Session / Séance 23-D

Computing and Visualizing Morphologically Sound DEMs in a Raster Environment

Juha Oksanen *and* **Olli Jaakkola**

Finnish Geodetic Institute, Cartography and Geoinformatics
Geodeetinrinne 2 (P.O. Box 15), FIN-02431 Masala, Finland
Tel. +358-9-295 550
Fax. +358-9-295 55 200
e-mail: juha.oksanen@fgi.fi, olli.jaakkola@fgi.fi

Abstract

The present study dealt with the problem of interpolating a continuous digital elevation model (DEM) from contour lines. Several methods were developed that make use of an interpolation algorithm. The basic idea behind the methodology has been the use of tools provided in a standard raster geographic information system (GIS) environment and the avoidance of problems related to the triangulation algorithms. The interpolation algorithm was developed using raster model operations with 3 different linear solutions. Separate morphological corrections for peaks and pits were calculated. The interpolated DEM can either be used as such for accurate spatial analysis, or generalized for coarser scale analysis, or smoothed for visualization of terrain. The generalization and visualization of DEMs forms another problem area, in which no single solution appears to achieve all the desired properties into a single DEM.

Several methods for visualizing the differences between various algorithmic solutions were used to reveal the potentials and drawbacks of different interpolation methods. The interpolation algorithms were tested against 2 DEM generation algorithms provided in a common GIS software. The results show that the interpolation algorithm developed results in minimum deviations to the original data. Morphological features are represented more naturally with the algorithm, and the resulting DEM is visually pleasing since triangulation artifacts are not present.

Introduction

The use of digital height, elevation, or terrain modeling has spanned several decades since the 1950s [e.g. Miller & Laflamme 1958]. Similarly, spatial processes have been modeled using terrain information derived from the digital elevation model (DEM), especially some processes that are directly affected by the terrain characteristics.

This research is restricted to derivation and manipulation of a regular GRID DEM from various data sources. The most important data source used here is contour lines, which originally were stored as elevation isolines. They are, at present, the only available DEM data source covering the whole of Finland. Since no limits are placed on the output DEM, the study concerns both the interpolation of the DEM from source data as well as smoothing or generalization of various coarser scale-level DEMs. Originally, no limits were set for the input data to enable use of multisource data integration, and some of the methods are still open for integrating various types of elevation data into one homogeneous output DEM. The interpolation methodology was based on a deterministic model in which no data uncertainty was considered, with 3 versions of linear interpolation algorithms. Comparisons of these algorithms have been made producing both accurate data for morphological analysis as well as visually pleasing models. The key issue is always accuracy, since no transformations are performed without first ensuring preservation of the original data values.

There are sound reasons for this research since present interpolation methods have several defects, of which some have been resolved with the methods developed. None of the present interpolation methods adapts to terrain morphology perfectly, each is variously defective. The aim is to preserve the important morphological features, e.g. peaks, pits, ridges, drainages, passes, and planes, as much as possible. The difficulty in this is that the contour lines do not explicitly contain all these morphological features; they are merely interpreted visually by the reader during the map-reading process. This visual ‘gestalt´ information should be explicitly stored on the computer. Preservation of the original data accuracy is also important for some types of analysis, since errors are also preserved in those data. This must be accepted since often it is not known where the errors are, and therefore the correction procedures are hard to carry through. Most present interpolation methods are also not sufficiently powerful for large DEM databases with smaller pixel sizes. These present methods do not consider interpolation as a purpose-driven task; thus the same DEM model should serve as a basis for all types of spatial analysis as well as for visualization of terrain morphology combined with other data sources. Often interpolation of a DEM is not seen as scale-driven process in which the input data, interpolation method, and output pixel size set ‘the scale limits’ for the DEM model. Evidently, the most accurate DEM should be seen as a data source for generalizing various DEMs to different scale levels, since one DEM cannot be combined with all types of data source at different scale levels. Lastly, since the introduction of new data-gathering methods, such as radar images and global positioning system (GPS) measurements, the present contour-based DEMs will be supplemented with different types of height values, and in the future we will have access to multisource input data. Most present interpolation methods do not allow full exploitation of multisource data, and therefore we deteriorate the original data quality of height points and isolines or flat areas, which are essentially different types of height value collections. Finally, no global interpolation solution is suitable for all purposes, which justifies the many interpolation procedures presented here.

