ABSTRACT

Nowadays, geographical information system (GIS) and remote sensing are emerging as powerful techniques widely applicable in natural resource management and development virtual models. Recent developments in remote sensing, aerial photography and GIS make it possible to detect changes and devise strategies based on these changes. The study focuses on using aerial photography for the detection of changes and effects of mining on geomorphology using the ArcGIS9 extension, Geostatistical Analyst. In addition, the distinctive surface topography of karst landscapes can be characterized in order to compare them with non-karst landscapes, and to determine geological and/or climatic conditions that are responsible for the observed terrain of Kinta Valley Limestone formation at Perak, Malaysia. Geostatistical analyses of the karstic terrain are used in order to distinguish between karst and non-karst area and karst area to observe the variation from the deterministic sample. In contrast, if the range is less, that means the average distance between two points that are similar in height is less and therefore there is more variation in the area. The average range for karst area is 435, while the average range for non-karst area is 690 meters. The difference between the major range and minor range which indicates the degree of anisotropy is more for
the karst area and this is an indicator of more variation in spatial structure and autocorrelation of the karst elevation.

KEYWORDS: Geostatistical analysis, geographical information system (GIS), aerial photography Karst, Malaysia

INTRODUCTION

Traditional management practices depended on human factor more than machines. Experts were asked to pay field visits to Karst sites; mapping was mostly done manually; urbanization expansion was out of planning and randomly distributed which deteriorate the quality and spatiality of natural resources; scale dimension of maps was small which was not practical for detailed investigations; development agendas were focusing on mining of Kinta resources as a reward business more than preserving them. For all these disadvantages, karst landforms were in short of preservation and maintenance. Their diminishing rate over sighted by human reflected inversely on the ecosystem balance. Nowadays, karst formations are under the mercy of many geotechnical issues and geohazard like sinkholes, weathering, dissolving and erosion, hydrological contamination and quality problems in such a way that human has been affected and endangered by those hazards. Therefore, effective, time and cost-effective planning was demanded to protect both the human and karst from involving in discord. Traditional management as discussed above had failed to balance between human expectation for welfare and natural conservation for sustainability through the life cycle of development.

Growing environmental problems, especially concerning geohazards (sinkholes, rock full, along with technological advances in Geographic Information Systems (GIS), have given rise to increased efforts by researchers, engineers, and planners to better understand the spatial distribution of karst features that characterize these regions. GIS applications enable researchers to objectively identify the conditions that trigger karst hazards.

Their incorporation into a GIS database allowed the subsequent statistical analysis of effects of environmental factors on landscape changes to quantify their impact. Parameters derived from the statistical analysis are then used to calibrate a spatially explicit model of anthropic dynamics. The models are potential for predicting future settlement changes and relating to actual management problems.

With Geostatistical Analyst, you can easily create a continuous surface, or map, from measured sample points stored in a point-feature layer, raster layer, or by using polygon centroids. The sample points may be measurements such as elevation, depth to the water table, or levels of pollution, as is the case in this tutorial. When used in conjunction with ArcMap, Geostatistical Analyst provides a comprehensive set of tools for creating surfaces that can be used to visualize, analyze, and understand spatial phenomena (ESRI, 2001).

Geostatistics

Geostatistical operation relies on statistical analysis. The conformity of sample data to the normal distribution that is necessary to produce justifiable results from the kriging is an important part of it. The analysis of outliers using the histogram of data, creating QQ plots for the quantiles
of sample data, transformations and bivariate data analysis to investigate the relationship of variables before kriging are some of the merits of statistics to the geostatistics (Wackernagel, 1995).

