Angular Range Analysis (ARA) and K-Means Clustering of Multibeam Echosounder Data for Determining Sediment Type

Fahrulian[1]*, Henry M Manik[1], Indra Jaya[1],and Udrekh[2]

[1]Department of Marine Science and Technology Faculty of Fisheries and Marine Sciences
Bogor Agricultural University, Jl. Agatis, Kampus IPB Dramaga Bogor, Indonesia16680
[2]Agency for the Assessment and Application of Technology
Jl. M.H.Thamrin 8, Jakarta, Indonesia 10340
Email: fahrulian@gmail.com

Abstract

Backscatter value was a key to determine seabed characteristic. Level of intensity or backscatter informed through seabed type. One approach was use to analyze seabed type based on the value of backscatter was Angular Range Analysis (ARA). ARA utilize influence of angle backscattering intensities. The aim of this research was to determine value of backscatter from the bottom and used to initial prediction of seabed. Extraction processes of raw data obtained by acoustic signal processing techniques. Analysis of backscatter data was conducted by using K-means method to look the proximity of the centroid backscatter value against other values. Backscatter intensity from this sites ranged from -41,93 dB to -27 dB. The range value divided into three major classes based on Wenworth scale classification. Substrate type in the study site consists of sand, silt, and clay. Grain size diameter of each type include: sand 0.122713 mm (phi = 3.02), silt 0.018171 mm (phi = 5.78) and clay 0.002690 mm (phi= 8.53). Based on this result, signal processing multibeam echosounder able to classify seabed backscatter values to determine the seabed type

Keywords: *Angular Range Analysis, multibeam echosounder, K-Means*

Introduction

Seabed mapping using multibeam echosounder is still being used around the world. Multibeam able to sounding with broad scope, high accuracy and can be used to determine characteristics of seabed type (Anderson *et al*., 2007; Anderson *et al*., 2008; Van Holliday *et al*. 2008; Fernandes and Chakraborty, 2009; Brown, Smith *et al*., 2011; Hasan *et al*., 2011; Manik, 2011). Backscatter was the key to determine the condition of the bottom waters. Backscatter intensity was obtained from the signal receiver provide preliminary information on sediments type in sounding location spatially (Jackson and Brigs, 1992; Fonseca and Mayer, 2007; Jackson and Richardson, 2007; Tegowski *et al*., 2009: Manik, 2011; Ruiz-Cortés and Dainty, 2012; Huang *et al*., 2013; Zhi *et al.,* 2014).

Signal from receiver can be analyzed further to determine size of sediment type (Jackson and Richardson, 2007). Characteristics from the bottom was became a very important thing (Park *et al*., 2011). One approach used to analyze sediment type, relationship between angle of incidence acoustic waves and backscatter received was Angular Range Analysis (Le Chenadec *et al.,* 2003; Boucher *et al*., 2003; Fonseca, 2009; Hasan *et al*., 2012; Hasan *et al*., 2014). Research about seabed type is very important to do. Determining of seafloor geomorphology and studies of benthic habitat can be analyzed from this research. In the military, information about seabed type used to steering of gear submarines and amphibious operation (Bentrem *et al.,* 2002).

ARA method was selected because it has a high angular resolution. From this reason, it can get full information of backscatter value along sounding process angle to angle of signal transmit (Fonseca and Mayer, 2007; Jackson and Richardson, 2007). ARA will produce backscatter value from seabed. Backscatters have to be separated based on intensity. One of the way to cluster data was K-Means. This method gives us effective and efficient way to cluster of data. Besides that, calculation process will be fast than the others relatively. K-Means apply a simple algorithm and it able to produce good results.

Using grab sampler in the deep sea is inconvenience for some reason desired point or site. This condition is usually caused by oceanographic factors, such as underwater currents. Underwater

*) Corresponding author

DOI: 10.14710/ik.ijms.21.4.177-184
Received : 07-01-2016
Accepted : 12-03-2016

acoustic methods do not experience this condition as well as the conditions brought on by the water grab. The aim of this study was to determine distribution of backscatter intensity along the survey lines and classify those values for the purpose of determining sediment type. It is also as an effort to optimize the use of instruments multibeam echosounder.

