PROBABILITY MAPS AS A MEASURE OF RELIABILITY FOR INTERVISIBILITY ANALYSIS

Dušan Joksić, Branislav Bajat

Digital terrain models (DTMs) represent segments of spatial data bases related to presentation of terrain features and landforms. Square grid elevation models (DEMs) have emerged as the most widely used structure during the past decade because of their simplicity and simple computer implementation. They have become an important segment of Topographic Information Systems (TIS), storing natural and artificial landscape in forms of digital models. This kind of a data structure is especially suitable for morphometric terrain evaluation and analysis, which is very important in environmental and urban planning and Earth surface modeling applications.

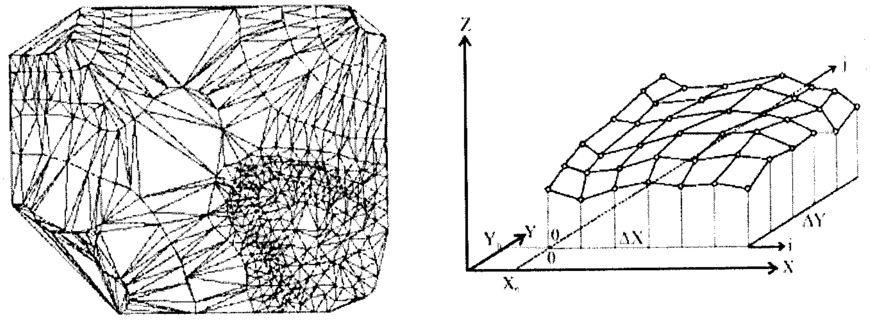
One of the most often used functionalities of Geographical information systems software packages is intervisibility or viewshed analysis of terrain. Intervisibility determination from analog topographic maps may be very exhausting, because of the large number of profiles that have to be extracted and compared. Terrain representation in form of the DEMs databases facilitates this task. This paper describes simple algorithm for terrain viewshed analysis by using DEMs database structures, taking into consideration the influence of uncertainties of such data to the results obtained thus far. The concept of probability maps is introduced as a mean for evaluation of results, and is presented as thematic display.

Key words: Digital elevation models, intervisibility analysis, probability maps.

INTRODUCTION

Gathering, maintenance and visualization of relief data are mostly used GIS (Geographical Information System) functions. This segment of GIS applications is well known as Digital Terrain Model (DTM). The DTM can be defined as 'any digital representation of the continuous variation of relief over the space' (Burrough (1986)). The process called visibility or intervisibility analysis recognizes that if you are located at a particular point on a topographic surface, there are portions of the terrain you can see (viewshed) and others you cannot see (DeMers (2000)). During the last decade this has become convenient tool for representing terrain surface, not only in GIS applications, but also in computer-assisted software for design. DTM bases are organized in different manners, like Triangular Irregular Networks (TIN) or in regular lattice (GRID). Term DEM is related to terrain height databases with regular grid structure or altitude matrix, and the term Digital Terrain Models (DTM) is mostly related to TIN structures in which the terrain surface is defined by the triangular vertices.

DEMs approach for data base structures is suitable for National Heights Data Bases which cover the whole country area as a part of the nation-wide spatial data bases available from the national mapping agencies. This kind of data bases are very useful in various application areas in GIS, like ecological studies, 3D urban mapping, environmental monitoring, landscape planning, geological analy-



a) TIN

b) GRID

Figure 1. DTM and DEM data structures

sis, civil engineering, floodplain analysis, risk maps etc. They are also very suitable in complex analysis, especially in superimposing other kind of data, like satellite images or remotely sensing information. The new and improved methods of an analysis allow users to process even more complex application tasks with combinations of geometrical, topological and thematic aspects using hybrid data, i.e. raster data as well as vector data (Glemser *et al,* 2000).

DTM concept has started as a stand-alone application intended for different engineering tasks. Nowadays, it is a part of the spatial databases, which store the natural and artificial landscape in the form of digital models. DEM are compulsory component of one category of spatial information systems known as Topographic Information System (TIS) (Kraus (1995)).

