# INCORPORATING VEGETATION IN VIEWSHED AND LINE-OF-SIGHT ALGORITHMS

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

## ABSTRACT

Computing intervisibility with line-of-sight profiles and viewsheds represents one of the important uses for digital topography, with applications for the military, communications engineers, landscape planners, and others. The computations must consider data resolution, quality, and format, and have generally used a single surface and neglected the effects of vegetation. Vegetation can potentially be included in a number of ways. (1) Separate grids from LIDAR or IFSAR contain the bare earth elevation (or bare earth and buildings) and the first return, and the difference between these provides the vegetation height. (2) Vector data sets contain vegetation information, which can be combined with the topography grids during analysis, probably most easily with a vector to grid conversion. (3) Gridded landcover data sets include vegetation categories, some of which include tree height data. Examples of these include NLCD, NBCD 2000, and LandFire. Most of these approaches will require probabilistic models with ranges of heights assigned to the vegetation categories based on analysis of local conditions, but with suitable development of algorithms will greatly improve on current intervisibility computations. (4) The increasing availability of LIDAR point clouds offers perhaps the greatest potential for improving intervisibility computations. Algorithms will have to deal with detecting buildings that will completely block line-of-sight, inferring a vegetation density from the sampled number of returns within the canopy, and developing probabilistic models based on the point cloud data. Regardless of the approach, viewshed results should be compared to recent high resolution imagery to verify that vegetation reported in digital data sets still represents reality. Commercial tree farms and orchards grow rapidly and can significantly impact intervisibility in many areas.

Key words: intervisibility, DEM, lidar, vegetation

## INTRODUCTION

Including vegetation in intervisibility computations requires suitable data and modifications to the algorithms to use the data. This paper describes an algorithm to use vegetation for line-of-sight and intervisibility, and evaluates various data sets currently available for use in this fashion. The algorithm has been implemented in MICRODEM (Guth, 2009), but could be adapted to any GIS software.

## ALGORITHM

The algorithm works with gridded digital elevation models (DEMs) which are assumed to represent a bare earth digital terrain model (DTM). The algorithm requires a second grid with the same spacing and extent as the DEM, and with the elevation of the vegetation at every point in the grid. If the user has both a DTM and a digital surface model (DSM), the vegetation grid will be the difference between the two grids. For speed in operation both grids must be loaded into memory. During intervisibility operations, if the user wants to include vegetation, the height from the vegetation grid will be added to the ground elevation from the DTM before performing intervisibility computations. Figures 1 and 2 show the results of the algorithm for a line-of-sight computation, with a visual depiction of the vegetation heights.

