BATHYMETRY EXTRACTION FROM SPOT 7 SATELLITE IMAGERY USING RANDOM FOREST METHODS

Kuncoro Teguh Setiawan[1,*], Nana Suwargana[1], Devica Natalia BR Ginting[1], Masita Dwi Mandini Manessa[2], Nanin Anggraini[1], Syifa Wismayati Adawiah[1], Atriyon Julzarika[1], Surahman[3], Syamsu Rosid[4], Agustinus Harsono Supardjo[4]

[1]Remote Sensing and Applications Center,
Indonesian National Institute of Aeronautics and Space (LAPAN)
[2]Departement of Geography,
Indonesia University
[3]Hidrography and Oceanography Center,
Indonesia NAVI
[4]Departement of Physics,
Indonesia University
*E-mail: kunteguhs@gmail.com

Received: 13 December 2018; Revised: 8 August 2019; Approved: 20 August 2019

Abstract. The scope of this research is the application of the random forest method to SPOT 7 data to produce bathymetry information for shallow waters in Indonesia. The study aimed to analyze the effect of base objects in shallow marine habitats on estimating bathymetry from SPOT 7 satellite imagery. SPOT 7 satellite imagery of the shallow sea waters of Gili Matra, West Nusa Tenggara Province was used in this research. The estimation of bathymetry was carried out using two in-situ depth-data modifications, in the form of a random forest algorithm used both without and with benthic habitats (coral reefs, seagrass, macroalgae, and substrates). For bathymetry estimation from SPOT 7 data, the first modification (without benthic habitats) resulted in a 90.2% coefficient of determination (R^2) and 1.57 RMSE, while the second modification (with benthic habitats) resulted in an 85.3% coefficient of determination (R^2) and 2.48 RMSE. This research showed that the first modification achieved slightly better results than the second modification; thus, the benthic habitat did not significantly influence bathymetry estimation from SPOT 7 imagery.

Keywords: *bathymetry, random forest, SPOT 7*

1 INTRODUCTION

Bathymetry is essential information about the depth of water and underwater topography. As an archipelagic country, Indonesia requires information on the depth of its waters to aid its position as a maritime axis in developing its economy and for reasons of national sovereignty. In achieving this there are constraints on updating the depth information for Indonesian waters needed to update water-depth maps (Manessa et al., 2016). This situation challenges Indonesian researchers to develop methods and technologies that are effective and efficient.

The Indonesian government, through the National Aeronautics and Space Agency, has been able to utilize remote sensing technology to obtain depth information from shallow waters. Remote sensing technology can obtain information spatially and temporally at a relatively lower price than the conventional direct-measurement method used to obtain such information. The conventional method also has the disadvantage of not being able to reach

http://dx.doi.org/10.30536/j.ijreses.2019.v16.a3085

very shallow waters (Jawak, Vadlamani, & Luis, 2015) and areas with reef bases (Kanno, Kiobuchi & Isobe, 2011). Research by Manessa et al., (2016) identified the potential for effective and efficient remote sensing technology to be used to compile and revise natural resource information. In addition, these technologies are useful for supporting resource planning and management. Other areas of research that can be supported by remote sensing technology are spatial planning, marine environment and aquaculture (Hell et al.2012). Remote sensing technology, especially optical imagery, works on the spectrum of electromagnetic waves by utilizing sunlight. The amount of sunlight penetrating a water object depends on the ability of the water to absorb sunlight. The greater the absorption capacity of the waters, the less likely it is that the water can be penetrated by sunlight. According to Lillesand and Kiefer (1994), the lowest absorption capacity of water lies in wavelengths of 400–600 nm. According to Jagalingam, Akshaya, and Hegde (2015), in clear water conditions, remote sensing technology can detect up to a depth of 30 m.