Concepts and hypothesis

The DEM is a regularly spaced matrix of height values with planimetric coordinates. It is differentiated from the Digital Terrain Model (DTM), which in addition to height points may include some higher-level information, e.g. on terrain discontinuities or other derived attributes [Burrough 1986]. The DEM consists of a 2.5 dimensional model in which planimetric locations are associated with elevation values from mean sea level determined by geoid height. The true 3-D model would need a model in which attribute information could also be associated with height, i.e. x, y for planimetric coordinates, z for elevation and a for attribute information. Normally the DEM is arranged with regularly spaced grids having a fixed pixel size, e.g. 10 m. Therefore the DEM is only an abstraction of continuous natural terrain.

Since the DEM is a regular arrangement of height values this involves calculating most of the primitive derived attributes using either 4- or 8-directional topology, i.e. using only orthogonal directions or in addition both diagonal directions. The most common derived attributes such as slope, aspect and curvature are calculated using this topology with a fixed pixel size. Therefore, the attributes derived are tied to a certain scale-level, and describe the characteristics of that scale level. The DEM is by no means a general description of the terrain, and it may have various height values according to the pixel size used, topology defined, and the neighborhood.

Several requirements for the use of interpolation methods were described in the study. Many of these requirements can be fulfilled, although all methods have their advantages and disadvantages. The 1st requirement is that the interpolation method should allow the use of multiple input data at varying density, points, lines, and areas with single or multiple elevation values. This is the requirement that most commonly used present interpolation methods find difficult to fulfill, since some of the essential morphological features are often lost from input data, such as exact contour values or elevation values of the lake surfaces. The 2nd requirement is that the deterministic linear algorithms assume, although should perhaps not, equal quality for the input data, therefore some type of prefiltering or error recognition procedure may be needed prior to actual interpolation. Thirdly, the computing cost should be relatively low to allow large areas and large data volumes to be interpolated simultaneously. Fourthly, the interpolation method should make no or minimum deviations to the original data, which requirement assumes that measured or interpreted height values are already better than interpolated values. The side effect is of course that errors in input data will be copied into the interpolated DEM. Fifthly, a contour-specific requirement may specify that height values between contour pairs should be within that pair of values and should change more or less linearly. Sixthly, the interpolation methods should produce good data for GIS analysis and should be capable of retaining some predefined topological or other properties important for certain types of spatial analysis. Seventhly, the interpolation method should produce good data for visualization purposes, noting that the output should be visually pleasing without too much emphasis on other applications. Fulfillment of all these requirements is analyzed in the ‘Results’.

Methods of DEM interpolation

Interpolation here is restricted to the creation of regularly spaced GRID DEM. The study concentrates on linear interpolation methods in which we have developed or compiled algorithmic solutions to simple linear interpolation problems. All solutions are raster based, i.e. input data need a translation to raster model prior to the actual interpolation. This will evidently introduce some error, either planimetric or height to the grid points, depending on whether nearest neighbor assignment or a more advanced method, e.g. bilinear, will be used. Since the original data were gathered as contours from the 1:10 000 scale paper copies, the pixel size of 5 m introduces unimportant small errors into the output DEM.

The interpolation methods may have flexible data sources as an input, and eventually a multisource data input for the DEM may become available. In Finland the nationwide DEM is interpolated from topographic map contours which have been used as input data for the DEM interpolation. Here we developed 3 versions of a simple linear interpolation algorithm. Essentially, they are all based on calculations of distances from known height values and on setting the interpolated value as a ratio of distances to known elevation values, e.g. to 2 nearest known elevation values using formula, where h = elevation value to be interpolated, h_1 and h_2 = two nearest known elevation values, d_1 and d_2 = distances to two nearest known elevation values:

$$h = h_1 \times d_2 / (d_1 + d_2) + h_2 \times d_1 / (d_1 + d_2) \quad (1).$$

Two of the solutions are based on this direct linear approach, using either euclidean distance (ICHI) or cost distance (LIDIC). The 3rd (SIFI) is an iterative mean filtering approach which has most of the properties of direct linear approach, although the linearity of the solution is also dependent on how many iterations are executed and on how parallel the contour lines are. The LIDIC and ICHI also involve a smoothing procedure to filter out the effect of the raster topology of neighboring pixels used in the calculation. None of the methods developed need any triangulation phase prior to the actual interpolation, thus avoiding some of the associated problems.

