18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Locally adaptive gridding of noisy high resolution topographic data

Hutchinson, M.F., John A. Stein, Janet L. Stein and Tingbao Xu

The Fenner School of Environment and Society, The Australian National University, Australian Capital Territory
Email: michael.hutchinson@anu.edu.au

Abstract: Topography plays a fundamental role in modulating land surface and atmospheric processes across a wide range of spatial scales (Hutchinson 2008). Thus digital elevation models (DEMs) have provided a key role in supporting mesoscale representations of surface climate as well as supporting finer scale representations of surface hydrology and catchment processes. The usual means of supporting these representations is a regular grid elevation model interpolated from a variety of data sources. The ANUDEM locally adaptive elevation gridding procedure (Hutchinson 2007) is commonly used to calculate these elevation models. Two key features of the method are its computational efficiency, allowing it to be applied to very large data sets, and a drainage enforcement algorithm that attempts to maintain connected drainage structure in the interpolated DEM, a key necessity for hydrological applications.

The ANUDEM procedure has been steadily upgraded over the last two decades to process a wide range of topographic data and to improve the quality of its underlying algorithms. In particular, the drainage enforcement algorithm has undergone considerable revision. Traditional data sources include point elevation data and stream line data. These have been augmented by contour elevation data, coastline data, cliff line data, stream distributaries and lake boundary data. All of these have been used to calculate Version 3 of the national 9 second digital elevation model, recently jointly released by the ANU Fenner School of Environment and Society and Geoscience Australia (2008).

Until recently the source elevation data have been relatively sparse and essentially error free, apart from coding errors. The ANUDEM procedure has applied an efficient multi-grid algorithm to underpin the basic interpolation of these sparse data according to a variety of roughness penalties that can be tuned to the different data sources. Thus sparse point elevation data has been interpolated using a roughness penalty composed of surface curvature and potential. This has been particularly effective in eliciting surface drainage structure from relatively sparse surface specific elevation data points. Contour data have been effectively interpolated using a minimum curvature roughness penalty augmented by a locally adaptive minimum profile curvature penalty. In both cases initialisation of elevations on data streamlines has relied on the error free nature of the source elevation data.

Recent elevation data sources are airborne and spaceborne platforms using laser and radar techniques. They differ from the traditional data sources in two fundamental ways. The data typically have high spatial resolution, less than 1 metre for airborne laser data, and all have significant elevation errors. The errors can be both systematic, as particularly evident in data obtained by spaceborne platforms, such as SRTM data, and sporadic. Sporadic errors are associated with surface features such as vegetation cover and man-made structures.

This paper describes key modifications to the ANUDEM procedure to effectively process these noisy, high resolution data sources. The multi-grid interpolation procedure is still effective in stably interpolating high resolution data. This is important in enabling effective application of drainage enforcement to the interpolated grid. The different errors in the source data can be specifically accommodated by smoothing the data according to the differing magnitudes of the variances of the errors of the source elevation data. This is particularly appropriate when the data have been subject to pre-processing. Finally, revision of the initialization of heights on data streamlines is required to prevent corruption of stream heights by noisy elevation values and to improve the computational efficiency of the initialization of distributary streamlines in the presence of dense source elevation data. Both are facilitated by the underlying stable multi-grid interpolation method.

Keywords: *Digital elevation model, multi-grid interpolation, data smoothing, drainage enforcement*

1. INTRODUCTION

Digital elevation models play a central role in environmental modelling across a range of spatial scales. The regular grid mode of representation has become the dominant form for digital elevation models used in these applications. This form is directly compatible with remotely sensed geographic data sources and can simplify terrain-based analyses, including assessments of spatial scale. A distinguishing feature for many applications, particularly those that operate at finer scale, is a primary requirement for information about terrain shape and drainage structure, rather than elevation. For this reason, elevation contours and streamlines have remained popular sources of primary topographic data. They can be used to construct fine scale digital elevation models by gridding methods that are locally adaptive to surface shape and drainage structure. Remotely sensed digital topographic data, from both airborne and spaceborne sensors, are an emerging source of fine scale digital elevation data. A major impetus for this development has been the goal of generating high resolution DEMs with global coverage.

