Regression kriging analysis for predicting the shallow depth water from Sentinel-2 satellite multi-spectral images, study area: Coastline of Florida, USA.

B G Dewanto 1,4, I Arifianto 2,4*, C Suhendi 3,4 and M Fittipaldi 4

1 Department of Geodetic Engineering, Faculty of Engineering, Universitas Gadjah Mada (UGM), Yogyakarta 55281, Indonesia.
2 Department of Geological Engineering Faculty of Engineering, Universitas Gadjah Mada (UGM), Yogyakarta 55281, Indonesia.
3 Geophysical Engineering, Institute Technology of Sumatera (ITERA), Lampung Selatan 35365, Indonesia.
4 Physical Science and Engineering Division, King Abdullah University of Science and Engineering, Thuwal 23955-6900, Kingdom of Saudi Arabia.

*E-mail: indra.arifianto@mail.ugm.ac.id

Abstract. The shallow depth water mapping has become important to the study of morphology and resources management of the coastal area. Moreover, for city or urban planning, it can be used for determining the proper location for seaport or tourist destination areas such as diving spots, coral reef monitoring, etc. However, the acquisition of shallow depth water data in a large area is somehow costly. In this paper, we represent an approach of mapping the high accuracy bathymetry with input from open satellite access data and some point measurement of bathymetry in the field (can be from LiDAR, Fathometer, Single Beam, or Multi-Beam). The study area is located in the Florida coastline, the USA that has satellite data from Sentinel-2 while the bathymetry is from a single beam survey. The method is combining satellite-derived bathymetry (SDB) with the regression kriging analysis, which shows a better depth water prediction compared to the SDB alone or the ordinary kriging method. The statistical result of the bathymetry shows the regression kriging has a better mean value, standard deviation and coefficient correlation compared to the true bathymetry value. Thus, this method can be utilized as an alternative method to map shallow depth water.

1. Introduction
Shallow depth water mapping is one of the important surveys to be conducted in the scientific study of the marine environment, geomorphology, coastal, and marine planning management [1]. Bathymetric information also very useful in the nautical navigation for shipping traffic safety in the seaport, this application is necessary to map the underwater sand bars, rock, shoals, reefs, and other marine feature that can be dangerous for a big ship in the shallow water docking [2]. In marine science, bathymetric data is important for modelling coastal storm surge, flooding, circulation of ocean water, and tsunami wave propagation. There is global bathymetric data derived from satellite marine gravity measurements like Geosat and ERS-1 satellite radar altimeters that can be accessed and downloaded free online such as from GEBCO, NOAA, or TOPEX website. This global bathymetric is cover shallow to the deep-water area, but the spatial resolution is from 400 meters to 4 km. This global bathymetric data resolution is not sufficient in many coastal applications that need detailed bathymetry information.
The traditional bathymetry acquisition is by shipborne echo sounding may not be feasible for several cases of shallow waters due to sound saturation, narrow survey swath, and inaccessible path for a big vessel [3]. The other popular shallow water bathymetry survey is the airborne survey by light detection and ranging systems (LiDAR). The airborne survey has the advantage of rapid and high accurate bathymetric measurement (4 m spatial and 20 cm vertical accuracy) in the shallow water area, but this survey is a relatively high-cost operation for a large area [4]. This condition aims for another approach that more cost-effective alternative, various passive remote sense multispectral images (i.e., satellite image) have been used to obtain water depth estimation in the shallow water area. The methodology applying satellite data known as satellite-derived bathymetry (SDB) either linear or log-linear regression have been started since 1980 such as [5], [6], [7], [3], [4], and [8]. Although this method has a good spatial resolution (4 m to 10 m), the vertical accuracy has a root mean square errors of around 2 meters [4] [9] [10].

A new approach to minimize the vertical accuracy based on the SDB method has been applied by some researchers from machine learning algorithms [10] to a simple prediction by kriging method as Best Linear Unbiased Estimator (BLUE) methods [9]. In this research, we want to apply regression kriging as the better prediction tool for shallow depth estimator using the multi-spectral image from satellite data. The idea is by combining deterministic approach from SDB with the stochastic variable is acquired from residual value from SDB and field samples, it is similar to [9] but instead of using IKONOS satellite and LiDAR as true data, we use Sentinel-2 satellite and single beam and fathometer survey as true data. The reason of using the Sentinel-2 is because it is an opensource data and it also has the SAR image which one or both reason is not available in the other platform. Since this method requires field samples, we choose to perform this method in the area that has open bathymetric data as true depth water for calibration. Florida coastline the USA has complete bathymetry survey from Fathometer, single beam to LiDAR data (figure 1). Thus, the objectives in this research; firstly, we want to implement the regression kriging in the Florida coastal area to estimate the bathymetry from multi-spectral satellite Sentinel-2 image. The more important thing, we want to compare the water depth estimation from regression kriging with the ordinary kriging and log-linear regression SDB method.