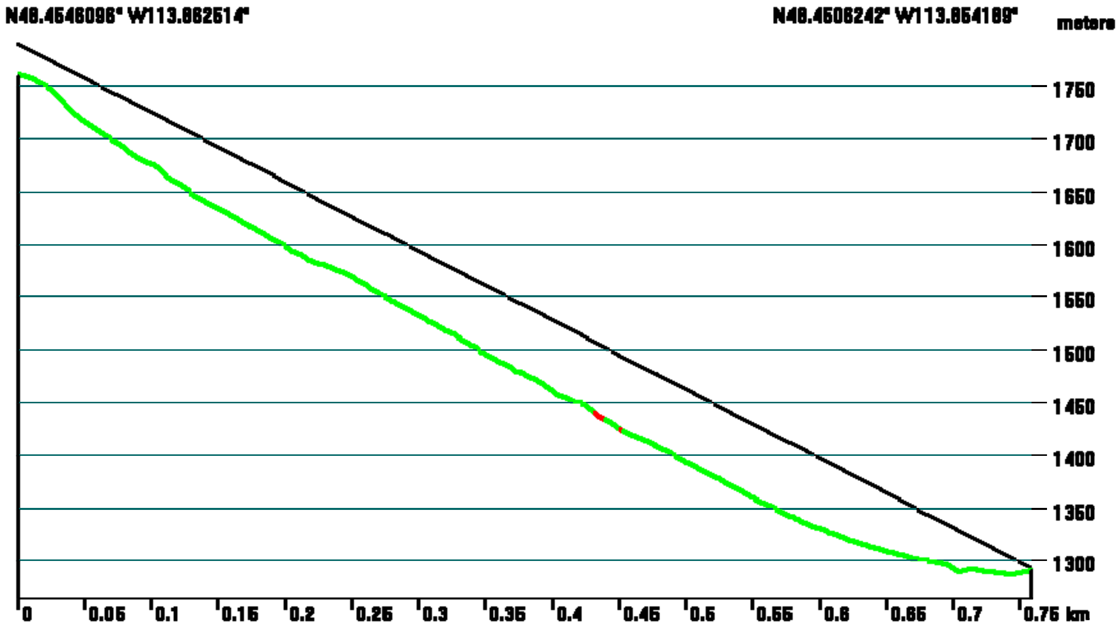

**Figure 1.** Line-of-sight without vegetation, showing almost the entire profile as visible (green) by the observer on the left.

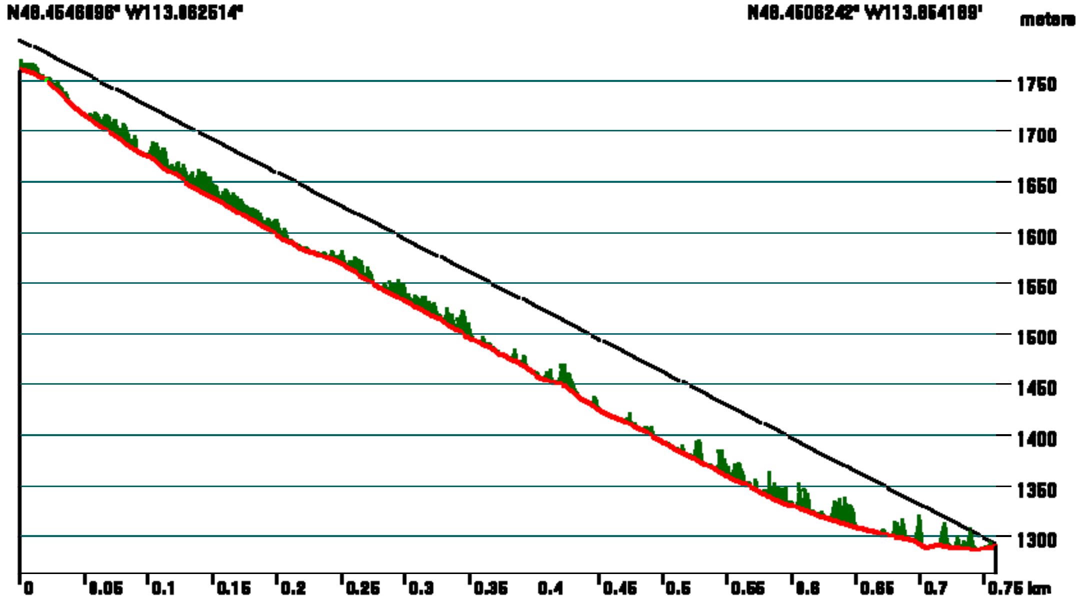

**Figure 2.** Line-of-sight with vegetation, showing almost all the profile as masked (red) for the observer on the left. This has the same geometry as Figure 1, and the user can rapidly toggle between the two views.

## DERIVING VEGETATION FROM ELEVATION DATA

Either vector or raster data could be used to prepare the vegetation grid. Vector polygon data must be converted to a grid with the same size and dimensions as the DTM, and the accuracy of the result will depend on the size of the polygons and how well they model the vegetation height. To my knowledge there are no suitable vector data sets with vegetation height readily available, so raster options might supply a better choice.

