Selecting Optimal Data-Fitting Model for Surveying and Geodetic Applications

Olayemi Joseph ODUMOSU, Fisayo Funmilola IDOWU and Chika Vincent OKOROCHA, Nigeria

Key words: Terrain Modelling, Geodetic Applications, Interpolation.

SUMMARY

This Paper evaluates degree of suitability of six spatial interpolation techniques in representing a set of discreet point values as a continuous surface. The various classes of interpolants (geostatistical, non geostatistical and combined) were all examined in producing a digital Elevation Model of Lagos State.

It was discovered that the Kriging and Radial Basis Functions provided the best visually acceptable results and gave the least RMSE. However, due to their long processing time, it is recommended that for large area projects with data points exceeding 1000, the TIN could be used as it requires a less processing time and when the data points are large and closely spaced provides an accuracy similar to kriging.
Selecting Optimal Data-Fitting Model for Surveying and Geodetic Applications

Olayemi Joseph ODUMOSU, Fisayo Funmilola IDOWU and Chika Vincent OKOROCHA, Nigeria

ABSTRACT
Geoid Computation, Topographic Surveying, Modelling of geo-hazards, datum transformation, image registration and all surveying and geodetic tasks require analytical interpolation of mid-point values having obtained certain bounding conditions via field observations especially as coverage area increases. Besides, field data is often accompanied by noise irrespective of the degree of refinement and accuracy involved in the data gathering process. While separation of noise from signal can be achieved via least squares collocation, interpolation of mid-point signal values requires that appropriate analytical interpolation technique is employed. Six interpolation models are herein presented and the optimal for terrain modelling determined based on the derived residuals and predictive error estimates.

INTRODUCTION:
In order to make any meaningful analysis as regards risk assessment, platform stability, terrain modelling and most survey related applications; spatial continuous data are required. Although, with the advent of space based observation techniques that produce raster-form products, continuous data can easily be obtained for every point with minimal computational effort they are costly and the accuracy of results obtainable from such system remains questionable. Besides, coupled with the difficulty it presents in precise point selection the method is not the best for spatial analysis where high level accuracy is required.

The geodesist is thus left with no other option than to acquire point data at selected locations and then seek for empirical means to fit the discreet points into a model such as to generate value at other spatial locations with least error. In other words, the geodesist is faced with the task of finding the mathematical model that best fits the data-set such as to allow prediction at other points with least error residual.

Several spatial interpolation techniques exist to solve these problems and they can generally be classified into three (3) categories namely: (1) non-geostatistical methods (2) geostatistical methods (multivariate or univariate) and (3) combined methods (Jin and Andrew, 2008). Each category having certain advantage over the others. Many factors including sample size, sampling design and data properties affect the estimations of the method and there are no consistent findings about how these factors the performance of the spatial interpolator therefore it is difficult to select an appropriate spatial interpolation method for a given input dataset (Burrough and McDonnells, 1998)
Five techniques are herein presented with a view of determining the best interpolation technique for terrain analysis with a given dataset.

2.0 MATHEMATICAL BACKGROUND:

The most basic mathematical formulation behind almost all spatial interpolation techniques is as given by Webster and Oliver, 2001:

\[ z(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \]  

(1)

Where \( z = \text{Estimated Value at the point of interest } x_0 \)

\( z = \text{Observed value at the sampled point } x_i \)

\( \lambda_i = \text{Weight assigned to the sampled Point} \)

\( n = \text{Number of sampled points used for the estimation} \)

Five common techniques are considered in this work covering all the categories as briefly described by the table below:

<table>
<thead>
<tr>
<th>S/N</th>
<th>INTERPOLATION METHOD</th>
<th>CATEGORY</th>
<th>SUB-CATEGORY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TIN</td>
<td>Non-Geostatistical</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>2</td>
<td>IDW</td>
<td>Non-Geostatistical</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>3</td>
<td>Radial Basis Function</td>
<td>Non-Geostatistical</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>4</td>
<td>Ordinary Kriging</td>
<td>Geostatistical</td>
<td>Univariate</td>
</tr>
<tr>
<td>5</td>
<td>Universal Kriging</td>
<td>Geostatistical</td>
<td>Multivariate</td>
</tr>
<tr>
<td>6</td>
<td>Trend Surface with Kriging</td>
<td>Combined</td>
<td>Not Applicable</td>
</tr>
</tbody>
</table>

2.1 Triangulated Irregular Network

This method uses a series of triangles based on a Delauney’s triangulation to join all sampled points together (Jin and Andrew, 2008). It creates a surface formed by triangles of nearest neighbour points. To do this, circumcircles around selected sample points are created and their intersections are connected to a network of non-overlapping and as compact as possible triangles as shown in Figure 1 below (QGIS User Guide).
Although, this method has a relatively faster processing time compared to other interpolation methods, it does not have the capacity to extrapolate beyond “Z” data range.