Figure 1. Florida USA coastline area with the bathymetric data and Sentinel-2 satellite image.

Figure 2. SDB concept (Su et al., 2008).
2. Methodology
Satellite observations include removing the bottom radiance from the measured water-leaving reflectance for bathymetry calculation [11]. The ratio of the radiance leaving the water surface to the downwelling irradiance just above the water surface is known as the reflectance R. The core principle of bathymetry derivation using satellite measurements is based on the dynamics of water column light pass attenuation, bottom reflection, and refraction (figure 2). The mathematical equation can be written as Equation 1:

\[ L_{\text{TOA}} = L_B + L_V + L_S + L_A \]  

(1)

where the radiance \( L_{\text{TOA}} \) located at the sensor. \( L_A \) is the atmospheric scattering, \( L_S \) is the reflection of optical energy from the water surface, \( L_V \) is the subsurface volumetric radiance, and \( L_B \) is the result of energy reflection from the seafloor (gives the information about bottom scattering water depth). Disaggregating the bottom and volumetric radiance from the total radiance is very important for deriving the depth of seafloor from satellite observation.

From the Sentinel 2 satellites images, we can get the information about Multi-Spectral Instrument (MSI) with 12 spectral bands. The main bands for the SDB consist of red, green, and blue bands with 5 m, 15 m, and 30 m as the resolutions, respectively. In this research, we only use the green and blue bands. Then, in the pre-processing step, the conversion from the reflectance to radiance (L) is needed which can be seen in Equation 2:

\[ L = \frac{\text{pixelValueBandY \times cos(incidenceAngle) \times solarIrradianceBandY \times phi \times d2}}{10000} \]  

(2)

where \( d2 \) is 1 divided by reflectance conversion coefficient. The figure 3 shows us about the conversion process in this study.

Kriging is often favoured by different spatial interpolation techniques because it enables one to calculate and manipulate the spatial autocorrelation and joint dependency between neighbouring observations to get an optimized approximation with the corresponding uncertainty estimate for the variable values at unsampled locations. The most widely used kriging variant, ordinary kriging (OK), is used in this analysis as the benchmark system for comparison.
\[ \tilde{z}(s) = \sum_{j=1}^{n} w_j z(s_j); \quad \sum_{j=1}^{n} w_j = 1 \]  

where \( \tilde{z}(s) \) is the estimated water depth at location \( s \), \( z(s_j) \) is the observed water depth at location \( s_j \), \( n \) is the number of water depth sample points used in the calculation, and \( w_j \) is the weight assigned to the sample point at \( s_j \). In conjunction with the secondary variable (auxiliary information), the spatial autocorrelation of the primary variable (densities) is sampled with the primary variable. In the case of RK, the water depth at a new unsampled location \( s \) is predicted by summing the predicted drift and residual (see figure 5).

\[ \tilde{z}(s) = \tilde{m}(s) + \tilde{e}(s) \]  

where the drift \( \tilde{m}(s) \) is estimated from the log-linear inversion model. Lyzenga (1978, 1981) derived water depth from inverting multispectral satellite images by using a log-linear inversion model (Equation 5), that assuming the ratio of the bottom reflectance between two spectral bands is constant for all bottom types within a given scene. \( L(\lambda_i) \) is the water-leaving radiance measured by a remote sensor for spectral band \( \lambda_i \). \( L_\infty(\lambda_i) \) is is the deepwater radiance for spectral band \( \lambda_i \). \( \alpha_i (i = 0, 1, ..., N) \) are the model parameters to be determined, and \( N \) is the number of spectral bands to be used.

\[ \tilde{m}(s) = \alpha_0 + \sum_{i=1}^{N} \alpha_i + \ln [L(\lambda_i) - L_\infty(\lambda_i)] \]  

and the residual \( \tilde{e}(s) \) are interpolated using the ordinary kriging method. \( e(s_j) \) explains the regression residuals from the log-linear inversion model and \( w_j(s) \) tell us about the weights. The histogram residual of measured bathymetry and residual error can be seen in figure 4a and 4b, respectively. Then, 4c shows us about the covariogram vs variogram of residual. Our research’s workflow of can be seen in figure 6.

\[ \tilde{e}(s) = \sum_{j=1}^{n} w_j(s) e(s_j); \quad \sum_{j=1}^{n} w_j(s) = 1 \]
Figure 4. Histogram residual of measured bathymetry (a), histogram residual of residual error (c), and covariogram vs variogram of residual (c).

