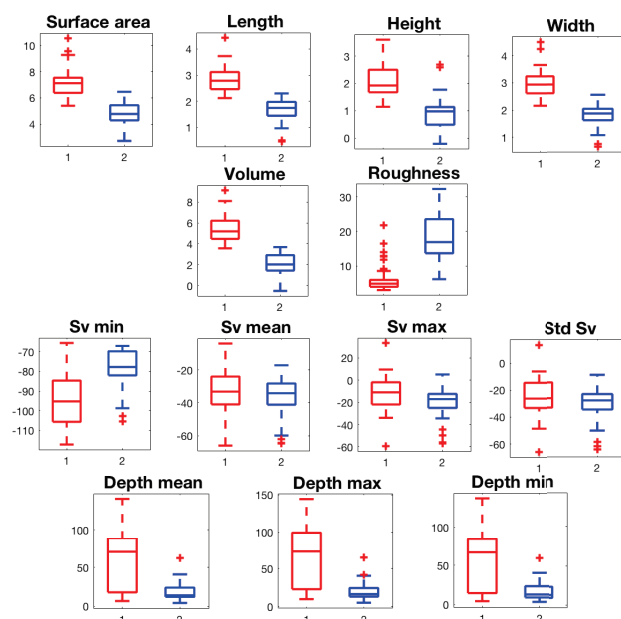e fish schools labeled with the colors blue and light blue respectively (clusters 2 and 3 in Fig. 7), while the morphological characteristics differ for the red fish schools (cluster 1 in Fig. 7). In the same figure, a slight difference can be seen between the energy characteristics of clusters 2 and 3 while greater similarities are found in some of the energy characteristics belonging to clusters 1 and 3.

4. Discussion

To the best of our knowledge, this work represents the first attempt to use the multi-beam data acquired in the water column in the Svalbard Islands to evaluate the presence of different fish species. For this area there is no previous knowledge relating to the acoustic patterns identified. The use of the morphological, energetic and bathymetric characteristics of the aggregations, along with clustering results, allowed to hypothesize the presence of three distinct clusters of fish schools. Two of them are quite different in terms of morphological and energetic variables. The presence of a third cluster, which differs from the other two only for some characteristics, suggests that this may refer to one of the species associated with one of the other two clusters, and that therefore it is the same species. Specifically, it is possible to hypothesize that in the three-cluster classification (Fig. 7), cluster 2 and cluster 3 correspond to the same species, showing the same morphological features and differing in terms of bathymetric and energetic ones (likely correlated to the specimens size). In Fig. 7 it is possible to see that cluster 3 prefers lower depths than cluster 2. This result, associated with the high variability of the same cluster in the energy characteristics and the period of the survey, makes us hypothesize that the species is present in a juvenile stage of
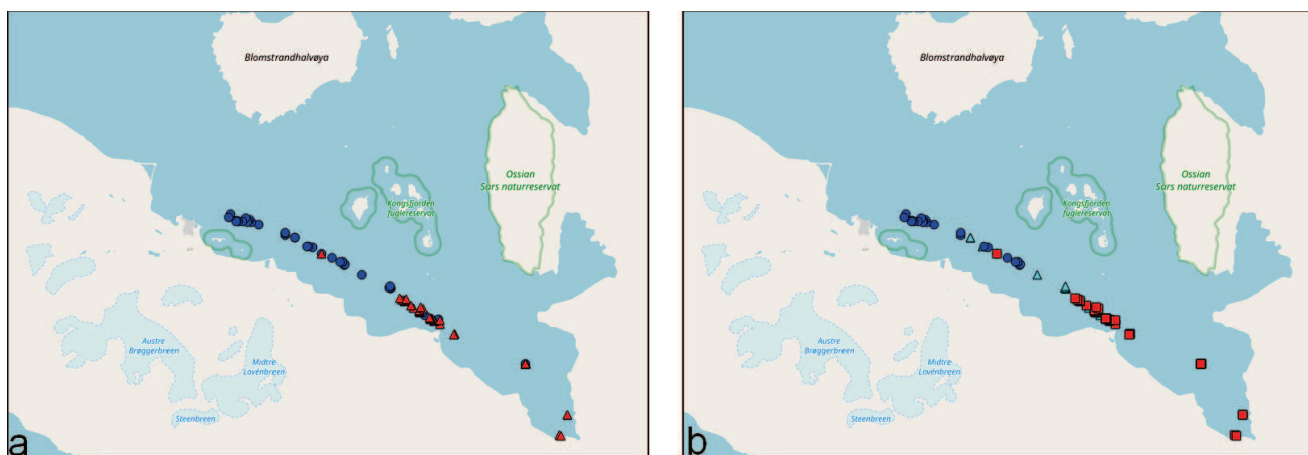
Figure 9: Spatial distribution of the clusters: figure a) shows the position of the patterns divided in k=2 clusters, figure b) shows the position of the patterns for k=3 clusters.

