SLOPE, REFLECTANCE, AND VIEWSHEDS ALGORITHMS FOR ARC-SECOND DIGITAL ELEVATION MODELS

Peter L. Guth, Professor
Department of Oceanography
United States Naval Academy
572C Holloway Road
Annapolis, Maryland 21402
pguth@usna.edu

ABSTRACT

Many common digital elevation models (DEMs) use geographic coordinates, including NED, DTED, SRTM, GDEM, and NEXTMAP, with horizontal spacings in the range of 1/9" to 3". Many GIS operations, such as slope/aspect, reflectance or hillshade mapping, and viewsheds, require geometric knowledge of the relationship between horizontal and vertical spacing. Published discussions implicitly assume a rectangular UTM or UTM-like grid because they refer to a single value for data spacing, and some commercial software either requires re-interpolation to UTM before algorithms will work or allows only a single factor to relate degrees to meters, ignoring the significant changes that occur with latitude. No matter how it is done, reinterpolation changes elevations, and cannot produce a better elevation surface. Geographic grids are not square and at best quasi-rectangular. The mathematics of using geographic coordinates should not deter software from doing computations in geographic space. Programs can compute a single y spacing, and for small areas a single x spacing. For larger areas, or at higher latitudes where the x spacing changes rapidly, the program can compute a different x spacing for each row in the data set which neither consumes significant storage nor increases processing time. Results with these formulas and geographic coordinates produce very similar results to those using reinterpolated UTM coordinates, but cannot be compared directly because reinterpolation produces a DEM with a different grid size, and nodes at different locations. Because date producers have embraced the benefits of seamless data in geographic coordinates, GIS software vendors should adapt algorithms to better manipulate this data.

INTRODUCTION

Slope and aspect computations provide one of the most useful derivative products from digital elevation models (DEMs). Slope and aspect affect mobility and insolation, so users in fields as diverse as the military and ecology care about the results. At least a half dozen distinctly different slope algorithms exist, and a host of studies (e.g. Carter, 1992; Florinsky, 1998; Guth, 1995; Hodgson, 1998; Jones, 1998; Zhou and Liu, 2004a, 2004b) have looked at slope algorithms and the effects of data resolution and precision on the computed results. However, except for a brief mention in Guth (1995, p.32), the published discussions implicitly assume a rectangular UTM or UTM-like grid for the DEM because they refer to a single value for data spacing. This paper will explore the implications of that assumption, and its importance when the best medium resolution (10-100 m post spacing), and some high resolution (3 m) DEMs all use arc-second spacing.

Many common DEMs use geographic coordinates:

- U.S Geological Survey's National Elevation Dataset (USGS NED, Gesch and others, 2002).
- National Geospatial-Intelligence Agency's Digital Terrain Elevation Data (NGA DTED, National Imagery and Mapping Agency, 2000).
- Shuttle Radar Topography Mission (SRTM, Slater and others, 2006).
- ASTER "30 m" GDEM (ASTER GDEM Validation Team, 2009).
- INTERMAP's NextMap IFSAR (INTERMAP, 2010).

These data sets have horizontal spacing from 30" (DTED-0 and SRTM-30) to 1/9" (limited quantities of NED). Much of this data is freely available on the WWW: 30" and 3" (SRTM) for the entire world, and 1" data for the world (GDEM) and United States (both SRTM and NED). Free means both easy to obtain and without cost.

Because of its ready availability, these DEMs have seen widespread usage world-wide.

Computations involving the geographic DEM, which relate distances in the horizontal and vertical directions, must deal with horizontal spacing in degrees, and vertical spacing in meters. This could be done in several ways:

1. Reprojecting the DEM to a projected cartesian system, such as UTM.
2. Using a single conversion factor to relate degrees to meters, and ignoring the differences between the x and y spacings.
3. Assuming the DEM has a rectangular framework, with constant spacing throughout the DEM but a different x and y spacing.
4. Assuming the DEM has a trapezoidal framework, with constant y spacing throughout the DEM but an x spacing that varies with latitude.
5. Assuming the DEM has a trapezoidal framework, with variable x and y spacing that varies with latitude.

Options 1 and 2 have been assumed, at least implicitly, by most GIS programs. This paper will explore the improvements possible with options 3 and 4, and argue that option 5 would not provide any additional improvement.

Geodetic formulas (Vincenty, 1975) can be used to compute spacing of DEMs with coordinates in latitude and longitude, and with current processor speeds, this does not create a significant time penalty. Appendix 1 shows adjustments to common slope algorithms to use variable x and y spacing, and the same adjustments can easily be made to reflectance or hill-shading algorithms. Guth (2004) described how to perform intervisibility computations using geographic DEMs.