Introduction of the GIS technology in spatial analysis has provided new capabilities in manipulating the spatial data, especially with data related to terrain features. One of the most often used GIS functionalities is viewshed or intervisibility analysis of terrain. A viewshed analysis can be defined as an analysis that "indicates not only what areas of a surface can be seen by one or more observers, but also, for any visible position, how many observers can see the position" (ArcView online help). Algorithms for this kind of spatial functionalities are developed both for GRID and TIN data base structures of DTMs. It is simple GIS function, where spatial units, which could be seen from one or more viewpoints, are coded as binary variable 1 or "true", in contrast to 0 or "false" for invisible parts. The collective distribution of all "true" spatial units is called the viewshed (Burrough&McDonell (2000)).

Viewshed terrain analysis has multipurpose practical usage. Applications include the location of fire towers, radar sites, radio, TV or telephone transmitters, path or route planning, navigation and orientation etc. Viewshed analysis is very suitable for evaluating urban environmental planning and development of new settlements, where it is used for discovering positions from which they can be seen, or at least for determination of visible routes for some locations, or for avoiding disturbance of natural landscape while designing huge constructions. This analysis is used also for setting up the best locations for watchtowers on forest terrain for the purposes of monitoring fire disasters or protecting endangered species. It is also useful in landscape architecture for defining less visible or totally uncovered space, or for planning in the least visible routes for displacement of army corps or for setting up radar systems for military purposes. In telecommunications, viewshed analysis is unavoidable part of project design for covering relay towers or for setting out locations for transceiver antennas, cellular telephone transmitters and receiving stations etc. Nowadays, this kind of spatial analysis is very popular tool in analyzing archaeological locations. Site location criteria are considered with respect to both the visibility from high viewshed locations, and the sheltering potential from wind, weather conditions, and other human and animal impacts, provided by low viewshed locations (Tripcevich (2002)).

Between June 1998 and December 1999, a very big survey via the World Wide Web (www) was been conducted among DEM users from various countries, organizations and industries. More than 200 participants from all over the world participated in this poll. One of the questions was related to purposes for which DMTs are used. Viewshed analysis is on the third place with 8 percent, after the hydrological analysis for catchment area delineation with 10.7%, and deriving drainage networks from DTMs with 8.6% (Wechsler (2000)).

ALGORITHMS FOR VISIBILITY COMPUTATIONS

Significant visibility structures for single viewpoint are the viewshed, which represents the position and a part of the terrain visible from given viewpoint, and the horizon, which expresses distal border line of the viewshed (De Floriani&Magillo (1999)). The simplest method to convey this analysis on a topographic map is to connect an observer location to each possible target in the coverage. Next step is to follow the ray from each target point back to the starting point, looking for the elevations that are higher. Those points would obscure the observers' view of what is behind it (Figure 2). Viewshed analysis from analog topographic maps may be very exhausting, because of the large number of profiles that have to be extracted and compared.

Viewshed analysis is a capability possessed by many GIS software packages with raster terrain data structures. Data output is usually in a raster format but some GIS software is capable to create vector output. Viewshed analysis performed in vector data format requires the use of a TIN data model. A wide range of procedures for viewshed analysis was developed for vector GIS software packages where terrain heights data are stored as TIN structures. One of those front-to-back approaches was developed and implemented by De Floriani, who used the fact that a spatial triangle could be hidden only by triangles in front of it. At each step, current horizon is maintained and used to determine visibility of new triangles (Figure 3).

The basic algorithm for generating a viewshed coverage from elevation data in raster structure is based on the assessment of the elevation difference of intermediate cells between the viewpoint and target point. The determination as to whether the target point can be seen from the viewpoint is accomplished by examining each of the intermediate pixels between the two cells, which allows determination of the 'line-of-sight'. If the land surface rises above the line-of-sight, the target is hidden. Otherwise, it is visible from the viewpoint.

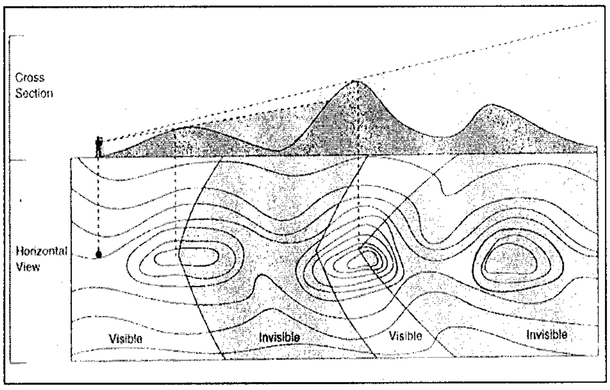


Figure 2. Viewshed analysis on a topographic map by ray tracing process (DeMers, 2000.)

