

Performance Evaluation of GIS Interpolation Techniques to Generate 3D Bed Surfaces Profiles of Lake Nubia

Mohamed Elsahabi, and Hickmat Hossen*

1 Civil Engineering Department, Faculty of Engineering, Aswan University, Egypt;

Received: 12/3/2023 Accepted: 26/4/2023

© Unit of Environmental Studies and Development, Aswan University

Abstract:

Interpolation process is essential whenever researchers are faced with a data shortage. Within the GIS, interpolation can be conducted using all kinds of environmental phenomena represented in the form of spatial data. Remote Sensing (RS) data represented in several Landsat ETM+ images, acquired in different dates, were processed by integration with the in-situ elevations data to achieve the interpolation process. As a result, this research aims to identify the best method for interpolating Lake Nubia's bed surface elevations, the Sudanese portion of Aswan High Dam Lake (AHDL). The three techniques tested in this study include Radial Base Function [RBF], Inverse Distance Weighting [IDW], and Ordinary Kriging [OK] technique. Also, this study focuses on estimating the accumulated sediment in Lake Nubia (LN) from 2004 to 2012 via the produced 3D profiles of the lake. Since the RBF provided the lowest values for both the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE), the findings clearly demonstrated the usefulness and superiority of the method. Moreover, results indicated that the amount of sediment calculated using the RBF method is close to that of the Aswan High Dam Authority (AHDA) using the cross-section method. These results indicate that the amount of sediment calculated using the RBF method is valid and reliable. It is recommended to apply this technique to efficiently create the 3D profiles of the lakes from the available elevations data using GIS/RS.

Keywords Aswan High Dam Lake, Interpolation Methods, Sediment Amount, Remote Sensing, GIS

1- Introduction

The science of determining depths and the physical characteristics of underwater features is known as Contemporary bathymetry. This science is primarily based on measurements from hydrographic surveying that were taken to determine how the bottom of an ocean, sea, river, lake, or any other water-related object on Earth was configured (Siljeg et al., 2015). The 3D bed surfaces profiles (bathymetric profiles) of lakes are used to monitor lakes (Curtarelli et al., 2015).

Corresponding author*: E-mail address: hickmat.hossen@aswu.edu.eg
To create 3D bed surfaces profiles, various techniques are being used, including topographic data-based, remote sensing-based, and spatial interpolation techniques based on gathered spatial data (bed elevation samples). Spatial interpolation techniques are the ones that are most frequently employed to produce 3D profiles among the techniques mentioned above. This study uses spatial interpolation techniques based on depth samples collected during hydrographic surveying.

The main focus of GIS is the management, aggregation and analysis of spatial data. The ArcGIS Spatial Analyst extension provides a toolkit for analysing and modelling such spatial data (Childs, 2004). Hydrographic survey data can be used to create profiles of bed surfaces, making interpolation efficient and practical.

Presently, there are various methodologies available to perform the spatial interpolation process. These methodologies are divided into two categories techniques: deterministic and Geostatistical (Siljeg et al., 2015; Childs, 2004). Inverse Distance Weight (IDW) and Radial Basis Functions (RBF), two deterministic interpolation methods, produce surfaces based on measured points or mathematical formulas. These interpolation techniques are used on a wide range of engineering applications, other than creating profiles for 3D bed surfaces. Examples of these applications include (a) detecting the spatial distribution of heavy metals in Soil (Li et al., 2021), (b) creating fertilizer application maps (Matcham et al., 2021), (c) stream flow predictions via the hydrological modelling (Felix and Jung, 2022) and (d) modelling of 3D rock surface profiles as a kind of geological modelling technology (Yin and Zhou, 2020). Furthermore, geostatistical interpolation techniques, such Kriging, are employed for more complex prediction surface modeling. The type of interpolated phenomenon, the nature of the interpolated surfaces, or the objective of the produced 3D profiles are just a few examples of the many variables that influence the choice of the interpolation method and its parameters (Fencik and Vajsablova, 2006).