In utilizing remote sensing technology to obtain information on depth, one of the challenges faced is in identifying the appropriate accuracy of the information collected. To address this, the utilization of image data with various resolutions and different extraction methods has been carried out by, among others, Pacheco, Horta, Loureiro, and Ferreira (2015); Jagalingam et al., (2015); Vinayaraj, Raghavan, and Masumoto, (2016) and Pushparaj and Hegde (2017) using Landsat 8 OLI data. Arya, Winarso, and Santoso (2016); Kanno et al. (2011) and Manessa et al., (2016) utilized images from SPOTs 6 and 7 in estimating bathymetry. The method of estimating bathymetry using Worldview data is used by Kanno, Tanaka, Kurosawa, & Sekine (2013); Yuzugullu and Aksoy (2014); Eugenio, Marcello, and Martin (2015); Manessa et al. (2016); Guzinski et al. (2016) and Hernandez and Armstrong (2016), all of which research makes use of images with better spatial resolution, namely, that provided by Worldview.

Manessa, Haidar, Hartuti, and Kresnawati (2017) conducted bathymetry mapping in shallow sea waters containing coral, seagrass, macroalgae and substrate cover. This approach is supported by research conducted by Budhiman, Winarso and Asriningrum (2013) which suggests that the taking of training samples from bottom water substrate has very different radians. In the present study, using random forest analysis the researchers examine and analyze the influence of each of the basal habitat objects in shallow waters on the extraction of bathymetric information. The four objects used in the modelling process are corals, seagrass, macroalgae and substrates. The purpose of this study is to investigate the effect of benthic habitat on bathymetry extraction using SPOT 7 satellite imagery and random forest methods for the Gili Matra Islands, Lombok, West Nusa Tenggara, and to determine the accuracy delivered by the model.

This research was conducted at this location because it contains both clear waters and the four basic habitat objects found in shallow marine waters: coral, seagrass, macroalgae and substrates. These conditions relate to the requirements that must be met when using remote sensing technology and to the model used. In addition, the location is a national marine conservation area, so the research can be used to support conservation management efforts as these relate to the produced bathymetry maps.

2 MATERIALS AND METHODOLOGY

2.1 Location and data

The research location was the shallow marine waters of Gili Matra, including the waters of Gili Trawangan, Gili Meno and Gili Air NTB (Figure 2-1). The data used in this study are SPOT 7 satellite image data with multispectral 6 m spatial resolution. Image recording time was 28 June 2018 at 10:12:49 Central Indonesian Time. The hydrographic survey data was collected during a field survey conducted from June 22 to 28, 2018, using a single beam echosounder and a differential global positioning system.

2.2 Method

This research was carried out in several stages; namely, measurement of depth data in situ with an echosounder, analysis of tidal data, and image processing. In-situ depth data is corrected with tidal data to obtain corrected depth data.

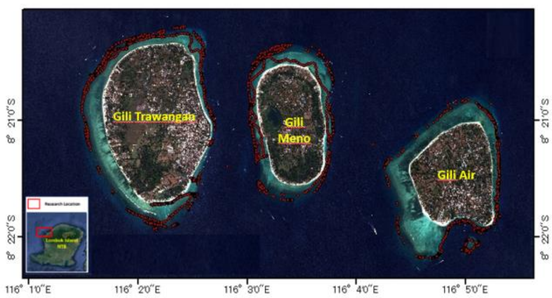

Figure 2-1: Research location image.

Image processing of satellite imagery includes radiometric and atmospheric correction. Bathymetry extraction is the process of determining sea depth information from remote sensing imagery. The bathymetry extraction in this study uses SPOT 7 satellite imagery processed using the random forest method. The bathymetry calculation is processed with modifications to field data usage.

The formula for bathymetry extraction using random forest methods is shown in Equation 2-1 (Manessa et al., 2017):

$$\hat{h} = \frac{1}{m}\sum_{j=1}^{m} W_j(X_{blue}, X_{blue}') + \frac{1}{m}\sum_{j=1}^{m} W_j\left(X_{green}, X_{green}'\right) + \frac{1}{m}\sum_{j=1}^{m} W_j(X_{red}, X_{red}') + \varepsilon \qquad (2\text{-}1)$$

where Wj (Xi, Xi') is a non-negative weight from training point i relative to new point x' in the same stage, and m is the number of stages.