Material and Methods

Survey was carried in west of Sumatra Island between Nias Island and Sumatera Island, on 1°39'0"N–1°40'30"North and 97°34'30"E–97°39' 0" East. Data were acquired by using Baruna Jaya 4 belongs to BPPT.

Extraction process of raw data multibeam was started by changing data format (.all) from specific types of instruments into a data format that was recognized by signal processing software. Some corrections were applied such as altitude, swath and navigation editor to remove influence of noise signal. The next process was to enter tide range from the location at same day. It was intended as correction depth of fatherly generate depth values corresponding to the current state of data retrieval. In addition, sound speed data in the location was needed. Both data were then put together and merge. The end of this signal processing was base surface. Information that can be obtained was the distribution value and the value of backscatter depth spatially

Angular Range Analysis (ARA)

The signals received from the seabed through many events or processes in period emitted and return to receiver. Multibeam is able to record from water column to near bottom surface. There are 191 beams arranged regularly and aperture. Signal intensity from nadir value will be great than nadir. Analysis effect of incident angle multibeam was originally adopted by Jackson *et al*. (1986) and developed into a method known as Geocoder (Fonseca and Calder, 2007; Fonseca and Mayer, 2007; Fonseca, 2009). Backscatter analysis was performed based on backscatter surface in covered area and backscattering volume. Integration both become key analysis backscatter value based angle function.

Backscattering = Surface backscattering + Volume backscattering

Surface backscattering :

$$\sigma_r(\theta, f) = F(\theta, f\ ; \xi, \rho(\xi,), v(\xi,), \delta(\xi,), \omega_2, \lambda)$$

where :

θ = Interface backscatter cross section per unit solid angle
F = Frequency
ρ = Ratio sediments to mass water density
v = Ratio sediment sound speed to water column
δ = Imaginary parameter ratio of the number of sound waves in sediment
ω_2 = Spectral strength of bottom relief
λ = Spectral exponent bottom waters
ξ = Gas volume / total volume of sediment

Volume Backscattering :

$$\sigma_v = \frac{5\left((\xi \Sigma_h + \sigma_2\,\alpha(\xi))c_w\left|1 - R^2(\theta)\right|^2 \sin^2(\theta)\right)}{40\pi f\,|P(\theta)|^2 \mathrm{Im}|P(\theta)|}$$

where :

Σ_h = Backscattering cross section in bubbles area
σ_2 = Scattering volume
R(θ) = Reflection coefficient
P(θ) = Function of energy loss
$\alpha(\xi)$ = Attenuation coefficient in dB/m
c_w = Sediment Sound speed

K-means clustering

The function of seabed classification is to see sediment type in the study site. Classification process conducted with numerical approach. Backscatter data was obtained by extraction process from GeoBaR then observed from smallest to largest. From this step, we have to determine number of class (K) based on Wenworth scale in 1983. Additional data that used for input parameters in the K-means clustering was grain size and *phi*. Algorithm that used in this process (MacQueen 1967):

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_1} \left(\|x_i - v_j\|\right)^2$$

where :

c_i = Number of data
c = Number of cluster
$\|x_i - v_j\|$ = Euclidean Distance between x_i and v_j

K-means clustering works by utilizing centroid of data (Telgarsky and Vattani 2010). The result from this process will became reference to grouping of other values in collecting data sets to be able to classify, clustered into new groups based on proximity. Data that have similar properties will be clustered closer to centroid point. This process is not enough one times, iteration or repetition until the classification process stable is needed. Calculate iterations to produce new centroid point can use by following algorithm (Lloyd 1982):

$$v_i = \frac{1}{|c_1|}\sum_{j=1}^{c_1} x_1$$

where :

c_i = Number of data i^{th}

Distance determination of multibeam data point to centroid apply Euclidean Distance algorithm. In practice, this algorithm calculates root mean square sum of difference between centroid point to other point (Balaji and Bapat 2007). If distance too far, the data will grouped into nearest centroid point. The equation can be written:

$$ED = \sqrt{\sum_{j=1}^{p}(a_{ij} - a_{hj})^2}$$

where :

