# COMPARATIVE ANALYSIS ON INTERPOLATION METHODS FOR BATHYMETRIC DATA GAPS

Joseph E. Acosta (1), Vicente Calag (1), Zarah Jean Diche (2), Christine Lou Lazaga (2), Trisha Mae Lopena (2), Daisuke Tatsuda (2), Gus Kali R. Oguis (2)

[1]Department of Mathematics, Physics, and Computer Science, College of Science and Mathematics, University of the Philippines Mindanao, Mintal, Davao City, Philippines 8022
[2]Project 5 Geo-informatics for the Systematic Assessment of Flood Effects for a Resilient (Geo-SAFER) Southeastern Mindanao, College of Science and Mathematics building, University of the Philippines Mindanao, Mintal, Davao City, Philippines 8022
Email: jeacosta@up.edu.ph

**KEY WORDS:** Bathymetry, Interpolation, Hydrocorrection

**ABSTRACT:**

Light Detection and Ranging (LIDAR) technology delivers high accuracy elevation values and ground features. However, the capability of this technology is inhibited in terms of its strength to penetrate certain surfaces. For instance, LIDAR is limited to the elevation values of the river water surface and not the elevation of its riverbed. Hence, topographic and bathymetric surveys are conducted to obtain an accurate set of elevation values for areas where the technology is unable to permeate. Bathymetric surveys are conducted using a scientific echosounder equipment, which utilizes sonar technology to determine the river depth relative to the water's surface by transmitting sound pulses and calculating the interval between the emanation and regress of a pulse with respect to time. Like in all remote sensing measurements, errors are inevitable. Noise and external factors that cause faulty or bad readings result to data gaps. Gaps in the gathered elevation data sets can only be identified during filtering, which is done after actual survey. In addition, covering the gaps back in the field would mean additional costs. This study aims to optimize data gathered by using different interpolation methods to simulate points in the data gaps. Inverse Distance Weighted (IDW), Spline, and Kriging methods are used to extrapolate the values within the gaps. Statistical calculations are shown to analyze the accuracy and efficiency of the results.

## 1. INTRODUCTION

Raw LIDAR-derived elevation models have high accuracy, elevation z-values for terrain and surface models. However, the considered elevation values are limited to the above water level measurement of topographical features. In order to enhance the capability of this technology to provide as much information in a data as possible, geodetic leveling activities are conducted. To obtain elevation values or underwater depth and map certain underwater features, bathymetric surveys are carried out. Hydrographic and bathymetric surveys are necessary for various kinds of studies, such as scour and stabilization, flood inundation and mapping, spill and fill, and other research studies. Various bathymetric survey techniques with corresponding survey-grade equipment sets are employed for different purposes of hydrographic measurements.

Integration of the data obtained from these hydrographic surveys are then implemented in order to hydrologically correct the depth of the specific water body. Similar to the variation of techniques employed during bathymetric survey, there are also a selection of bathymetric interpolation methods to cater to the variation of bathymetric survey methods, bathymetric points density, and various types of water bodies. The integration of both elevation datasets in topography and bathymetry is vital to the completion of the digital elevation model.

The objective of the study is to compare several interpolation methods of river bathymetry data by studying the calibration bathymetry points and calculating the root mean square error of the generated bathymetry elevation model to the validation bathymetry points. This study deals with a comparison among spatial interpolation methods for computing elevation or z-values in data gaps of bathymetric data used to measure the elevation of the riverbed in meters above sea level (MASL), then integrated to the light detection and ranging (LIDAR) derived digital elevation model, which has a 1 meter resolution, to rectify the elevation of the riverbed by interpolating the topographic surface with the elevation values of obtained by survey-grade equipment. Manual surveys are conducted in order to gather bathymetric data which represents the elevation values which can be randomly distributed and may sometimes carry erratic values and/or not carry sensible information at all due to instrument limitations. Hence, interpolation of the values of the points in the necessary segments are performed in order to predict the missing values, in lieu of its error or absence, using the neighboring sets of points.