The main disadvantage of the TIN interpolation is that the surfaces are not smooth and may give a jagged appearance. This is caused by discontinuous slopes at the triangle edges and sample data points. In addition, it is generally not suitable for extrapolation beyond the area with collected sample data points.

Besides, the method is not suitable when sample data points are few (data less than 300 points).

2.2 Inverse Distance Weight

This interpolation method is very versatile, easy to program and fairly accurate under a wide range of conditions (Lam, 1983). This method assumes that influence of the variable entered on the map decrease with the increase of the distance from its sampling site (Dreskovic and Dug, 2012) and is given by the equation:

\[ P_i = \frac{\sum_{j=1}^{G} P_j / D_{ij}^n}{\sum_{j=1}^{G} 1 / D_{ij}^n} \]  
Equ. 2

Where \( P_i = \text{Property @ location } i \)

\( P_j = \text{Property @ sample location } j \)

\( D_{ij} = \text{Distance from } i \text{ to } j \)

\( G = \text{No of sampled locations} \)

\( n = \text{Inverse distance weighting power} \)
The main factor affecting the accuracy of IDW is the value of the power parameter (Isaaks and Srivastava, 1989). Weights diminish as the distance increases, this results in a local spatial interpolation. Besides, because of the structure of the process algorithm, it is obvious that interpolation quality will reduce if distribution of sample points is uneven. Also, maximum and minimum values can only occur at sample data points which often results on small peaks and pits around the sample data.

IDW is referred to as “Moving Average” when $p = 0$ (Brus et al, 1996; Hosseini et al, 1993; Laslett et al, 1987), “Linear Interpolation” when $p = 1$ and “Weighted Moving Average” when $p$ is not equal to 1 (Burrough and McDonnell, 1998).

![Figure 2: Conceptual view of IDW interpolation technique.](image)

2.3 Radial Basis Functions (RBF)

This is closest to the Kriging technique and is a flexible interpolation method. It gives the best overall interpolation of most datasets (Surfer help file). It is an exact interpolator that utilises basic kernel functions that are analogous to Variograms in Kriging (Ojigi, 2011). Types of RBF include: Inverse Multiquadratic, Multilog, Multiquadratic, Natural cubic Spline and Thin plate Spline.

2.4 Ordinary Kriging

Kriging involves interactive investigation of spatial behaviour of data analysed before selecting the best method of assessment for derivation of output area (Oliver, 1989). Spatial variation is quantified by semi – variograms which is calculated from number of the data input point sets.

After selecting the semi-variogram, it is possible to use smaller size of the grid cell in creation of the actual output grid (Dreskovic and Dug, 2012).

The estimation of an ordinary kriging is based on the formula below:

$$Z(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) + [1 - \sum_{i=1}^{n} \lambda_i] \mu(s_0)$$
Where \( Z(s_0) \) = Predicting Location

\[
\lambda_i = \text{Unknown weight of the measured value of pairs of points @ ith location}
\]

\( Z(s_i) \) = Measured values of pairs of points @ ith location

\( n \) = number of measured values.

Here the \( \mu(s_0) \) (mean of samples within a search window) here replaces the ordinary \( \mu \) of the simple kriging which is assumed to be constant for the whole domain and calculated as average of the data. Also the Ordinary Kriging forces \( 1 - \sum_{i=1}^{n} \lambda_i \) to be equal to zero.

Although, all kriging methods are time-consuming, they are flexible and optimum for almost all dataset even when few data points are provided. A major advantage of kriging is that it can extrapolate for values beyond the “z” data range.

2.5 Universal Kriging

This is otherwise known as kriging with a trend. It is an extension of the Ordinary Kriging by incorporating the local trend within the neighbourhood search window as a smoothly varying function of the co-ordinates (Jin and Andrew, 2008).