ED = Euclidean Distance

a_{ij} = Point at i,j

a_{hj} = Centroid each point in class h,j

Sediment grain size

Estimation of sediment types using acoustical wave based on size of each type of sediments. Sediment with hard characteristics will have a greater backscatter than fine textured sediment types. This condition became the basic determining type of sediments in this research. An algorithm applied measuring of sediment type based on backscatter results. Seabed type was calculated through the negative logarithm base 2 from diameter in millimeters unit. Mathematically can be written:

$$Mz\ \phi\ (phi) = -\ Log_2\ (d/d_0)$$
$$= -3,32\ Log_{10}(d/d_0)$$

where :

d = Grain size of sediment (mm)

d_0 = Reference value (1 mm)

Result and Discussion

Geobar and mosaic from backscatter strength

GeoBaR created by an algorithm which effective enough to estimate, classify and delineate size of sediment type. This is the main attraction for users to process and analyze acoustic signal that successfully recorded. Geocoder method is not only

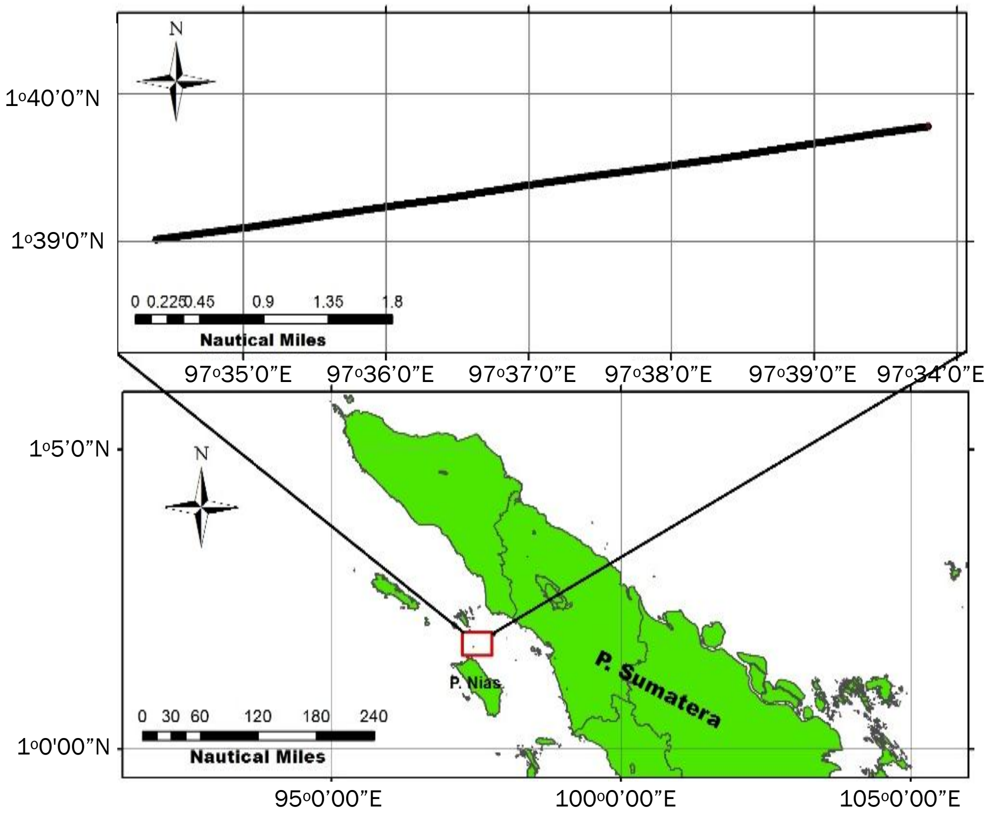

Figure 1. Multibeam echosounder survey sites

applied in multibeam data, but also side scan sonar basically. GeoBaR obtained separately from line surveys. If we have only one line survey, GeoBaR still be formed to look at the distribution of intensity values.

Backscatter data created in image and classed by learning machine. The color distribution was seen from the proximity of 8 bit digital number. Values that are in a range will go into a certain color class. If wide range value was given, the type of color result will be small. If small range given, it will produce the kind of color. But kind of range colors is not main thing success in determining classification of seabed type.