The set of random variables, \( Z(x_1), Z(x_2), \ldots \) constitute a random function, a random process, or stochastic process (Figure 1). The set of actual values of \( Z \) that comprise the realization of the random function are known as a regionalized variable (Wackernagel, 1995). In spatial data for \( n \) pairs of observations, \( z_{i,1}, z_{i,2}, i = 1,2,\ldots,n \), of two variables, \( z_1 \) and \( z_2 \) the covariance is given by Formula 1:

\[
\hat{C}(z_1, z_2) = \frac{1}{n} \sum_{i=1}^{n} (z_{i,1} - \bar{z}_1)(z_{i,2} - \bar{z}_2),
\]

where \( \bar{z}_1 \) and \( \bar{z}_2 \) are the means of \( z_1 \) and \( z_2 \), respectively. If the units \( i = 1,2,\ldots,n \), on which the observations were drawn at random, then \( \hat{C}(z_1, z_2) \) estimates the population covariance without bias.

Stationarity means that the distribution of the random process has certain attributes that are the same everywhere. However constancy of mean is not a sensible assumption. To solve this problem we can take this view that whereas in general the mean might not be constant, it would be so for small \( h \) at least (Jahanshir, 2006). Further if we replace the covariance with the variance of differences as measures of spatial relations, it will like the covariance depend on the lag \( h \) and not on absolute positions of the samples, which is given by Formula 2:

\[
\gamma(h) = \text{var}[Z(x) - Z(x + h)] = E[(z(x) - Z(x + h))^2] = 2\gamma(h)
\]

This constitutes the intrinsic hypothesis (Webster, 2001). The quantity \( \gamma(h) \) is known as the semivariance at lag \( h \), which is half of variance.

Formula (3) is the mathematical function which is being used for the theoretical assumptions. For a set of data that has been sampled:

\[
\hat{\gamma}(h) = \frac{1}{2mh} \sum_{i=1}^{m(h)} \left(z(x_i) - z(x_i + h)\right)^2
\]

where \( m(h) \) is the number of pairs of data points separated by the particular lag vector \( h \) (Goovaerts, 1999). The possible semi-variograms that are being calculated over the possible lag intervals will give the experimental variogram cloud (Figure 2A). As shown, the clustered points will not give much information about the structural behavior of the phenomenon under study.

More frequently we sum up the semi-variograms for specific lags and show them as the average points (Figure 2B). Furthermore, for the purpose of prediction we need to have complete information about the phenomenon in all lag distances which will then help to predict for unsampled areas. Therefore we model the variogram clouds i.e. pass a line with least square error distance through the points. This model has some characteristics as \( C_0 \), nugget variance is indicator of variation less than the lag 0 which is not possible; also, the value of variogram at lag
0 must be zero. Therefore we consider the nugget as error of the measurement although the term comes from the gold mining industry where the nuggets of gold are not continuous.

Figure 1: The concept of regionalized and random variables (Jahanshir, 2006)

Figure 2: Variogram cloud for lag class 8 “a” and average lag of all classes (B) for electrical conductivity data set cockpit (Jahanshir, 2006).

C+C₀ is the sill, which indicates the global variance of the phenomenon if we where to use the classical statistics to measure variance of all the data samples. R is the range (the lag in which the sill has been obtained). The geostatistical range is an important aspect, before range occurs in the graph of variogram there is spatial dependency between the samples.

The relationship decreases while the value of h increases, when h>R, there will be no more relativity among the samples. R measures the spatial dependence of samples. The soil data points that are located within the distance of the lag to the unknown can be interpolated. So the value of the lag when the variogram reaches its sill can be used as reasonable interval to sample the soil.
This concept has been adopted during the design of optimum systematic sampling (Hangsheng et al., 2004).

Modeling the Variogram

Selecting a few lag numbers and big lag sizes may lead the analysis to a wrong conclusion that will ignore the short range variation inside the field. Too many lag sizes will make modeling a formidable task as there are too many average semivariances in the plot that must be considered when we pass a line through them.

In theory it is always possible to have a better model with minimizing the least square error between predicted and measured or, by adding more parameters to the models, i.e. making the model more complex (Wakernagel, 1995).

The behavior near the origin is also important; the parabolic approach of model curve to the origin shows the deterministic variation or trend and drift. Based on Webster (2001), most important functions for modeling of the variogram in the environmental data are spherical, exponential, and circular and to some extent pentaspherical model.