The 1st interpolation method is the most complicated of the developed methods and be termed as Linear Inverse Distance Interpolation from Contours (LIDIC). It is a variant of the simple inverse distance weighting (IDW) method modified especially for contour data. The principles are explained e.g. by Fukue et al. (1990), and the method eventually combines 2 different interpolation approaches: contour halving providing the flexibility and linear distance interpolation providing the accuracy and morphology. LIDIC first involves the conversion of contours to the raster mode, and the finding and repairing of contour parts where they are only diagonally connected. LIDIC then creates a temporary surface by contour halving, selects contour elevation range, and creates cost surfaces out of each contour within the elevation range. The elevation values are calculated using the following formula, where h_j = point to be interpolated, h_i = elevation values of regular/auxiliary contours and lakes, d_{ij} = cost distances to contours and lakes and n = number of elevation values between elevation ranges:

$$h_j = \frac{\sum_{i=1}^{n} \left(\frac{1}{d_{ij}} * h_i\right)}{\sum_{i=1}^{n} \frac{1}{d_{ij}}} \qquad (2).$$

The DEM calculated from the elevation range is merged with the contours, and the next elevation range is selected until the highest interval is chosen. The peaks and pits are set separately to the LIDIC, and added to the interpolated DEM by calculating euclidean distances inside the peak and pit contours and assigning half of the contour range to the maximum distance point. The respective linear values are calculated between the maximum distance point and the contour line.

The 2nd interpolation method is an Iterative Contour Halving Interpolation (ICHI). Although it partly resembles the Fukue solution, the ICHI approach is not only limited to the contour values, but is a general solution for iteratively filling in the terrain using a simple linear halving procedure from known height values. It does not have as rigid requirements for the input contours, since height points will do equally well. The solution uses Medial Axis Transformation (MAT) algorithm [e.g. Tang 1992], of which the normal medial axes are given new values during iterative interpolation by setting the value as a mean of the 2 nearest neighboring elevation values. It uses euclidean allocation for finding the places for half the distance to the known values and sets the new elevation value by simply using the following formula, where h_j = point to be interpolated and h_1 and h_2 = elevation values of 2 nearest known points:

$$h_j = (h_1 + h_2) / 2 \qquad (3).$$

The calculated DEM values are combined with the original elevation values, and in the following ICHI iteration all the new data together with existing data are used for calculating euclidean allocation for again halving the space, and calculating the new values with the same formula until all the space is filled with known values. In the ICHI the peaks and pits are also set separately and added to the interpolated DEM similarly as in the LIDIC.

The 3rd interpolation method is a Simple Iterative Filtering Interpolation (SIFI). The SIFI is also not only limited to the contour values, but is a general solution for iteratively filtering the entire DEM. The SIFI simply uses the same formula 3 for calculating new values for the unknown center pixel within a simple 3 x 3 filter. The filter is run so many times that the interpolation iteratively fills the entire terrain. The SIFI also needs an addition for the peak and pit values prior to initiating the iteration procedure. Both the ICHI and SIFI interpolation methods are open for additional morphological information, e.g. for adding river networks with values interpolated from contours as giving the exact hydrological structure to the DEM. This results in a truly realistic nature for the interpolated DEM surface, assuming that the known elevation values from contours fit well together with the river network.

The interesting point in the filtering approach is that it is a solution for generalizing the DEM. The number of iterations may first be altered and therefore produce a series of generalized DEMs having differing numbers of iteration runs. The other solution is to increase the radius of the filter and therefore also produce a series of generalized DEMs with different radii. The difficulty in generalization is that not all the properties of the original DEM may be preserved, and therefore we must choose the properties that we do want to preserve. Evidently, the SIFI does not provide a sufficiently flexible generalization procedure, and therefore the study on generalization of DEMs continues.The reference DEMs used as comparative models in quality testing were generated by using 2 standard methods provided by the ARC/INFO. A triangulation-based model (RasterTIN) was triangulated using the Delaunay method with shoreline data added as hard breaklines [ARC/INFO User's Guide 1992]. Since the model was rasterized by linear interpolation no other morphological data were used. TIN models are very widely used, because they are considered to be very accurate and triangulation adapts well to data sources with highly varying density. TINs conserve every measured data point directly and incorporation e.g. of shorelines, streamlines and fault lines is quite easy [Petrie 1990], however triangulation itself does not provide a good solution for calculating elevation values for these morphological features. The other reference DEM was interpolated using Hutchinson's spline-based TOPOGRID, which was developed especially for models used in hydrological analyses. In TOPOGRID iterative finite difference interpolation should overcome the undershoot and overshoot errors of the unmodified minimum curvature interpolated surface [Hutchinson 1989], although some errors are still present. A drainage enforcement algorithm, which was used in reference model interpolation, attempts to guarantee freedom of DEM from spurious sinks.