Blending more than one image will produce a mosaic. This image used to classify or segment several classes of seabed based on the texture and attributes of each class. Mosaic was a picture of sea floor that storage information from seabed features such as intensity in time series or beam average. Mosaics can be used as initial information to see distribution of intensity values sounding results. Figure 2 is mosaic result intensity from seabed. Backscatter intensity from this site ranged from -41.93 dB to -27 dB. These values are obtained from ratio between power transmitted and power received by the transceiver. Based on the distribution of these values, we can observe that higher intensity tend in nadir area. This is because signals are still in aperture or known as the narrow beam.

Clustering result from backscatter

K-means clustering chosen because this method is quite simple to use and also a appropriate method used in amounts of data such as multibeam echosounder data (Ellingsen, 2002). Iteration process begins with the classification and determination of the centroid point at random. Variables used in this clustering process consist of three input, namely: backscatter, grain size of sediment and phi. The number of data can be extracted as much as 101.248 points throughout

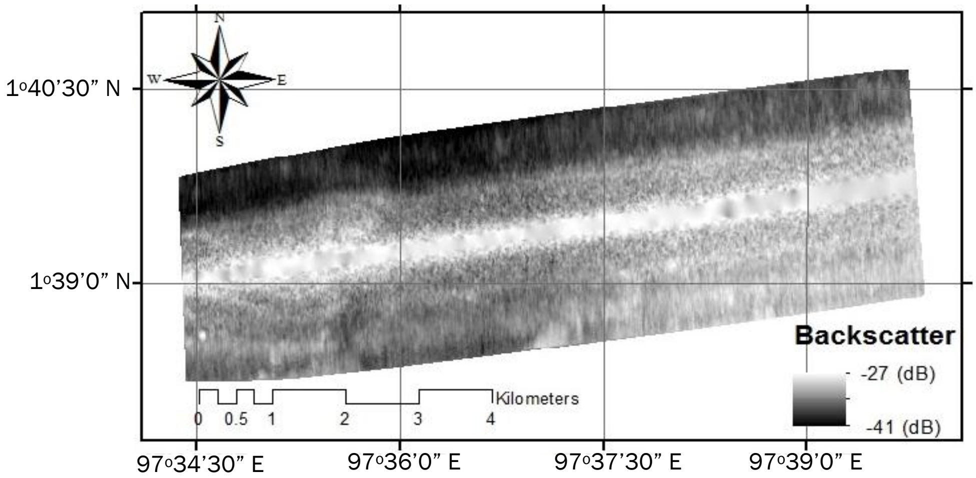

Figure 2. Backscatter mosaic (BS) in research site

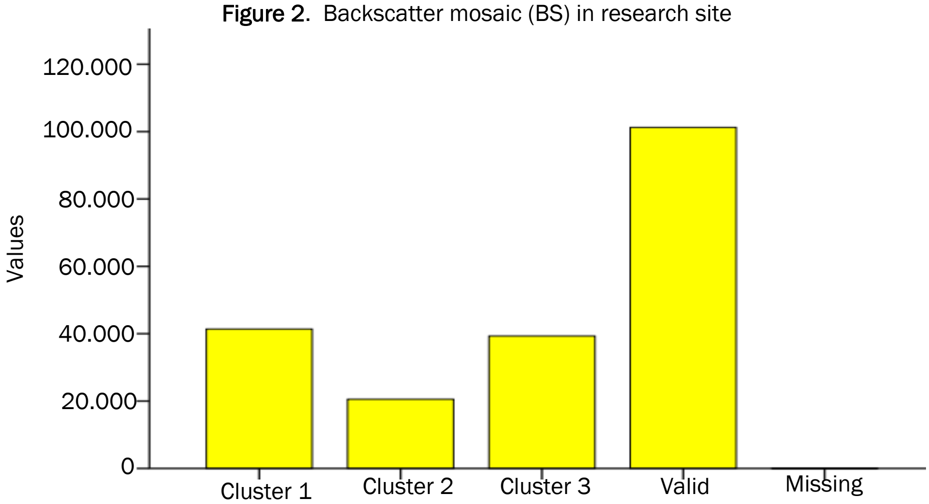

Figure 3. Histogram of clustering multibeam data

the area. Figure 3showed statistical calculation data presented in histogram. From 101248 point data, it was obtained: Class 1 amounted to 41.367 data cluster, class 2 320.549 data point and 39.332 in class 3. Total data which successfully processed in K-means are same with original data obtained through the extraction process raw data. In other words, all data valid and processed using K-means clustering.