This has recently been achieved with the completion of the 3 second (90 m) DEM for the globe obtained from the Shuttle Radar Topography Mission (USGS 2005). These data have two generic limitations. The sensor could not measure ground elevations underneath dense vegetation cover or man-made structures, leading to sporadic errors in vegetated areas of up to 10m. Secondly, all measured data had significant random errors that depended on the inherent limitations of the observing instrument, as well as surface slope and roughness (Harding *et al.* 1994). The product specification for the SRTM data is that 90% of the elevations have error within ±16m. These errors require appropriate filtering, without degrading shape and drainage structure, to maximise the utility of the data in environmental applications, particularly in areas with low relief or with significant surface cover.

The ANUDEM locally adaptive gridding procedure (Hutchinson 2007) can be used to construct digital elevation models from digital elevation contours, surface specific point elevations and streamlines so that the elevation models preserve terrain shape and drainage structure. Grid resolution can be optimised to match the true information content of the source data and to maximise the quality of primary terrain parameters derived from the interpolated DEM. These elevation sources are all characterised by being relatively sparse and having negligible elevation error. SRTM data, on the other hand, are relatively dense and have significant elevation error. Applying the ANUDEM procedure successfully to these data raises a number of issues, particularly with reliable initialisation of elevations on data streamlines in a way that is consistent across the borders of neighbouring map tiles in low relief areas. These issues are examined by developing and applying a modified version of ANUDEM to 1 second SRTM data for a low relief area in the Murray-Darling Basin. The data had been pre-processed to largely remove errors due to vegetation cover (Gallant and Dowling, *private communication*). Drainage structure was enforced in the filtered DEM by using corrected GEODATA 250K streamline data (Geoscience Australia 2006) in association with automated drainage enforcement algorithm. The corrected streamline data were as produced for Version 3 of the 9 second DEM (ANU Fenner School of Environment and Society and Geoscience Australia, 2008)

As for traditional data sources, the process of producing an accurate DEM requires careful attention to the accuracy of the source data and the quality of the interpolated DEM. A prime requirement for hydrological applications is that the filtered DEM accurately represents shape and drainage structure. Shaded relief views and the number of remaining depressions in the filtered DEM are simple and effective shape-based measures of DEM quality that do not require the existence of separate reference elevation data. Remaining depressions can also be readily plotted to greatly assist in detecting and remedying data errors. The number of remaining depressions was the prime measure of DEM quality in the production of the national 9 second DEM. It was also a key measure of progress here in the development of the modified locally adaptive gridding procedure.

Once finalised, this methodology will be used in partnership with CSIRO Land and Water, Geoscience Australia and the Bureau of Meteorology to produce a new 1 second DEM for Australia with an accurate representation of shape and drainage structure. This will eventually upgrade the coordinated national coverage currently provided by Version 3 of the 9 Second DEM. Earlier versions of this DEM have already supported continent-wide hydrological modelling and related analyses (Hutchinson *et al*. 2000, Gallant and Dowling 2001, Stein *et al*. 2002).

2. KEY FEATURES OF ANUDEM

The ANUDEM program can process arbitrarily many different input data files, each of arbitrary size. The only size limit imposed by the program is the size of the fitted DEM which needs to be stored in the memory of the computer running ANUDEM. Each data file may be one of eight types:-

1. Point elevation data
2. Sink point data
3. Streamline data
4. Boundary polygon data
5. Contour line data
6. Lake boundary data
7. Cliff line data
8. Mask boundary data

The program reads input data points from each input data file, trims the data to the user-specified map limits and then generalises the data to the user-specified grid resolution. Point elevation data are generalised by using the average elevation of up to 100 data points per grid cell and discarding any remaining points. Line data are generalised by accepting at most one line data point per grid cell, and, in the case of stream line and contour line data, removing unnecessary kinks.

2.1. Underlying interpolation algorithm

The program interpolates the accepted elevation data onto a regular grid by minimising the sum of a user-specified roughness penalty and a weighted sum of squares of the residuals from the elevation data of the surface represented by the grid. This is best described by first defining an appropriate statistical model for the observed elevation data. Each elevation data value z_i at location x_i, y_i is assumed to be given by

$$z_i = f\left(x_i, y_i\right) + \varepsilon_i \qquad \left(i = 1, ..., n\right) \tag{1}$$

where f is an unknown suitably smooth bivariate function of horizontal location represented as a finite difference grid, n is the number of data points and ε_i is a zero mean error term with standard deviation w_i. For accurately surveyed elevation data the standard deviation is dominated by the natural discretisation error of the finite difference representation of f. Assuming that each data point is located randomly within its corresponding grid cell, the standard deviation of the discretisation error is given by

$$w_i = hs_i / \sqrt{12} \tag{2}$$

where h is the grid spacing and s_i is the slope of the grid cell associated with the i th data point (Hutchinson 1996). The function f is then estimated by solving for the regular grid finite difference approximation to the bivariate function f that minimises