Figure 5. Map of estimated bathymetry using log linear regression (a) and map of the residual R using OK method (b).

- Convert reflectance to radiance.
- Neglect the deep water-band (see C. Fiener, 2013)
- Estimating the variogram and covariogram model
- Do OK to estimate the residual map
- Cut the land part
  - RK vs OK
  - RK vs GEBCO
  - Estimating $z_i$
    $\tilde{z}(s) = \hat{m}(s) + \tilde{e}(s)$

Figure 6. Research workflow.
3. Result

In our research, the radiance value from the land is not used. In the other word, we only use the radiance value from the water area. Jagalingam (2015) used the NIR for separating the radiance of water from the land. From their study, the NIR value anomaly detected in land area, while the smooth result shown in water area.

For comparison to our result, we use grid data from 15 arc-second GEBCO from various method and NOAA Bathymetry data acquired from Fathometer method. The GEBCO data comes from the combination between single beam echosounder, lidar, and satellite images surveys that acquired in 1983. Then, the fathometer data was coming from which was collected in 1930.

In the GEBCO and NOAA data, we also did the gridding process by using OK method. The results of these two processes can be seen in figure 7. The water depth resolution of GEBCO is 400 m, then for the fathometer survey is 150 m and 500 m.

Figure 7. Water depth data acquired from GEBCO (a) and fathometer survey NOAA (b). The water depth calculation from Sentinel 2 satellite image by using log linear method can be seen in figure 8a. It has 22 m as the resolution and compared to the GEBCO and fathometer data, it has better comparison in shallow depth water. Then, we also tried to use the OK method for estimating the water depth from the satellite image. The result shows us that the water depth estimation is much better than by using the log linear, especially in deeper water. The OK method result is shown in figure 8b.
In the next step, we used the result from the log linear and OK estimations to calculate the water depth using the RK method (see figure 8c). From the result, we can see that it looks similar with the water depth detection from OK. We will explain more details about the difference between log linear, OK, and RK estimations results in the discussion section.

Overall, ordinary kriging and regression kriging maps has high similarity to the GEBCO and fathometer data. These two methods show more detail maps compare to the log linear map. The water depth is explained by the blue to red colours that means deep to shallow bathymetry, respectively.

4. Discussion

Based on the bathymetric profile (figure 9), the East-West profile clearly shows that the ordinary kriging cannot capture the slope of the beach or the near shore it is due to no data constrain in the land area, more over the range of variogram is too far from the nearest data point. While from the satellite derive bathymetry (SDB) profile shows that the slope is still capture although it steeper than the real data from NOAA fathometer or GEBCO data where the inaccuracy of SDB is caused by no bathymetry data in the near shore therefore we do not have good gradient at the near shore area. The regression kriging method profile show the best estimation in the near shore area, where it accounts for the SDB bathymetry plus the residual error from the regression of the OK method although the result still quite steep [9].

The other interesting feature in the SDB is that the SDB failed to estimate a basinal area in the middle of the area (figure 8) where the profile also shows that the bathymetry just dipping toward the east or left side (figure 9). This is maybe due to several problem such as atmosphere condition, sunlight correction, deep water correction or there is habitat or substrate reflectance which cause the SDB cannot estimate the water depth accurately [12]. In this case we are neglected the atmosphere and deep-water correction because choose the clear satellite image and relatively shallow water less than 30 meter where the blue light still can penetrate clear down to 30 m depth. The possible reason for this is we do not perform most of the corrections, however the result is still acceptable, and we argue that in this basinal area error might be due to bottom radiance which integrates the information about water depth and bottom characteristics including habitat and substrate.
Figure 9. Water depth cross section of the GEBCO, fathometer, log linear, ordinary kriging, and regression kriging, for the section map see figure 8(c).