Methods of DEM quality evaluation

Evaluation of the DEM quality is not a trivial task. Normally, accuracy of the DEM may be based on basic parametric statistics, such as mean error and standard deviation or as root mean squared error (RMSE). They give a general description about the accuracy, but no information on the spatial structure of error or the usefulness of DEM for certain specific applications. Therefore, there is a need for more precise measures for describing the spatial and morphological accuracy of the DEM, which is important for use of the DEM in various applications as already stated. The spatial randomness or clustering may be described by using some measure of spatial autocovariance. The problem with any spatial autocovariance statistic is that it gives only a single value for the entire area, but no indication of morphological clusterness. The common statistics also give only an absolute term for the entire terrain, but not a measure adjusted to the relative differences of elevations at various places, and therefore they give no indication of the relative importance of errors at various terrain types. Similarly, it is important to separate the interpolation errors from the other sources of error: errors in input contours, points and areas with elevation, measured and used checkpoints, or used raster or vector data model. Although the accuracy of the DEM is given only as height value, it must be remembered that the elevation errors also indicate the planar errors. Furthermore, some trials for analyzing errors in the derived DEM attributes, such as slope in connection with mean error have been performed [e.g. Koppen formulas by Imhof 1982]. There is evidently a need for more precise accuracy measures for deriving information from DEMs with different types of spatial analysis, not only simple measures such as slope and aspect, but also complex measures such as profiles, streams, ridges, catchment areas, wetness indices, viewsheds, and irradiances.Yet another area of error analysis is the demand for accuracy from the point of view of application. Some of the problems involve determining the general requirements for accuracy (e.g. maximum desirable errors), the critical types of errors (e.g. overshoots, undershoots, topology of the DEM model), and where the critical points of the model are located (e.g. based on morphological division). Examples of application demands for hydrological analysis may given as the critical points, such as sinks, pass points, errors in derived slopes or aspects, or demands for visibility analysis [e.g. Fisher 1993].

Visualization of the errors needs yet further handling. Traditional DEM visualization methods are normally developed for qualitative data, e.g. differences in contour lines expressed in colors. Although hill-shading is a powerful method for visualizing common artifacts of the DEM, it is not capable of expressing the morphology in detail. The quantified error images, shaded or not, do not fully express the differences in DEMs in detail, if they are not classified. The error images may be calculated using test points, contours, or another DEM as masters. The data of interpolated master-models may be compared against slave-models by subtracting the values and visualizing the subtracted values with the methods described Jaakkola & Sarjakoski 1993. Similar methods are also to some extent applicable to comparing interpolated models against contours [Wood 1996] or comparing models to ground truth points, which is the only statistically valid comparison. The relative comparisons, however, are important for giving more detailed images of morphological differences between the models.

Testing of interpolation methods

Evaluations of quality have been made to the fully completed DEMs, although the interpolation methods themselves are composed of several steps. In the comparison these steps are treated as a single interpolation method with one output. Evidently, as stated previously, future research should include comparisons of single methods with multiple parameters adjusted more toward a single-application purpose, since one model may not fulfill all the requirements for multiple applications.

A number of problems are related to the testing of interpolation methods [see Carrara et al. 1997]. The 1st, and perhaps not yet fully understood problem is that no single DEM exists that is suitable for all purposes, a collection of DEMs adjusted to the different terrain types and to the different types of usage will always be needed. Likewise, the absolutely best interpolation method cannot be found for all the purposes. Terrain morphology includes simple general statements such as flat is flat, gentle is gentle, and steep is steep which should be represented in a DEM similarly, e.g. so that the grid topology of the number of direct neighbors (4 or 8) used in the calculation does not deteriorate morphological terrain characteristics. Some problem related to generalization of the terrain surface are: how generalized is the DEM used, how much of the terrain characteristics is preserved, and which characteristics are preferred. This matter is not discussed as thoroughly, although it is evident that during the generalization process only certain characteristics may be preserved, and that the generalization method must be adjusted carefully.