## 2. REVIEW OF RELATED LITERATURE

### 2.1 Digital Elevation Model (DEM)

The LIDAR-derived digital elevation model (DEM) is classified into two models – digital terrain model (DTM) and digital surface model (DSM) – which showcases different topographical features. DTM refers to the topographic configuration of the bare Earth (Chen, Gao, & Devereux, 2017). DSM contains elevation values of the features found on the surface of the Earth, including both man-made and natural objects (Chen, Gao, & Devereux, 2017); (Maune, Kopp, Crawford, & Zervas, 2007). Using a matrix structure in a raster (grid) format, elevation values and topological relations between points in grid cells are recorded to form the DEM (Ramirez, 2006). The resolution of a grid DEM is equivalent to the grid size of the DEM, which reflects the ground distance (Liu, 2008).

### 2.2 Bathymetry

Hydrographic surveys are conducted to acquire data from and/or involving water surfaces. These hydrographic surveys focus on the measurement and data collection of the bottom of any form of waterbody, such as oceans, lakes, and rivers (NOAA, 1976). Bathymetric surveys are directed towards obtaining data, specifically, elevation, in this case, the elevation of the riverbed. Bathymetric surveys are customarily performed using acoustic echo sounding equipment, which can return accurate depth profiles (Gao, 2009) which is calculated from the interval between the return times of the pulses on the surface (Klemas, 2011). However, acoustic echo sounding is bound to certain limitations. Among its constraints with efficiency and accessibility (Gao, 2009), acquisition of bathymetric data by echo sounding on shoal waters poses difficulties due to certain environmental conditions and technical considerations (Tronvig, 2005).

### 2.3 Interpolation

Interpolation is a mathematical process of approximation, which determines a set of values for parameters or points given the values of its neighboring data (Mitas and Mitasova, 2005). In the geographic information systems (GIS) environment, interpolation methods are programmed to predict values given a set of discrete or continuous data. Interpolation has practical uses in data management, specifically known data alongside missing data, where long-term cycles are known

(Kaya, 2014). However, no specific interpolation methods are strictly prescribed for use on bathymetric data (Curtarelli, Leão, Ogashawara, Lorenzzetti, & Stech, 2015).

**2.3.1 Inverse distance weighting (IDW):** The inverse distance weighting or IDW interpolation method is a local neighborhood approach which makes use of the values of its nearest neighbors by distance to derive a set of neighboring values (Watson and Philip, 1985). IDW builds its basis on the premise that values at the data gaps (unsampled locations) can be estimated using the weighted average values of the points at a certain neighboring distance (Mitas and Mitasova, 2005), given that these weights are inversely proportional to a given distance (Watson, 1992). IDW is calculated as:

$$Z_j = \frac{\sum_{i=1}^{n} \frac{Z_i}{(h_{ij} + \delta)^{\beta}}}{\sum_{i=1}^{n} \frac{1}{(h_{ij} + \delta)^{\beta}}} \tag{1}$$

where $Z_j$ is the unknown value to be interpolated, $Z_i$ is the known value, $\beta$ is the weight, $\delta$ is the smoothing factor, and $h_{ij}$is the separation distance, calculated as:

$$h_{ij} = \sqrt{(\Delta x)^2 + (\Delta y)^2} \tag{2}$$

where $\Delta x$ and $\Delta y$are the distances between the unknown point $j$ and the known point $i$ according to reference axes (Mitas and Mitasova, 2005).

Using IDW interpolation in Arcmap requires the Power variable. Its purpose is to determine the influence of the sample points used to determine the value. Its value must be greater than zero. The higher the Power, the more influence the nearest points have on the value. (University of Namur Department of Geography, n.d.)

**2.3.2 Kriging:** Kriging interpolation is a geostatistical approach which takes the values of neighboring points and their respective locations as basis for estimation of the values of the specific points at a location (Kiš, 2016); (Longley, Goodchild, Maguire, & Rhind, 2010) primarily revolving around the principle that point values near sampled locations should be assigned a greater weight in approximating the values for prediction in unsampled locations to improve its accuracy (Kiš, 2016), with the assumption that the distance (with respect to the location) between point values in sampled locations have a spatial correlation that can be a basis for the variation of the surface (Childs, 2004).Kriging interpolation method was originally designed for approximations in the mining industry (Tang, 2005), developed by Georges Matheron and Daniel Krige, with principle on the theory of regionalized variables (Kerry & Hawick, 2005). Ordinary Kriging (OK) is a commonly utilized kriging method and is referred to as best linear unbiased estimator (Kiš, 2016). The Kriging algorithm is expressed as:

$$Z(S_0) = \sum_{i=1}^{n} \lambda_i Z(S_i) \tag{3}$$

where $n$ is the number of values, $\lambda_i$ is the weight for the measured value at the $i^{\text{th}}$ location, and $S_0$represents the location of the value to for prediction.