life. This can be justified by the fact that some fish species show different behaviors in relation to the life stage. From the analysis of the box plots, in cluster 3 a greater range between the values of $S_{v\ max}$ and $S_{v\ min}$ is visible more than in the other clusters. This characteristic is in agreement with literature studies that observed that juvenile individuals form very compact schools in the central part and more dispersed in the peripheral/external part [44]. Therefore, the similarities of the morphological characteristics between clusters 2 and 3 lead to the hypothesis that the species identified by both is only one, the Atlantic cod. In fact, it is known from literature that the biological behavior of this species is different in its main life stages, as juveniles and as adults. Furthermore, this species appears to be among the most present species [24] in the central area of the fjord as a species capable of tolerating a high range of variability in both salinity and temperature [18]. From the study of the biological characteristics and behavior of this species it can be highlighted how the morphological and energetic characteristics as well as the bathymetric preferences support this hypothesis. This species reproduces in a period between December and June [14], [21], this would justify the presence of juveniles in our sampling period (August). Fisher et al. (2017) [24] stated that in the period of August 2014 a great abundance of cod was recorded in shallow waters with an average standard length between 40 and 80 mm, while adult individuals up to 300 mm have been observed sporadically in shallow waters. These individuals, once having reached an average length of 100-125 mm, appear to abandon shallow water habitats by migrating to deeper habitats or offshore [24]. Our results is in agreement with literature data as regards the greater presence of Atlantic cod in the area compared to the others [33], [53], [7]. Furthermore, to corroborate this it must be considered that this species, even if tolerant to low temperatures, is not able to withstand the low salinity values present in the proximity of the glacier. In fact, it can easily be assumed that cluster 1, located in the innermost part of the fjord, is the other species, endemic to the area, and namely the Polar cod. Among the species considered so far, the latter appears

to be the only one able to tolerate low salinity values typical of a marine area in which there is a strong dilution of fresh water coming from the melting of the Kronebreen glacier. The depth preference is also in line with what reported in literature. Marine fish population can react in different ways to environmental changes; unfavorable environmental conditions could impact on the population dynamics leading to changes in abundance [58] or in the depth and/or geographical distribution [13], [47], [20]. The intensity and extent of such responses depends on a number of factors such as the thermal tolerance, the migratory capacity and the ability to adapt to different environmental conditions. Nonetheless, the adaptation rate is population and species-specific [30], [48]. Furthermore, new species (often referred as alien species) entering in a specific ecosystem could lead to deep modification in community structure, even causing local extinctions. The case of Atlantic cod (*Gadus morhua*) competing with the Polar cod (*Boreogadus saida*) is one such example. The increasing abundance of Atlantic cod, observed in the Kongsfjorden area, is considered as one of the effects of the ongoing environmental changes.

5. Conclusions

The results of this work show that the dataset is composed, from a numerical point of view, of three clusters while, from a biological point of view, the species associated with the identified patterns are likely two, and one of them with different life stages and different energetic characteristics. In the light of the results obtained, we can sustain that it is possible, starting from raw acoustic data, to distinguish fish schools in relation to some morphological, energetic and bathymetric positional parameters of the aggregations, in particular study areas such as the one analyzed in this context. The results obtained are in agreement with other studies based on biological data published for the same study area. Therefore, this method, based on the analysis of acoustic data, has proved to be very useful for the recognition of species and can be used to support biological research on target species.

The main advantage lies in the fact that the analyzed characteristics may allow to obtain a classification of the species on the basis of their biological behavior, without using invasive sampling methods, thus reducing the fishing effort. This work opens up different scenarios on the possibility of using this data for identifying fish species and therefore represents a starting point for possible future works in this area and in other restricted areas.

Acknowledgement

References

[1] Aronica, S., Fontana, I., Giacalone, G., Lo Bosco, G., Rizzo, R., Mazzola, S., Basilone, G., Ferreri, R., Genovese, S., Barra, M., Bonanno, A., 2019. Identifying small pelagic mediterranean fish schools from acoustic and environmental data using optimized artificial neural networks. Ecological Informatics 50, 149–161.