CHARACTERISTICS OF ARC-SECOND DEMS

Figure 1 shows the x and y spacing for a 1" geographic coordinates DEM as a function of latitude. Other DEM spacings would be simple multiples of these values, and data sets like DTED which varied the x spacing with latitude would have a sawtooth pattern, but both would share the fundamental characteristics. Poleward of about 10˚ latitude, where the x and y spacings differ by less than 1%, the x spacing will be less than the y spacing, and the difference increases dramatically with latitude. While the y spacing does change slightly with latitude, the change is less than 1% from the equator to 75˚ latitude. Figure 2 shows the percentage difference in the x spacing at the northern and southern edges of a 1˚ DEM, such as a single cell of DTED, SRTM, or GDEM.

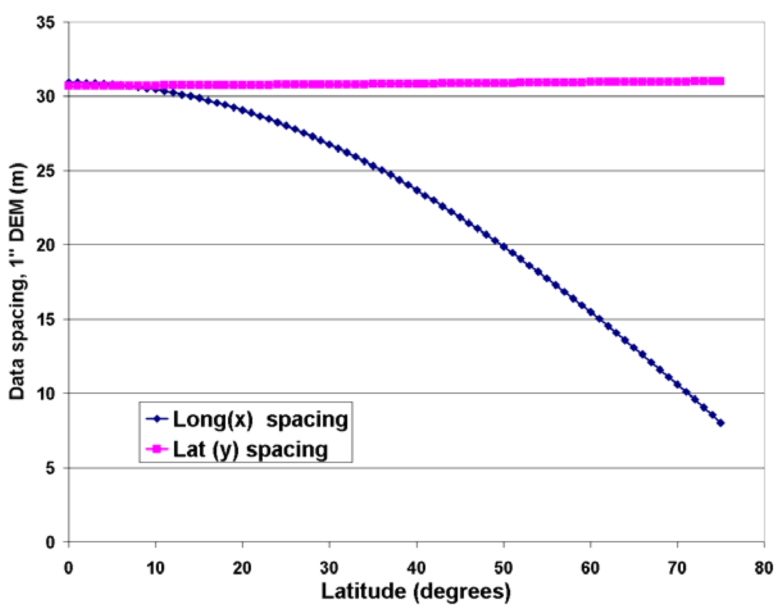

Figure 1. Spacing in meters for a 1” DEM as a function of latitude.

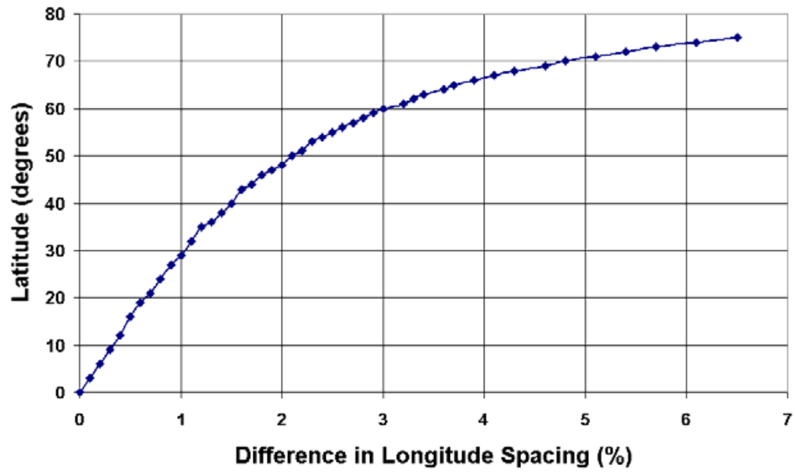

Figure 2. Percent difference in E-W data spacing from the top to the bottom of a 1º cell of a lat/long DEM.

Figure 3 shows the maximum difference in computed slopes, using a single average x spacing for the entire DEM versus computing the spacing at every row in the DEM, which we define as the true value. This used a merge of SRTM data, and computed the slope at every point in the DEM using the two algorithms. The error percentage is defined as the difference between the two computed slopes divided by the true value. The magnitude of the error depends on the orientation of the slope vector (the aspect), but this data set had enough points to produce a linear curve to show the maximum error in the most unfavorable direction. Slope errors approach 14% over this area. This merge covers a large area, but users are currently working with 3" DEMs over areas this large.

