Automatic classification techniques for type of sediment map from multibeam sonar data

R Zakariya¹, M A Abdullah¹, R Che Hasan² and I Khalil¹
¹School of Marine and Environmental Sciences, University Malaysia Terengganu, Terengganu, Malaysia.
²UTM Razak School of Engineering and Advanced Technology, University Teknologi Malaysia (UTM), Jalan Semarak 54100 Kuala Lumpur, Malaysia.

*Azhafiz88@gmail.com

Abstract. Sediment map can be important information for various applications such as oil drilling, environmental and pollution study. A study on sediment mapping was conducted at a natural reef (rock) in Pulau Payar using Sound Navigation and Ranging (SONAR) technology which is Multibeam Echosounder R2-Sonic. This study aims to determine sediment type by obtaining backscatter and bathymetry data from multibeam echosounder. Ground truth data were used to verify the classification produced. The method used to analyze ground truth samples consists of particle size analysis (PSA) and dry sieving methods. Different analysis being carried out due to different sizes of sediment sample obtained. The smaller size was analyzed using PSA with the brand CILAS while bigger size sediment was analyzed using sieve. For multibeam, data acquisition includes backscatter strength and bathymetry data were processed using QINSy, Qimera, and ArcGIS. This study shows the capability of multibeam data to differentiate the four types of sediments which are i) very coarse sand, ii) coarse sand, iii) very coarse silt and coarse silt. The accuracy was reported as 92.31% overall accuracy and 0.88 kappa coefficient.

1. Introduction
Fauna communities harbored different space that suits their lifestyle in order to survive. In the aquatic or marine world, benthic communities are the one that has to meet this condition in order to survive. Benthic communities are organisms that spend their lives on, in or near the ocean floor or sediment [1]. Different benthic organisms have different preferences on the habitat that they lived on which specifically based on the sediment type which is called zonation [1]. Example of this zonation can be seen at coastal area from the intertidal zone to sub tidal zone where sediment types are different in each zone.

Since sediments have different size and texture [2], it would provide all sorts of information and data for scientists. Besides, physical data, the distribution of sediment on certain area could also provide biological data on species that live there. This means that area of different sediment distribution provides different habitat based on the organisms adaptation or specificity.
Multibeam echosounder is a device with the ability to map the seafloor in large scale by emitting active sonar. Multibeam echosounder can sounding with broad scope, high accuracy and can be used to determine characteristics of seabed type [4]. One of the methods is to use backscatter or intensity returns from underwater acoustic sonar such as multibeam echosounder to remotely gather seafloor sediment types [5]. Backscatter can be used to identify hard and soft seafloor surfaces which can later be used to predict sediment classes. An analysis of the amplitude of the return wave sound allows for information on the structure and hardness of the base waters, which is used to classify sediment Strong types that return to the hard surface (rock, gravel) and weak signals returning to a finer surface (silt, clay). This is because the greater the impedance of a medium the larger the reflection coefficient.

There are many benefits that could be obtained by using multibeam echo-sounder in marine science applications. First, it helps to decrease the expenses and required less time for fieldwork. The data obtained which specifically on sediment mapping can be used for future studies. Stephens and Diesing (2014) stated that a sediment map of high resolution is in demand for better environmental assessment and management. On top of that, it can also be used by a third party like Department of Fisheries so that they can conduct optimum yet effective research on locality of fish and invertebrate in the area. Department of Environment also will benefit from sediment mapping as the information can be used to scrutinize the sediment distribution, somehow related to pollution study. The aim of this study to determine types of sediment map using Multibeam sonar data.

2. Materials and methods
2.1 Study area
Pulau Payar is situated far apart from the coast of Kedah state, between Pulau Langkawi and Pulau Pinang. The study area covers part of the north of Pulau Payar. The coverage of this survey area is approximately 239385 m². The area is selected as the study area due to the availability of bathymetry data and backscatter data to classify the sediment map characteristics. The height range of this area is 17 meter up to 25.76 meter. Figure 1 shows the location of the study area.