The line-of-sight computation is repeated for all target pixels from a set of viewpoints, and the set of targets which are visible from the viewpoints from the viewshed coverage (fig.4).

Practical algorithm for viweshed analysis consists of following queries (figure 5):

```
IF  AD>AE
    OR DF>EF
    OR ∠ACB <∠AEF
    OR ∠BAC <∠FAE
    OR AC/BC >AE/FE
THEN C visible from point A
    ELSE C invisible from point A
```

where:

A view point with given height offset (men's or construction's height),

B target point in horizontal plane with corresponding terrain height difference **BC** above point **A**,

F target point in horizontal plane with corresponding terrain height difference **FE** above point **A**.

Given formulas refer to terrain above the horizon of viewpoint. When viewed terrain is below the horizon, conditional statements would be opposite (instead of 'less' we will use 'more' and vice versa).

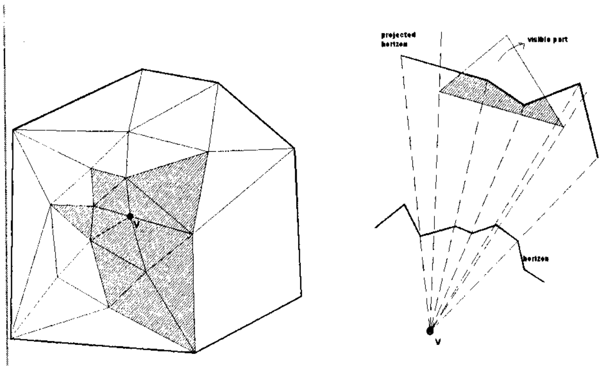


Figure 3. Front-to-back approach; Intervisibility algorithm for TIN data base structure; (left) a set of processed triangles (shaded) with current horizon; (right) determination of the visibility of one triangle by projecting the current horizon on the triangle itself (the parts above the projected horizon are visible). (De Floriani&Magillo, 1999)

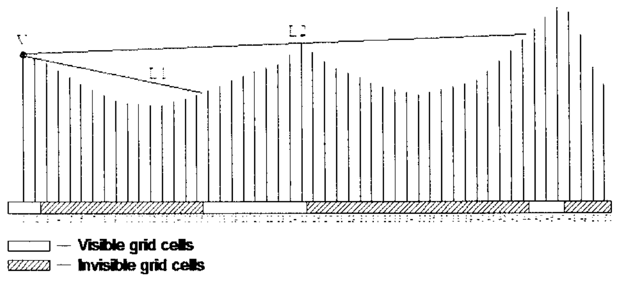


Figure 4. Line of sight approach for grid databases.

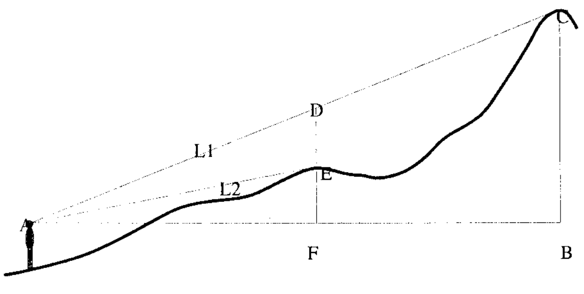


Figure 5. Viewshed analysis considering above given algorithm

PROBABILITY MAPS

The DEM is a model of the terrain surface, and like other models, this kind of data are subject to error. Terrain heights in DEM databases are considered as point features in which heights components are stored as attributes. This concept is well known as 2.5 GIS data model. Attribute accuracy is the main element of quality of such spatial databases. Accuracy can be defined as closeness of agreement between the test result and the accepted reference or 'true' value, where reference values are data sources of higher accuracy. However, a true value or location is a luxury not often available in spatial data.