$$\sum_{i=1,n} \left[\left(z_i - f\left(x_i, y_i\right)\right) / w_i\right]^2 + \lambda J\left(f\right) \tag{3}$$

where $J\left(f\right)$ is a measure of the roughness of the function f in terms of first and second derivatives (Hutchinson 1989) and λ is a positive number called the smoothing parameter. The smoothing parameter λ is normally chosen so that the weighted residual sum of squares in equation (3) is equal to n. This can be achieved with an approximate Newton-Rhapson method coupled with the iterative solution of f (Hutchinson 2000). The spatially varying weights in the residual sum of squares in (3) is a <u>locally adaptive</u> feature that can only be achieved with an iterative interpolation method for which the slopes of the grid cells are available as the iterative solution proceeds.

The program employs a simple multi-grid method to minimise equation (3). This calculates grids at successively finer resolutions, starting from an initial coarse grid, until the final, user-specified grid resolution. For each grid resolution, the accepted data points are allocated to the grid and the grid values are calculated by Gauss-Seidel iteration with over-relaxation (SOR method) subject to the user-specified roughness penalty, while simultaneously respecting ordered chain constraints associated with the drainage enforcement algorithm and data streamlines and breakline conditions across all data cliff lines.

Iteration terminates at each successive grid resolution when the user-specified maximum number of iterations (normally 20) has been reached. Starting values for the first coarse grid resolution are set to the average elevation of all elevation data points. Starting values for each successive finer grid are simply interpolated from the preceding coarser grid. On completion of the iterations, the program calculates all sink points remaining in the fitted grid and writes a detailed summary to an output log file. The remaining sink points are written to an output file for plotting to aid in detection and correction of input data errors. Stream line data, as incorporated onto the grid, and other diagnostics, are also written to output point and line diagnostic files for further assessment of the gridding process.

2.2. Drainage enforcement algorithm

The drainage enforcement algorithm attempts to remove all sink points that have not been identified as input sink data. This imposed global drainage condition has been found in practice to be a powerful condition that can significantly increase the accuracy of a fitted digital elevation model, especially in terms of its drainage properties (Hutchinson 1989). The global drainage condition minimises the need for detailed manual editing of interpolated elevation grids to remove spurious drainage features.

The essence of the drainage enforcement algorithm is to find for each sink point the lowest adjacent saddle point that leads to a lower data point, sink or edge. Provided a conflicting elevation data point has not been allocated to the saddle, the algorithm then enforces a descending chain condition from the sink via the intervening saddle to the lower data point, sink or edge. This action is modified by the systematic application of a user-supplied elevation tolerance to adjust the strength of drainage enforcement in relation to both the accuracy and density of the input elevation data.

Drainage enforcement is also obtained by incorporating stream line data. This is useful when more accurate placement of streams is required than can be calculated automatically by the program. The program checks for closed loops in data stream lines and appropriate diagnostics are produced. Side conditions are also set for each stream line. These ensure that the stream line acts as a breakline for the interpolation conditions so that each stream line lies at the bottom of its associated valley. The program has also been extended to accept stream distributaries. These abound in many low relief areas of the continent and were incorporated in Version 3 of the national 9 second DEM. Elevations on streamlines have been initialised for each successive grid resolution by linearly interpolating along streamlines between descending elevation points located on the streamlines. Data points on streamlines that are higher than upstream data points are normally removed. This process becomes problematic when the elevation data have significant error, giving rise to occasional upstream data points that are too low. These can remove accurate downstream elevation points and give rise to streamlines that are too low with respect to the neighbouring landscape.

3. APPLICATION TO CLEANED SRTM DATA

One second SRTM data with vegetation and other artefacts removed (Gallant and Dowling, *private communication*) were obtained for a 1 degree square in the Gwydir Valley. This valley has extensive areas of low relief and a large number of stream distributaries. The streamlines generally flow from the east to the west. The data were processed in western and eastern halves to examine the stability of the process, particularly with respect to the derivation of elevations on streamlines at the edges defined by the central north-south boundary line. This is critical in building a DEM with continent-wide consistent drainage structure from a series of overlapping map tiles. A portion of the cleaned SRTM data is shown in shaded relief in Figure 1. The noise in the data in this low relief region is sufficient to mask the underlying shape and drainage structure.