Multiple raster solutions are possible. The first approach would be to take the difference between two independent grids, one a DSM and the other a DTM. Within the United States, the National Elevation Dataset (NED) is supposed to be a DTM, and the Shuttle Radar Topography Mission (SRTM) is supposed to be a DSM. Thus a difference map should be a vegetation height map. Figure 3 shows an example profile in Maryland, and in the case it appears that SRTM is in fact generally higher than NED. However, there is an obvious major anomaly at about the 2.25 km point along the profile, and several other similar areas. Figure 4 shows a histogram of the differences in the region across which the profile was drawn, and while there are more positive values (with SRTM higher than NED), the overall distribution looks more like a random, normal distribution centered on zero, with a few very large negative anomalies. Figure 5 shows a similar histogram for southern Nevada, a region with much sparser vegetation, with the same results that the assumptions lead to large areas of negative vegetation height. Gallant and Read (2009) discussed the difficulty in trying to remove vegetation from SRTM data, and their experience confirms that the vegetation signature in SRTM data is not simple. Users wanting to use the SRTM - NED difference as a proxy for vegetation height should proceed with caution.

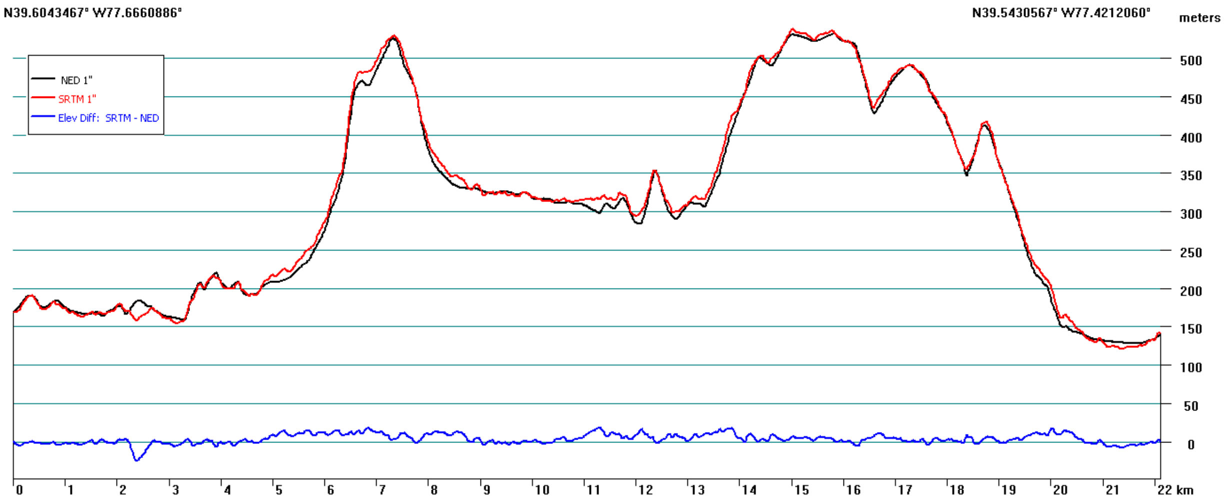

**Figure 3.** Profile in Maryland with NED, SRTM, and their difference which should be positive and related to vegetation height.