The formula 4 for circular variogram is:

\[
\gamma(h) = c \left\{ 1 - \frac{2}{\pi} \cos^{-1} \left( \frac{h}{a} \right) + \frac{2h}{\pi a} \sqrt{1 - \left( \frac{h}{a} \right)^2} \right\} \text{ for } h \leq a \\
\gamma(h) = c \text{ for } h > a
\]

The parameters of c and a, are sill and range, the gradient of this model at the origin is \(4c/a\). It is variogram of conditionally negative semidefinite (CNSD) and it is valid for one, two and three dimensional data.

The formula 5 for spherical model is:

\[
\gamma(h) = c \left\{ \frac{3h}{2a} \left[ 1 - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right] \right\} \text{ for } h \leq a \\
\gamma(h) = c \text{ for } h > a
\]

The spherical model is the 3 dimensional analogue of the circular model. It is the default variogram model for most of geostatistical modeling packages. When the curve is somewhat more gradual than the spherical model, then the penta-spherical model maybe used, which is given by Formula 6:
The other important model is exponential model, which is given by Formula 7:

\[
\gamma(h) = c \left[ 1 - \exp \left( -\frac{h}{r} \right) \right]
\]

with sill \( c \), and a distance parameter, \( r \), that defines the spatial extent of the model. The function approaches its sill asymptotically and so it does not have a finite range.

The other models include Gaussian model, Whittle’s elementary correlation, linear and other unbounded models like the power function. In general it has been shown that there is no usage for these models in the environmental studies (Webster, 2001).

Anisotropy differs from the global trend in that it cannot be described by a physical process and modeled by a mathematical formula (polynomials). The cause of the anisotropy (directional influences) in the semi-variogram is not usually known so it is modeled as a random error. Even without knowing the cause, anisotropic influences can be quantified and accounted for (ESRI, 2001). With presence of anisotropy, it is easiest to build the variogram model with reference to the axes of anisotropy.

The key factor to distinguish between the isotropic and anisotropic variation is to examine the shape of variogram model, if for all considered directions, the degree of variation is the same, the sill and range of variation is the same, the phenomenon under study will be considered as isotropic. If in the other hand the range or sill of variogram models show obvious change with direction, then for the geostatistical operation to continue, there is a need to take into account the direction. Reisa (2002) shows a good example of considering the anisotropy in the variogram calculations.

**Kriging**

Kriging is the linear sum of the value of nearest points and their associated weights. There are many methods of kriging that are suited for specific situations (Goovaerts, 1997). Among the most common methods are ordinary and simple kriging.

The difference between the simple and ordinary kriging is consideration of the mean. Sometimes we know the mean of the random variable from previous experiences. In these circumstances we should use that knowledge to improve our estimates, and we can do so by simple kriging (Wackernagel, 1995). However the ordinary kriging assumes that the mean is unknown. The kriged estimate is the weighted average of the data, which is given by Formula 8.

\[
\hat{Z}(x_0) = \sum \lambda_i z(x_i) + \left[ 1 - \sum \lambda_i z(x_i) \right] \mu
\]
where \( \lambda_i \) are weights and \( \mu \) is the mean of regionalized variable. In ordinary kriging to ensure that the estimate is unbiased, the weights are made to sum to 1, and therefore the second part of the right hand side of formula will be deleted. In practice there will be some limitations that must be taken into account. One of them is the number of nearest points \( z(x_i) \). Nearest points will carry more weights than distant ones. Their relative properties depend on the positions of sampling points and on the variogram. The larger the nugget variance, the smaller are the weights of the points that are nearest to target point or block. Clustered points carry less weight individually than isolated ones at the same distance. If there is a trend in spatial process i.e. deterministic variation, then one must try to detrend the data set, trying to remove the short range and long range deterministic variations (Grohmann, 2004). The variogram however will be estimated using the estimates of residuals \( (u(x)) \).

The residuals are given by

\[
\varepsilon(x) = Z(x) - u(x)
\]  

(9)

In Eq. 9, \( \varepsilon(x) \) constitutes the random process. To estimate the variogram we need to be able to separate trend from the random process.