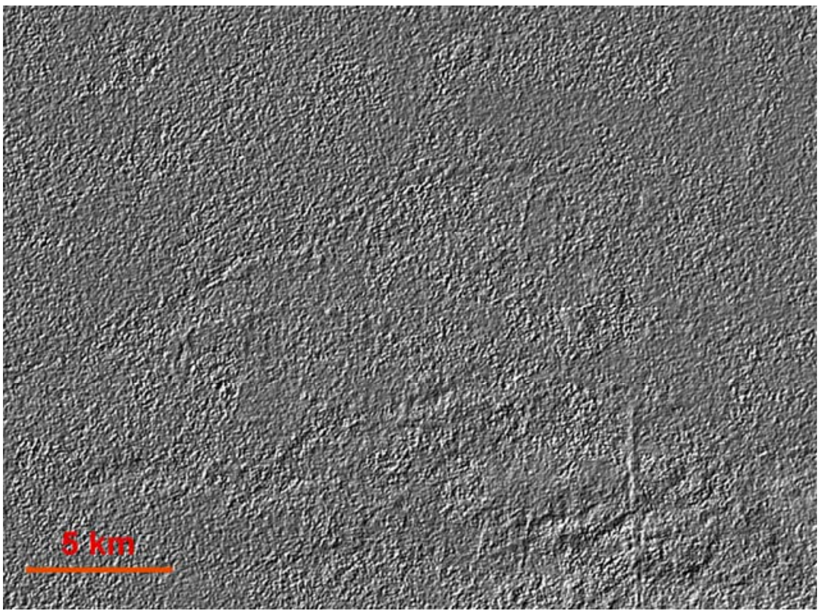

Figure 1 Shaded relief view of cleaned SRTM data.

3.1. Data smoothing with drainage enforcement

A first estimate of the standard elevation error of the cleaned SRTM data was 2m, although there were points with sporadic errors exceeding 10m. ANUDEM was applied to these data with drainage enforcement using corrected GEODATA 250K streamline data and applying data smoothing consistent with a vertical standard error of 2m. The result is shown in shaded relief in Figure 2 for the same portion shown in Figure 1. The program has been clearly effective in constructing a coherent valley structure in this portion of the 1 degree square. Further testing is required to assess whether this level of data smoothing is appropriate. A minor north-south artefact in the data remains in the south-east of both Figure 1 and Figure 2.

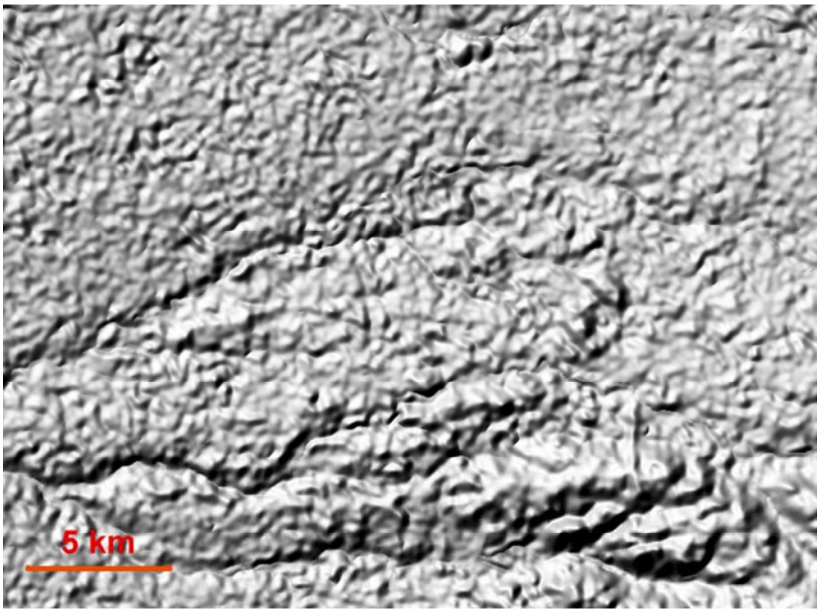

Figure 2 SRTM filtered by ANUDEM with drainage enforcement and 2m data smoothing.

3.2. Height initialisation of streams

The existing streamline initialisation procedure to both halves of the 1 degree grid square and a detail is shown in Figure 3, where each half was gridded as a separate tile. The streamline on the right hand side has an elevation at the central line that is 9m below the elevation of the streamline beginning on the left hand side. The different colours indicate elevation contours at 1m intervals, confirming the overall low relief of this landscape.