Iteration process illustrate of repetition to find stable centroid point. The initial iteration processed from set of data randomly. Each object that has same value to centroid point will group by it self. Every iteration will produce different centroid point value than previous. Centroid point is value obtained from sum of the data in one cluster divided number of data. The number of iterations applied in this study was 19 times. Number of iteration will not same one each other. If result of changing centroid point was nothing it means the process was stable, Figure 4 showed history of iterations performed used K-means clustering. In each class, it appears that the same pattern. The initial value of iteration process tends to be high, it is because the algorithm would take random initial value. This process showed change of iteration short relatively, it means range of seabed type. Based on the results of the classification performed backscatter values, the obtained three main classes that represent sediment type along the survey line. The three classes are: sand, silt and clay. Highest of these classes was sand. Physically, sand has more compact and dense than silt and clay. Grain size of sand was greater than the other sediment types. For grain size category, sand has the smallest value. Grain size which in phi unit has valuation upside compared to other parameter. Phi greater value was used to describe a class of sediments which have characteristics very delicate and very small diameter for example clay.

Angular response analysis and backscatter curve

Acoustic signals received from the bottom of waters were sound energy that converted into electrical energy. Signal has many events such as attenuation, absorption, scatter, transmission and some other possibilities (Fonseca and Mayer, 2007). This causes power received becomes smaller when turn back from the bottom of sea floor. However, intensity of the back can still be analyzed as an object identifier signal in bottom waters. In this study, every 1 line survey was divided into 15 patches. Totally there are 30 patches along the survey line. Each patch has different type of sediment, and for this reason, intensity of each also different. Figure 5 is a relationship between backscatter in decibel unit and angle of incidence.