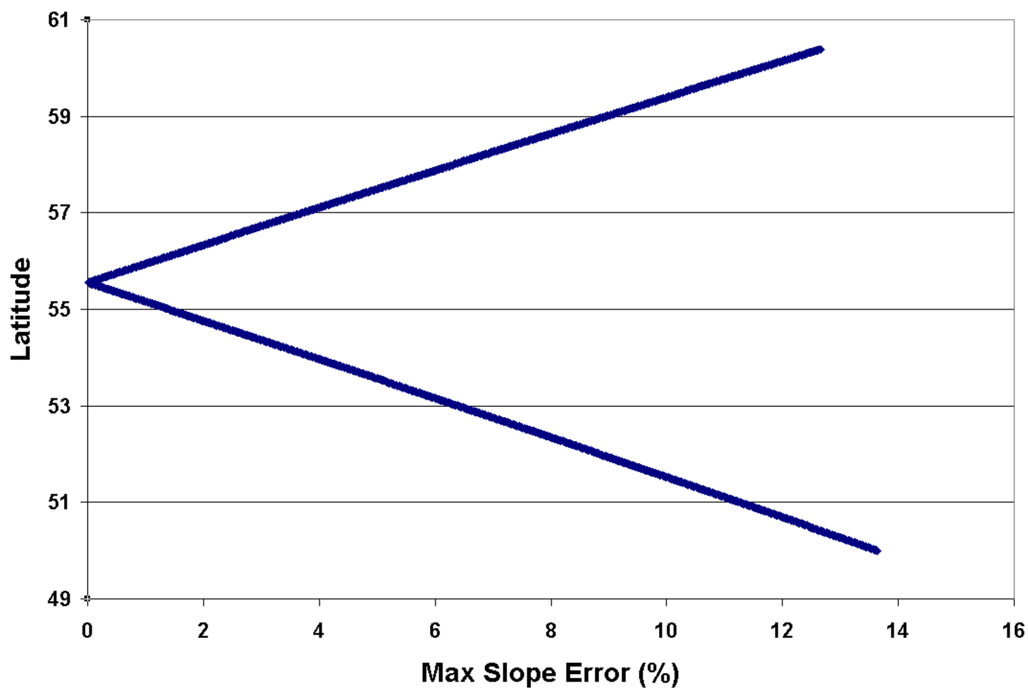

Figure 3. Maximum error in the value of the slope computed using a constant E-W data spacing equal to the value in the center of a DEM, for a DEM centered near 55º N, compared to using the actual spacing at each row of the DEM.

EFFECTS OF REINTERPOLATION ON SLOPE COMPUTATIONS

Reinterpolating a geographic DEM to UTM produces new elevations at points different from those in the original DEM. There will be a different number of points in the reinterpolated DEM, and it will not be possible to directly compare points in the original and reinterpolated DEMs. Comparisons will have to either use the nearest point in the other grid, which could be displaced by half the data spacing, or interpolate a value. Reinterpolation cannot create a better surface (unless the goal is to filter, for example to remove noise); at best reinterpolation can result in negligible differences, and at worst it will introduces errors.

Figure 4 shows the average slope for 6 regions in the western United States, computed for NED DEMs from the USGS Seamless Server. Each DEM was reinterpolated to the UTM resolution commonly given for the arc second spacing, using a bilinear interpolation in MICRODEM (Guth, 2008). Table 1 shows the results for the DEM from Washington state, with the maximum slope in the DEM and the standard deviation of the slope distribution included as well as the mean slope. All data sets show that as the spacing becomes smaller, the slopes become larger, and the UTM reinterpolations are slightly less steep than the original arc second DEMs.

