RUGIPO RESEARCH AND TECHNICAL JOURNAL

ESTIMATION OF ELEVATIONS OF UNOBSERVED LOCATIONS USING SIMPLE KRIGING METHOD

Authors
Olomoshola, O. Olanitori, A. and Olawole, S. A.
Surveying & Geoinformatics Department,
Rufus Giwa Polytechnic, Owo, Nigeria.
olomosholapatrique@gmail.com

Abstract:

Geo-statistical interpolation predicts values for cells in a raster from a limited number of sample data points. It is the process of manipulating spatial information to extract information and meaning from the original data (Chang, 2006). The selection of an appropriate estimation method is one of the fundamental decisions in resource estimation. The effects of selecting an inappropriate estimation method can lead to 50% error in the estimate (Dominy et al., 2002). There are different methods of interpolating spatial data such as Inverse Distance Weighted (IDW) Trend, Kriging, etc, and the choice of these estimation methods is based on the spatial correlation of the spatially distributed objects. For example, the inverse distance method is unbiased but does not minimize the estimation variance, while kriging is subject to certain conditions, such as providing the best estimate possible by a linear combination of the available weighted data as well as minimizing the error variance of the estimate.
This paper presents a study of the application of a linear geostatistical estimation technique; Simple kriging, in estimation of elevations of unobserved locations. Included in the study is a discussion on variography and its necessity in resource estimation. The theory of kriging as well as the simple kriging equations is discussed. Simple Kriging estimation techniques such as Histogram (frequency distribution), Normal Quantile-quantile (QQ) plots, Trend analysis, and Semi-variogram/Covariance were used to validate the elevations of the observed data. The elevation data were further analyzed using Cross validation method, and the result indicated that Simple Kriging method of interpolation is suitable for the elevation data observed.

Keywords: Geo-statistics, Interpolation, Kriging, Variogram, Estimation.

Introduction

Geostatistics refers to statistics of the earth, such as in geography and geology, and interpolation refers to the method of constructing new data points within the range of a discrete set of known data points (Kress, 1998). Geo-statistical interpolation predicts values for cells in a raster from a limited number of sample data points. It is the process of manipulating spatial information to extract information and meaning from the original data (Chang, 2006). It can be used to predict unknown values of any geographic point data, which can be elevation, rainfall, levels, gravity values etc. In other words, it can be defined as the process of using points with known values or sample points to estimate values at other unknown points.

It is an assumption that spatially distributed objects are spatially correlated, which means, things that are

close together tend to have similar characteristics. Spatial interpolation is used to create surfaces with densely populated x, y and z data and the output surface depends on the *Z*data supplied (elevation, rainfall, levels, gravity values etc.), to produce a surface in form of grids (raster), which can be used to depict the topography of a land.

A suitable interpolated method can be used to optimally estimate the values at those locations where no measurements were taken. The results of the interpolation analysis can then be used for analysis that cover the whole area for modelling.

There are different methods of interpolating spatial data such as Inverse Distance Weighted (IDW), Natural Neighbor Inverse Distance Weighted (NNIDW), Spline, Pointinterp, Trend, Kriging etc., in which Kriging method of interpolation was used to interpolate the elevation data of the study area.

Aim and Objectives

Aim

The aim of this study is to analyze the use of simple Kriging of interpolation method in estimating elevation values for the unobserved location of the study area.

Objectives:

i. To acquire elevation data at 25m interval.
ii. To interpolate elevation values to unobserved locations and produce a surface map using simple Kriging method of interpolation.
iii. To analyze the accuracy of the unpredicted values.

Study Area

The study area (Fig. 1.1) is located within Obanla Campus, North Central part of the Federal University of Technology, Akure in Akure South Local Government Area, Ondo State, Nigeria, with boundary coordinates of 735557.402mE, 800158.202mN; 735975.729mE, 808072.918mN; 735925.192mE, 807530.562mN and 735386.834mE, 807721.119mN, in Minna Datum, which has an approximate Area of 25 Hectares.

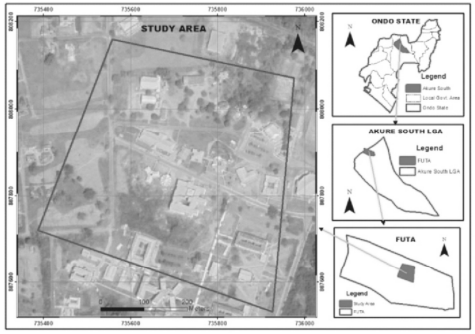


Fig. 1.1: Study Area Map

Conceptual framework and literature review

Spatial interpolation method (kriging)