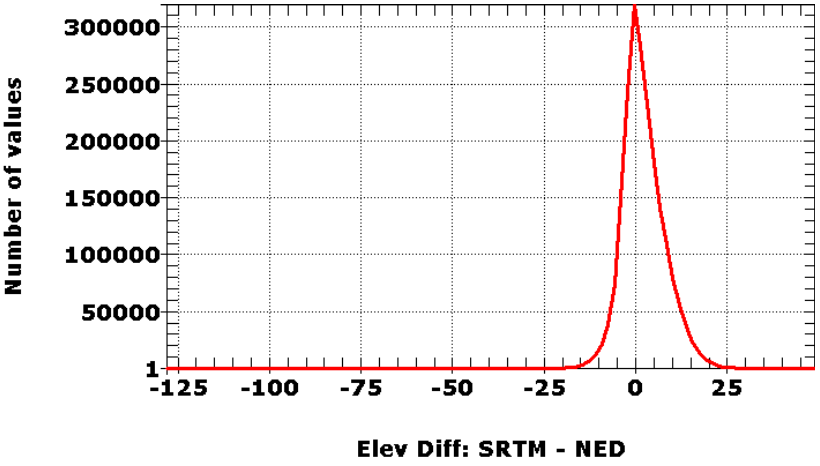

**Figure 4.** Histogram of elevation differences between SRTM and NED for a region in Maryland.

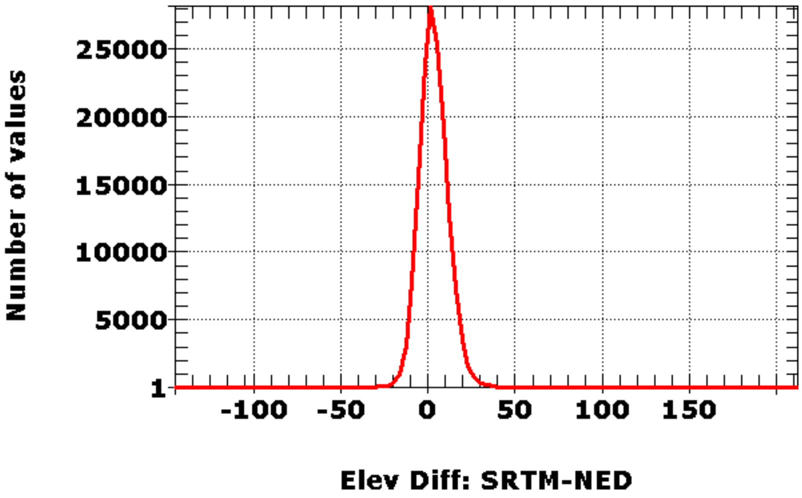

**Figure 5.** SRTM-NED difference for those areas classified as forests in NLCD-2001 in southern Nevada.

One problem with the difference between NED and SRTM is likely to be that they were collected independently, with different methodologies, and at widely different times. LIDAR data could overcome many of these issues, with the bare earth DTM and first return DSM collected at the same time and separated during processing. To investigate the utility of using the difference between the bare earth and first returns as the vegetation height, I looked at 686 DTM/DSM grid pairs from the National Center of Airborne Laser Mapping (NCALM, http://calm.geo.berkeley.edu/ncalm/index.html) and the OpenTopography Portal (http://www.opentopography.org/). The difference between the two should represent the vegetation or building height, which will have equivalent effects on intervisibility. Among these 686 data sets, the percentage of points with a negative vegetation height ranged from 0.2 to 35.3% of the data, with a mean of 6.26%. The largest negative value was -282.3 m (clearly a blunder), but 88 of the data sets had at least one point with a negative vegetation height exceeding -10 m, and one had 0.4% of the data points with a negative vegetation height exceeding -10 m. This suggests that a large fraction of LIDAR grids have problems with either the DTM or the DSM, and that existing quality control does not catch these points (a simple check during processing would be that the DTM elevation is less than the DSM elevation, the DSM elevation should replace the DTM elevation). In the absence of any better alternative, users could just set all the derived negative vegetation heights to zero. In addition, the maximum vegetation height also contains errors and gross blunders. Figure 6 shows a histogram of the calculated maximum vegetation height for the 686 grid pairs, with the maximum value at 1024.8 m (another blunder). The maximum vegetation height exceeds 100 m for 84 of the DEMs, and many of these points are clearly artefacts when viewed in the original DEM. These could be edited, with all points above a maximum assumed vegetation height either set to the maximum, or to zero if they are assumed to be blunders in processing including stray returns from birds or other random objects.