Many researches comparing and contrasting interpolation techniques have been conducted as a result of the choice of interpolation methods. There has not yet been a thorough investigation into the superiority of one strategy over another (Rodriguez, 2015). Among those studies that discuss the comparison of the interpolation methods are (Hu et al., 2004; Arun, 2013; Bello and Stefanoni, 2007; Erdogan, 2009; Naoum and Tsanis, 2004; Merwade et al., 2006; Aykut et al., 2012; Meng et al., 2013; Thanh et al., 2020; Arseni et al., 2019).

Considering the above, this study first focuses on evaluating two spatial methods (deterministic and geostatistical) for generating a reliable 3D bed surface for Nubia Lake. Secondly, we will estimate the amount of accumulated sediment in Nubia lake from 2004 to 2012 via the obtained 3D profiles by the spatial interpolation method with the best performance. This is expected to help improve the management and operation of the lake.

2 Materials and Methods

2.1 The Study Area

Due to its 500 km length south of Aswan High Dam, AHDL is one of the largest artificial lakes in Africa. With a length of about 350 km, most of this lake is located in Egypt and is referred to as Lake Nasser. With a length of 150 km, it is known as Lake Nubia (LN) on the Sudanese side (Ali, 2006). The chosen study area is LN. It is situated upstream of the Aswan High Dam, between latitudes 21° 02' 00" and 22° 00' 00" North. According to Figure 1, the northern part of
the study area is much wider than the southern part whose length is almost two-thirds of the length of the lake (Lake Nubia), (Elsahabi et al., 2016). Moreover, this area is characterized by high depth and bed relief.

![AHDL map containing Landsat ETM+ scene of study area.](image)

2.2 **In-Situ Data**

2.2.1 **Hydrographic Survey Data**

During field trips funded by (AHDA and NRI), hydrographic data were collected using an echosounder. The east-north lake geometry (E, N, Z) was used to create a 3D bed surface. Hydrographic survey data for the study area from 2004 and 2012 were used (NRI, 2012).

2.2.2 **Water Levels Data**

To determine the water surface levels at the time the satellite imagery was taken, water levels upstream of AHD are collected from AHDA gauge stations at various times throughout the year (MALR, 2012).

2.3 **Remote Sensing Data**

This study used three Landsat ETM+ images (path/row = 175/045). For GeoTIFF products, three images were downloaded from his GLCF website at various times (September 2000, March 2006, and March 2009) (systematic correction) (GLCF, 2014). The lake boundaries are extracted from the captured images. Three satellite images were captured: one in March 2006 at water level (173 m), one in March 2009 at water level (176.60 m), all at lake level (178 m) in September 2000.

2.4 **Methodology**

The adopted methodology to achieve the goals of the present work is shown in Figure 2 and described in the following subsections.
2.4.1. Water Surface Areas Extraction

It was decided to use the unsupervised classification method of Landsat images in this study because it is thought to be the most effective method for identifying water textures (Elsahabi et al., 2016; Elsahabi et al., 2015). Using WGS84, UTM Z36N, a series of scattering points (x, y, and z) were used to form surface geometry.

2.4.2. Data Exploratory

After merging the points extracted from the satellite images with the hydrographic survey points, it is necessary to check the maximum and minimum level values for all years to exclude the unrealistic values using ArcGIS software. The Exploratory Spatial Data Analysis (ESDA) (Curtarelli et al., 2015; Rodriguez, 2015) is the process of analyzing the merged points dataset statically and spatially using ArcGIS software prior to the interpolation process. This process is achieved to examine the accuracy of the dataset (Rodriguez, 2015).