**2.3.3 Spline:** Spline interpolation is a piecewise polynomial interpolation, which approximates values by using mathematical functions and splines to fit values into several fixed points with values (Ikechukwu, Ebinne, Idorenyin, & Raphael, 2017) while minimizing the curvature of interpolated surface (Childs, 2004). In comparison to IDW, the spline interpolation method is designed to consider point values outside a minimum-maximum range of values in the sample data during the process of estimation (Liu, 2008) which carries the advantage of this method in predicting values in ridges and valleys (Childs, 2004). Spline interpolation algorithm is represented as:

$$S(x, y) = T(x, y) + \sum_{j=1}^{n} \lambda_j R(r_j) \tag{4}$$

where $n$ is the number of points, $\lambda_j$ are coefficients found by the solution of a system of linear equations, and $r_j$ is the distance from *(x, y)* to $(x_j, y_j)$.

There are two types of Spline interpolation in Arcmap: Spline and Spline with Barriers. Spline only requires the Weight factor. The value has to be greater than zero. At the minimum 0.1, Spline will try to closely match the data, and at greater values, the more smooth the fit will be. (Smith, 2015) Another, called, Spline with Barriers, utilizes breaklines in order to constrain the influence of closer points that are considered coincident points (ArcGIS, 2016).

## 3. METHODOLOGY

### 3.1 Topography and Bathymetry Dataset

The bathymetric dataset used for this study is the length of Dapnan River located in the Municipality of Baganga, Davao Oriental, Philippines. The bathymetric points were obtained using South S86 and Trimble Survey Grade Global Navigation Satellite System (GNSS) receivers, in a combination of zigzag, cross-section, and centerline manner along the length of the river with a total stretch of approximately 22 km. The DTM used for integration with the combination of surveyed and interpolated bathymetric data was acquired using LIDAR technology, with a 1 meter by 1 meter resolution, on 2014 by the Data Acquisition Component of the Disaster and Risk Exposure Assessment for Mitigation (DREAM) Program.

### 3.2 Bathymetric Data Integration

To conduct the study, a continuous stream of actual bathymetric data was selected as a baseline for comparison. The chosen study area is a portion of Dapnan River with ideal conditions for the baseline, with a length of 1 km. The survey path covers 1.29306 km, traversing the center as well as some embankments in a zigzag manner. The area is home to an uninterrupted stream of 213 sample data points.

**3.2.1 Data Gap:** An artificial data gap was made with 51 continuous points of the 213 total were chosen to act as the data gap. These were situated in the middle of the sampled location, with 81 points before it and another 81 points after. This was made to ensure that there are sufficient value samples for each side of the gap, as basis for the prediction of values in the data gap during the application of the different interpolation methods and algorithms.

**3.2.2 Interpolation Methods:** In this study, three interpolation methods were used to extrapolate the values for the gap. The best results – those closest to the actual values – for each method was used as the actual values for bathymetric data integration. Table 1 shows the common base parameters required for all the methods. IDW method used five (5) different values in the Power parameter: 0.25, 0.5, 2, 3, and 6 for both With Barriers and Without Barriers technique. Kriging method was implemented in five different Semivariogram Models, namely: Spherical, Circular, Exponential, Gaussian, and Linear. Spline method was implemented using both Regularized and Spline with Barriers techniques with values: 0.1, 0.5, and 1. The parameter with smallest Root Mean Square Error (RMSE) is chosen.

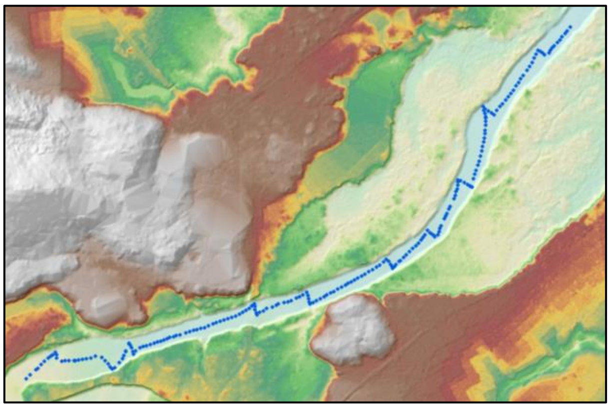

Figure 1. The bathymetric points superimposed to the DTM

### 3.2.3 Comparison of Interpolation Methods:

Using the results from 3.2.2, each result was integrated into the LiDAR DTM using IDW. A comparison is made by comparing their Root Mean Square Error (RMSE) and Standard Deviation values. The sample bathymetric points, including those points with values derived from interpolation, were integrated to form an interpolated surface. The resulting interpolated surface of bathymetric data points were analyzed and compared in terms of RMSE, calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \tag{5}$$

where *n* is the population and *x* is the deviation of the elevation values in comparison (Chai and Draxler, 2014).