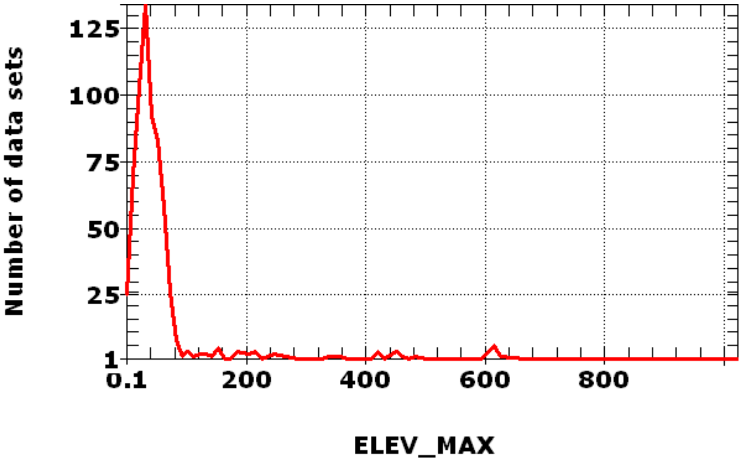

**Figure 6.** Maximum vegetation heights from the difference between 686 LIDAR DSMs and DTMs.

# VEGETATION INVENTORIES

Three gridded data sets could potentially be used to get vegetation heights. The National Land Cover Data (NLCD, Homer and others, 2007) classifies surface cover into about 30 categories, based on analysis of Landsat imagery. The algorithm allows assigning a vegetation height to selected categories, which can include buildings. The 30 m size of the NLCD pixels, and the ability of a single height per category, limit the accuracy of this method. The NLCD could also be used to clip the SRTM-NED vegetation estimates, and remove any areas not classified as vegetation (Figure 7).