In the case of spatial data it refers to data of larger scale and resolution or usage of instrument of higher accuracy or more recent measurements. All those imply that definition of accuracy is relative and for that reason term accuracy is substituted by uncertainty. Many words with similar connotation are used to express term uncertainty, like reliability, confidence, accuracy, error etc. This term is especially suitable for reporting the assertiveness of variables where true values are unknown. Under these circumstances it is not feasible to calculate exactly the error of the appraisal. A useful step in assessing the uncertainty is to consider the factors by which the error is influenced (Isaaks&Srivastava (1989)).

Two DEM-related uncertainty categories are commonly recognized. The primary is data model-based uncertainty, resulting from differences between the form of data model and actual elevation surface. The second is data-based uncertainty, referring to differences between the elevation of location specified in the data set and actual elevation at that location (Shortridge (2001)).

DEM altitudes are given as interval data, and root mean square error (RMSE) or standard deviation could be used as an uncertainty metrics. The RMSE calculated from residuals between models heights and 'ground truth' points is commonly used measure of the accuracy for the DEMs products (Caruso (1987)). Another powerful measure is the standard deviation of residuals, obtained from discrepancies between model heights and terrain heights at the particular locations.

$$\sigma_z = \sqrt{\frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1}} \quad , \quad \bar{d} = \frac{\sum d_i}{n} \qquad (1)$$

where:

d_i- residual between height in DEM and 'true' height $d_i = z_{DEM} - z_t$,

n- number of control points.

An accuracy evaluation often uses field survey data, which is based on the Global Positioning System (GPS) because of its high accuracy.

Performing control measurements is a standard procedure for evaluating the quality of DEMs products. DTM vendors often provide this kind of data associated with DTM products as meta data. DEMs users are able to perform control measurements if they are not acquainted with lineage of data. Locations of the control points have to be randomly chosen (Li (1991)).

Monte Carlo simulation approach is very often used for studying the propagation of DEM errors. The main idea of this approach is to use conventional statistical tools for evaluating error propagation by producing a number of realizations of DEMs for same terrain area. Each realization is obtained by adding "error fields" to initial DEM. An "error field" is surface of random values with a zero mean and standard deviation equivalent to the supposed range of error. Physically "error field" represents a matrix of values produced by random generator algorithm with same dimensions (rows and columns) as initial DEM. Errors in spatial data are spatially autocorrelated. The error field could also be modeled with various correlation functions if spatial structure of errors is known. Spatial structure of the error may be determined by spatial autocorrelation measures or variograms of the error fields.

For each realization of DEM, we can achieve thematic display of viewshed map, which is regarded as a distribution of possible realization within which the true values lie. This approach is known as "stochastic imaging", or the modeling of spatial uncertainty through alternative, equiprobable, numerical representations (maps) of spatially distributed phenomena (Journel (1996)). Stochastic simulations supply a series of random likely maps using stochastic modeling methods from mathematical statistics. This method does not guarantee that "real" map is generated, but it provides a bound within which we can affirm that the true map lies.

The probability of any DEMs grid cell being visible from viewpoint is given by (Fisher (1999)):

$$p\left(X_{ij}\right) = \frac{\sum_{k=1}^{n} x_{ijk}}{n} \qquad (2)$$

where,

$p\left(X_{ij}\right)$ is the probability of a cell at row i and column j in the grid raster DEM being visible;

x_{ijk} is the value at the grid cell of the binary-coded viewshed in realization k (1 for "true" and 0 for "false"), such that k takes values 1 to n (the number of simulations).

The image produced in such a way consists of the pixels with $p\left(X_{ij}\right)$ values, and it is known as probability map. It is used as a measure of "attribute" accuracy. Namely, for all thematic displays of some phenomena where quantitative parameters are used, probability map is the best indicator for reliability of such analysis.

CASE STUDY AREA

DEM was produced by digitizing contour layers from two adjacent sheets of the topographic maps of scale 1:5000, with contour interval of 5 m, with total area of 13.5 square kilometers for the research purpose. Test location is the resort area of Zlatibor in southwestern Serbia with minimum height of 850 m and maximum height of 1174 m. This area is hilly plateau, with the exception of the west and northwest part with greater terrain slopes.

Digitized polylines with height attributes were broken into vertices and by using the Douglas-Peucker algorithm for polyline simplification to 51847 points reduced such big amount of obtained points. A DEM was produced in two steps: An initial TIN was produced using a Delaunay triangulation, being subsequently converted into regular grid with 10 m resolution (350 rows by 400 columns).