To analyze the effect of benthic habitats on bathymetry accuracy, the field data used was modified. The first modification was to use all in-situ depth data without regard to benthic habitat. The second modification was the use of internal depth data separated into benthic habitats consisting of coral reef, seagrass, macroalgae and substrates. The modifications were compared to find the best accuracy value from the methods used.

2.3 Accuracy test

The accuracy of the model is calculated using the coefficient of determination R^2 and RMSE (root mean square error). The calculation process is carried out by random cross-validation experiments using 70% in-situ data with 100 repetitions. All calculation processes are carried out with R32 software. Equations 2-2 and 2-3 are used to calculate the coefficient of determination R^2 and RMSE:

$$R^2 = 1 - \sum_i\left(h_i - \hat{h}_i\right)^2 / \sum_i\left(h_i - \bar{h}\right)^2 \qquad (2\text{-}2)$$

$$\text{RMSE} = \left(\sum_{i=1}^{n}\left(h_i - \hat{h}_i\right)^2 / n\right)^{0.5} \qquad (2\text{-}3)$$

where

h = in-situ depth

$\hat{h}$ = extraction depth from SPOT 7

$\bar{h}$ = mean of in-situ depth

n = number of data

The next validation process is determining the fulfilment of accuracy standards based on IHO S44 (IHO, 2008). Bathymetric data from multiple linear models were analyzed using field data and calculated order of accuracy based on IHO-S44 standards consisting of special order, orders 1A and 1B, and second order. The criteria used were the values of total vertical uncertainty (TVU).

Table 2-1: Maximum value of TVU with 95% trust rate (Source: IHO, 2008).

Order	**A**	**B**
Special order	0.250	0.0075
Order 1A	0.500	0.0130
Order 1B	0.500	0.0130
Second order	1.000	0.0230

$$TVU = \pm \sqrt{a^2 + (b \; x \; d)^2} \qquad (2\text{-}4)$$

where:
- a = uncertainty coefficient that does not depend on depth
- b = uncertainty coefficient that depends on depth
- d = depth

There are two kinds of errors that can affect depth uncertainty; namely, errors that depend on depth and those which do not depend on depth. Equation 2-4 is used to calculate the maximum TVU. The parameters A and B for each order are shown in Table 2-1.

3 RESULTS AND DISCUSSION

Bathymetry determination at first modification uses 3254 items of in-situ measurement data. The result has a coefficient of determination R^2 of 0.902 and an RMSE value of 1.15 m. The bathymetry results from the first modification are displayed in the scatterplot of the in-situ data and the model results are shown in Figure 3-1.

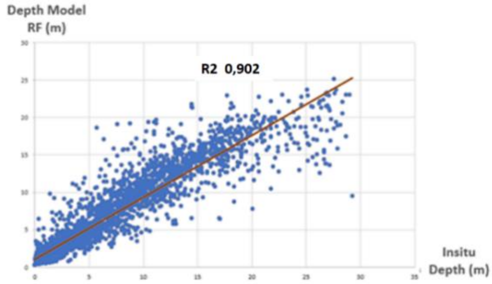

Figure 3-1: Scatterplot of first modification.

The result accuracy determination from the depth model was calculated at seven depth intervals. Calculation of efficiency is based on the IHO S44 standards shown in Table 3-1.

Table 3-1: Results of first modification accuracy (TVU, IHO S44).

Depth data (m)	Data no.	Order (%) Sp.	1A/ 1B	2	Ex. (%)	Error (m)
< 1	395	35	28	20	17	0.90
1–2	565	35	23	25	17	0.95
2.1–5	909	29	26	22	23	1.21
5.1–10	600	11	13	20	55	2.33
10.1–15	402	9	10	19	62	2.53
15.1–20	244	9	8	16	67	2.79
> 20	139	3	1	1	95	6.68