![Figure 1. location of the survey area in Pulau Payar, Kedah](image)

2.2 Acoustic data
Acoustic data were acquired in Pulau Payar on 13th to 17th January in 2017. Multibeam echo sounder system (MBES) R2Sonic 2020 was used in this study for collecting bathymetry image and backscatter image. The average speed for survey vessel is 5 knots. This instrument produces 256 beams equiangular arrayed over an arc of 130° and used frequency at 400 kHz. Differential Global Position System (DGPS)
Trimble 461 was used for system navigation and heading (gyro compass). The vessel’s motion was quantified using DMS10 motion sensor with 0.05° accuracy for heave, roll and pitch. Real time speed of sound for MBES transducer used Valeport Mini sound velocity sensor (SVS). Software QINSy v8.0 was used during a survey for data logging, real time quality control, display and navigation. The sound velocity profile (SVP) collects from cast away-CTD used for measuring sound velocity during data acquisition to correct the effect of refraction.

Software Qimera version 1.4 was used for data processing on bathymetry raw data. The bathymetry raw data processing follows step correction for sound velocity variation though the water column, correction for tide in water depths, correction of motion sensor information for data bathymetry and remove for spike or noise [6]. FMGeocoder Toolbox or FMGT version 7.4 was used to process the raw MBES backscatter and obtain backscatter mosaic.

2.3 Ground truth data

At the sampling site (figure 2a), ponar grab was used to collect the sediment sample because it was suitable and handy. The grab was tied to a rope prior to sampling. At each sampling point, the grab was deployed until it hits the bottom, collect the sediment and finally was retrieved. The sediment was scoped into a plastic bag, labelled and sealed. The sediment samples were brought back to the laboratory for further analysis.

The sediment sample analysis was carried out using dry and wet sieving techniques to determine the grain size of the sediment. Using dry sieving techniques, it is suitable for coarse fraction and particles with diameter greater than 63 μm. Approximately; a 100g of dried sample was weighed. The sample was then placed on top of the sieve which was stacked with other sieves below it with different mesh size. A series of sieves was shaken using sieve shaker for 15 and 20 minutes. Post to the shaking process, each sieve containing the trapped samples was weighted. For wet sieving the laser diffraction method (using a Particle Size Analyzer (PSA) CILAS 1180) was employed to obtain the statistical value.

The type of the sediment classes was based on sediment grain size [7-9]. In sedimentological calculation, mm unit is commonly used as a metric unit, using to several geological terms i.e. mean size. Mean size ($X_\Phi$) is an indicator of sediment size were measured based on its weight. The decreasing of a mean size value indicates an increase in sediment grain size and vice versa. The value of the mean was obtained either from the calculation of the gravity centre under the grain size curve or the gravity centre of the cumulative distribution curve. The classification systems of sediment grain size typically as Table 1.

| Grain size class | Grain diameter (mm) | Descriptive terminology |
|------------------|---------------------|------------------------|
| 4 – 2            | Granule             |
| 2 – 1            | Very coarse sand    |
| 1 – 0.5          | Coarse sand         |
| 0.5– 0.25        | Medium sand         |
| 0.25– 0.125      | Fine sand           |
| 0.125– 0.0625    | Very fine sand      |
| 0.0625 – 0.031   | Very coarse silt    |
| 0.031– 0.0156    | Coarse silt         |
| 0.0156 – 0.0078  | Medium silt         |
| 0.0078 – 0.0039  | Fine silt           |
| 0.0039 – 0.0020  | Very fine silt      |
| > 0.0020         | Clay                |
2.4 Classification
According to [2], both techniques of supervised and unsupervised have been extensively proposed in benthic studies. These include the easy-to-implement maximum likelihood and ISO cluster classifiers found in most GIS packages [10 – 11]. In this research, supervised technique has been chosen to classify the data. The selected data have been classified by using maximum likelihood supervised.