## 4. RESULTS AND DISCUSSION

The results were computed using the stated parameters in 3.2.2 and are presented in Tables 2, 3, and 4 together with their Root Mean Square Errors (RMSE) and Standard Deviations. The barriers was observed to have improved the results for IDW and Spline methods. Comparing the results, the parameters that will be used with the respective interpolation method are: (1) IDW with barriers at the 6th Power; (2) Linear semivariogram model for Kriging, and; (3) Spline with barriers at 0.1 Smoothing Factor. The approximated values derived from the three interpolation methods applied to identical dataset were incorporated to the existing values obtained from ground survey. Table 5 outlines the results of each method, with RMSE, average, and standard deviation values for the validation points comprising 20% of the bathymetric data points. In each of the three interpolation methods, the best fit technique according to RMSE value, with respect to the individual parameter settings and models, are compared for analysis. Figure 2 and 3 shows a graph of the trend line of IDW method; Figures 4 to 8 shows a line graph of each semivariogram model of the Kriging method compared to the original dataset; and Figure 9 shows the graphical representation of the interpolated values using Spline method.

Table 1. Common parameters present in all methods

| Parameter | Constant Value |
|---|---|
| Cell Size | 1 |
| Number of Points | 12 |
| Search Radius (for IDW and Kriging methods) | Variable |

Table 2. Parameters and results for IDW method

| Power | Without Barriers | | With Barriers | |
|---|---|---|---|---|
| | RMSE | Standard Deviation | RMSE | Standard Deviation |
| 0.25 | 0.45004 | 0.43700 | 0.40128 | 0.36770 |
| 0.5 | 0.44985 | 0.43622 | 0.36100 | 0.33262 |
| 2 | 0.44718 | 0.43101 | 0.27528 | 0.27357 |
| 3 | 0.44460 | 0.42766 | 0.26476 | 0.26705 |
| 6 | 0.42954 | 0.41193 | 0.25406 | 0.25399 |

Table 3. Parameters and results for Kriging method

| Semivariogram Model | RMSE | Standard Deviation |
|---|---|---|
| Spherical | 0.39040 | 0.31398 |
| Circular | 0.39035 | 0.31370 |
| Exponential | 0.39498 | 0.32073 |
| Gaussian | 0.44242 | 0.32147 |
| Linear | 0.39030 | 0.31342 |

Table 4. Parameters and results for Spline method

| Value | Regularized Spline (Weight) | | Spline with Barriers (Smoothing factor) | |
|---|---|---|---|---|
| | RMSE | Standard Deviation | RMSE | Standard Deviation |
| 0.1 | 9.64757 | 6.94313 | 0.48298 | 0.35641 |
| 0.5 | 6.67182 | 6.57728 | 0.65400 | 0.39609 |
| 0 | 4.93911 | 4.97314 | 0.65526 | 0.39365 |

Table 5. Validation results for each method

| Method | RMSE | Average | Standard Deviation |
|---|---|---|---|
| IDW with Barriers (Power: 6) | 0.19764 | -0.03543 | 0.19674 |
| Spline with Barriers (Smoothing Factor: 0.1) | 0.19669 | -0.03633 | 0.19559 |
| Ordinary Kriging (Linear semivariogram) | 0.19646 | -0.04266 | 0.19404 |

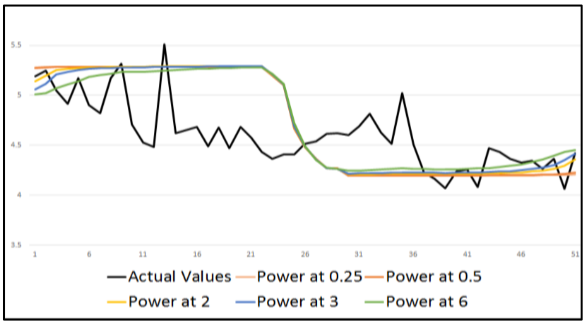


Figure 2. Inverse Distance Weighting Interpolation Method (Without Barriers)