Kriging is a geostatistical interpolation technique that considers both the distance and the degree of variation between known data points when estimating values in unknown areas. A kriged estimate is a weighted linear combination of the known sample values around the point to be estimated. Kriging procedure generates an estimated surface from a scattered set of points with z-values. Kriging assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variation in the surface. The Kriging tool fits a mathematical function to a specified number of points, or all points within a specified radius, to determine the output value for each location. Kriging is a multistep process which includes exploratory statistical analysis of the data, variogram modeling, creating the surface, and (optionally) exploring a variance surface. Kriging is most appropriate when you know there is a spatially correlated distance or directional bias in the data. It is often used in estimating topo surface.The predicted values are derived from the measure of relationship in samples using sophisticated weighted average technique. It uses a search radius that can be fixed or variable. The generated cell values can exceed value range of samples, and the surface does not pass through samples.
Kriging is similar to IDW in that it weights the surrounding measured values to derive a prediction for an unmeasured location. The general formula for both interpolators is formed as a weighted sum of the data:

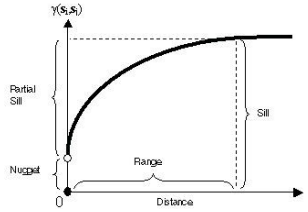


Fig. 2.1: Anatomy of a Typical Semivariogram.

The variance of the difference increases with distance, thus, the semivariogram can be thought of as a dissimilarity function. The height that the semivariogram reaches when it levels off is called the sill. It is often composed of two parts: a discontinuity at the origin, called the nugget effect, and the partial sill; added together, these give the sill.The nugget effect can be further divided into measurement error and microscale variation. The nugget effect is simply the sum of measurement error and microscale variation, and, since either component can be zero, the nugget effect can be composed wholly of one or the other. The distance at which the semivariogram levels off to the sill is called the range.

Covariance Function

The covariance function is defined to be;

$$\mathbf{C(si, sj) = cov[Z(si), Z(sj)]}$$

Where cov. is the covariance.

Covariance is a scaled version of correlation. Thus, when two locations, si and sj, are close to each other, they will be similar, and their covariance (a correlation) will be large. As si and sj get farther apart, they

become less similar, and their covariance becomes zero. This can be seen in figure 2.2 below, which shows the anatomy of a typical covariance function.

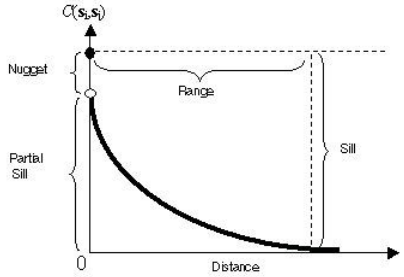


Fig. 2.2: Anatomy of a typical covariance function.

Materials and Method

Data Acquisition

The data set were obtained from field observation using South 350R Total Station instrument to obtain point elevation data (x, y and z co-ordinate). The x and y co-ordinates are used to fix the position of the points in space, the z co-ordinates denote the height. The point elevation data were observed at 25m interval. The following equipment were used in the execution of this project:

Hardware Components

- South 350R Total Station instrument and all accessories
- Tracking rods/ Prism Reflectors
- Computer (Acer Aspire E 14)

Software Components

- South Download Manager (Manufacture Software)
- Notepad.
- Microsoft word/ Excel
- ArcMap 10.5

Data Exploration

Data exploration is the first step in data analysis and typically involves summarizing the main characteristics of a dataset. This was carried out on ArcMap 10.5, using the geostatistical tools which are as follows:

Histogram

Histogram (frequency distribution) provided a one-variable description of the observed data. The frequency distribution displays how often the observed values fall within certain intervals or classes. In frequency distribution, the median and mean of the observed values must be close to validate the

observed value.

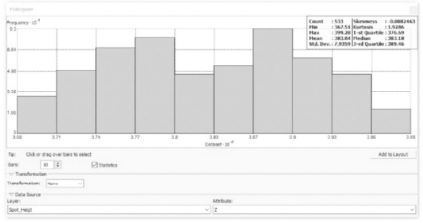


Fig. 3.1: Histogram of the observed data

From Fig 3.1 above it was observed that the mean of the observed data was 383.04m and the median was 383.18m, which shows that both values are close, which validates the correctness of the observed data.

Normal Quantile-quantile (QQ) Plot

Quantile-quantile (QQ) plots are an exploratory tool used to assess similarity between the distribution of one numeric variable and a normal distribution, or between the distributions of two numeric variables. If the distributions of the compared quintiles are identical, the plotted points will form a straight 45-degree line. The farther the plotted points deviate from a straight line, the less similar the compared distributions.