The problem with ground truth is important from the point of view of how representative the points are. Whether the point is really on the surface or above or below the surface is actually dependent on the scale level at which we define the surface, e.g. how detailed changes on the surface are accounted for. The problem with sampled ground truth points is emphasized when the terrain is composed of different morphological characteristics. Not only should a statistically sufficient number of sample points be available, but also a sufficient number of sample points for each morphological feature. Ground truth should reflect the sampled terrain landscape in its full morphological diversity.

The test data are composed of several point samples. Some of the control points were already stored from the previous ground campaigns and were used as the basis for test data. The dataset includes various types of sampled data, with GPS measurements, tachymetric measurements, base triangulation points, etc. This data set is then supplemented with ground measurements using GPS equipment to derive a better representation of the terrain morphology. Test data were gathered from a 175 km^2 area, although at the time of writing only a smaller comparison has been completed. In addition to independent ground truth some direct comparisons with other interpolated data sources should be done to emphasize the relative differences among various interpolation methods.

Results

Since at the time of writing only the preliminary results have been available, these results can be briefly summarized in Table 1.

Table 1. General conclusions from comparison of various interpolation methods. LIDIC is the linear inverse distance interpolation from contours, ICHI is the iterative contour halving interpolation, SIFI is the simple iterative filtering interpolation and RasterTIN is the rasterized TIN model. Requirements for interpolating a DEM include: 1) it should allow multiple input data at varying density, 2) it should assume equal quality for the input data, 3) it should have low computing cost, 4) it should make no or minimum deviations to the original data, 5) it should hold the elevation values between contours within a contour range, 6) it should produce data suitable for GI analysis, and 7) it should produce data suitable for GI visualization.

	Requirements:						
Interpolation methods:	*1*	*2*	*3*	*4*	*5*	*6*	*7*
LIDIC	poor	good	poor	good	good	good	good
ICHI	good	good	good	good	good	fair	fair
SIFI	good	good	good	good	good	fair	poor
RasterTIN	good	good	good	fair	good	poor	poor
TOPOGRID	good	poor	fair	fair	fair	fair	good

The immediate impression is that the differences are moderate, which can be seen from Table 1 and Figure 1. RasterTIN visually differs more from the others. It can also be seen that the LIDIC is limited from its input data and has a higher computing cost. Other differences are small and can be seen in more detailed in Figure 1. LIDIC gives the best description of terrain morphology while adding no undesired extra forms, whereas TOPOGRID is also adequate, but sometimes (since it uses splines) adds artificial pits to the flat terrain near steep slopes. ICHI is generally somewhat less expressed in ridges and drainages, but together with TOPOGRID makes it much more possible to include additional information from other data sources. SIFI is a relatively simple approach and needs further improvement to better describe morphological features. SIFI and especially ICHI have low computing costs, and can be used for large databases.

Discussion

Some suggestions for continuing research may be drawn from the results. It is evident that each interpolation method offers its own benefits. Of the 3 methods developed LIDIC is perhaps the most accurate, but since it accepts limited input data, it likewise offers limited use for continuing multisource data interpolation. ICHI makes it possible to include additional data (stream network and other GIS sources) and should be studied further. SIFI is presently less sensitive to morphology but as a general filtering approach makes it possible to generalize coarser DEMs, using a conditional filtering approach rather than simple filtering. The overall quality of the 3 methods is acceptable: they all offer quite error-free DEMs for spatial analysis, assuming that input data are also error-free. Research on DEM interpolation is continuing at the Finnish Geodetic Institute, since the above results to some extent can still be considered to be preliminary.