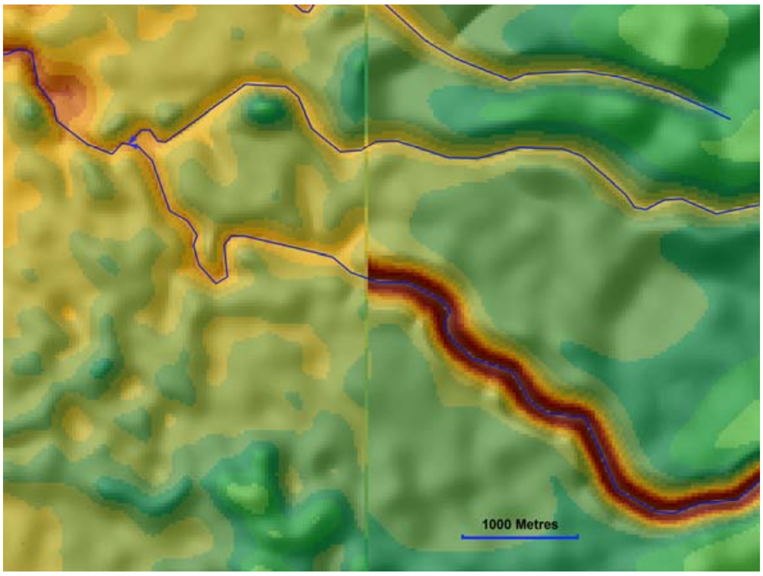

Figure 3 Adjacent map tiles with mismatching stream heights.

The streamline initialisation procedure in ANUDEM was revised to take initial streamline heights from the preceding coarser grid in the multi-grid interpolation process, rather than the error prone actual data point heights that lie on the stream lines. The program then ensures that all initial heights on the streamlines linearly descend down each stream segment between successive stream junctions and/or disjunctions. This takes account of the many distributaries (over 300) in this degree square. The revised initialisation appears to give rise to stable elevations on streamlines, as indicated in Figure 4.

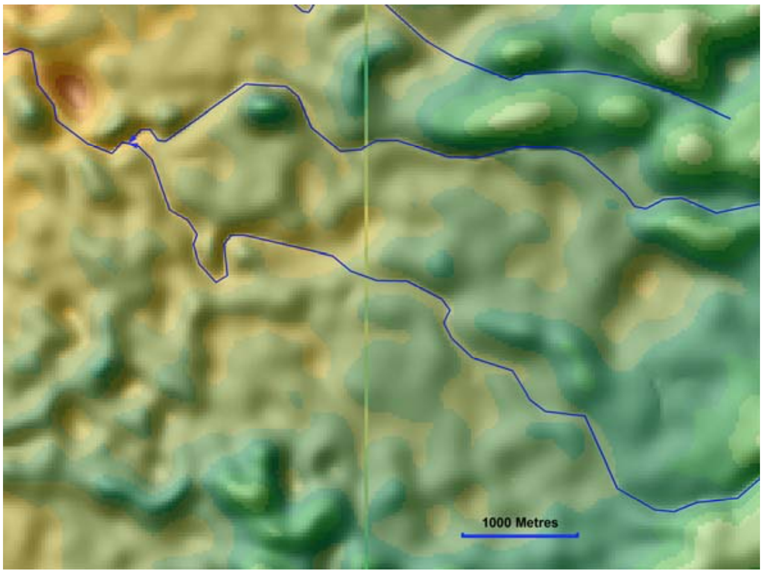

Figure 4 Adjacent map tiles with matching stream heights.

The remaining mismatch in elevations of the streamlines at the central boundary line is less than 20cm and is visually undetectable in this figure. In a larger exercise the remaining minor mismatches could be readily addressed by gridding each tile with small overlaps and smoothly blending the edges. Again, further testing and tuning is required to finalise this procedure. It appears that the streamlines should lie slightly lower than that depicted in Figure 4. This can be readily implemented within the multi-grid interpolation scheme.

4. DISCUSSION AND CONCLUSION

Revisions to the initialisation of streamlines in the ANUDEM multi-grid elevation interpolation procedure appear to have addressed former problems in interpolating DEMs from noisy SRTM data in low relief regions with extensive distributary streamline networks. Further work is required to fine tune the amount of data smoothing applied and to further refine the initialisation of elevations on streamline networks.

An additional intention is to strengthen the automated drainage enforcement algorithm in the presence of spatially dense noisy data, such as SRTM data and other remotely sensed elevation data, to remove the large numbers of spurious depressions that tend to remain in low relief landscapes, such as the one shown here.