In practice kriging is not a push bottom process; various factors must be considered from the number of neighborhood samples, the variogram model and the positions of the samples. Kriging the irregularly point data is different from regular one. The clustered points and their effects must be considered thoroughly. The results of kriging must be examined in appropriate GIS software. Kriging can be done for point (punctual) or for block, for either case the kriging variance which is the mean square of difference between the predicted and actual must be the least. Various methods like cross-validation or jack-knifing will be valuable to evaluate the results (Webster, 2001).

The advantage of kriging over other interpolation methods is that, it tries to give a sense of error by representing the error or expected value of error variance of predicted \( Z(x_i) \) minus the real value, which is given by Formula 10:

\[
Error = E[\hat{Z}(x_i) - Z(x_i)] = 0.
\]  

(10)

This value should be zero theoretically and from this equation the value of estimation variance can computed, which is given by Formula 11:

\[
Var(\hat{Z}(x_i)) = E[(\hat{Z}(x_i) - Z(x_i))^2]
\]  

(11)

For any estimated value, there will be a variance associated with, therefore a map of error variance also can be computed the same as \( Z \) values.
Methods of Geostatistical Analysis

The variography of the data points involves the creation of a semi-variogram, which is the first step in investigating the nature of spatial variation (Lyew-Ayee et al., 2006). The semi-variogram quantifies the assumption that points near to each other tend to be more similar than points farther apart from one another (Johnston et al., 2001), measuring the strength of statistical correlation as function of distance (Lyew-Ayee et al., 2006).

Geostatistical investigations of the variography of elevation data from the study areas were conducted using the ArcGIS9 extension, Geostatistical Analyst (ESRI, 2001). The extension allows users to not only use kriging to interpolate a continuous surface, but also to visually explore the data and manipulate parameters before interpolating. The extension created a semi-variogram surface as a typical example shown in Figure 3, as well as a semi-variogram plot. The semi-variogram surface also revealed the effects of directional influences on spatial variation – anisotropy. As such, the extension is highly interactive and versatile. This allowed for quick experimentations with other theoretical models used to fit the experimental variogram, as well as checking for anisotropy at the click of a button.

Figure 3: Semi-variogram surface of data from Karst sample

Additionally, the extension allows users to perform cross-validation of the model, to assess how well values were predicted relatively to known sample locations. This is achieved by withholding the value at a known location and predicting the value at the same location. As a result, the predicted value could be directly compared to a known value, thereby assessing how accurate the model was. This may be plotted on a graph of predicted values versus the measured values.

The resulting scatter plot may be fitted with a linear regression line. An accurate prediction model will produce a plot where the points are tightly scattered around the fitted regression line with a slope of 1:1. FigureError! Reference source not found. 4 shows the results of a cross-validation performed on Karst sample at Gunung Terendum. From the semi-variogram surface, the extent of spatial autocorrelation and directional effects on spatial variation are shown. The blue and green areas are where the semivariance values are low; the orange areas indicate increasing dissimilarity with distance.
Figure 4: Typical output of the cross-validation of geostatistical parameters.

In this plot, all points cluster closely around the blue line-of-best-fit (barely visible in the top right-hand corner of the graph). A line with a slope of 1:1 is also drawn for comparison (broken line, also barely visible in the top right-hand corner of the graph). It is clear that the prediction model used to create the Karst sample at Gunung Terendum DEM was reasonably accurate; there were no points that drastically deviate from either the line-of-best-fit, or the 1:1 comparative line.

Geostatistical analysis of the terrain for Kinta Valley draped on the geological map shown in Figure 5, in order to distinguish between karst and non-karst area and to observe the variation from the deterministic sample areas which have been selected from the 2004 and 1981 surfaces. The study area has been selected for the distribution of limestone hills in Kinta valley that represent the karst topography.

Geostatistical Analysis of Terrain Results

The parameters of DTM statistics include the range of elevation and the mean elevation, as well as the variance and the standard deviation of the terrain. Statistics itself is useful, but the nature of raster data as a mathematically continuous surface allows for the derivation of numerical values at any point on the surface. Once created, these derivations provide a wealth of information about the characteristics of the terrain.