The ESDA within ArcGIS can be completed using the Geostatistical Analyst toolbar through two kinds of graphs, the first called the histogram and the second one known as the Normal Quantile-Quantile (Normal QQ) plot.
Histograms should not contain values that are too far from the mean and median. The presence of outliers in a dataset can skew the distribution and cause problems with some interpolation methods. These outliers are detected by regular QQ charts. The skewness coefficient should be close to zero, which means that the mean and median are approximately the same. If your data are not normally distributed, you can fix some of the irregularities by applying a transformation (Rodriguez, 2015).

2.4.3. Spatial Interpolation Process

Utilizing three interpolation techniques (two deterministic and one geostatistical) available on the ArcGIS Geostatistical Analyst toolbar, a spatial interpolation process is carried out on the current survey. The interpolation methods evaluated in this study were:

Inverse Distance Weighting (IDW), Radial Base Function (RBF), Kriging technique (OK). The only geostatistical technique among them is OK. These techniques were selected because they are the most commonly used methods for interpolating bathymetric data in various publications around the world (Bello and Stefanoni, 2007; Merwade et al., 2006; Azpurua and Ramos, 2010; Merwade, 2009; Li and Heap, 2008).

The following provides an overview about the used interpolation methods (with abbreviations):

Deterministic methods: inverse distance weighting (IDW), RBF (radial basis function)-spline with tension (SWT).

Geostatistical method: ordinary kriging (OK).

- RBF is a precise interpolation technique that, conceptually, fits a surface using measured depth values while reducing the surface's overall curvature. This technique is demonstrated for computing smooth surfaces from numerous data points (Curtarelli et al., 2015). The RBF interpolation has several types, which slightly varied in their applications. This study will concentrate on RBF-spline with tension (SWT) method. When the measurement points do not change rapidly or are close together, spline functions give smoother results (Rodriguez, 2015).

- IDW is especially useful for narrow data sets where other fitting techniques may be subject to error. This process is very flexible when it comes to interpolating unclustered datasets. Furthermore, this approach directly implements the presumption that the value of an attribute for an unsampled location is a weighted average of the known data points in the vicinity of the unsampled location (Garnero and Godone, 2013).

- Kriging can calculate the level of uncertainty. Kriging does not represent trends and assumes spatially correlated points, a normal distribution, and data stationarity. Kriging interpolation comes in a number of slightly distinct forms, each with a variety of uses. Yes. We will concentrate on the typical kriging (OK) type in this study. The OK is an effective geostatistical method for describing spatial patterns, extrapolating the values of important variables at unsampled locations, and calculating surface uncertainty or modeling the error. Semivariogram fitting was done interactively during OK interpolation by adjusting the threshold, range, delay, and nugget values (Curtarelli et al., 2015).

To assess the interpolation techniques, we tested different parameters of each method as follows:
- For IDW, we experimented with various semiaxis lengths, neighbor counts, and performance metrics.
- For RBF, we investigated different semi-axis lengths, neighbor counts, and kernel functions for the RBF method.
- For OK, we tested various semivariogram adaptations (Curtarelli et al., 2015). This step is essential to identify the best parameters for each interpolation technique to be used in optimizing the interpolation results.

### 2.4.4 Validation of the Interpolation Methods

The performance of each interpolation method was evaluated by comparing the deviation of the estimates from the measured data using the cross-validation method (Bello and Stefanoni, 2007; Elsahabi and Negm, 2015) implemented in the ArcGIS Geostatistical Analyst Toolbar. Such methods remove samples from the dataset, apply an interpolation method to the remaining dataset, and after interpolation the result (the predicted value of the removed depth data) is compared with the samples removed from the entire dataset. (measured value). At the end of this procedure, we performed a performance comparison between interpolation methods using the following statistical indicators:

- $R^2$ indicates the degree of correlation between measured and estimated depth values.
- ME is used to determine the level of bias in the estimate.
- MAE provides a precise measurement of the magnitude of the error.
- RMSE gives an error size measurement that is sensitive to outliers.
- RI captures the relative improvement of the best method compared to other interpolation methods.