[2] Ball, G., Hall, D., 1965. ISODATA: A novel method of data analysis and pattern classification. Technical Report. Stanford Research Institute.

[3] Beale, E., 1969. Euclidean Cluster Analysis. Scientific Control Systems Limited.

[4] Bergstad, O.A., Jørgensen, T., Dragesund, O., 1987. Life history and ecology of the gadoid resources of the barents sea. Fish Res. 5, 119–161.

[5] Bonanno, A., Goncharov, S., Mazzola, S., Popov, S., Cuttitta, A., Patti, B., Basilone, G., Di Nieri, A., Patti, C., Aronica, S., Buscaino, S., 2006. Acoustic evaluation of anchovy larvae distribution in relation to oceanography in the cape passero area (strait of sicily). Chemistry and Ecology 22, S265–S273.

[6] Bonanno, A., Zgozi, S., Basilone, G., Hamza, M., Barra, M., Genovese, S., Rumolo, P., Nfate, A., Elsger, M., Goncharov, S., Popov, S., Mifsud, R., Bahri, T., Giacalone, G., Fontana, I., Buongiorno Nardelli, B., Aronica, S., Ceriola, L., Patti, B., Ferreri, R., Colella, S., G., V., Mazzola, S., 2015. Acoustically detected pelagic fish community in relation to environmental conditions observed in the central mediterranean sea: a comparison of libyan and sicilian–maltese coastal areas. Hydrobiologia 755, 209–224.

[7] Brand, M., Fischer, P., 2016. Species composition and abundance of the shallow water fish community of kongsfjorden, svalbard. Polar Biology 39, 2155–2167.

[8] Buscaino, G., Picciulin, M., Canale, D.E., Papale, E., Ceraulo, M., Grammauta, R., Mazzola, S., 2020. Spatio-temporal distribution and acoustic characterization of haddock (*Melanogrammus aeglefinus*, gadidae) calls in the arctic fjord kongsfjorden (svalbard islands). Sci. Rep. 10.

[9] Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Communications in Statistics 3, 1–27.

[10] Chapman, L.C.J., Hawkins, A.D., 1973. A field study of hearing in the cod, *Gadus morhua*. J. comp. Physiol. 85, 147–167.

[11] Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., 2014a. Nbclust: An r package for determining the relevant number of clusters in a data set. Joutnal of Statistical Software 61, 1–36.

[12] Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., 2014b. Nbclust: an r package for determining the relevant number of clusters in a data set. Journal of statistical software 61, 1–36.

[13] Cheung, W.W.L., Watson, R., Pauly, D., 2013. Signature of ocean warming in global fisheries catch. Nature 497, 365–367.

[14] Cohen, D.M., Inada, T., Iwamoto, T., Scialabba, N., 1990. Fao species catalogue. vol. 10. gadiform fishes of the world (order gadiformes). an annotated and illustrated catalogue of cods, hakes, grenadiers and other gadiform fishes known to date. FAO Fish. Synop 125, 442.

[15] Craig, P.C., Griffiths, W.B., Haldorson, L., McElderry, H., 1982. Ecological studies of arctic cod (boreogadus saida) in beaufort sea coastal waters. Canadian Journal of Fisheries and Aquatic Sciences. 39, 395–406.

[16] Dalpadado, P., Bogstad, B., Eriksen, E., Rey, L., 2009. Distribution and diet of 0-group cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) in the barents sea in relation to food availability and temperature. Polar Biol. 32, 1583–1596.

[17] Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 1, 224–227.

[18] Drinkwater, K.F., 2005. The response of atlantic cod (gadus morhua) to future climate change. ICES Journal of Marine Science 62, 1327–1337.

[19] Duda, R.O., Hart, P.E., 1973. Pattern Classification and Scene Analysis. John Willey and Sons.

[20] Dulvy, N.K., Rogers, S.I., Jennings, S., Stelzenmüller, V., Dye, S.R., Skjoldal, H.R., 2008. Climate change and deepening of the north sea fish assemblage: A biotic indicator of warming seas. J. Appl. Ecol. 45, 1029–1039.

[21] FAO 2004-2021, 2021. Cultured Aquatic Species Information Programme. Gadus morhua. URL: http://www.fao.org/fishery/culturedspecies/Gadus_morhua/en.

[22] Fernandes, P.G., Gerlotto, F., Holliday, D.V., Nakken, O., Simmonds, E.J., 2002. Measuring skewness and kurtosis. Acoustic applications in fisheries science - the ICES contribution - ICES Marine Science Symposia 215, 483–492.