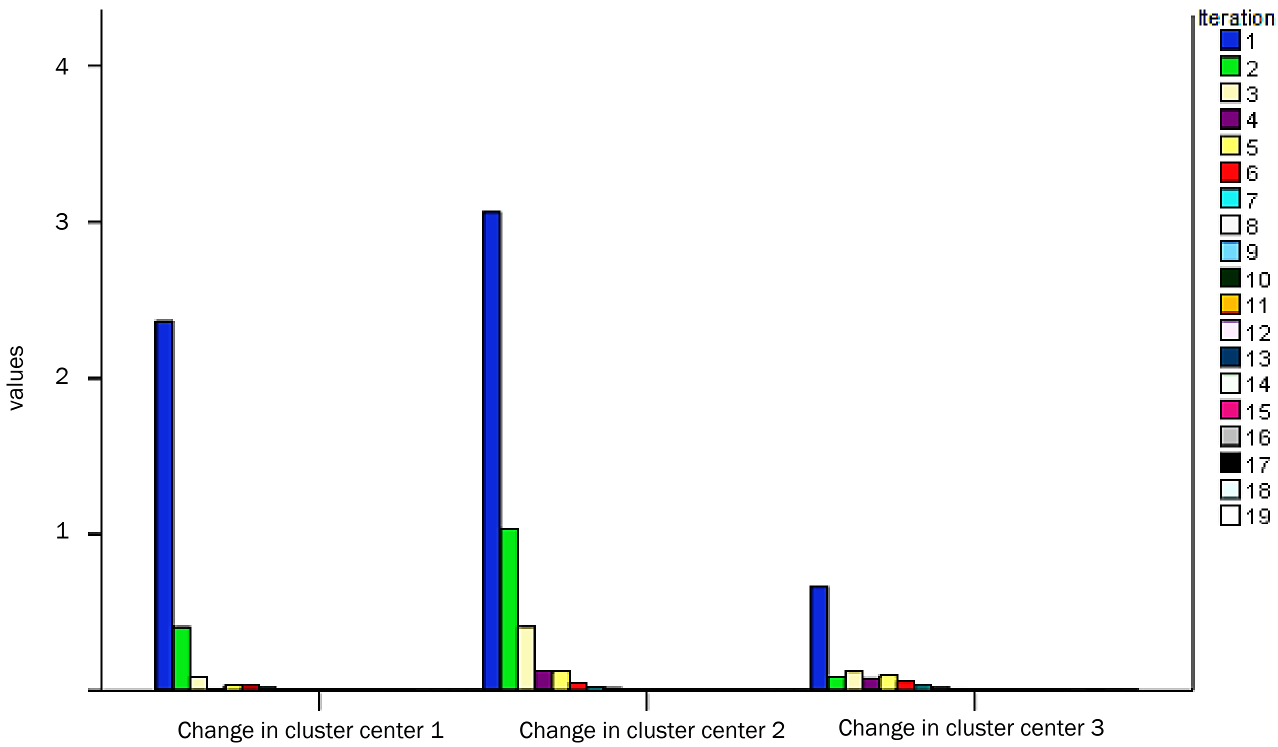

Figure 4. Changes of middle value on the classification process

Backscatter classification results from seabed can be divided into three major classes, they are sand, silt and clay. Although in real condition, they are possibility combination from 3 type of sediment. This study only focused on segmentation by optimizing ability of mutibeam echosounder. In determining the seabed type, ARA method through applied principle of homogeneity half swath and estimation of the average measurement backscatter values obtained from sweep area. Angle response will make a border of a sediment that will generate the kind of sediment of different composition.

Backscatter value from sand is higher than, silt and clay, because sand is harder and denser with -27 dB. When acoustic signal reach on sand, return of acoustic signal still relatively high compared with other types and graph obtained by backscatter intensity was also higher. Intensity from silt sediment tend weaker than sand because physically, silt has finer than sand. Signal regarding the type of sediment is more likely to have absorption compared to that in the waste. Sediment that has low intensity is clay with -29,64 dB. This sedimentary structure finer than sand and silt. In Wenworth classification scale, clay is the most delicate kinds of sediments from all types of sediments with -35,45 dB. Almost all the energy emitted from transducer absorbed into sediments. Based on this condition, intensity from this type of sediments is very small. Figure 5 showed backscatter level of each main types of sediment. Based on ARA and clustering data using K-Means, the final result was produce map of sediment type spatially. In area of research there are three kind of sediment type consist of: sand, silt and clay. From the map the dominant type was sand, which is almost 40%. Second was silt with 22%. The other one was clay with 38%. Distribution of sediment type degraded form a fairly regular pattern. We can see

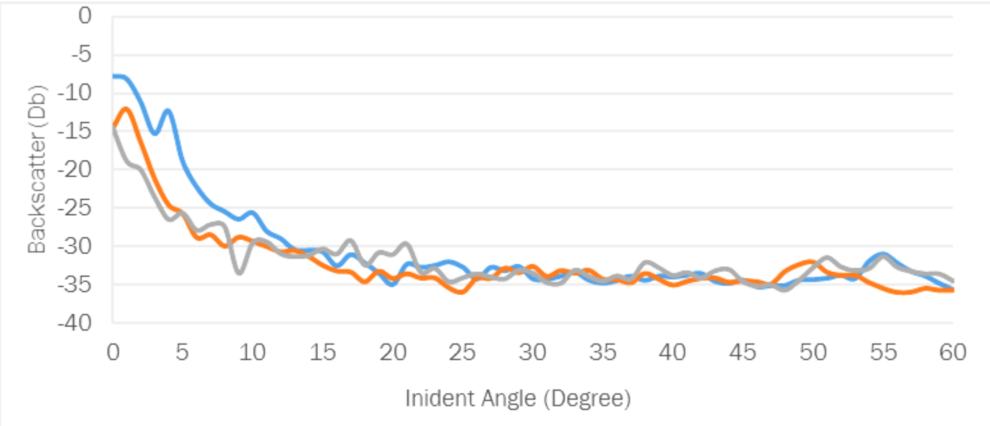

Figure 5. Backscatter level of each main types of sediment
Note: _____ = Sand, _____ = Silt, _____ = Clay