The equation for calculating each statistics indicator is shown in Table 1.

Finally, the three tested spatial interpolation methods were compared based on the statistical indicators' computed values. It should be taken into consideration that based on the cross-validation results, the best parameterization for each interpolation technique is determined.

#### Table 1. Statistical indicators for accuracy assessment of the Interpolation methods (Bello and Stefanoni, 2007)

<table>
<thead>
<tr>
<th>Concept</th>
<th>Name</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root Mean Square Error</td>
<td>RMSE</td>
<td>$\sqrt{\frac{1}{n} \sum(Mes - Est)^2}$</td>
</tr>
<tr>
<td>Relative Improvement</td>
<td>RI (%)</td>
<td>$\frac{100(RMSE_{current} - RMSE_{best})}{RMSE_{best}}$</td>
</tr>
<tr>
<td>Mean Error (Bias)</td>
<td>ME</td>
<td>$\frac{1}{n} \sum(Mes - Est)$</td>
</tr>
<tr>
<td>Mean Absolute Error</td>
<td>MAE</td>
<td>$\frac{1}{n} \sum</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>$R^2$</td>
<td>$\frac{\sum(Est - Avg.Mes)^2}{\sum(Mes - Avg.Mes)^2}$</td>
</tr>
</tbody>
</table>
2.4.5. Creation of The 3D Bed Surfaces

The original lake 3D bed surfaces for the years 2004 and 2012, which were obtained using the spatial interpolation method with the best performance, were used to create the map of changes over this time period from 2004 to 2012.

2.4.6. Establishing Maps of Changes

The (cut / fill) tool in ArcGIS Software was used to overlay the two newly created bed surfaces to create the map of changes, which depicts the zones of change (sediment / erosion). To estimate the rates of erosion and sedimentation in the study area from 2004 to 2012, this map of change was created.

3 Results and Discussion

3.1 Exploratory Data Analysis

Figure 3 depicts the spatial distribution of the bathymetric routes (collected elevation points), which were carried out during field surveying missions, for the northern part of the study area (located between latitudes 21°44'30" and 22°00'00" N upstream the Aswan High Dam).

![Fig.3 The hydrographic survey routes without (1) and with (2) the water surfaces points.](image)

When selecting an interpolation method, it's crucial to consider the normal distribution of the point dataset. In order to examine the normality in the distribution of such points, the Statistical analysis of the bathymetric point’s dataset (about 242,811 depth sample) is performed by establishing the Histogram and the QQ plot through the ArcGIS Geostatistical. Figure 4 shows the Histogram and the Normal QQ plot for the collected points dataset of year 2012 (the tested data for interpolation methods). When the histogram is examined, the distribution of the points that represent the analyst tool bar elevations is clearly skewed to the right with a skewness factor of 0.34. Additionally, it is obvious that the minimum, mean, median, and maximum of the statistical distribution of the points vary slightly. Despite being skewed and not reflecting each
other equally, the points dataset is thought to be relatively normally distributed, so no transformation or modification is needed (the irregularities are not great enough).

Fig. 4 The statistical variation plots (a) the Normal Q-Q plot for year 2012; (b) the Histogram for year 2012

The normal QQ plot's representation of the dataset contrasts it with the standard normal (Gaussian) distribution, which is represented by a solid line in this graph. This plot shows that the points dataset is spatially auto-correlated, has little to no stationarity, and does not have any trends that need to be removed.

Figures 3, 4a, and 4b, which show examples of the spatial and statistical variation found in the points dataset for the study area, demonstrate how most interpolation techniques can be applied to this dataset. This shows that all of the interpolation techniques used in this study's elevation data are valid.