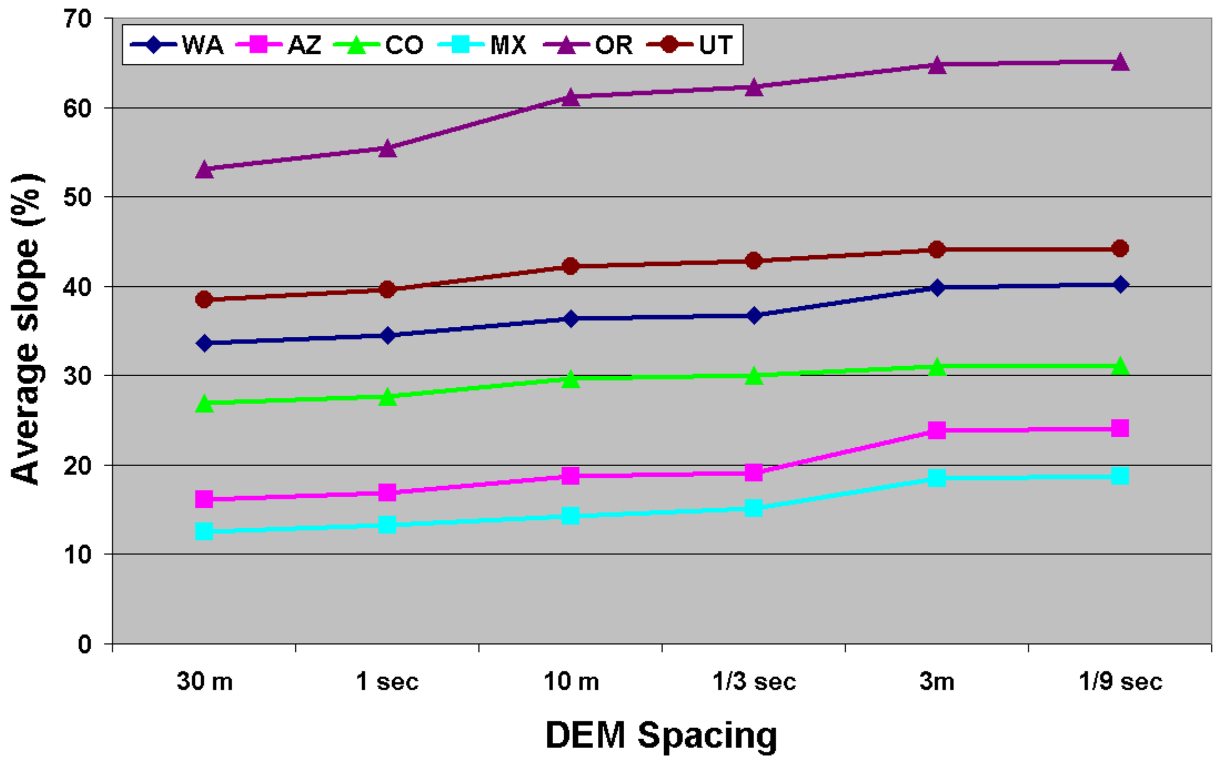

Figure 4. Average slope for 6 regions, from NED (1", 1/3", and 1/9") and reinterpolated UTM DEMs

Table 1. Elevation statistics for 6 DEMs of Mt St Helens (NED and UTM reinterpolations)

	30 m	1 sec	10 m	1/3 sec	3m	1/9 sec
Max slope(%)	213.99	240.06	331.91	377.61	915.01	985.65
Mean slope (%)	33.65	34.48	36.35	36.70	39.87	40.18
Slope std dev	25.06	25.77	27.54	27.93	31.84	27.93

Table 2 shows the results of a series of reinterpolations on the steepest DEM shown in Figure 2. For this area in Oregon, 1/9" corresponds with an average spacing about 2.9 m (2.41x3.43 m in x and y respectively). Larger UTM spacings (3-5 m) have smaller maximum and mean slopes, and a smaller standard deviation, while smaller spacings of 1-2 m have large values.

Table 3 compares the results from MICRODEM and ArcGIS for the Oregon 1/9" NED. Using the DEM without reprojection, ArcMap produces generally similar values for the maximum and average slope, but a much smaller standard deviation. The three reprojected DEMs produce average slopes and slope standard deviation very similar to the geographic results from MICRODEM, with a much smaller maximum slope. The results from the geographic DEM in ArcMap probably reflect the fact that ArcGIS uses a single scaling factor to adjust geographic coordinates to meters, which cannot account for the differences in the x and y spacings.

Table 2. Slope comparisons for 1/9" NED DEM in Oregon, compared to UTM reinterpolations

DEM Spacing	Max slope (%)	Average slope (%)	Slope Std Dev
1 m, MICRODEM bilinear	1115.93	65.65	30.29
2 m, MICRODEM bilinear	888.87	65.26	29.50
1/9" (~2.41x3.43 m)	854.29	65.17	29.26
3 m, MICRODEM bilinear	707.35	64.83	28.72
4 m, MICRODEM bilinear	580.38	64.39	28.02
5 m, MICRODEM bilinear	516.20	63.94	27.40

Table 3. Slope comparisons for 1/9" NED DEM in Oregon, with MICRODEM and ArcGIS

DEM Spacing	Max slope (%)	Average slope (%)	Slope Std Dev
1/9", MICRODEM	854.29	65.17	29.26
1/9", ArcMap	794.16	65.64	25.72
3 m, MICRODEM bilinear	707.35	64.83	28.72
3 m, ArcMap bilinear	713.69	64.87	28.86
3 m, ArcMap cubic	733.57	65.11	29.25

Figure 5 shows the 1" NED DEM and the reinterpolated 30 m UTM DEM. The two slope maps look very similar (5A and 5C), but the difference map (5B) shows many points with slopes ±20% different in the two DEMs. Figure 6 shows a profile through this DEM, with the topographic profiles, the slopes, and the slope differences. In the topographic profile, the reinterpolated 30 m DEM misses a number of the peaks and valleys. Consequently the slopes at these points differ.