2.6 Trend Surface With Kriging

TSA is fitted to the data, which describes the large scale (global) spatial variability; the residuals from TSA are then modelled using ordinary Kriging. The final estimates are the sum of the kriged residuals and the estimated trend surface (Wang et al, 2005).

3.0 DATA USED

A total of 216 points being part of the Lagos State second-order control network were used in the model formulation. The control points selected being part of the ZTT 14 – 30 Series covering most parts of Lagos State has an even spatial distribution across the state.
Figure 3: Showing Control – Network used for the Terrain – Analysis.

4.0 RESULTS AND ANALYSIS:

The dataset were plotted using the models earlier described and the results are as presented below:

Figure 4: Contour Map generated from TIN.

Figure 5: Surface Map (Wireframe) generated from TIN.
Figure 6: Contour Map generated from IDW

Figure 7: Surface Map (Wireframe) generated from IDW

Figure 8: Contour Map generated from Ordinary Kriging.
Evaluation of the various interpolation techniques reveal certain empirical values which are herein used to judge the overall accuracy of a particular technique for terrain analysis.
parameters considered include; Root Mean Square Error, Standard Error, Processing Time and Visual Appearance.

Table 2: Result Analysis

<table>
<thead>
<tr>
<th>S/N</th>
<th>TECHNIQUE</th>
<th>RMSE</th>
<th>STD. ERROR</th>
<th>TIME (Secs)</th>
<th>APPEARANCE</th>
<th>REMARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TIN</td>
<td>37.8935</td>
<td>0.2523</td>
<td>0.01</td>
<td>Poor</td>
<td>Not Suitable</td>
</tr>
<tr>
<td>2</td>
<td>IDW</td>
<td>44.4564</td>
<td>0.0208</td>
<td>0.01</td>
<td>Poor</td>
<td>Not Suitable</td>
</tr>
<tr>
<td>3</td>
<td>Kriging</td>
<td>37.5752</td>
<td>0.2416</td>
<td>0.09</td>
<td>Very Good</td>
<td>Good</td>
</tr>
<tr>
<td>4</td>
<td>Radial Basis Function (RBF)</td>
<td>37.6112</td>
<td>0.2459</td>
<td>0.07</td>
<td>Good</td>
<td>Good</td>
</tr>
</tbody>
</table>

5.0 CONCLUSION:

The Kriging and Radial Basis function are the best interpolation techniques for terrain analysis and modelling. However, consequent upon the long time it takes to process them, it is advisable that when the available dataset exceeds 1000 points, the TIN or “minimum curvature” should be used as they both produce accuracy close to kriging when a large dataset is available.

REFERENCES:


Dreskovic and Dug, (2012); Application of the inverse Distance weighting and Kriging methods of the spatial interpolation on the mapping the annual amount of Precipitation in Bosnia and Herzegovina.


Jin Li and Andrew D. Leap (2008), A review of Spatial Interpolation Methods for Environmental Scientists)

BIOGRAPHICAL NOTES:

Mr Odumosu, Joseph Olayemi  Is a first Class Graduate of the Federal University of Technology, Minna (2010) and he’s presently a Postgraduate research Student of the department of Surveying and Geoinformatics, University of Lagos, Nigeria. With several field experiences in Engineering Surveying and Country mapping, he is presently an academic staff of Department of Surveying and Geoinformatics, Federal University of Technology Minna. His research interest is in Spatial Analysis and Simulations of Earth Geodynamics.

Miss Idowu, Funmilola Fisayo  Is a Graduate of the Federal University of Technology, Minna (2010) and she’s presently an academic staff at the Polythecnic Ibadan.

CONTACTS

Name: Olayemi Joseph, ODUMOSU  
Institution: Federal University of Technology, Minna  
Address: Surveying and Geoinformatics Department, Federal University of Technology, Minna.  
City: Minna, Niger-State  
Country: Nigeria  
Tel. +2348065916462  
Fax: Nil  
Email: odumossu4life@yahoo.com  
Website: Nil

Name: Fisayo Funmilola, IDOWU  
Institution: The Polythecnic Ibadan  
Address: Surveying and Geoinformatics Department, Ibadan Polythecnic  
City: Ibadan, Oyo - State  
Country: Nigeria  
Tel. +2348166707515  
Fax: Nil  
Email: funmiforeverfirst@yahoo.com  
Website: Nil