The visual representation shows that all of the Karst selected areas in the years 1981 and 2004, can be statically fit into spherical model rather than that of other models, which show very small or no nuggets. This indicates non-existence of microstructure and measurement error.
Figure 5: Location samples for statistical terrain analysis

Semi-variograms of elevation data karst area of 2004 in areas b and c show that the spatial structure in those areas are following the same pattern compared to area a as shown in Figure 6. While the model for 1981 data is showing the same pattern as the semi-variogram clouds show more conformity to the spherical model Figure 7. The variogram model helps to interpret the spatial variability of the phenomenon under study. In this case the variogram models from 2004 and 1981 show the variation of elevation by the distance. The conformity of variation of elevation data to the spherical model for all of the study areas, is an indicator of existence of spatial autocorrelation.
Figure 6: Semi-variograms of elevation data karst area in 2004.

While observing for the area “a” the variogram model shows dramatic change between two data sets. The sill has been increased and the range is decreased. If the range of variogram model decreases, it shows that the variation in elevation has been increased. This could be the result of human intervention or natural causes.

Figure 7: Semi-variograms of elevation data karst area in 1981

In case of “b” and “c” areas, the variogram model does not show changes over time, but it is apparent that the range and sill have been changed. The spread of semi-variogram points shows that the model tends to be asymptotic in 2004, although the statistical analysis shows that the
spherical model is a fit for this scatter plot, but this cannot cover the fact that extreme change in
the elevation could result due to human intervention. This has caused variogram model a definite
sill while its range is difficult to determine. If the model is considered asymptotic like the
exponential model as reported (Webster and Oliver, 2001) there will be no sill recognizable and
the range could only be estimated, not calculated.

Table shows the characteristics of variogram models that are fitted to the elevation data in
karst areas in 2004. If we consider the isotropic structure first, the range of variogram model
which shows the degree of spatial autocorrelation is changing from 430 to 490m from areas “a” to
“c”. This steady is an indicator of steadiness of autocorrelation between the three areas. The
higher the range, the less variation is expected in a given area. Therefore, the variation in
elevation in area “b” is higher than area “a” and “c” respectively, although the amount of change
can be considered the same for all three regions.

Table 1: Summary of Geostatistical modeling for karst 2004

<table>
<thead>
<tr>
<th></th>
<th>2004 Karst “a”</th>
<th>Karst (b)</th>
<th>Karst “c”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>isotropy</td>
<td>Anisotropy</td>
<td>isotropy</td>
</tr>
<tr>
<td>Major Range</td>
<td>441.64</td>
<td>865.18</td>
<td>432.61</td>
</tr>
<tr>
<td>Minor Range</td>
<td></td>
<td>336.88</td>
<td>421.26</td>
</tr>
<tr>
<td>Anisotropy differences</td>
<td></td>
<td>528.29</td>
<td>823.32</td>
</tr>
<tr>
<td>Partial Sill</td>
<td>1844.57</td>
<td>1876.16</td>
<td>1158.63</td>
</tr>
<tr>
<td>Direction</td>
<td></td>
<td>359.8</td>
<td>312.9</td>
</tr>
<tr>
<td>Nugget</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lags Size</td>
<td>73.66</td>
<td>73.66</td>
<td>105</td>
</tr>
</tbody>
</table>

For the year 1981, the range starts from 437 to 515 km. It is obvious that the higher range
means less variation for the 1981 surface. This indicates the extent of elevation changes between
1981 and 2004. This could be the result of quarrying in the area. In case for area “b”, the range is
smaller this shows more variation in elevation. It is observed that the sill is higher in respect to a
specific distance. This shows that there exist fluctuations in the field. Results in Tables 1 and 3
shows that the values of sill increase from 1981 to 2004. There is no nugget associated with all of
the models; therefore all of the models have authentic value of sill which is not mixed up with the
error. In terms of variogram statistic the range has been decreased and the sill has been increased
in the year 2004 comparing to the variogram data of the same regions in 1981.