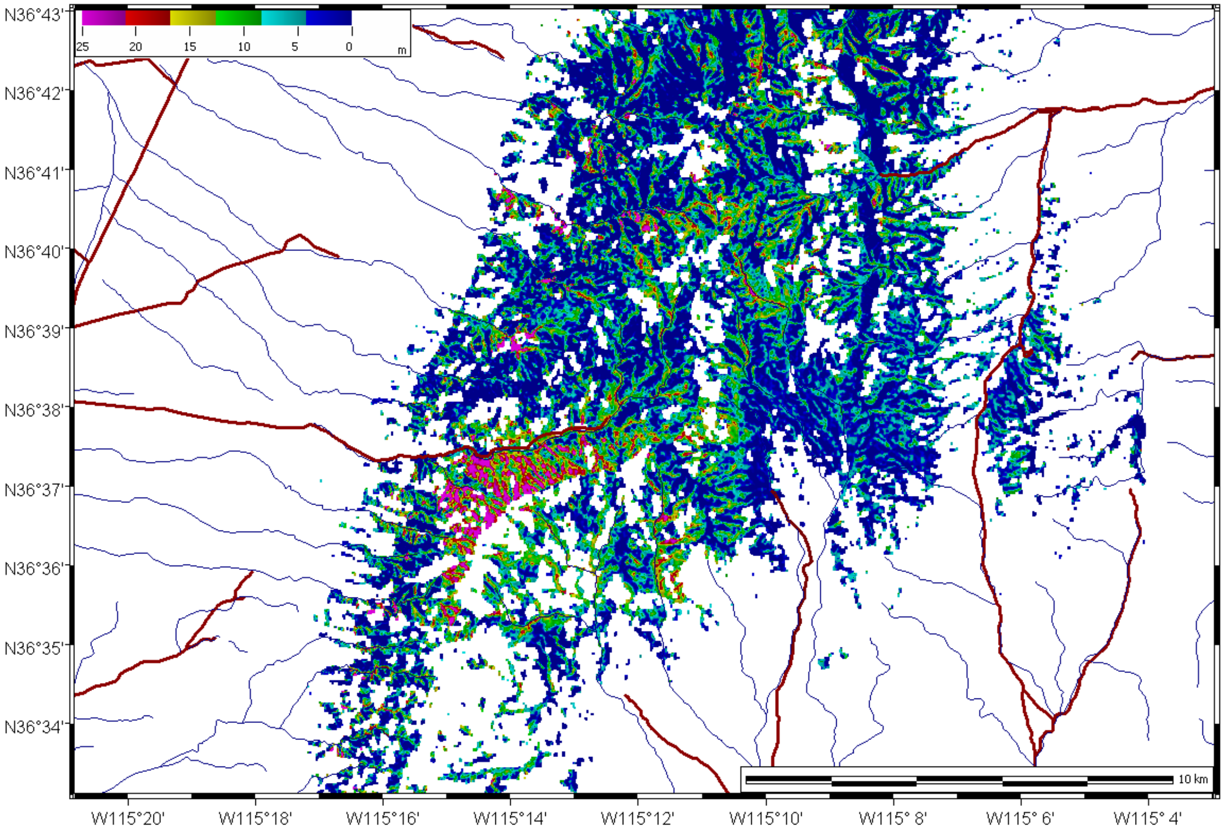

**Figure 7.** SRTM-NED difference with the colors clipped between 0-25 m for those areas classified as forests in NLCD-2001 in southern Nevada.

The LANDFIRE data set (http://www.landfire.gov/index.php) has a series of data sets of potential use:

1. Existing vegetation height, which groups trees into 5 height categories, and shrubs into 4 categories. These have wide ranges (e.g. 10-25 m, or 25-50 m). This layer also includes the NLCD categories for non-vegetated areas (Figure 8).
2. Forest canopy height, in a half dozen categories.
3. Canopy coverage, the percentage of the pixel with forest cover, in 10 categories.
4. Existing vegetation type: the terrestrial ecological system category.

While none of these layers provides vegetation height as an absolute estimate, and the data set was designed for modeling fire risks and burning dangers, the data set has the potential for use in how vegetation will interfere with intervisiblity.

The National Biomass and Carbon Dataset 2000 (NBCD 2000, http://whrc.org/proposals/ALOS_Google/index.htm; Walket et al., 2007) generated a 30 m dataset of the basal area-weighted (BAW) canopy height. These are continuous values, rather than the binned categories used in Landfire, and based on SRTM, NED, the LANDFIRE project, and NLCD-2001. Figure 9 shows the BAW heights computed in southern Nevada.