3.2 Comparison of Spatial Interpolation Methods

The outcomes of using the three spatial interpolation techniques on the dataset of depth samples are shown in Figure 5. Table 2 displays the parameters for each method that produced the best results. These parameters are regarded as the fundamental criteria that influence the quality of the results produced by interpolation methods.
Fig. 5 Bed surface profile of year 2012 obtained from the different spatial interpolation methods: (a) Radial Basis Function-RBF; (b) Inverse Distance Weighting-IDW; (c) Ordinary Kriging-OK.

Table 2. The parameters of the spatial interpolation methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF</td>
<td>Neighbors = 32; length of semi-axis = 10000; kernel function = spline with tension</td>
</tr>
<tr>
<td>IDW</td>
<td>Neighbors = 48; length of semi-axis = 10000; power = 2</td>
</tr>
<tr>
<td>OK</td>
<td>Neighbors = 64; length of semi-axis = 1200; lags = 12; lag size = 200; semivariogram = Stable</td>
</tr>
</tbody>
</table>

Figure 5 demonstrated, qualitatively, that all three interpolation techniques used to calculate depth samples produced accurate results and enabled the creation of significant 3D feature models, such as the 3D bed surface profile of Nubia Lake.

3.3 Accuracy Assessment (Cross-Validation)

The outcomes of the cross-validation process are displayed in Figure 6 and Table 3.

Fig. 6 Results of cross validation analysis used to compare the interpolation methods: (a) Radial Basis Function-RBF; (b) Inverse Distance Weighting-IDW; (c) Ordinary Kriging-OK.
Table 3. Cross-validation results

<table>
<thead>
<tr>
<th>Interpolation method</th>
<th>ME (m)</th>
<th>MAE (m)</th>
<th>RMSE (m)</th>
<th>R²</th>
<th>RI (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBF</td>
<td>−0.0008</td>
<td>0.22</td>
<td>0.46</td>
<td>0.998</td>
<td>0.00</td>
</tr>
<tr>
<td>IDW</td>
<td>−0.0015</td>
<td>0.31</td>
<td>0.59</td>
<td>0.991</td>
<td>28.30</td>
</tr>
<tr>
<td>OK</td>
<td>0.009</td>
<td>0.59</td>
<td>1.07</td>
<td>0.889</td>
<td>132.61</td>
</tr>
</tbody>
</table>

When compared to the other two interpolation techniques, the RBF method had a higher correlation with in situ measurements ($R^2 = 0.998$), a lower MAE (0.22 m), and a lower RMSE (0.46 m) of the three interpolation techniques that were tested. The OK method, on the other hand, produced the worst outcomes, with a RMSE greater than 1 m and a correlation coefficient of 0.889. With an RMSE less than 0.6 m and a correlation coefficient higher than 0.99, the IDW method produced outcomes that were most similar to those of the RBF method. Only the Ok method ($ME = 0.009$ m) displayed a general tendency to overestimate the depth; the RBF and IDW methods ($ME$ lower than 0.002 m) slightly underestimated the depths at Nubia Lake.

The other two methods examined in this study, with the exception of the OK method, displayed low RMSE values. The RMSE values for the IDW and OK methods were about 28% and 132% higher, respectively, when compared to the RMSE value of the RBF method. The high sample density can be credited for the good performance of RBF and IDW. This shows and guarantees that for high density samples, deterministic methods like IDW and RBF show significantly greater improvements in prediction than geostatistical methods like OK.

Table 2 also displays the best techniques relative improvement when compared to the others. RBF method allowed at least a reduction of 132.61% in the error compared with the OK method and a reduction of 28.3% in the error compared with the IDW method.

Moreover, it must be kept in mind that because its estimates are unbiased and have low variance, the RBF technique has an inherent advantage over other interpolation techniques.

Other studies have reported similar results (Bello and Stefanoni, 2007; Meng et al., 2013; Azpurua and Ramos, 2010; Elshahabi and Negm, 2015), when comparing different interpolation methods, revealing that the deterministic methods also presented a better performance. By comparing the result of the current research regarding the accuracy of the method used to create the 3D bed surfaces profiles with another research (Elsahabi et al., 2018), it was noted that the RBF method gives more accurate results in the northern part of the current study area (wide part) than in whole lake (Lake Nubia) as in the current research.