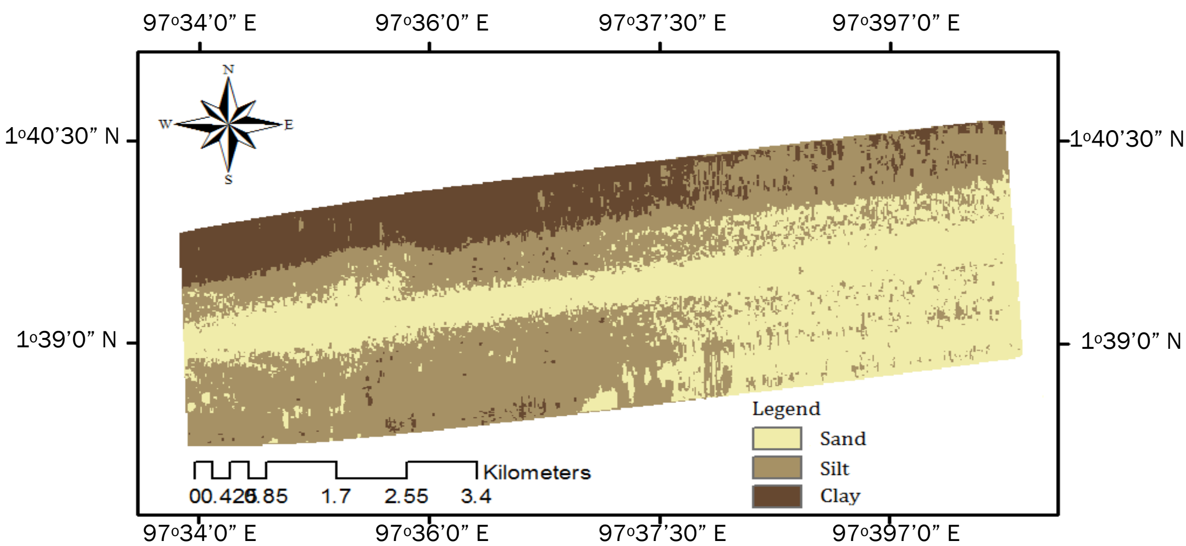

Figure 6. Map distribution of sediment type in research site

quite clearly visible location of each type of sediment type. This information is very important for a variety type of sediment clustered in a certain region based on their respective sediment. Figure 6 show map distribution of sediment classes from the sediment type. Based on the foregoing, multibeam echosunder proved quite powerful and can be optimized in its use.

Validation of sediment type data from this research was quite difficult to do. Deep water become boundary factor to get coring of sediment type. For this reason, using literature become a way to obtain validation process. There are some similar research using multibeam echososunder for mapping and explored sea bottom. Table 1 show comparison of backscatter for validation data.

Table 1. Comparison of backscatter from each sediment type

No	Sediment Type	Dufek	SIMRAD	This research
1	Sand	-27,00 dB	-30,00 dB	-27,00 dB
2	Silt	-31,00 dB	-31,00 dB	-29,64 dB
3	Clay	-34,00 dB	-36,00 dB	-35,45 dB

Conclusion

Backscattering strength obtained based on the signal processing is quite varied. Characterized by the distribution of different values and spatially spread along the survey line. These values describe the information about condition and seabed type of sediments. Based on the classification results, obtained that there are three main classes of seabed types, sand, and clay. Based on the results, multibeam echosounder can be optimized not only for bathymetry but also can be used to determine seabed type

Reference

Anderson, J.T., Holliday, V., Kloser, R., Reid, D. & Simard, Y. 2007. Acoustic seabed classification of marine physical and biological landscapes. *International Council for the Exploration of the Sea*. p.198 doi: 10.1093/icesjms/fsn061.

Anderson, J.T., van Holliday, D., Kloser, R., Reid, D.G. & Simard, Y. 2008. Acoustic seabed classification: Current practice and future directions. *ICES J. Mar.Sci.* 65(6):1004-1011. doi:10.1093/icesjms/fsn 0 6 1.

Balaji, R. & Bapat, R. 2007. On euclidean distance matrices. *Linear algebra and its applications.* 424(1):108-117.

Brown, C.J., Smith, S.J., Lawton, P. & Anderson, J.T. 2011. Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Est. Coast. Shelf Sci.* 92(3):502-520. doi:10.1016/j.ecss.2011.02. 007.

Hasan, R.C., Ierodiaconou, D. & Laurenson, L., 2012. Combining angular response classification and backscatter imagery segmentation for benthic biological habitat mapping. *Est. Coast. Shelf Sci 97*:1-9. doi:10.1016/j.ecss.2011.10.004.