The data in Tables 1 and 2 can be analysed in the variogram models that can be predicted the
specific variance based on specific distance. However, in this research we are technically
interested in using the variogram model characteristics to predict the autocorrelation for the
future. Therefore, variogram characteristics can be used to form regression models to predict the range of spatial autocorrelation for the future.

The result of regression analysis for the isotropic range of two dates is shown in Table. This result is based on considering the range for 2004 a dependent variable and range for 1981 dependent variables. However, the analysis showed that the r square is the same when we consider the range of 1981 as independent variable. Based on the result for the range (Figure 8), the regression model will be:

\[
\text{Range of 2004} = 122.56 + 0.692 \times \text{Range of 1981} \quad (\text{R square 0.7})
\]  

Figure 8: Regression model for isotropic range

Residual analysis for the regression line shows that the model could predict the range for 2004 with high accuracy and the residual plot is normal. Since the R square of the model is high, this model could be used to predict the range for the future samples.

Table 2: Summary of Geostatistical Results for karst 1981

<table>
<thead>
<tr>
<th>1981</th>
<th>Karst “a”</th>
<th>Karst (b)</th>
<th>Karst “c”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anisotropy</td>
<td>Anisotropy</td>
<td>Anisotropy</td>
<td>Anisotropy</td>
</tr>
<tr>
<td>Major Range</td>
<td>485.028</td>
<td>952.517</td>
<td>437.986</td>
</tr>
<tr>
<td>Minor Range</td>
<td>-</td>
<td>373.666</td>
<td>-</td>
</tr>
<tr>
<td>Anisotropy differences</td>
<td>-</td>
<td>578.851</td>
<td>-</td>
</tr>
<tr>
<td>Partial Sill</td>
<td>1590.8</td>
<td>1624.4</td>
<td>1119.5</td>
</tr>
</tbody>
</table>
Table 3: Summary output for isotropic range analysis

Regression Statistics

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0.879</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R Square</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R Square</td>
<td>0.544</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard Error</td>
<td>20.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOVA

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>SS</td>
<td>MS</td>
<td>F</td>
<td>Significance F</td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>1</td>
<td>1471.312</td>
<td>1471.312</td>
<td>3.390063</td>
<td>0.316748</td>
</tr>
<tr>
<td>Residual</td>
<td>1</td>
<td>434.0073</td>
<td>434.0073</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>1905.319</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Residual Output

<table>
<thead>
<tr>
<th>Observation</th>
<th>Predicted range 2004</th>
<th>Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>458.5261</td>
<td>-16.8861</td>
</tr>
<tr>
<td>2</td>
<td>425.9415</td>
<td>6.668506</td>
</tr>
<tr>
<td>3</td>
<td>479.7924</td>
<td>10.21757</td>
</tr>
</tbody>
</table>

In Table 3 and 5, non-karst areas show that the amount of range is highly different from karst areas, although for regions (e) and (f) the range is the same but for region (d) the range is about three folds higher than the other regions. This result shows more continuity of spatial autocorrelation and therefore less variation of elevation for region (d) and more amount of autocorrelation of elevation data for overall non-karst areas.

Table 4: Summary of Geostatistical Results for non-karst 2004

<table>
<thead>
<tr>
<th>2004</th>
<th>Non-Karst (d)</th>
<th>Non-Karst (e)</th>
<th>Non-Karst (f)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Isotropy</td>
<td>Anisotropy</td>
<td>Isotropy</td>
</tr>
<tr>
<td>Major Range</td>
<td>1274.23</td>
<td>1274.23</td>
<td>431.728</td>
</tr>
<tr>
<td>Minor Range</td>
<td>-</td>
<td>934.636</td>
<td>-</td>
</tr>
</tbody>
</table>
It can be seen that the range is steadier for karst than non Karst, the major range for isotropic models is the same with minor differences and this reveals the fact that the range where the extent is less similar is less in average than non-karst areas (Figure 9 and 10).