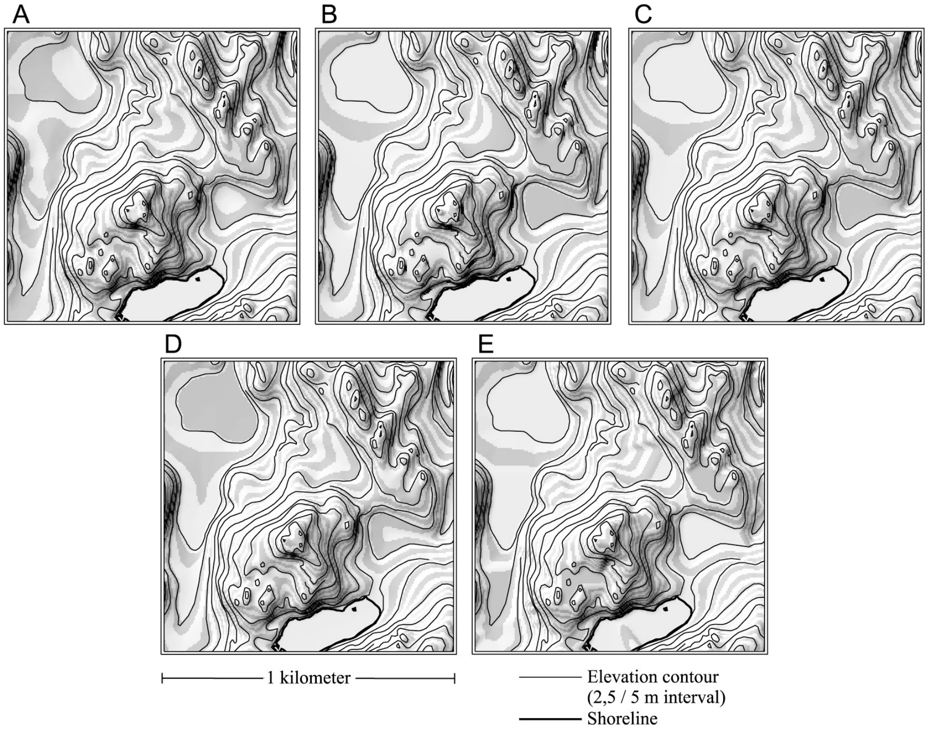

Figure 1. Filtered DEMs created with 5 different interpolators: A - spline-based TOPOGRID, B - simple iterative filtering interpolation (SIFI), C - iterative contour halving interpolation (ICHI), D - linear inverse distance interpolation from contours (LIDIC) and E - rasterized TIN model. Elevation contours were used as an input data for interpolating DEMs. All models have 5-m pixel size and each meter of elevation change is represented by distinct gray shades (e.g. 50 to < 51m lighter gray, 51 to < 52m darker gray, 52 to < 53m lighter gray, etc.)

References

ARC/INFO User's Guide (1992). Surface Modeling with TIN. *Environmental Systems Research Institute Inc.*

Burrough, P.A. (1986). *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon Press, Oxford.

Carrara, A., G. Bitelli & R. Carla (1997). Comparison of techniques for generating digital terrain models from contour lines. *International Journal of Geographical Information Science*, 11(5), 451-473.

Fisher, P.F. (1993). Algorithm and implementation uncertainty in viewshed analysis. *International Journal of Geographical Information Systems*, 7(4), 331-347.

Fukue, K., Kuroda, Y., Shimoda, H. & Sakata, T. (1990). Simple DEM generation method from a contour image. *Proceedings of the Symposium on Cartographic And Data Base Applications of Photogrammetry And Remote Sensing*, May 15-18, 1990, Tsukuba, Japan, 347-355.

Hutchinson, M. (1989). New procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology* 106, 211-232.

Imhof, E. (1982). *Cartographic Relief Presentation*. Walter de Gruyter, Berlin.

Jaakkola, O. & T. Sarjakoski (1993). Visualization of the quality of satellite-based land-use classification. *Surveying Science in Finland*, 11(1-2), 61-81.

Miller, C. L. & R. A. Laflamme (1958). The digital terrain model: theory and application. *Photogrammetric Engineering* 25, 433-442.

Petrie, G. (1990). Modelling, interpolation and contouring procedures. In Petrie & Kennie *(Eds.). Terrain Modelling in Surveying and Civil Engineering*, McGraw-Hill, Inc. New York.

Tang, L. (1992). Automatic extraction of specific geomorphological elements from contours. *Proceedings 5th International Symposium on Spatial Data Handling*, Charleston, USA, 554-566.

Wood, J.D. (1996). *The geomorphological characterisation of digital elevation models*, PhD Thesis, University of Leicester, U.K.