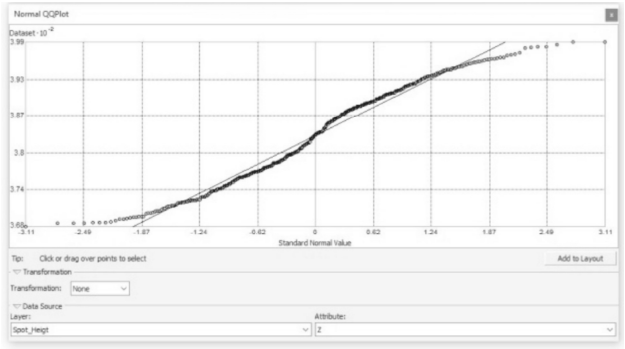


Fig. 3.2: Normal Quantile-quantile (QQ) Plot of the observed data

From table 3.2 above, the straight line is 45 degree, and the dots are the observed data. The dots are located around the straight line, which shows that the datasets follow the standard normal distribution.

Trend Analysis

Trend analysis provides a three-dimensional perspective of the data. The location of the sample points was plotted on the x, y plane, and above each sample point, the value is given by the height of a stick in Z dimension. In trend analysis, values are projected on to the x, z plane and y, z plane are scatter plots.

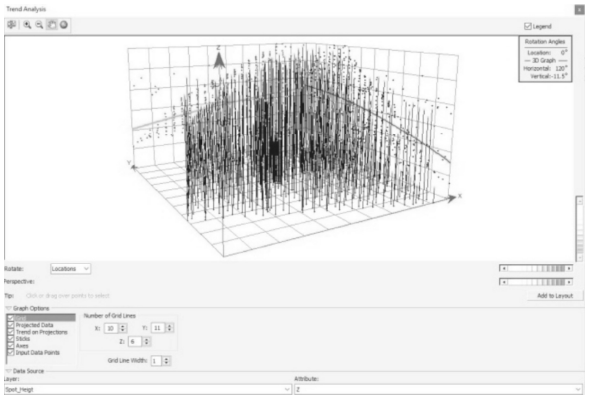


Fig. 3.3: Trend analysis of the observed data

Fig 3.3 shows that the trend is flat and less prominent as it does not really form a U shaped.

Semivariogram/Covariance Cloud

This shows the empirical semivariogram and covariance values for all pairs within a dataset and plots, then, as a function of distance that separates the two locations. It is used to examine the local characteristics of spatial autocorrelation within a dataset and look for local outliers.

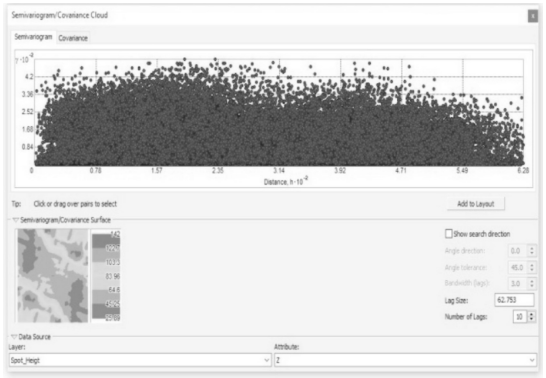


Fig. 3.4: Semivariogram/covariance of the observed data

Fig. 3.4 shows that there is an autocorrelation which allows models to be fitted for spatial relationship.

Results and Analyses

This chapter presents the results and analyses of the processed data.

Spot Height

The spot heights map showing the points elevation recorded as depicted in Fig. 4.1

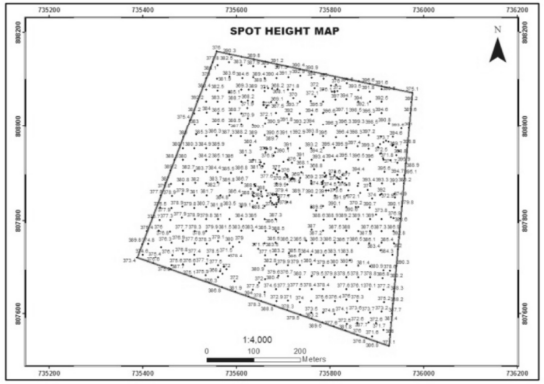


Fig. 4.1: Spot Height Map

Areas with no spot heights in Figure 4.1 are areas covered with buildings. The point with the least elevation is 367.512 m and the point with the highest elevation is 399.278m.

Interpolated Surface

Fig. 4.2 shows interpolated surfaces that have values at every point across their extent, produced from simple Kriging interpolation method. It shows ranges of elevation in which the least range of elevation is within 367.896m and 373.177m, and the highest range is within 394.299m and 399.087m.