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
1981 & Non-Karst “a” & Non-Karst (b) & Non-Karst “c” \\
\hline
Isotropy & Anisotropy & Isotropy & Anisotropy & Isotropy & Anisotropy \\
\hline
Major Range & 493 & 745.511 & 478.148 & 1244.59 & 493.629 & 390.194 \\
\hline
Minor Range & - & 389.348 & - & 451.408 & - & 482.112 \\
\hline
Anisotropy & - & 356.163 & - & 793.182 & - & 482.112 \\
\hline
Partial Sill & 322.93 & 326.9 & 147.14 & 134.31 & - & 39.9 \\
\hline
Direction & - & 40.9 & - & 307.7 & 0 & 0 \\
\hline
Nugget & 0 & 0 & 0 & 0 & 73.592 & 73.592 \\
\hline
Lags Size & 105 & 105 & 105 & 105 & 493.629 & 872.306 \\
\hline
\end{tabular}
\caption{Summary of Geostatistical Results for non- karst 1981}
\end{table}

Figure 9: Semi-variograms of elevation data non-karst area in 2004.
If the range is high, then the variation is less and it means we won’t see so many differences in the phenomenon that we are studying. In contrast, if the range is less, that means the average distance between two points that are similar in height is less and therefore there is more variation in the area. The average range for karst area is 435, while the average range for non-karst area is 690 meters. The difference between the major range and minor range which indicates the degree of anisotropy is more for the karst area and this is an indicator of more variation in spatial structure and autocorrelation of the karst elevation.

Figure 10: Semi-variograms of elevation data non-karst area in 1981

The result of the analysis for the samples from digital terrain data set that have been taken, can predict the future variation in the elevation of the area and the elevation indicator of human interference as the results of surface mining (quarry).

Primary Elevation Characteristics

Table shows the mean elevation and slope of the study areas, as well as their standard deviations. Observing the mean elevation alone does not give any idea about the nature of a surface, whether it is karstic or not, nor does it give any idea about the degree of roughness present in the terrain.

Table 6: Mean and Standard Deviation Values of Elevation and Slope of the Study Areas

<table>
<thead>
<tr>
<th></th>
<th>2004</th>
<th>Karst “a”</th>
<th>Karst “b”</th>
<th>Karst “c”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>131.03</td>
<td>186.39</td>
<td>116.72</td>
<td></td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>71.70</td>
<td>110.77</td>
<td>59.82</td>
<td></td>
</tr>
</tbody>
</table>
By looking at the standard deviation of elevation, however, a clearer picture may be obtained. The Cockpit Country study areas have standard deviations of elevation ranging from 59.82 at sample “c” to 110.77 at sample “a”, Gunung Terendum located in Ipoh with areas of tower karst relief, shows higher standard deviations of elevation. In these areas, extensive planar surfaces result in topography, with more variation than the ‘rouglier’ surfaces of the other study areas. Values at Rapat Gunung (with a standard deviation of 71.70) lie mid-way between those determined from the karst sample study areas.

### Karst Surface Investigation

A large amount of data concerning DTM of karst was analyzed using geostatistics for 1981 and 2004. The results showed the range of elevation, mean elevation and the semi-variogram surface, which also revealed the effects of directional influences on spatial variation in a selected sample.

The semi-variogram surface provides 2D images of the spatial behaviour of time serial images of 1981 and 2004 for selected areas. It allows identifying the direction of maximum grade continuity (anisotropy) and to understand the short scale grade variability (nugget effect) based on land use modification and mining activities. Figure 11 shows semi-variogram surface for the first selected sample that can be used to provide important information about a dominant orientation alignment and TIN elevation. This alignment leads to a directional variation of the effects of anisotropy to define human activities, drainage system, and major water inflow and its velocities. Change of terrain TIN elevation between 1981 and 2004 can be defined for each pixel in the semi-variogram surface.

The semi-variogram surface Figure 12 of 2004 shows the maximum and mean length of flow path, due to the lowest pixel values. The difference between the major range and minor range for selected samples is large, which indicates the degree of anisotropy is high based on variation in spatial structure, and autocorrelation of the karst elevation. Land use modification and mining activities had changed the directional of mine water drainage and major water inflow, which increases the possibility of pollution in the area. From the geo-hazard map result shows the tow data samples lies under high and very high risk zone as shown in Figure 13 (Sample a and c), these factors and hazard risk map have been increased the potential of sinkholes due to the increasing weathering.