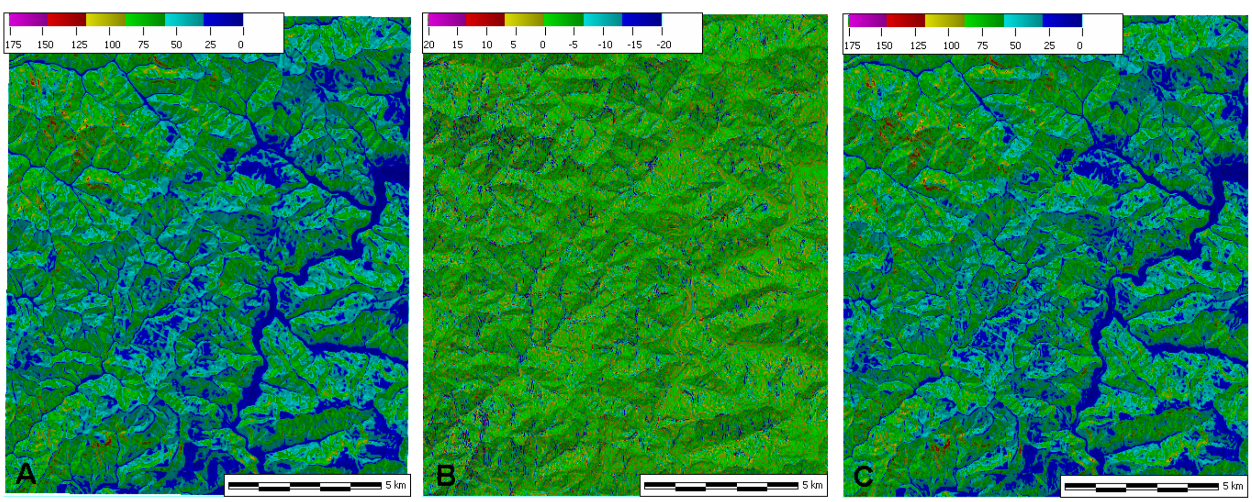

Figure 5. Oregon DEM. A. Slope map from 30 m re-interpolated DEM. B. Difference between slopes from the geographic and UTM DEMs. C. Slope map from original 1" NED DEM.

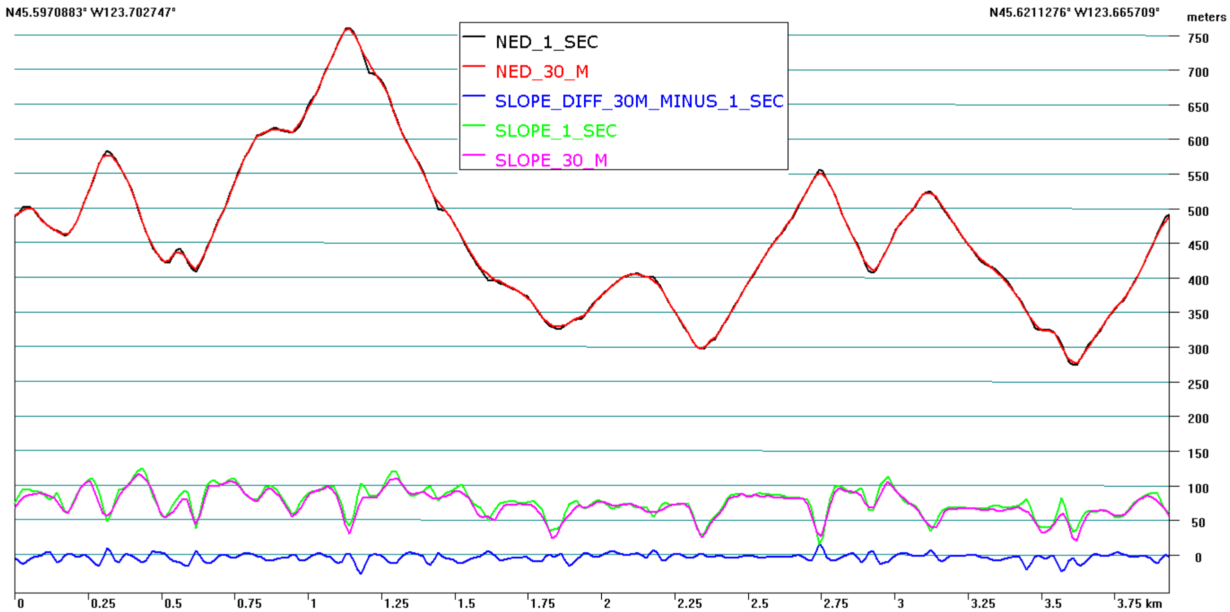

Figure 6. Profile through Oregon 1" and 30 m DEMs, with elevations, slopes, and difference between slopes.

Figure 7 shows the difference between the slopes from the 1/9" NED and the reinterpolated 3 m UTM DEM. Lines of large slope differences, both positive and negative, occur when the slope changes along ridge crests and in the valley bottoms. Figure 8 shows a profile through this DEM, with the topographic profiles, the slopes, and the slope differences. In the topographic profile, the reinterpolated 3 m DEM appears to closely follow the 1/9" NED, but the slopes show that the reinterpolated DEM smoothes out a number of the changes in slope, and underestimates those slopes.