Sp. = special; Ex. = excluded

Table 3-1 shows that for 3254 items of field data divided into seven depth intervals, the accuracy values vary. At a depth of less than 1 m, 395 data items are accurate to 0.9 m. Accuracy results are grouped into four orders: 35% special order accuracy, 28% orders of accuracy 1A and IB, 20% second order accuracy, and 17% excluded. At depths of 1 to 2 m, accuracy is 0.95 m for 565 items of data: 35% special order, 23% orders of accuracy 1A and 1B, 25% second order accuracy, and 17% excluded. At depths of 2.1 to 5 m there is accuracy to 1.21 m for 909 data items, consisting of 29% at special order accuracy, 26% in orders of accuracy 1A and 1B, 22% at second order accuracy, and 23% excluded. At depths of 5.1 to 10 m there is accuracy to 2.33 m for 600 data items, consisting of 11% at special order accuracy, 13% in orders of accuracy 1A and 1B, 20% at second order accuracy, and 55% excluded. At depths of 10.1 to 15 m there is precision to 2.5 m for 402 items of detailed depth data consisting of 9% at special order accuracy,10% at orders of accuracy 1A and 1B, 19% at second order accuracy, and 62% excluded. At depths of 15.1 to 20 m there is accuracy to 2.79 m for 244 data items, consisting of 9% at special order accuracy, 8% in orders of accuracy 1A and 1B, 16% at second order accuracy, and 67% excluded. At a depth of more than 20 m there is accuracy to 6.68 m for 139 data items, consisting of 3% at special order accuracy, 1% at orders of accuracy 1A and 1B, 1% at second order accuracy, and 95% excluded.

The results of TVU shown in Table 2-1 can be seen in the form of histogram distribution in Figure 3-2. From Figure 3-2 it can be seen that, using the first random forest modification method at a depth of less than 5 m, less than 25% of results of TVU are in the excluded order, with more than 75% being spread across special order, orders 1A and1B and second order accuracy.

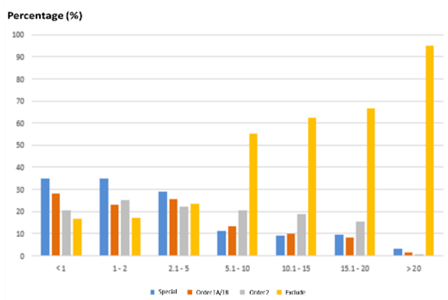

Figure 3-2: Distribution of TVU accuracy, first modification.

The calculation of the determination of the second modified bathymetry uses in-situ depth data separated according to benthic habitat; namely, coral, seagrass, macroalgae and substrates. The results of processing the bathymetry estimation with the second modification uses 3254 items of in-situ depth data separated by benthic habitat and produces coefficient of determination R^2 of 0.853 and RMSE value of 1.62 m.

The bathymetry results from the second modification are displayed in the scatterplot for the in-situ data, shown in Figure 3-3. From Table 3-2, the results of accuracy based on the IHO S44 standard show that for the 3254 data divided into seven depth intervals, the values of accuracy vary. At a depth of less than 1 m, accuracy of 0.89 m for 395 data items is obtained, consisting of 31% special order accuracy, 29% orders of accuracy 1A and 1B, 22% second order accuracy, and 18% excluded.

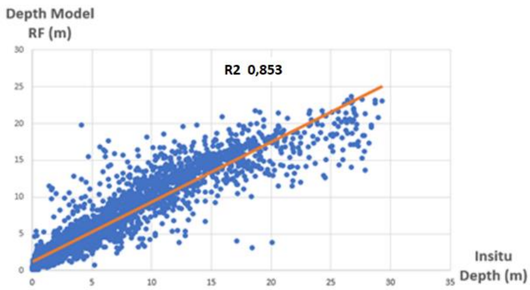

Figure 3-3: Scatterplot of second modification.