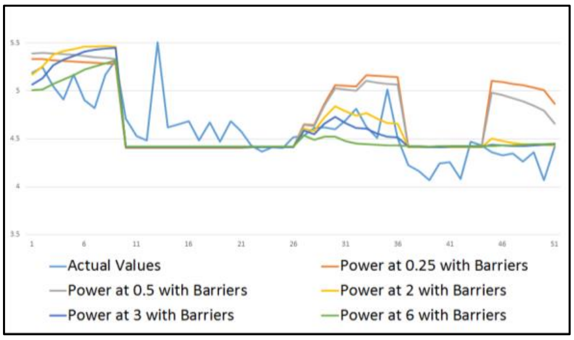


Figure 3. Inverse Distance Weighting Interpolation Method (with barriers)

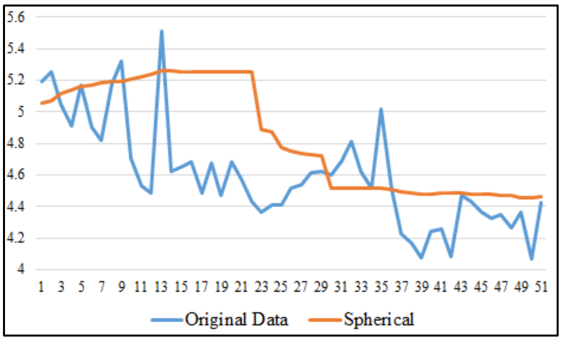


Figure 4. Kriging Interpolation Method (Spherical Semivariogram Model)

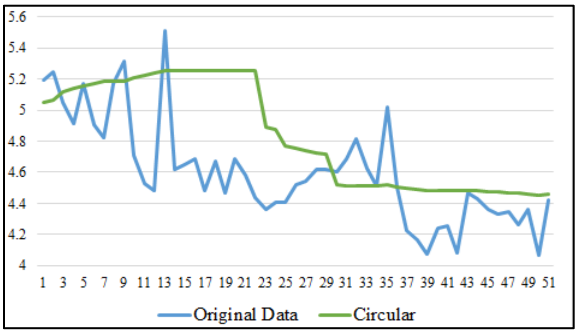


Figure 5. Kriging Interpolation Method (Circular Semivariogram Model)

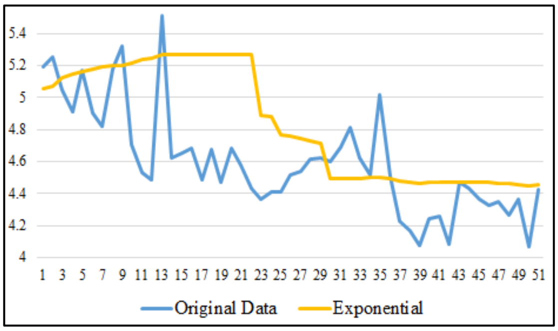


Figure 6. Kriging Interpolation Method (Exponential Semivariogram Model)

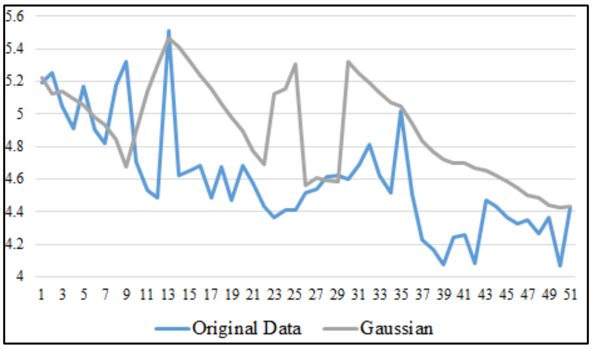


Figure 7. Kriging Interpolation Method (Gaussian Semivariogram Model)

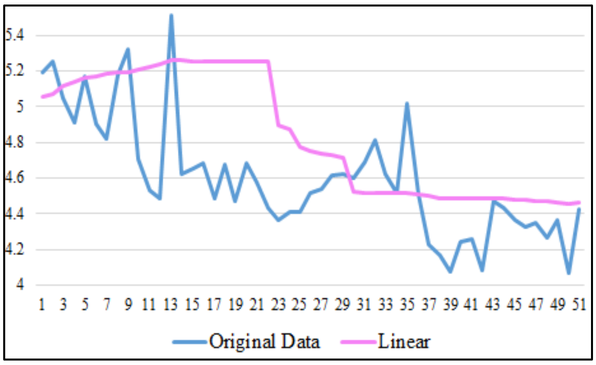


Figure 8. Kriging Interpolation (Linear Semivariogram Model)