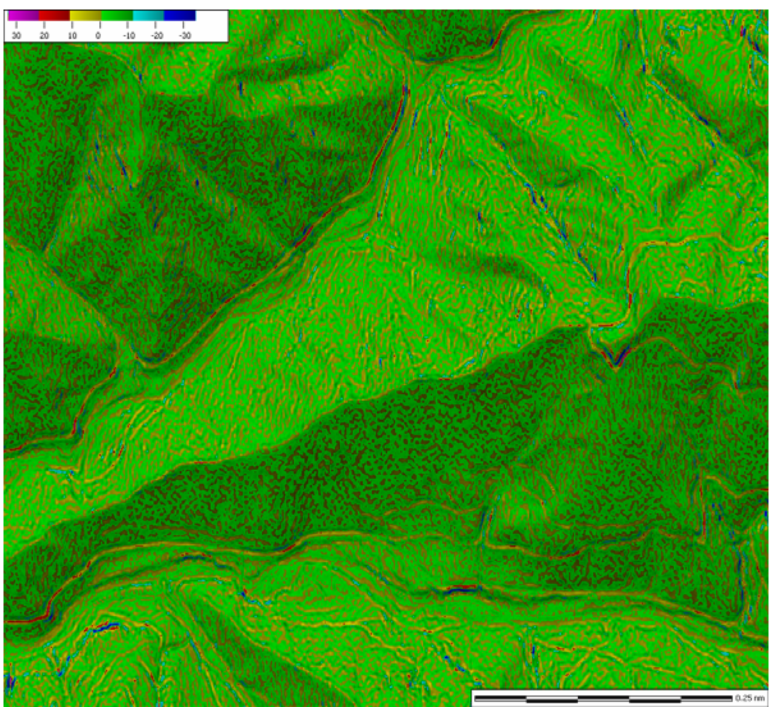

Figure 7. Slope difference (Slope from 3 m DEM - Slope from 1/9" DEM), overlaid on reflectance map from the Oregon DEM.

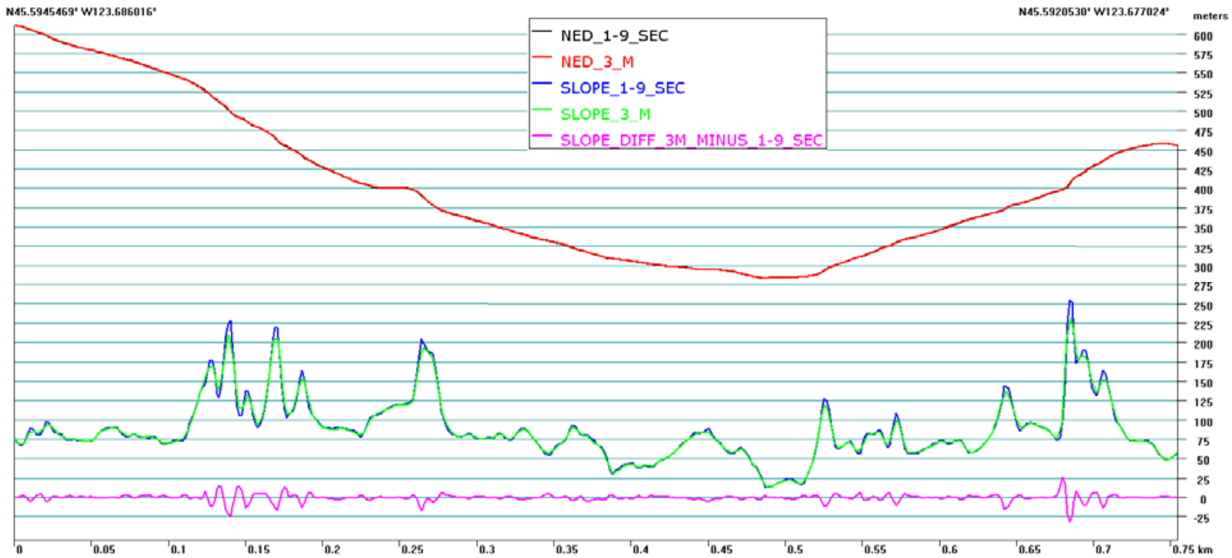

Figure 8. Profile through Oregon 1/9" and 3 m DEMs, with elevations, slopes, and difference between slopes.

Figure 9 shows the aspect distribution computed with the Oregon 1/9" and 3 m reprojected UTM DEMs in ArcGIS and MICRODEM. ArcGIS clearly does not account for different x and y spacings in the geographic DEM, and computes aspects incorrectly. Table 3 showed that the slope distribution from ArcMap for the geographic DEM differed substantially from that computed by MICRODEM or by either program for the UTM reinterpolations, suggesting that MICRODEM probably produces more reliable results for an arc second DEM. The UTM reprojection from ArcGIS agrees with both computations from MICRODEM.