Supervised classification requires class-describing information that must be as accurate, representative and complete as possible. Such information on the class characteristics might be acquired by training classifiers (e.g. maximum likelihood, k-nearest neighbour, classification tree, random forest, support vector machine and Bayesian network) with ground samples. Maximum likelihood classification method [11 – 12] was used in the present study. This method was chosen as it assigns class to the pixels. Data of sediment type from the dry and wet sieving techniques will be used as guideline to classify certain pixels in order to know the distribution of the sediment based on the pixels characteristic which correlated with different type of sediment.

The backscatter variable is the use of automatic classification were backscatter, mean, mode, median. The statistic data derived from the analysis of variance in backscatter in FMGT version 7.4 software. All information collected will be compared with each other based on the backscatter data to know how reliable the data, based on the error matrix. Plus, kappa coefficient will also be used to determine the accuracy to evaluate the reliability of the data. Kappa coefficient compares the expected accuracy with the observed accuracy of the ground truth data.

A kappa value of 1 implies complete agreement, while a value less than 0 implies more disagreement than a random assignment of classes to one of the maps [13]. Kappa considers all the cells of an error matrix and thus incorporate more information [14 – 17]. Besides that, Kappa is suitable for comparison between different error matrices because it removes chance agreement [18 – 19]. Accuracy of more than 80% marked the reliability of the classified data and concludes it as accurate.

3. Result and Discussion
3.1 Acoustic analysis
The results of the bathymetry map processing show in figure 2 (a) the depth value of natural reef (rock) ranged from 17.626 to 25.760 meters with composition in depth less than 20 meters indicated on natural reef (rock). The average depth in the dominated depth ranges from 20 to 25 meters.

Backscatter map of natural reef (rock) mapped using MBES is shown in Figure 2(b). The backscatter strength, in decibels (dB) ranges from -38 to -80 dB presented as a greyscale image. High backscatter strength can be seen on white part of the map while weak or low backscatter strength focused on darker region on the map. High decibel indicates coarse particles as rough surface reflects sound wave acts on it. Low decibel on the other hand indicate soft surface as some of the sound waves are being absorbed.

3.2 Type of sediment analysis
The analysis of particle size of the total number of samples for each sedimentary class was: very coarse sand is 5 samples, coarse sand is 3 samples, very coarse silt is 4 samples and coarse silt is 1. The mean value in phi indicates the size of particle diameter where higher values exhibit smaller diameter or fine particles and vice versa. Overall all phi values can be classified into 4 types of sediment, i.e. very coarse sand, coarse sand, very coarse silt and coarse silt. The high sediment grain diameter size is 1.57 mm that was very coarse sand for station 13 and the low sediment grain diameter size is 0.02 that was coarse silt. Therefore, it can be stated that all situations it is a particle that is different from the size of the smooth and rough. Table 2 shows the results of classification type of sediment sample wit multibeam data.
Table 2. Result of classification type of sediment sample with multibeam data.

| Station | Depth (m) | Intensity (dB) | Grain diameter (mm) | Classification       |
|---------|-----------|----------------|---------------------|----------------------|
| 1       | 23.738    | -59.450        | 1.25                | Very coarse sand    |
| 2       | 24.081    | -63.351        | 0.63                | Coarse sand         |
| 3       | 23.629    | -64.771        | 0.02                | Coarse silt         |
| 4       | 23.330    | -62.291        | 0.91                | Coarse sand         |
| 5       | 24.308    | -61.231        | 1.42                | Very coarse sand    |
| 6       | 23.815    | -66.541        | 0.04                | Very coarse silt    |
| 7       | 23.876    | -63.710        | 1.01                | Very coarse sand    |
| 8       | 24.568    | -65.831        | 0.04                | Very coarse silt    |
| 9       | 24.738    | -65.121        | 0.06                | Very coarse silt    |
| 10      | 22.689    | -59.100        | 1.00                | Very coarse sand    |
| 11      | 24.703    | -64.411        | 0.88                | Coarse sand         |
| 12      | 24.772    | -65.121        | 0.04                | Very coarse silt    |
| 13      | 20.761    | -58.040        | 1.57                | Very coarse sand    |

Figure 2. Acoustic data of the study area: (a) multibeam bathymetry with the positions of the sediment stations (b) relative backscatter intensity image at natural reef (rock) in Pulau Payar.