Table 3-2: Results of second modification accuracy (TVU, IHO S44)

Depth data (m)	Data no.	Order (%) Sp.	1A/ 1B	2	Ex. (%)	Error (m)
< 1	395	31	29	22	18	0.89
1–2	565	36	24	24	16	1.17
2.1–5	909	27	22	25	26	1.47
5.1–10	600	12	12	22	55	2.33
10.1–15	402	11	10	22	57	2.34
15.1–20	244	7	10	14	68	2.92
> 20	139	0	0	2	98	6.64

Sp. = special; Ex. = excluded

At depths of 1 to 2 m, there is precision to 1.17 m for 565 data items, consisting of 36% special order accuracy, 24% orders of accuracy 1A and 1B, 24% second order accuracy, and 16% excluded. At depths of 2.1 to 5 m there is precision to 1.47 m for 909 data items, consisting of 27% special order accuracy, 22% order of accuracy 1A and 1B, 25%

second order accuracy, and 26% excluded. At depths of 5.1 to 10 m there is accuracy to 2.33 m for 600 data items, consisting of 12% at special order accuracy, 12% orders of accuracy 1A and 1B, 22% second order accuracy, and 55% excluded. At depths of 10.1 to 15 m there is accuracy to 2.34 meters for 402 data items, consisting of 11% special order accuracy, 10% orders of accuracy 1A and 1B, 22% second order accuracy, and 57% excluded. At depths of 15.1 to 20 m there is accuracy to 2.92 m for 244 data items, consisting of 7% special order accuracy, 10% orders of accuracy 1A and 1B, 14% second order accuracy, and 68% excluded. At a depth of more than 20 m, accuracy to 6.64 m for 139 data items consists of 0% special order accuracy, 0% orders of accuracy 1A and 1B, 2% second order accuracy, and 98% excluded. The accuracy results based on depth intervals show that the largest special order value was achieved for the interval between 1 and 2 m, with an error of 1.17 meters.

The results of TVU from Table 3-2 above can be seen in the form of histogram distribution in Figure 3-4. It can be seen that, using the second random forest modification method at depths of less than 5 m resulted in 26% and 74% TVU entering the exclude order scattered in the order special, order 1A/1B and order 2. Therefore, it can be seen that the random forest method in the second modification results in decreased TVU accuracy values compared to the first modification.

The results of the two modifications used for field data indicate that the separation of field data by observing primary objects produced an R2 value of determination which decreased from 90.21% to 85.3% and RMSE value which increased from 1.15 m to 1.62 m. In addition, the TVU accuracy results at intervals of less than 1 m decreased, for special order accuracy from 36% to 31%, and excluded values increased from 17% to 18%. For accuracy in each interval, there is a decrease in the level of efficiency in each of the areas between the first modification and the second modification.

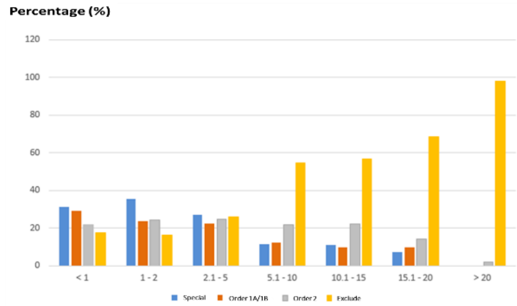

Figure 3-4: Distribution of TVU accuracy, second modification.

Depth-data distribution from the two modification methods is shown in Figure 3-5. It shows that the extraction depth of the first modification is always better than the second modification, with both following the same pattern. Both modifications are relatively good at depths of less than 5 m, as evidenced by the difference between the extraction results of the depth being quite small.

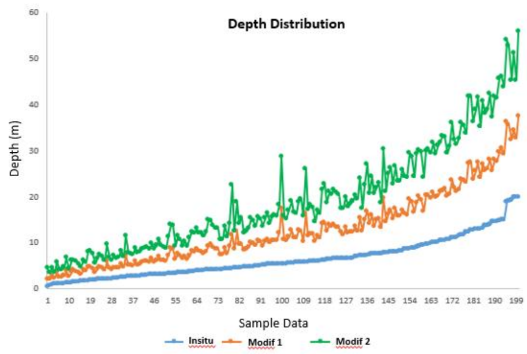

Figure 3-5: Depth distribution.