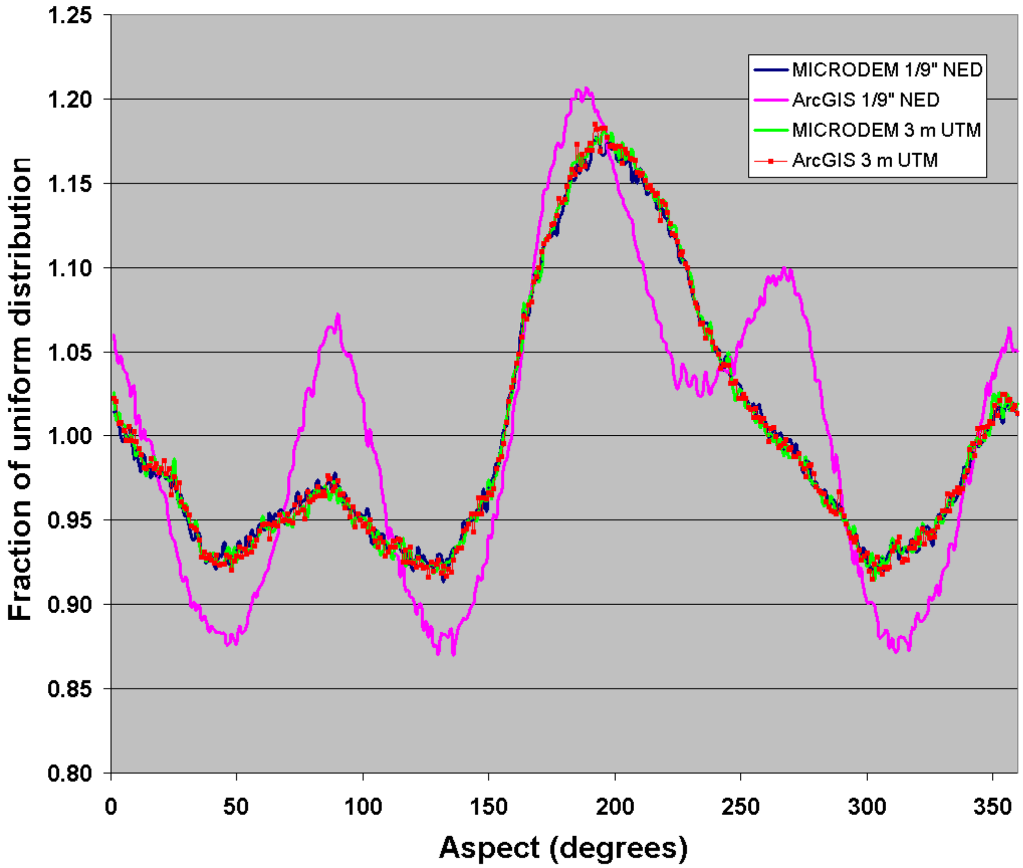

Figure 9. Aspect distribution of the Oregon 1/9" and 3 m UTM DEMs computed with MICRODEM and ArcGIS.

DISCUSSION

Comparing the results of the geographic algorithm with a UTM algorithm is not simple, for several reasons:

1. Reinterpolation produces grid nodes at different locations, so the two DEMs compute slope and other parameters at different locations.
2. Reinterpolation produces a grid with a different number of grid nodes. Some places the grid nodes in the original and reinterpolated DEMs will be close together, and elsewhere they will be separated by half the grid spacing interval.
3. Most reinterpolation algorithms probably smooth the original DEM, so the overall statistics will be different.

Despite these challenges, it is clear that the computations done in geographic coordinates can closely approximate those done with reinterpolations of the DEM. Since there is no theoretical justification to reinterpolate, the differences introduced by reinterpolation are at best insignificant and at worst a degradation of results. Since the computations can be done with geographic coordinates, GIS software should adapt their algorithms to use the geographic coordinate DEMs now being produced.

The algorithms most in need of adjustment are slope/aspect and intervisibility/viewshed, because these produce numeric results which will affect decision making. Reflectance/hillshade results produce graphical results and a visual depiction of the terrain, and users will probably not see much difference with simplifications in computing the different horizontal spacing in the x and y directions. However, the required modifications to the algorithms are so minor, and have to be done for the slope and aspect computations, that they should be done for this code as well.

CONCLUSION

GIS software does not have to reinterpolate geographic DEMs for use with slope/aspect, hillshades, and viewsheds. Reinterpolation can introduce unwanted artifacts in the DEM, takes unnecessary processing time, and requires unnecessary storage space. The required modifications to slope, aspect, hillshade, and viewshed algorithms require little effort, and have negligible impact on computation times. GIS software should compute accurate results for geographic DEMs without reinterpolation. If the mapping agencies producing DEMs think geographic coordinates provide the best solution for DEMs, GIS software should follow suit.

ACKNOWLEDGMENTS

This study used the freeware GIS MICRODEM (Guth, 2008).

REFERENCES

ASTER GDEM Validation Team, 2009, ASTER global DEM validation summary report: 28 p. Available online at http://www.ersdac.or.jp/GDEM/E/image/ASTERGDEM_ValidationSummaryReport_Ver1.pdf

Carter, J.R., 1992, The effect of data precision on the calculation of slope and aspect using gridded DEMs: *Cartographica*, 29(1:22-34.

Florinsky, I.V., 1998, Accuracy of local topographic variables derived from digital elevation models: *International Journal of Geographical Information Science,* 12(1): p.47-61.

Gesch, D.B., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., and Tyler, D., 2002. The national elevation dataset: *Photogrammetric Engineering & Remote Sensing*, 68(1):5-11.

Guth, P.L., 1995, Slope and aspect calculations on gridded digital elevation models: Examples from a geomorphometric toolbox for personal computers: *Zeitschrift fur Geomorphologie N.F. Supplementband* 101:31-52.

Guth, P.L, 2004, The Geometry of Line-of-Sight and Weapons Fan Algorithms: in Caldwell, D.R., Ehlen, J., and Harmon, R.S., eds., *Studies in Military Geography and Geology*, Dordrecht, The Netherlands, Kluwer Academic Publishers, chapter 21, p.271-285.