The anisotropy and variations in spatial structure in the area causes flow velocities in this area to increase compared to the velocities in other parts allowing little time to warn downstream with high degree of erosion and weathering. Groundwater flows through conduits so that there is great opportunity for groundwater to move rapidly through conduits and fractures especially after a rainstorm. Extracting limestone rocks from 1981 to 2004 might have increased the velocity of the water and groundwater flow. Figure 13 shows part of cockpit limestone pavement in the Wang in Gunung Rapat. Land use modification and mining activities had changed the direction of mine water drainage and major water inflow, which increases the possibility of pollution in the area.

<table>
<thead>
<tr>
<th></th>
<th>Sample a</th>
<th>Sample b</th>
<th>Sample c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>0.56</td>
<td>0.46</td>
<td>0.670</td>
</tr>
<tr>
<td>Median</td>
<td>111.49</td>
<td>169.33</td>
<td>163.13</td>
</tr>
<tr>
<td>Mean Slope (o)</td>
<td>17.22</td>
<td>27.47</td>
<td>24.28</td>
</tr>
</tbody>
</table>
These factors had increased the potential of sinkholes due to the increasing weathering. It may also be due to the presence and density of fractures, possibility of major water inflow and flow direction as shown in Figures (11, 12, and 13), from karst aquifers; in addition, the erosion may have been caused by limestone dissolution rates, which had high occurrence in this area.

**Figure 11:** TIN and semi-variogram surface sample “a” from 1981 to 2004
Figure 12: TIN and semi-variogram surface sample “c” from 1981 to 2004
Figure 13: Geohazard map samples surface investigation location “a” and “c” area
CONCLUSIONS

As a conclusion, geostatistical analysis used in this study for investigation the nature of spatial variations of karst terrain using the ArcGIS extension (Geostatistical Analyst). The data have been collected from aerial photography of 1981 and 2004 for six samples. In 1981, the range starts from 437 to 515 km. it is obvious that the higher range means less variation for 1981 surface. This indicates the extent of elevation changes between 1981 and 2004 at Gunung Terendum. This could be the result of quarrying in the area. Encase for area b, the range is smaller which shows more variation in elevation. It is observed that the sill is higher in respect to a specific distance. This shows that there exist fluctuations in the field. The anisotropy and variations in spatial structure in the area causes flow velocities in this area to increase compared to the velocities in other parts allowing little time to warn downstream with high degree of erosion and weathering. Groundwater flows through conduits so that there is great opportunity for groundwater to move rapidly through conduits and fractures especially through rainstorms. The results show the limestone rock extraction from 1981 to 2004 might the reason of increasing the velocity of the water and groundwater flow. Those factors that increased the potential of sinkholes due to the increasing weathering, the presence and density of fractures, the possibility of major water inflow from karst aquifers, and the erosion caused by limestone dissolution rates that are highest in this area. Landuse modification and mining activities are changed the directional of mine water drainage and major water inflow, which increases the possibility of pollution in the area

ACKNOWLEDGEMENTS

We appreciate the support of eScience Fund from Ministry of Science, Technology and Innovation Malaysia for providing the research grant for Modelling of Karst Land Form for Engineering Purpose”, to Mr. Ibrahim (Ph.D candidate, Universiti Putra Malaysia), and to all staff of Mountainous Terrain Development Research Center for their support throughout this study.

REFERENCES

4. Jahanshir, E. 2006. GIS-Based Soil Sampling Methods for Precision Farming of Rice, Univrsiti Putra Malaysia, Malaysia
   http://cat.inist.fr/?aModele=afficheN&cpsidt=16576237

Published online in Wiley InterScience. http://www.colorado.edu/geolsci/gtucker/preprints/lyew-ayee_etal2007cockpit.pdf

   Journal of Geochemical Exploration 77: 45-63.