Figure 3. Classified sediment map based on integration between backscatter image, statistic backscatter image and ground truth data using maximum likelihood classification.
3.3 Sediment map accuracy

Figure 3 displays sediment map made using integration of data from both ground truth (sediment type) and multibeam (decibel of backscatter) data. Comparing with map in Figure 2 (b) the accepted map which represents as the final map for the present study is the second map (Figure 3) as it is more accurate. Five sampling stations were chosen for the accuracy assessment consists of ST 3, 6, 9, 11 and 13.

The error matrix of user accuracy for the classification method used is shown in Table 3. Commission error is possibly made when classifying the map based on the selected pixel which did not match the ground truth data. From the table below, very coarse sand was misinterpreted as coarse sand and very coarse silt. Table 4 shows the accuracy assessment for the classification method used. The overall accuracy is 92.31%, while the Kappa coefficient is 0.88. The greater confidence in the classification with the value closer to 1 for k range between 0 and 1.

| Table 3. Error matrix of user accuracy for maximum likelihood classification method |
|---------------------------------|----------|----------|----------|----------|----------|
| Class              | Very coarse sand | Coarse sand | Very coarse silt | Coarse silt | Total    |
| Very coarse sand   | 5         | 0         | 0          | 0         | 5        |
| Coarse sand        | 0         | 3         | 0          | 0         | 3        |
| Very coarse silt   | 0         | 0         | 4          | 0         | 4        |
| Coarse silt        | 0         | 1         | 0          | 0         | 1        |
| Total              | 5         | 4         | 4          | 0         | 13       |
| User accuracy (%)  | 100       | 100       | 100        | 0         |          |

Table 4. Accuracy assessment of maximum Likelihood classification method

| Maximum Likelihood | Overall Accuracy | Kappa Coefficient |
|--------------------|------------------|-------------------|
| Maximum Likelihood | 92.31            | 0.88              |

Overall, from a total of 23.9383 km², the area is mostly covered by coarse sand is 13.457 km²; secondly, the area is mostly covered by very coarse sand is 6.7478. Meanwhile, very coarse silt and coarse silt comprise only about 3.7142 km² and 0.0193 km² respectively. Besides understanding the hydrodynamic effect of the sediment distribution, mud (silt loam) area also serves to understand the health status of the sea. The fine size of the silt loam makes it vulnerable for pollutants to be absorbed. Through the knowledge of sediment transport, researcher can assess the rate of dispersion of the pollutants. Besides, different substrate composition might harbor different benthic habitat. Table 5 shows the estimated total coverage for sediment.

| Table 5. Estimated total coverage for sediment area (km²) |
|---------------------------------------------------------|
| Sediment Type         | Area (km²) |
|-----------------------|------------|
| Very coarse sand      | 6.7478     |
| Coarse sand           | 13.457     |
| Very coarse silt      | 3.7142     |
| Coarse silt           | 0.0193     |
| Total                 | 23.9383    |
4. Conclusion
Analysis of size sediment of the sediment sample reveals that natural reef (rock) contains at least 4 types of sediment. The sediment was classified based on its texture from triangle plot, thus classifying the sediment as very coarse sand, coarse sand, very coarse silt and coarse silt. Furthermore, multibeam echosounder analysis helps to establish a backscatter image with values measured in dB that reflects its strength. Integration of both data creates the sediment map using maximum likelihood classification. In a nutshell, the map helps to understand the distribution of sediment along with the understanding of the physical parameter where hydrodynamic influence like tides, water velocity and depth stratified the sediments. The coarse sand is also has the highest coverage area of at natural reef (rock), followed by very coarse silt, very coarse sand and coarse silt. The overall accuracy and kappa coefficient obtained which are 92.31% and 0.88 respectively shows promising results. Future works on another site are suggested, especially in different sediment bottom type.