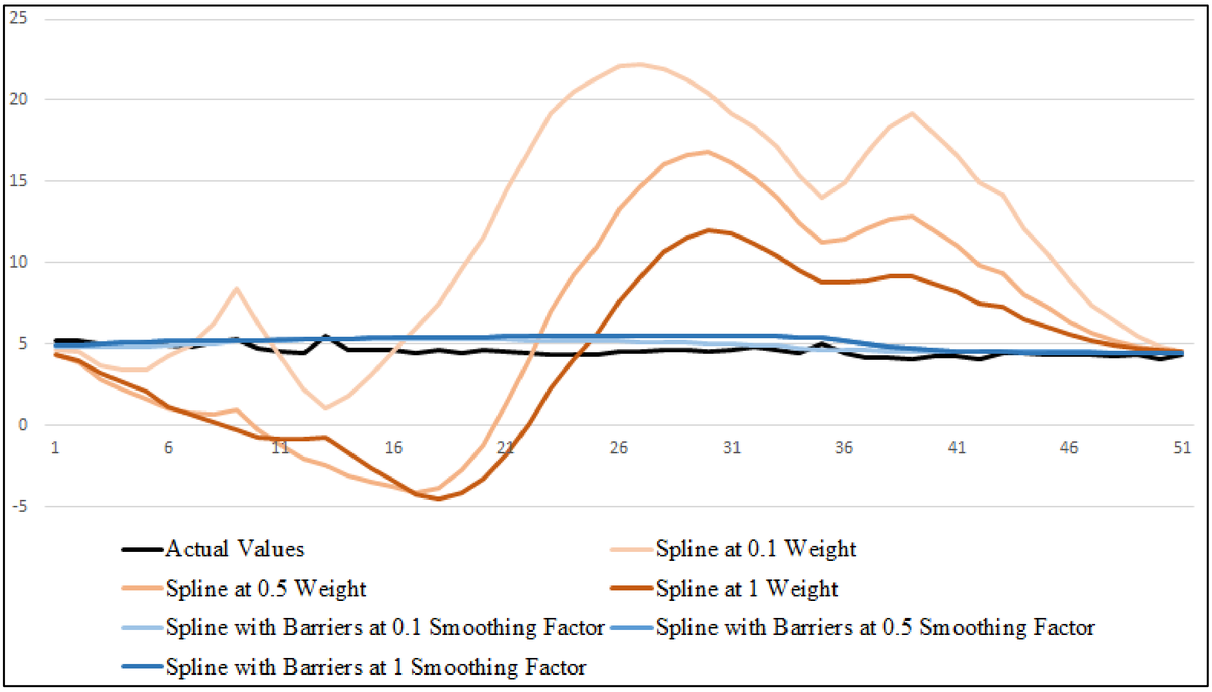


Figure 9. Spline Interpolation Method

## 5. CONCLUSION AND RECOMMENDATION

The bathymetric data integration of the combination of survey-acquired values and predicted values by interpolation techniques is feasible and necessary in the absence of values, considering mishaps during data gathering and field survey. However, specific interpolation techniques and algorithms must be employed to suit the dataset's behavior and application. This study limited its comparison to three interpolation methods and findings show that their performance is nearly similar (Table 5).  Given the RMSE values, the average error computed is at 19 cm, which is within the range of acceptable RMSE of 20 cm. This study can be further developed by comparing the methods to different rivers with varying characteristics and considering other interpolation methods that are applicable to geospatial datasets, while modifying their parameter settings and observing their deviations as these parameters vary.

## ACKNOWLEDGMENT

The authors extend their gratitude for the support of the Department of Science and Technology – Philippine Council for Industry, Engineering and Emerging Technology Research and Development (DOST – PCIEERD), the University of the Philippines Mindanao, and their colleagues in Geo-SAFER Southeastern Mindanao.

## REFERENCES

ArcGIS, 2016. Spline with Barriers. Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/spline-with-barriers.htm.