Hasan, R.C., Ierodiaconou, D., Laurenson, L. & Schimel, A., 2014. Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping. *PloS one*, *9*(5):p.e97339. doi: 10.1371/journal. pone.0097339.

Hasan, R.C., Ierodiaconou, D., Rattray, A., Monk, J. & Laurenson, L., Applications of multibeam echosounder data and video observations for biological monitoring on the south east Australian continental shelf. ISG & ISPRS. Shah Alam, Malaysia.

Fernandes, W & Chakraborty. 2009. Multi-beam backscatter image data processing techniques employed to em 1002 system. Pages 93-99. in Pillai PRSS, M.H, ed. Proceedings of the International Symposium on Ocean Electronics (SYMPOL-2009).

Fonseca. 2009. Angular range analysis of acoustic themes from Stanton Banks Ireland: A link between visual interpretation and multibeam echosounder angular signatures. *Applied Acoustics.* 70(10):1298-1304. doi:10.1016/ j.apacoust.2008.09.008.

Fonseca, L., & Mayer, L. 2007. Remote estimation of surficial seafloor properties through the application angular range analysis to multibeam sonar data. *Mar. Geophysical Res.* 28(2):119-126. doi:10.1007/s11001-007-90 19-4.

Fonseca, L., & Calder, B. 2007. Clustering acoustic backscatter in the angular response space. Proceedings of the US Hydrographic Conference. Norfolk, VA.

Huang, Z., Siwabessy, J., Nichol, S., Anderson, T. & Brooke, B., 2013. Predictive mapping of seabed cover types using angular response curves of multibeam backscatter data: Testing different feature analysis approaches.

Continental Shelf Res. 61:12-22. doi:10.10 16/j.csr.2013. 04.024.

Jackson, D. & Brigs, K. 1992. High-frequency bottom backscattering roughness versus sediment volume scattering. *J. Acoust. Soc. Am.* 92(2): 962- 977.

Jackson, D. & Richardson, M. 2007. High-frequency seafloor acoustics. Springer Science & Business Media.

Jackson, D.R., Winebrenner, D.P. & Ishimaru, A. 1986. Application of the composite roughness model to high-frequency bottom backscattering. *J. Acoust. Soc. Am.* 79(5):1410-1422.

Le Chenadec, G., Boucher, J.M., Lourton, X. & Augustin, J.M.. 2003. September. Angular dependence of statistical distributions for backscattered signals: modelling and application to multibeam echosounder data. In *OCEANS 2003 Proceedings* 2:897-903

Lloyd, S.P. 1982. Least squares quantization in pcm. *Information Theory, IEEE Transactions on.* 28(2):129-137.

Lurton X. 2010. An introduction to underwater acoustics. Verlag Berlin Heidelberg Springer.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Pages 281-297. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability: Oakland, CA, USA.

Manik, H.M. 2011. Underwater acoustic detection and signal processing near the seabed INTECH Open Access Publisher.

Park, Y., Lee, S. & Jung, S. 2011. Characterization of backscattering signal of 300 khz multibeam echo sounder. Proceeding of Symposium on Ultrasonic Electronics.

Ruiz-Cortés, V. & Dainty, C., 2012. Backscattering measurements from double-scale randomly rough surfaces. *J. Opt. Soc. Am A*, *29*(6):1154-1160.

Tegowski, J., Gorska, N., Kruss, A., Nowak, J. & Blenski, P. 2009. Analysis of single beam, multibeam and sidescan sonar data for benthic habitat classification in the southern Baltic Sea. In *Proc. 3rd Int. Conf. Exhibition on Underwater Acoustic Measurements: Technol. & Results* pp. 21-26

Telgarsky, M. & Vattani, A. 2010. Hartigan's method: K-means clustering without voronoi. *Int. Conf. on Artificial Intelligence and Stat.* pp 820-827

Zhi, H., Siwabessy, J., Nichol, S.L. & Brooke, B.P. 2014. Predictive mapping of seabed substrata using high-resolution multibeam sonar data: A case study from a shelf with complex geomorphology. *Mar. Geol.* 357:37-52. doi: 10.1016/j.margeo.2014.07.012.