Beside from qualitative analysis, we also perform statistical analysis to convincing that the bathymetry estimation result has good accuracy. We use the NOAA Fathometer acquired in 1930 as the real bathymetry data because it has close direct measurement data point, and with assumption there is no significant change in seafloor surface due to current or storm activity since it has depth below the fair-weather wave base (figure 9). The other reason is GEBCO data has resolution of 400 meter which is bigger than the NOAA data set, moreover it also has some North-South trace or relict in the bathymetry data due to many data sources accounted in the gridding process (figure 7a). The statistic we performed is the mean of the profile, the correlation coefficient that we compare to the true data (NOAA Fathometer) and the standard deviation of the bathymetry. As we can see that the RK method has better accuracy from SDB and OK method. However, the OK and RK have statistically close value which mean if we increase the number of the bathymetry data the result of OK and RK will be the same. In the other side, the OK can perform better than RK for the data set with strong spatial autocorrelation in the primary variable even when the correlation with a strong second variable. The RK method may be more suitable for water depth estimation with high variable in water depth (low spatial autocorrelation) and might be not useful for the area without too much variation in the water depth (high spatial autocorrelation).

5. Conclusions
We successfully perform bathymetry estimation with good accuracy using Regression Kriging method (RK) compiled with Satellite Derived Bathymetry method (SDB) from an open online database. The accuracy of the bathymetry estimation from SDB-RK is dependent on the number and the distribution of the calibration depth point where the SDB suppose not to be standalone without calibration data or the result will have very low accuracy. The Bathymetry acquired from SDB-RK is qualitatively and quantitatively better than bathymetry prediction only from Ordinary Kriging (OK) or SDB where the more real bathymetry data the result of OK will be similar to RK and the accuracy of SDB is dependent on the heterogeneity spatial of the shallow water condition, such suspension clay-sand...
material, pelagic micro algae, also weather when the image is taken. Furthermore, SDB-RK result in our area is comparable with direct measurement (LiDAR or Multibeam) survey.

Acknowledgments
We would like to say thank you to Prof. Sigurjon Jonsson and Dr. Ibrahim Hoteit as the instructors in the Data Analysis course in Earth Science and Engineering Department, King Abdullah University of Science and Technology so that this paper can be done as the final project of the course. We also would appreciate NOAA and GEBCO organization for opening the bathymetry data into public and providing the dataset available online.

References
[1] Finkl C W, Benedet L and Andrews J L 2005 Submarine geomorphology of the continental shelf off Southeast Florida based on interpretation of airborne laser bathymetry. *Journal of Coastal Research*, 21:1178–1190.
[2] Jupp D L B 1989 Background and extension to depth of penetration (DOP) mapping in shallow coastal waters. *Conference on Remote Sensing Application Projects Using microBRIAN Image Processing System*. pp.IV.2.1-19, 12 refs, illus.
[3] Stumpf R P, Holderied K and Sinclair M 2003 Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnology and Oceanography* 48:547–556.
[4] Su H, Liu H, and Heyman W D 2008 Automated Derivation of Bathymetric Information from Multi-Spectral Satellite Imagery Using a Non-Linear Inversion Model, *Marine Geodesy*, 31:4, 281-298.
[5] Lyzenga D R 1981 Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data. *International Journal of Remote Sensing* 2:71-82.
[6] Lyzenga D R 1985 Shallow-water bathymetry using combined lidar and passive multispectral scanner data. *International Journal of Remote Sensing* 6:115–125.
[7] Philpot W 1989 Bathymetric mapping with passive multispectral imagery. *Applied Optics* 28:1569–1578.
[8] Jagalingam P, Akshaya B J and Hegde A V 2015 Bathymetry Mapping Using Landsat 8 Satellite Imagery. *Procedia Engineering* 116: 560 – 566.
[9] Su H, Liu H and Wu Q 2015 Prediction of Water Depth from Multispectral Satellite Imagery—The Regression Kriging Alternative, *IEEE Geoscience and Remote Sensing Letters*, Vol. 12.
[10] Misra A, Vojinovic Z, Ramakrishnan B, Luijendijk A and Ranasinghe R 2018 Shallow water bathymetry mapping using Support Vector Machine (SVM) technique and multispectral imagery, *International Journal of Remote Sensing*, 39:13, 4431-4450
[11] Jensen J R 2007 Remote sensing of the environment: An earth resource perspective, 2nd ed. Upper Saddle River, NJ: Prentice Hall.
[12] Lyzenga D R, Malinas N P and Tanis F J 2006 Multispectral bathymetry using a simple physically based algorithm. *Geoscience and Remote Sensing, IEEE Transactions on* 44:2251–2259.