View point was settled at the southwest part of the test area "Zlatibor" , with Y = 7 394 967 m, X = 4 842 665 m coordinates in the national grid. Heights of vegetation and objects were not taken into account in this analysis.

Thematic display of viewshed for initial DEM is given at figure 7. The number of visible grid cells is 30.7% of the total test area, or 414.8 hectares.

A GPS survey of control points has been carried out. Calculated σ_z for data set was approximately 1.25 m, and the obtained result is in accordance with an expected accuracy of the cartographic source data used for DEM production (Merchant (1987)).

Error fields simulations were carried out with estimated σ_z, under correlated conditions. 25 simulated DEMs were used to obtain probability map. For each simulated DEM, viewshed map was calculated, and accordingly

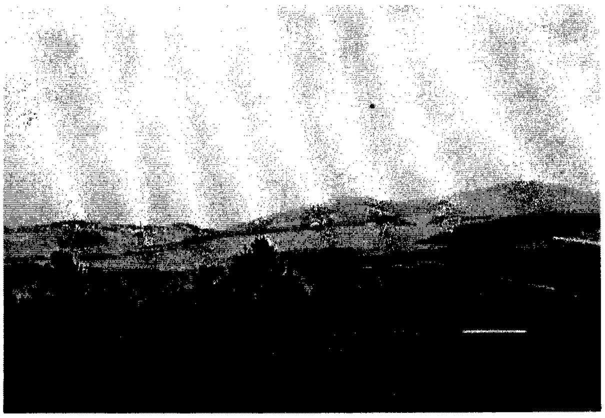


Figure 6. Panorama of the case study area "Zlatibor

the procedure given above, probability for each map unit (grid cell) was produced. Probabilities of viewshed are given in the map legend (figure 8).

Contour of viewshed probability area is very similar with outline of the viewshed area acquired with initial DEM. Different levels of probabilities for viewshed lie within that contour.

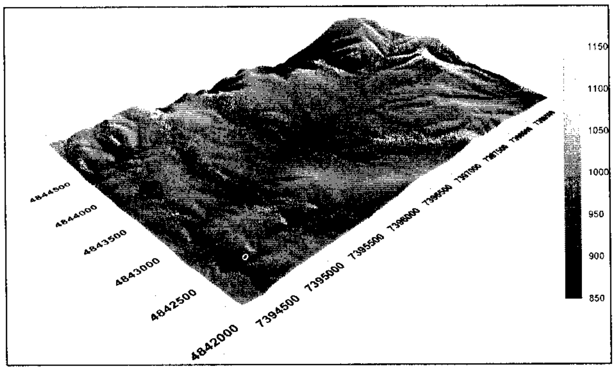


Figure 6. Panorama of the case study area "Zlatibor"

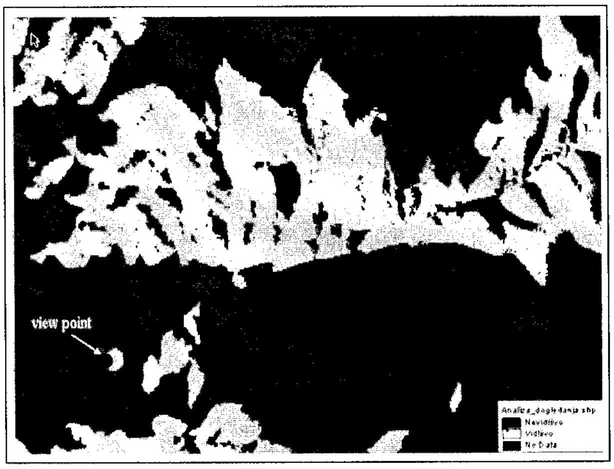


Figure 7. Thematic display of viewshed analysis for the test area

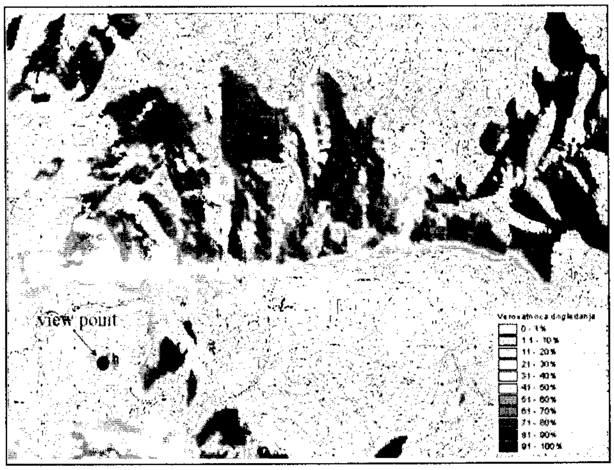


Figure 8. Probability map of viewshed for test area "Zlatibor"