3.4 Creation of the 3D Bed Surfaces

The 3D bed surfaces profiles are created for the years 2004 and 2012 from the RBF interpolation method, as this method gave the best performance compared with the other interpolation methods. These profiles are produced in order to estimate the amount of sediment in the study area. The created bed profiles are presented in Figure 7.
Three broad change categories were used to generalize the differences in bed levels between the created 3D bed surfaces. Map of change for the period (2004-2012) is presented in Figure 8.

3.5 **Amounts of Sediment and Erosion**

Based on statistics for change classes in the change map, Table 4 estimates the rates of erosion and sedimentation from 2004 to 2012. This Table clearly shows that, for this time period, sedimentation outpaces erosion by about 0.85 billion m$^3$.

**Table 4.** Sedimentation and erosion amounts from (2004-2012).

| Time period | amount of Sediment (B.m$^3$) | Erosion amount (B.m$^3$) | Total change amount (B.m$^3$)  
= (1)–(2) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2004-2012</td>
<td>1.241</td>
<td>0.391</td>
<td>0.850</td>
</tr>
</tbody>
</table>
3.6 Comparisons

The sediment amount estimated by AHDA was equal to 0.971 billion m$^3$ from 2004 to 2012 (Arseni et al., 2019), while the total amount of sediment estimated using the RBF technique during the same period was approximately 0.85 billion m$^3$ (see table 5). Which means that the amount of sediment calculated by the method used in this study is close to the amount of sediment calculated by the AHDA. These results indicate that the amount of sediment calculated using the RBF method is valid and reliable.

Table 5. The amount of sediment calculated by the method used in this study (RBF method) and the conventional method (cross-section method) in the period from 2004 to 2012.

<table>
<thead>
<tr>
<th>Time period</th>
<th>Amount of sediment by GIS/RS method (Bm$^3$)</th>
<th>Amount of sediment by AHDA method (Bm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004–2012</td>
<td>0.850</td>
<td>0.971</td>
</tr>
</tbody>
</table>

4 CONCLUSION

The creation of Nubia Lake's 3D bed surfaces (3D profiles) is essential for assisting decision-makers in putting best management practices for this lake into practice. In this way, we evaluate the performance of two different deterministic and one geostatistical interpolation methods for creating the 3D bed surface profile of the lake. The evaluation process was achieved through the Cross-Validation technique. Finally, the amount of sediment deposited in the lake in the period from 2004 to 2012 via the produced 3D profiles of the lake from the best performance interpolation method is estimated. The results qualitatively, showed that all tested interpolation method in this study were able to create the 3D bed profiles of the lake. By another mean, visually, all tested interpolation methods yielded similar results. When compared to the IDW and OK methods, the RBF interpolation method showed the best quantitative performance, with lower RMSE (0.46 m), lower MAE (0.22 m), and a higher correlation coefficient (0.998). The results also revealed that the Kriging method is the least accurate, resulting with at least an increase of 132.6% in the error compared with the RBF Method. Above all, the 3D bed profiles derived from the RBF method were the most suitable for estimating the amount of sediment in the study area. In addition, the results show that the amount of sediment computed by the RBF method is about 13% less than the amount of sediment computed by the AHDA method, based on supplementary cross-sections. The RBF method gives more accurate results in regions that are homogeneous in their geometries than in regions that have two parts that are different in its geometry. The authors highly recommended to apply the RBF technique in areas with similar conditions to efficiently create the 3D bed profiles of these areas. Moreover, it is recommended to study the effect of the geometry of lakes on the accuracy of the RBF technique in further research.

References


MALR (2012). The Ministry of Agriculture and Land Reclamation, Egypt, the General Authority for AHDL Development, AHDL levels (1978 to 2012).