[23] Fey, D.P., Węsławski, J.M., 2017. Age, growth rate, and otolith growth of polar cod (boreogadus saida) in two fjords of svalbard, kongsfjorden and rijpfjorden. Oceanologia 59, 576–584.

[24] Fischer, P., Schwanitz, M., Loth, R., Posner, U., Brand, M., Schröder, F., 2017. First year of practical experiences of the new arctic awipevcosyna cabled underwater observatory in kongsfjorden, spitsbergen. Ocean Sci. 13, 259–272.

[25] Giacalone, G., Lo Bosco, G., Barra, M., Bonanno, A., Buscaino, G., Noormets, R., Nuth, C., Calabrò, M., Basilone, G., Genovese, S., Fontana, I., Mazzola, S., Rizzo, R., Aronica, S., 2021. Pattern classification from multi-beam acoustic data acquired in kongsfjorden. Lecture Notes in Computer Science 12666, 55–64.

[26] Groeneveld, R.A., Meeden, G., 1984. Measuring skewness and kurtosis. Journal of the Royal Statistical Society: Series D (The Statistician) 33, 391–399.

[27] Halkidi, M., Batistakis, Y., Vazirgiannis, M., 2001. On clustering validation techniques. Journal of Intelligent Information Systems 17.

[28] Halkidi, M., Vazirgiannis, M., Batistakis, Y., 2000. Quality scheme assessment in the clustering process, in: Zighed, D.A., Komorowski, J., Żytkow, J. (Eds.), Principles of Data Mining and Knowledge Discovery, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 265–276.

[29] Hartigan, J.A., 1975. Clustering Algorithms. 99th ed., John Wiley and Sons, Inc., USA.

[30] Hoegh-Guldberg, O., Bruno, J.F., 2010. The impact of climate change on the world's marine ecosystems. Science 328, 1523–1528.

[31] Hollowed, A.B., Planque, B., Loeng, H., 2013. Potential movement of fish and shellfish stocks from the sub-arctic to the arctic ocean. Fish. Oceanogr. 22, 355–370.

[32] Hop, H., Gjøsæter, H., 2013. Polar cod (boreogadus saida) and capelin (mallotus villosus) as key species in marine food webs of the arctic and the barents sea. Marine Biology Research 9, 878–894.

[33] Hop, H., Pearson, T., Hegseth, E.N., Kovacs, K.M., Wiencke, C., Kwasniewski, S., Eiane, K., Mehlum, F., Gulliksen, B., Wlodarska Kowalczuk, M., Lydersen, C., Weslawski, J. M. Cochrane, S., Gabrielsen, G.W., Leakey, R.J., Lønne, O.J., Zajaczkowski, M., Falk Petersen, S., Kendall, M., Wängberg, S., Bischof, K., Voronkov, A.Y., Kovaltchouk, N.A., Wiktor, J., Poltermann, M., Prisco, G., Papucci, C., Gerland, S., 2002. The marine ecosystem of kongsfjor-

den, svalbard. Polar Research 21, 167–208.

[34] Hubert, L., Levin, J., 1976. A general statistical framework for assessing categorical clustering in free recall. Psychological Bulletin 83, 1072–1080.

[35] Innangi, S., Bonanno, A., Tonielli, R., Gerlotto, F., Innangi, M., Mazzola, S., 2016. High resolution 3-d shapes of fish schools: A new method to use the water column backscatter from hydrographic multibeam echo sounders. Applied Acoustics 111, 148–160.

[36] Krzanowski, W., Lai, Y.T., 1988. A criterion for determining the number of groups in a data set using sum-of-squares clustering. Biometrics 44, 23.

[37] Lonne, O.J., Gulliksen, B., 1989. Size, age and diet of polar cod, boreogadus saida (lepechin 1773), in ice covered waters. Polar Biol. 9, 187–191.

[38] McClain, J.O., Rao, V.R., 1975. Clustisz: A program to test for the quality of clustering of a set of objects. Journal of Marketing Research 12, 456–460.

[39] Miligan, G.W., Cooper, M.C., 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50, 159–179.

[40] Milligan, G., 1980. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. Psychometrika 45, 325–342.

[41] Minelli, A., Tassetti, A.N., Hutton, B., Pezzuti Cozzolino, G.N., Jarvis, T., Fabi, G., 2021. Semi-automated data processing and semi-supervised machine learning for the detection and classification of water-column fish schools and gas seeps with a multibeam echosounder. Sensors 21, 2999.