It is evident that the random forest method used with the second modification in shallow marine waters has no effect in increasing the accuracy of the produced bathymetry. When compared with previous research by Setiawan et al. (2018) using analysis of multi linear regression, the second modification provides an increase in the coefficient of determination R^2 value from 0.721 to 0.786 and decreases the RMSE amount from 3.3 m to 2.9 m. Manessa et al.'s (2017) random forest methods, also carried out in Gili Trawangan waters of Gili Meno and Gili Air, produce coefficient of determination R^2 of 0.45. The study used SPOT 6 image data gathered in 2013.

The decrease in the coefficient R^2 and the resulting RMSE value indicate that other factors influence the results of the bathymetry estimation.

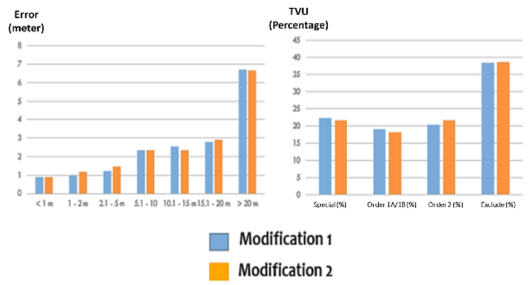

Figure 3-6: Error results.

The bathymetry extraction using the random forest method for SPOT 7 data was carried out by modifying the data usage field. The first modification is to use all in-situ depth data without regard to benthic habitat objects, producing better depth information, which can be seen in Figure 3-7.

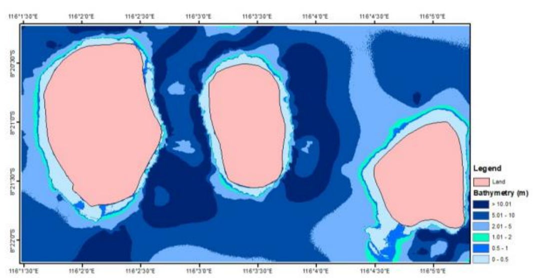

Figure 3-7: Bathymetry information from first modification.

The value of R^2 produced in this study using the first modified random forest method is 0.902 and RMSE is 1.57 m. while the R^2 using the second modified random forest method is 0.853 and RMSE is 2.48 m.

4 CONCLUSION

SPOT-7 satellite imagery can extract bathymetry using the random forest method in the shallow sea waters of Gili Trawangan, Gili Meno, and Gili Air in West Nusa Tenggara Province. The extraction process uses the random forest method by making two modifications, resulting in a decrease in the coefficient of determination R^2 from 0.902 to 0.853 and an increase in the RMSE value from 1.57 m to 2.48 m. The second modification, separating the depth field data based on the cover of the benthic habitat into the coral, seagrass, macroalgae, and substrates, does not improve the accuracy of the results of the bathymetry determination based on SPOT 7 satellite data.

ACKNOWLEDGMENTS

The author expresses his gratitude for the support provided by the Ministry of Research, Technology, and Higher Education through the National Innovation System Research Incentive Programme in 2018 and Remote Sensing Applications Centre, LAPAN.

REFERENCES

Arya, A., Winarso, G., & Santoso, A. I., (2016). Ekstraksi Kedalaman Laut Menggunakan Data SPOT 7 di Teluk Belangbelang Mamuju (Accuracy assessment of satellite derived bathymetry using Lyzenga method and its modification using SPOT 7 data at the Belangbelang Bay Waters, Mamuju). *J Ilm Geomatika, 22*, 9–19.

Budhiman, S., Winarso, G., & Asriningrum, W. (2013). Pengaruh Pengambilan Training Sample Substrat Dasar Berbeda pada Koreksi Kolom Air Menggunakan Data Penginderaan Jauh. *Jurnal Penginderaan Jauh, 10*(2), 83–92. LAPAN. Indonesia.