CONCLUSION

In the GIS environment, viewshed analysis is recognized as a helpful system component, contributing to interpretation of spatially related phenomena and complex data analysis that take the GIS a step beyond two-dimensional polygonal overlay analyses. Most applications of viewshed analysis are based only on topographic surfaces, but in some cases the topographic surface has forest covers with heights which differ to relief elevations. To get plausible results, it would be better to use Digital Surface Model (DSM). While DTM or DEM describe the earth surface in the sense of the "bald earth" without human artifacts such as buildings or bridges and without vegetation, DSMs include points on buildings and vegetation as well as terrain points.

Very often, outputs provided by GIS functions are assumed as exact results. Like other spatial data bases, DTM data bases can be affected by errors. It is very important to examine the effect that they may have on data analysis and modeling of spatial data. DTM data are often used in analysis without quantifying the effects of these errors. Probability maps based on Monte Carlo simulation techniques are very suitable for evaluating the impact of DEM error on viewshed analysis. These kinds of maps can be used as an accompanying material with exact thematic displays for any kind of displays which are the results of the GIS analyses, serving as a measure of reliability.

References

1. Burrough, P. (1986). Principles of Geographical Information Systems for Land Resource Assessment. Oxford University Press, New York, NY:194 pp.
2. Burrough, P., McDonell R (2000). Principles of Geographical Information Systems. Oxford University Press, New York, NY: 333 pp.
3. Caruso V.M. (1987) Standards for digital elevation models. Proceesings of the ACSM/ASPRS Annual Convention, Baltimore, pp.159-166.
4. De Floriani L., Magillo P. (1999) Intevisibility on terrains. - In: Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W.: Geographical Information Systems, Vol 1. John Wiley&Sons. pp. 543-556.
5. DeMers M. (2000) Fundamentals of Geographic Infoprmation Systems. John Wiley&Sons, New York, NY: 498 pp.
6. Fisher P. (1999) Models of uncertainty in spatial data- In: Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W.: Geographical Information Systems, Vol 1. John Wiley&Sons, pp. 191-205.
7. Glemser, M., Klein, U., Fritsch, D., Strunz G. (2000) Complex Analysis Methods in Hybrid GIS Using Uncertain Data. In: GIS - Geo-Informationssysteme, Wichmann Verlag,

Heidelberg. No. 2, pp. 34-40.

8. Isaaks E., Srivastava R. (1989) An Introduction to Applied Geostatistics. Oxford University Press, New York, NY: 561 pp.

9. Journel, A. (1996). Modelling uncertainty and spatial dependence: stochastic imaging. International Journal of Geographical Information Systems, Vol. 10, No. 5, pp. 517-522.

10. Kraus K.(1995) From digital elevation model to topographic information system. In: Week, D. Fritsch and D. Hubbie (ed.): 45th.Photogram metric week, Stuttgart, pp. 277-285.

11. Li Z. (1991) Effects of Check Points on the Reliability of DTM Accuracy Estimates Obtained from Experimental Tests. PERS, 57, (10), pp. 1333-1340.

12. Merchant, D.C. (1987). Spatial accuracy specification for large scale topographic maps, PERS 53, pp. 958-961.

13. Shortridge A.(2001) Characterizing Uncertainty in Digital Elevation Models., In: Hunsaker C.T., Goodchild M.F., Friedl M.A., Case T.J. (ed.): Spatial Uncertainty in Ecology, Springer-Verlag New York, pp.238-257.

14. Tripcevich N. (2002) Viewshed Analysis of the Ilave River Valley. Unpublished UC Santa Barbara Anthropology Master's paper.

15. Wechsler S.P. (2000) Effect of Digital Elevation Model (DEM) uncertainty on topographic parameters, DEM scales and terrain evaluation, Ph.D. Dissertation, New York.