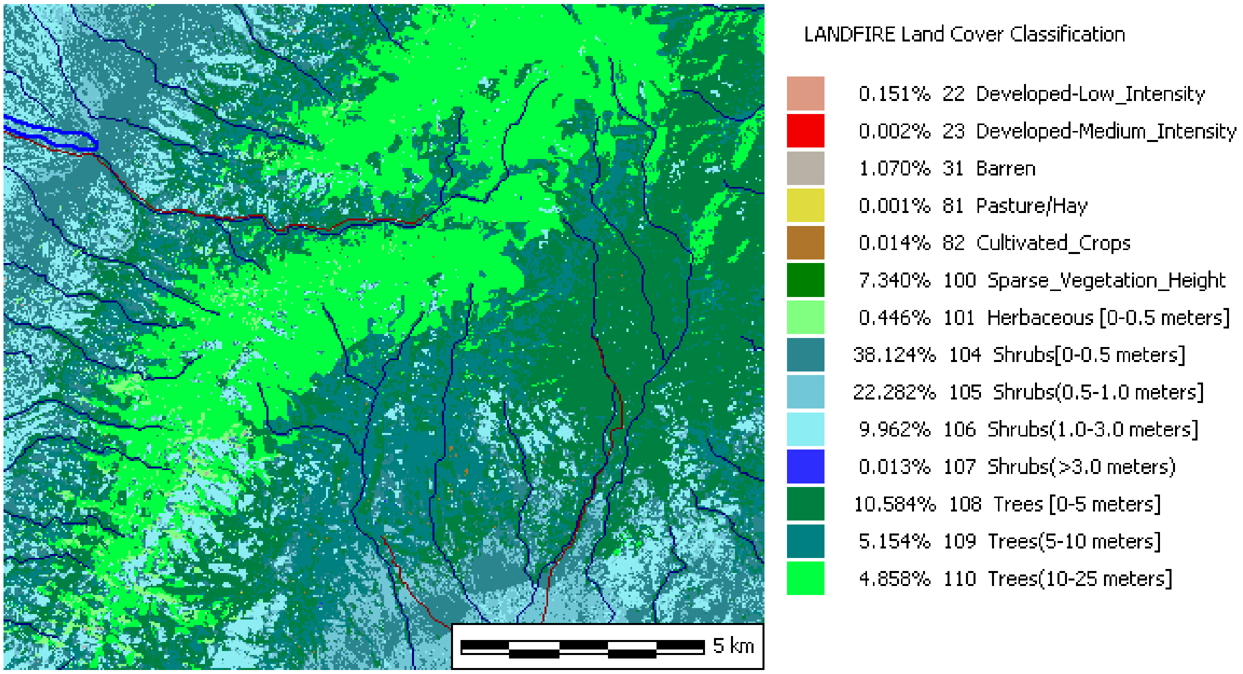

**Figure 8** Landfire existing vegetation height classification. The classification used a limited number of categories.

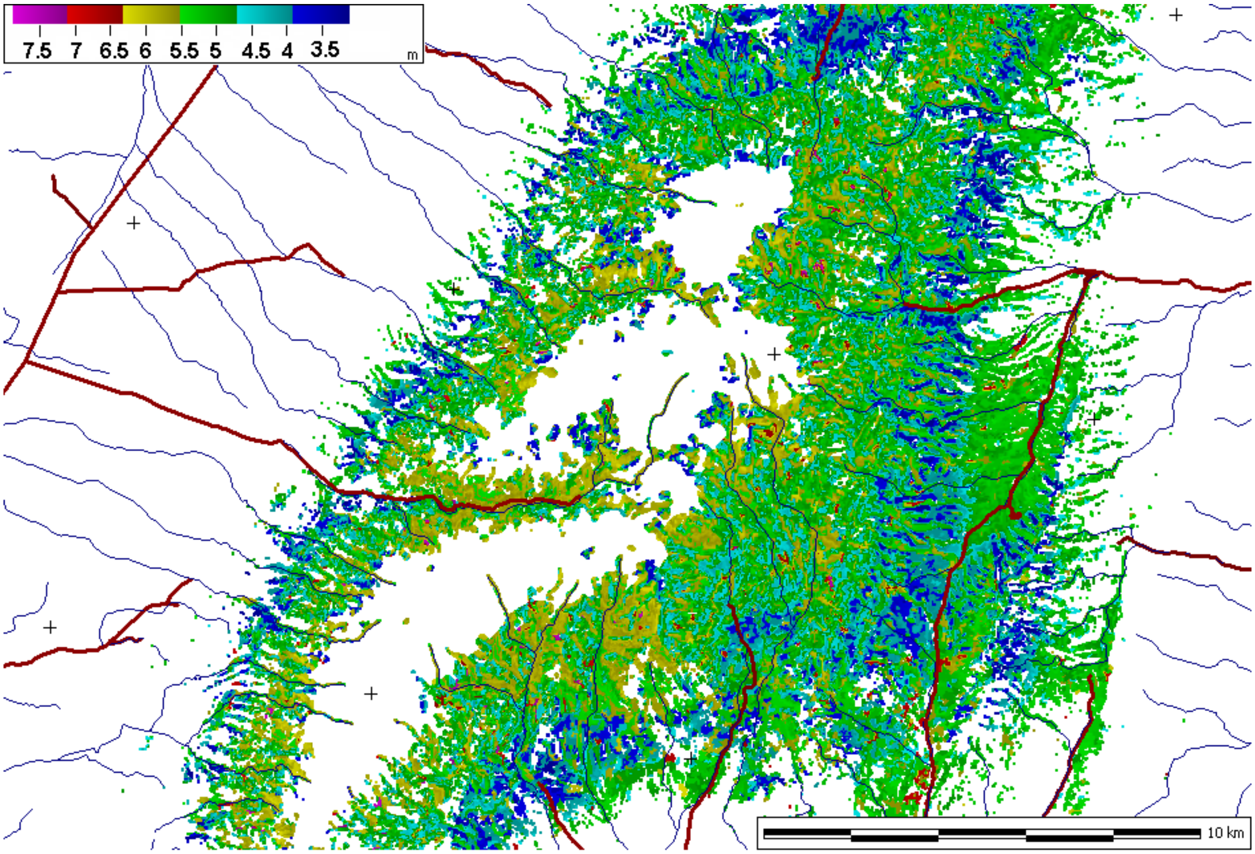

**Figure 9.** NBCD 2000 BAW.

# DISCUSSION

None of the existing land classification or vegetation indexes provide an ideal estimate of tree heights for intervisibility computations. All will require additional efforts to standardize data processing and quality control efforts, for instance to insure that when both first return and bare earth LIDAR grids are created, that the first return values (DSM) are not below the bare earth values (DTM). In addition, if any of these methods are used to estimate vegetation heights, the values should be compared with current imagery to insure that recently planted or rapidly growing vegetation has not changed the region.

The development of future algorithms should consider two enhancements for intervisiblity computations:

1. Where LIDAR data is available, it might prove best to replace a reliance on derived grids and use the raw point cloud instead. Algorithms could estimate the density of vegetation, and run probabilistic estimates to include visibility through light canopy.
2. If the vegetation and biomass data sets are used, the percent canopy coverage could also be used for probabilistic models for vision into the edges of forests, or though sparse forest such as in present in much of the western United States.