Acknowledgements
The authors gratefully to the Ministry of Education (MyBrain 15), University Malaysia Terengganu (UMT) School of Marine and Environmental Sciences and Universiti Teknologi Malaysia (UTM) Razak School of Engineering and Advanced Technology, for research facilities and financial assistance provided for this project.

References
[1] Castro, P. and Huber, M.E. (2010). Marine Biology (8th Edition) . New York: McGraw Hill.
[2] Prothero, D.R. & Schwab, F. (1996). Sedimentary Geology: An introduction to sedimentary rocks and stratigraphy. W.H. Freeman and Company, New York, pp. 1-25.
[3] Brown, C.J., Todd, B., Kostylev, J., Pickrill, R.A., V.E., 2011. Image-based classification of multibeam sonar backscatter data for objective surficial sediment mapping of Georges Bank, Canada. Continental Shelf Research 31, S1 10- S1 19.
[4] 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.
[5] Lecours, V., Dolan, M. F. J., Micallef, A., & Lucieer, V. L. (2016). Characterising the ocean frontier: A review of marine and coastal geomorphometry. Hydrol. Earth Syst. Sci. Discuss., 2016(March), 1–66. http://doi.org/10.5194/hess-2016-73
[6] Huang, Z., 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, 37e52.
[7] Blott, S. J., Pye, K. (2001) - GRADISTAT: a grain size distribution and statistics package for the analysis of unconsolidated sediments. Earth Surface Processes and Landforms, 26, 1237-1248.
[8] Udden, J. A. (1914) - Mechanical composition of clastic sediments. Bulletin Geological Society of America, 25, 655-744.
[9] Wentworth, C. K. (1922) - A scale of grade and class terms for clastic sediments. Journal of Geology, 30, 377-392
[10] Brown CJ, Collier JS (2008) Mapping benthic habitat in regions of gradational substrata: An automated approach utilising geophysical, geological, and biological relationships. Estuarine, Coastal and Shelf Science 78: 203–214.
[11] Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L., Versace, V.L., 2011. Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations. Cont. Shelf Res. 31, S28eS38
[12] Che Hasan, R.C., Ierodiaconou, D., Rattray, A., Monk, J. & Laurenson, L. 2011, Applications of multibeam echosounder data and video observations for biological monitoring on the south east Australian continental shelf. ISG & ISPRS. Shah Alam, Malaysia

[13] Rositter, D.G., 2014. Statistical methods for accuracy assessment of classified thematic maps. Department of Earth System Analysis

[14] Rosenfield, G. H., and K. Fitzpatricklins. 1986. A coefficient of agreement as a measure of thematic classification accuracy. Photogrammetric Engineering and Remote Sensing 52 (2):223-227.

[15] Fung, T., & E. Ledrew. 1988. The determination of optimal threshold levels for change detection using various accuracy indices. Photogrammetric Engineering and Remote Sensing 54 (10):1449-1454.

[16] Dicks, S. E., and T. H. C. Lo. 1990. Evaluation of thematic map accuracy in a land-use and land-cover mapping program. Photogrammetric Engineering and Remote Sensing 56 (9):1247-1252.

[17] Jansen, L. L. F., & F. J. M. van der Wel. 1994. Accuracy assessment of satellite derived land-cover data: a review. Photogrammetric Engineering and Remote Sensing 60 (4):419-426.

[18] Congalton, R.G., Oderwald, R.G & Mead, R.A., 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. Photogrammetric Engineering and Remote Sensing 49 (12):1671-1678.

[19] Congalton, R.G., R.G Oderwald, & Mead, R.A., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of. Environment 37 (1):35-46.