[42] Moline, M.A., Karnosky, N.J., Brown, Z., Divoky, G.J., Frazer, C.A., Jacoby, C.A., Torres, J.J., Fraser, W.R., 2008. High latitude changes in ice dynamics and their impact on polar marine ecosystems. Ann. N.Y. Acad. Sci 1134, 267–319.

[43] Morrow, J.E., 1980. The freshwater fishes of alaska. Alaska Northwest Pub. Co , pp. 248.

[44] Muiño, R., Carrera, P., Iglesias, M., 2003. The characterization of sardine (sardina pilchardus walbaum) schools off the spanish-atlantic coast. ICES Journal of Marine Science 60, 1361–1372.

[45] Muus, B., Nielsen, J.G., 1999. Sea fish. Scandinavian Fishing Year Book. Hedehusene, Denmark.

[46] Nuth, C., Kohler, J., König, M., Deschwanden, A., Hagen, J., Kääb, A., Moholdt, G., Pettersson, R., 2013. Decadal changes from a multitemporal glacier inventory of svalbard. The Cryosphere 7, 1603–1621.

[47] Perry, A.L., Low, P.J., Ellis, J.R., Reynolds, J.D., 2005. Climate change and distribution shifts in marine fishes. Science 308, 1912–1915.

[48] Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L., Levin, S.A., 2013. Marine taxa track local climate velocities. Science 341, 1239–1242.

[49] Ponomarenko, V.P., 1968. Some data on the distribution and migrations of polar cod in the seas of the soviet arctic. Rapports et Proce's-Verbeaux des Re´unions Conseil International l´Exploration de la Mer 158.

[50] R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.r-project.org/.

[51] Rass, T.S., 1968. Spawning and development of polar cod. Rapports et Proce's-Verbeaux des Re´unions Conseil International l´Exploration de la Mer 158.

[52] Ratkowsky, D., Lance, G., 1978. A criterion for determining the number of groups in a classification. Australian Computer Journal 10, 115–117.

[53] Renaud, P.E., Berge, J., Varpe, O., Lonne, J., Nahrgang, J., Ottesen, C., Hallange, I., 2012. Is the poleward expansion by atlantic cod and haddock threatening native polar cod, boreogadus saida? Polar Biology 35, 401–412.

[54] Renaud, P.E., Carroll, M.L., Ambrose, W.G.J., 2008. Effects of global warming on arctic sea-floor communities and its consequences for

higher trophic levels. In: Duarte CM, Agusti S (eds) Impacts of warming on polar ecosystems FBBVA Press, Madrid, 139–175.

[55] Rousseeuw, P., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20, 53–65.

[56] Sasgen, I., Wouters, B., Gardner, A.S., King, M.D., Tedesco, M., Landerer, F.W., Dahle, C., Save, H., Fettweis, X., 2020. Return to rapid ice loss in greenland and record loss in 2019 detected by the grace-fo satellites. Commun Earth Environ 1.

[57] Simmonds, J., MacLennan, D.N., 2008. Fisheries acoustics: theory and practice. John Wiley & Sons.

[58] Simpson, S.D., Jennings, S., Johnson, M.P., Blanchard, J.L., Schön, P.J., Sims, D.W., Genner, M.J., 2011. Continental shelf-wide response of a fish assemblage to rapid warming of the sea. Curr. Biol. 21, 1565–1570.

[59] Snoeijs-Leijonmalm, P., Gjøsæter, H., Ingvaldsen, R.B., Knutsen, T., Korneliussen, R., Ona, E., Skjoldal, H.R., Stranne, C., Mayer, L., Jakobsson, M., Gårdfeldt, K., 2021. A deep scattering layer under the north pole pack ice. Progress in Oceanography 194, 102560.

[60] Svendsen, H., Beszczynska-Møller, A., Ove Hagen, J., Lefauconnier, B., Tverberg, V., Gerland, S., Børre Ørbøk, J., Bischof, K., Papucci, C., Zajaczkowski, M., Azzolini, R., Bruland, O., Wiencke, C., Winther, J.G., W., D., 2002. Polar cod (boreogadus saida) and capelin (mallotus villosus) as key species in marine food webs of the arctic and the barents sea. Polar Res 21, 139–166.

[61] Tibshirani, R., Guenther, W., Hastie., T., 2001. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society Series B .

[62] Tynan, C.T., Ainley, D.G., Stirling, I., 2009. Sea ice: A critical habitat for polar marine mammals and birds. D.N. Thomas, G.S. Dieckmann (Eds.), Sea Ice (second ed.), Wiley-Blackwell, Oxford.