Eugenio, F., Marcello, J., & Martin, J. (2015). High resolution maps of bathymetry and benthic habitats in shallow-water environments using multispectral remote sensing imagery. *IEEE Trans Geosci Remote Sens 53*, 3539–3549. doi: 10.1109/ TGRS.2014.2377300

Guzinski, R., Spondylis, E., Michalis, M., Tusa, S., Brancato, G., Minno L., & Hansen, L. B. (2016). Exploring the utility of bathymetry maps derived with multispectral satellite observations in the field of underwater archaeology. *Open Archaeol 2*, 243–263. doi: 10.1515/opar-2016-0018

Hell, B., Broman, B., Jakobsson, L., Jakobsson, M., Magnusson, A., & Wiberg, P. (2012). The use of bathymetric data in society and science: A review from the Baltic Sea. *AMBIO 41,* 138–150.

Hernandez,, W. & Armstrong R., (2016). Deriving bathymetry from multispectral remote sensing data. *J Mar Sci Eng 4*, 8. doi: 10.3390/jmse4010008.

IHO (2008). *IHO standards for hydrographic surveys* (5th ed.). Special Publication No. 44, Monaco.

Jagalingam, P., Akshaya, B. J., & Hegde, A. V. (2016). Bathymetry mapping using Landsat 8 satellite imagery. *Procedia Engineering 116* (2015), 560–566.

Jawak, S. D., Vadlamani, S. S. & Luis, A. J. (2015). A synoptic review on deriving bathymetry information using remote sensing technologies: Models, method, and comparisons. *Advances in Remote Sensing, 4*, 147–162.

Kanno, A., Koibuchi, Y., & Isobe, M. (2011). Shallow water bathymetry from multispectral satellite images: Extensions of Lyzenga's method for improving accuracy. *Coastal Engineering Journal, 53*(4), 431–450.

Kanno, A., Tanaka, Y., Kurosawa, A., & Sekine, M. (2013). Generalized Lyzenga's predictor of shallow water depth for multispectral satellite imagery. *Mar Geod, 36*, 365–376. doi: 10.1080/01490419.2013.83997

Lillesand, T. M. & Kiefer, R. W. (1987). *Remote sensing and image interpretation* (2nd ed.). Canada.

Manessa, M. D. M., Kanno, A., Sekine, M., Ampou, E. E., Widagti, N., & As-syakur, A. R. (2016). Lyzenga multispectral bathymetry formula for Indonesian

shallow coral reef: Evaluation and proposed generalized coefficient. In: C. H. Bostater, X. Neyt, C. Nichol, & O. Aldred, (Eds). *Proc. Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions*. Edinburgh, UK: SPIE.

Manessa M. D. M., Haidar, M., Hartuti, M., & Kresnawati, D. K. (2017). Determination of the best methodology for bathymetry mapping using SPOT 6 imagery: A study of 12 empirical algorithms. *International Journal of Remote Sensing and Earth Sciences*, *14*(2) 127–136.

Pacheco, A., Horta,, J., Loureiro C., Ferreira, O. (2015). Retrieval of nearshore bathymetry from Landsat 8 images: A tool for coastal monitoring in shallow waters. *Remote Sens Environ. 159*, 102–116. doi: 10.1016/j.rse. 2014.12.004

Pushparaj, J., & Hegde, A. V, (2017), Estimation of bathymetry along the coast of Mangaluru using Landsat-8 imagery. *Int J Ocean Clim Syst 8*, 71–83. doi: 10.1177/17593 131166

Setiawan, K. T., Manessa, M.D.M., Winarso, G., Anggraini, N., Giarrastowo, G., Asriningrum, W., ... Supardjo, A. H. (2018). Bathymetry estimation Of SPOT 7: Case study of Gili Matra West Nusa Tenggara waters. *Journal of Remote Sensing and Digital Image Data Processing*, *15*(2). doi: 10.30536/j.pjpdcd.2018.v15.a3008

Vinayaraj, P., Raghavan, V., & Masumoto, S. (2016). Satellite-derived bathymetry using adaptive geographically weighted regression model. *Mar Geod 39*, 458–478. doi: 10.1080/01490419.2016.12452

Yuzugullu, O. & Aksoy, A. (2014). Generation of the bathymetry of a eutrophic shallow lake using WorldView-2 imagery. J. *Hydroinformatics 16*(50). doi: 10.2166/ hydro.2013.133