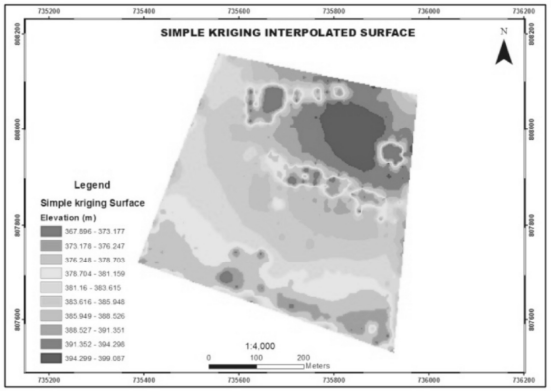


Fig. 4.2: Simple Kriging Interpolated Surface.

Cross Validation of Interpolated Surface

Cross validation shows how a model used for interpolation predicts the values at unknown locations. This serves as diagnostics that indicate whether the model and its associated parameter values are reasonable by comparing the predicted value to the observed value and useful information about the quality of the model.

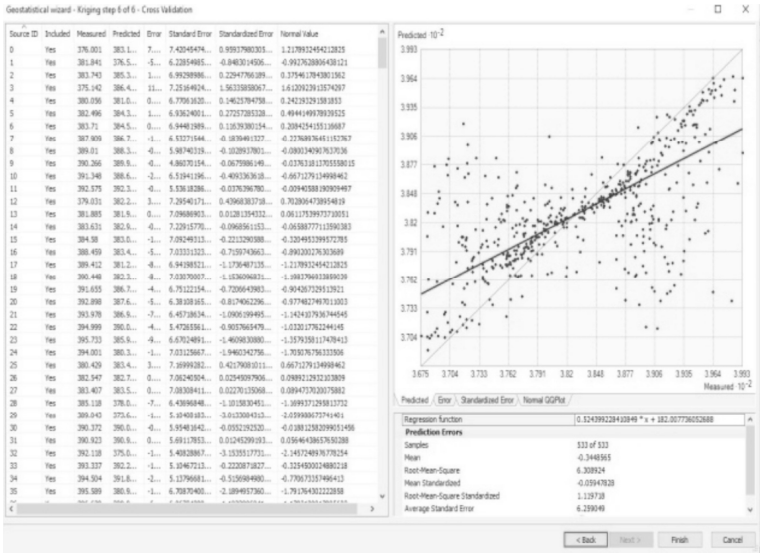


Fig. 4.3: Cross validation of Interpolated Surface.

Fig 4.3 shows the following;

i. The mean prediction error is -0.3448565, which close to zero (0)
ii. The Root Mean Square standardized error is 1.119718, which is close to one (1)

Conclusion

Simple Kriging has been successfully used to carry out the interpolation of point elevations obtained by field observation, by estimating elevation values in unknown areas. Data exploration such as Histogram (frequency distribution), Normal Quantile-quantile (QQ) plots, Trend analysis, and Semivariogram/Covariance were carried out to validate the elevation data observed.

The elevation data were further analyzed by cross validation method which showed that the mean prediction error was -0.3448565, which was close to zero (0), and also the Root Mean Square standardized error was 1.119718, which was also close to one (1), indicating that Simple Kriging method of interpolation is suitable for the elevation data observed.

References

Akima, H. (1970). *A new method of interpolation and smooth curve fitting based on local procedures*. *Journal of Association for Computing Machinery,* 17, pp589–602.

Armstrong, M. (1998). *Basic linear geostatistics*. Springer Verlag, Berlin (Germany), pp 153.

Burrough, P.A. & McDonnell, R.A. (1998). *Principles of Geographical Information Systems; Spatial information systems and geostatistics*. Oxford University Press. Oxford. UK, pp 333.

Chang, K. (2006). *Introduction to Geographic Information Systems*. Boston: McGraw-Hill: Higher Education.

Dominy, S.C., Noppe M.A. & Annels, A.E. (2002). *Error and uncertainty in mineral resource and ore reserve estimation: the importance of getting it right*. *Explor. Min. Geol*., 11(1-4), pp. 77-98.

Kress, R. (1998). *Numerical Analysis*. Springer, New York. http://dx.dol.org/10.1007/978-1-4612-0599-9.

Olea, Ricardo, A. (1999). *Geostatistics for Engineers and Earth Scientists*. Kluwer Academic.

Phillips, D. L, Dolph, J. & Marks, D. (1992). *A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain*. Agricultural and Forest Meteorology 58: 119–141.

Price, D. T. McKenney, D. W, Nalder, I. A, Hutchinson, M. F. & Kesteven, J. T. (2000). *A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agricultural and Forest Meteorology,* 101, pp 81–94.

Williams, C. K. I. (1998). *Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond. Learning in Graphical Models*. pp. 599–621. doi:10.1007/978-94-011-5014-9_23. ISBN 978-94-010-6104-9.