Chai, T. & Draxler, R. R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? - arguments against avoiding RMSE in the literature. In: Geoscientific Model

Development, 7, 1247-1250. Göttingen, Germany: Copernicus Publications. https://doi.org/10.5194/gmd-7-1247-2014.

Chen, Z. Gao, B. Devereux, B. 2017. State-of-the-art: DTM generation using airborne LIDAR data. In: Sensors 2017, 17 (1), 150. Multidisciplinary Digital Publishing Institute, Basel, Switzerland. https://doi.org/10.3390/s17010150.

Childs, C. 2004. Interpolation surface in ArcGIS spatial analyst. ArcUser, 32-35.

Curtarelli, M., Leão, J., Ogashawara, I., Lorenzzetti, J., & Stech, J., 2015. Assessment of spatial interpolation methods to map the bathymetry of an amazonian hydroelectric reservoir to aid in decision making for water management. In: ISPRS International Journal of Geo-Information, 4(1), 220-235. doi:10.3390/ijgi4010220.

Gao, J., 2009. Bathymetric mapping by means of remote sensing: methods, accuracy and limitations. In: Progress in Physical Geography, 33(1), 103-116. doi:http://dx.doi.org/10.1177/0309133309105657.

Ikechukwu, M. N., Ebinne, E., Idorenyin, U., & Raphael, N. I., 2017. Accuracy assessment and comparative analysis of IDW, spline and kriging in spatial interpolation of landform (topography): an experimental study. In: Journal of Geographic Information System, 9, 354-371. https://doi.org/10.4236/jgis.2017.93022.

Kaya, E., 2014. Spline interpolation techniques. Journal of Technical Science and Technologies, 2(47).

Kerry, K., & Hawick, K., 2005. Kriging interpolation on high-performance computers. International Conference on High-Performance Computing and Networking. HPCN-Europe 1998: High-Performance Computing and Networking, pp. 429-438.

Kiš, I.M., 2016. Comparison of ordinary and universal kriging interpolation techniques on a depth variable (a case of linear spatial trend), case study of the Šandrovac Field. In: Rudarsko Geolosko Naftni Zbornik, 31(2), 41-58.

Klemas, V., 2011. Beach profiling and LIDAR bathymetry: An overview with case studies. Journal of Coastal Research, 27(6), 1019-1028.

Liu, X., 2008. Airborne LIDAR for DEM Generation: Some Critical Issues. Centre for GIS, School of Geography and Environmental Science, Monash University, Clayton, Victoria 3800, Australia: University of Southern Queensland.

Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D. W., 2010. In: Geographic information systems and science, 3rd ed. Hoboken, New Jersey, USA: John Wiley & Sons, Ltd.

Maune, D. F., Nayegandhi, A., 2007. Digital elevation models. Digital Elevation Model Technologies and Applications: The DEM Users Manual. The American Society for Photogrammetry and Remote Sensing.

Mitas, L. & Mitasova, H., 2005. Spatial interpolation. In: Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W., (Eds.), Geographic Information Systems: Principles, Techniques, Management and Applications, 2nd ed., pp.481-492. John Wiley & Sons, Inc, Hoboken, New Jersey.

National Oceanic and Atmospheric Administration, 1976. NOAA: Hydrographic manual, 4th ed. US Department of Commerce, Rockville.

Ramirez, J. R., 2006. A new approach to relief representation. surveying and land information science, 66(1), pp.19-25.

Smith, T., 2015. Tips for Using Arcmap. Notebook for Spatial Data Analysis. University of Pennsylvania School of Engineering and Applied Science. Retrieved from https://www.seas.upenn.edu/~ese502/NOTEBOOK/Part_IV/Using_ARCMAP.pdf.

Tang, T., 2005. Spatial statistic interpolation of morphological factors for terrain development. GIScience and Remote Sensing, 42(2), pp. 131-143.

Tronvig, A., 2008. Near-shore bathymetry. In: Hydro International. Lemmer, Netherlands: Geomares Publishing.

University of Namur Department of Geography, n.d., Interpolating Raster Surfaces. Retrieved from http://webapps.fundp.ac.be/geotp/SIG/InterpolMethods.pdf.

Watson, D.F. and Philip, G.M., 1985. A refinement of inverse distance weighted interpolation. In: Geoprocessing, pp. 315-327.

Watson, D.F., 1992. Contouring: a guide to the analysis and display of spatial data. Oxford, United Kingdom: Pergamon Press.