Guth, P.L., 2008, Geomorphometry in MICRODEM, In Hengl, T., Reuter, H.I. (eds), *Geomorphometry: concepts, software, applications*. Developments in Soil Science Series, Elsevier, ISBN-13: 978-0-12-374345-9, p.351-366.

Heerdegan, R.G., and Beran, M.A., 1982, Quantifying source areas through land surface curvature and shape: *Journal of Hydrology*, 57(3-4): 359-373.

Hodgson, M.E., 1995, What cell size does the computed slope/aspect angle represent?: *Photogrammetic Engineering & Remote Sensing*, 61(5):513-517.

Hodgson, M.E., 1998, Comparison of angles from surface slope/aspect algorithms: *Cartography and Geographic Information Systems*, 25(3):173-185

Horn, B.K.P., 1981, Hill shading and the reflectance map: *Proceedings of the IEEE*, 69(1):14-47.

INTERMAP, 2010, NEXTMap countrywide mapping programs: http://www.intermap.com/nextmap-3d-mapping-program

Jones, K.H., 1998, A comparison of algorithms used to compute hill slope as a property of the DEM: *Computers & Geosciences*, 24(4):315-324.

National Imagery and Mapping Agency, 2000, Performance specification digital terrain elevation data (DTED): MIL-PRF-89020B, 45 p. Available online at http://www.nga.mil/ast/fm/acq/89020B.pdf.

Slater, J.A., Garvey, G., Johnston, C., Haase, J., Heady, B., Kroenung, G., and Little, J., 2006, The SRTM data "finishing" process and products: *Photogrammetric Engineering & Remote Sensing*, 72:237-247.

Verdin, K.L., Godt, J.W., Funk, C., Pedreros, D., Worstell, B., Verdin, J., 2007, Development of a global slope dataset for estimation of landslide occurrence resulting from earthquakes: Colorado: U.S. Geological Survey, Open-File Report 2007-1188, 25 p.

Vincenty, T. 1975. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations.

Survey Review. 23(176):88-93.
Zevenbegen, L.W., and Thorne, C.R., 1987, Quantitative analysis of land surface topography: *Earth Surface Processed and Landforms*, 12(1):47-56.
Zhou, Q., and Liu, X., 2004, Analysis of errors of derived slope and aspect related to DEM data properties: *Computers & Geosciences*, 30(4):369-378.
Zhou, Q., and Liu, X., 2004, Error analysis on grid-based slope and aspect algorithms: *Photogrammetic Engineering & Remote Sensing*, 70(8):957-962.

Appendix 1. Slope algorithm for DEMs in geographic coordinates.

Define data spacings in meters for the DEM, once when opening the DEM:

- AverageYSpace: differences with latitude can be ignored.
- AverageXSpace: for a small area, differences can be ignored. The larger the area, and the higher the latitude, the greater the errors in using this.
- UseXSpace: array with spacing for each row of the DEM, and all computations use the value for the row where the computation is done. There is no significant performance penalty.

The point and its neighbors are defined by:

Znw	**Zn**	**Zne**
Zw	**Z**	**Ze**
Zsw	**Zs**	**Zse**

Then compute:

Slope = sqrt(sqr(dzdx) + sqr(dzdy));
Asepect = GetAspect(dzdx,dzdy) (arctan function, with corrections to get compass azimuth);

Partial derivatives estimated for each algorithm by:

Eight Neighbors Unweighted (3FD, Heerdegen and Beran, 1982)
dzdx = (zne + ze + zse - zsw - zw - znw) / 6 / UseXSpace;
dzdy = (znw + zn + zne - zsw - zs - zse) / 6 / AverageYSpace;

Four Neighbors (2FD, Zevenbergen and Thorne, 1987)
dzdy = (zn - zs) * 0.5 / AverageYSpace;
dzdx = (ze - zw) * 0.5 / UseXSpace;

Eight Neighbors Weighted (3FDWRSD, Horn 1981 method)
dzdy = 0.125 * ((znw + 2 * zn + zne) - (zsw + 2 * zs + zse)) / AverageYSpace;
dzdx = 0.125 * ((zne + 2 * ze + zse) - (znw + 2 * zw + zsw)) / UseXSpace;

Eight Neighbors Weighted By Distance
dzdy = 1 / (4 + 2 * √2) * ((znw + √2*zn + zne) - (zsw + √2*zs + zse)) / AverageYSpace;
dzdx = 1 / (4 + 2 *√2) * ((zne + √2*ze + zse) - (znw + √2*zw + zsw)) / UseXSpace;

Frame Finite Difference (FFD, Chu and Tsai 1995 Conference Proceedings cited in Zhou and Liu, 2004)
dzdy = (znw - zsw + zne - zse) * 0.25 / AverageYSpace;
dzdx = (zse - zsw + zne - znw) * 0.25 / UseXSpace;

Simple Difference (SIMPLE-D, Jones 1998)
dzdy = (z - zs) * 0.5 / AverageYSpace;
dzdx = (z - zw) * 0.5 / UseXSpace;