The increasing development of complex data sets covering large area such as the entire United States suggests that all GIS operations must consider the possibility of using additional data sets for more accurate results. While we will have to wait for the development of fully adequate datasets with vegetation data, the algorithms in GIS software need to begin considering the effects of vegetation (e.g. Llobera, 2007).

# ACKNOWLEDGMENTS

This work used the freeware MICRODEM program (Guth, 2008), available at http://www.usna.edu/Users/oceano/pguth/website/microdemdown.htm

# REFERENCES

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J.N. VanDriel, and J. Wickham, 2007. Completion of the 2001 National Land Cover Database for the conterminous United States, *Photogrammetric Engineering & Remote Sensing*, vol.73, no.4, p.337-341.

Gallant, J., and A. Read, 2009. Enhancing the SRTM data for Australia, in Purves, R., Gruber, S., Hengl, T, and Straumann, R. (eds.), *Proceedings of Geomorphometry 2009*, 31 August to 2 September 2009, University of Zurich, Switzerland, p.149-154; available online at http://www.geomorphometry.org/gallantread2009geomorphometry.

Guth, P.L., 2008. Geomorphometry in MICRODEM, in Hengl, T., Reuter, H.I. (eds), *Geomorphometry: Concepts, Software, Applications*, Developments in Soil Science Series, Elsevier, p.351-366.

Llobera, M., 2007. Modeling visibility through vegetation, *International Journal of Geographical Information Science*, 21 (7): 799-810.

Walker, W.S., J.M. Kellndorfer, E. LaPoint, M. Hoppus and J. Westfall, 2007. An empirical InSAR-optical fusion approach to mapping vegetation canopy height, *Remote Sensing of Environment,* Volume 109, Issue 4, p. 482-499.