Close examination of the corrected GEODATA 250K streamline data has revealed remaining direction errors that have only come to light when applied at 1 second (about 30m) resolution. These errors had little impact at the 9 second (about 250m) resolution. This has prompted a major revision of the ANUDEM program to automatically detect, and reasonably correct, all direction errors in input vector streamline networks before the data are used to enforce drainage in a fitted DEM. This has involved developing search procedures to efficiently locate all connections in the input vector network. This is leading to a final comprehensive revision of the GEODATA 250K streamline network that will help it to remain a valuable natural data resource.

However, despite the quality of the GEODATA 250K streamline network at 9 second scale, its use at the finer 1 second scale remains problematic. It is the only continent-wide, topologically consistent, vector streamline network available for Australia, but the spatial location of the streamlines are consistent with their source scale of 1:250K. Thus their location errors are around 100 to 200 m, and on occasion up to 500 m.

This can lead to serious conflicts with the real terrain structure at 1 second (30m) resolution. These streamline data can be used as an interim supplement to current fine scale elevation gridding efforts, but the case for a new high resolution, topologically consistent, national streamline network is overwhelming in the light of new high resolution noisy elevation data sets that require appropriately coordinated data smoothing and drainage enforcement to realise their full potential.

ACKNOWLEDGMENTS

The 1 second resolution SRTM data for Australia was provided by Defence Imagery and Geospatial Organisation (DIGO) and CSIRO.

REFERENCES

ANU Fenner School of Environment and Society and Geoscience Australia (2008), *GEODATA 9 Second DEM and D8 Digital Elevation Model and Flow Direction Grid, User Guide*. Geoscience Australia, pp 1-43. http://www.ga.gov.au/image_cache/GA11644.pdf

Gallant, J.C. (2001), Topographic scaling for the NLWRA sediment project. *CSIRO Land and Water Technical Report* 27/01. http://www.clw.csiro.au/publications/technical2001/tr27-01.pdf

Gallant, J.C. and Dowling, T.I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, 39(12): 1347.

Geoscience Australia (2006), *GEODATA TOPO–250K Data User Guide, Version 3*. Canberra, Australia.

Harding, D.J., Bufton, J.L. and Frawley, J. (1994). Satellite laser altimetry of terrestrial topography: Vertical accuracy as a function of surface slope, roughness and cloud cover. *IEEE Transactions on Geoscience and Remote Sensing* 32: 329-339.

Hutchinson, M.F. (1989), A new method for gridding elevation and stream line data with automatic removal of pits. *Journal of Hydrology* 106: 211-232.

Hutchinson, M.F. (1996), A locally adaptive approach to the interpolation of digital elevation models. *Proceedings of the Third International Conference/Workshop on Integrating GIS and Environmental Modeling*. National Center for Geographic Information and Analysis, Santa Barbara, CA, CD.

Hutchinson, M.F. (2000), Optimising the degree of data smoothing for locally adaptive finite element bivariate smoothing splines. *ANZIAM Journal* 42(E): C774-C796.

Hutchinson, M.F. (2007), *ANUDEM Version 5.2.2*. Fenner School of Environment and Society, Australian National University, Canberra. http://fennerschool.anu.edu.au/publications/software/anudem.php

Hutchinson, M.F. (2008), Adding the Z-dimension. In: J.P. Wilson and A.S. Fotheringham (eds), *Handbook of Geographic Information Science*, Blackwell, pp 144-168.

Hutchinson, M.F. and Dowling, T.I. (1991), A continental hydrological assessment of a new grid-based digital elevation model of Australia. *Hydrological Processes* 5: 45-58.

Hutchinson, M.F. and Gallant, J.C. (2000), Digital elevation models and representation of terrain shape. In: J.P. Wilson and J.C. Gallant (eds), *Terrain Analysis*. John Wiley & Sons, New York, 29-50.

Hutchinson, M.F., Stein, J.L. and Stein, J.A. (2000), *Derivation of nested catchments and sub-catchments for the Australian continent.* Centre for Resource and Environmental Studies, Australian National University, Canberra. http://cres.anu.edu.au/outputs/programs.html

Stein, J.L., Stein, J.A. and Nix, H.A. (2002), Spatial analysis of anthropogenic river disturbance at regional and continental scales: identifying the wild rivers of Australia. *Journal of Landscape and Urban Planning* 60:1-25.

USGS (2005), *Shuttle Radar Topography Mission – Mapping the World in Three Dimensions*. http://srtm.usgs.gov/