Journal of the Arkansas Academy of Science

Volume 66

Article 16

2012

Comparing Methods for Interpolation to Improve Raster Digital Elevation Models

J. C. Guarneri
University of Arkansas at Monticello, guarneri@uamont.edu

R. C. Weih Jr.
University of Arkansas at Monticello

Comparing Methods for Interpolation to Improve Raster Digital Elevation Models

J.C. Guarneri[1,2] and R.C. Weih, Jr.[1, 2]

[1] *Arkansas Forest Resources Center, University of Arkansas at Monticello, School of Forest Resources, Spatial Analysis Laboratory, 110 University Court, Monticello, Arkansas 71656*

[2] Correspondence: guarneri@uamont.edu

Abstract

Digital elevation models (DEMs) are available as raster files at 100m, 30m, and 10m resolutions for the contiguous United States and are used in a variety of geographic analyses. Some projects may require a finer resolution. GIS software offers many options for interpolating data to higher resolutions. We compared ten interpolation methods using 10m sample data from the Ouachita Mountains in central Arkansas. We interpolated the 10m DEM to 5m, 2.5m, and 1m resolutions and compared the absolute mean difference (AMD) for each using surveyed control points. Overall, there was little difference in the accuracy between interpolation methods at the resolutions tested and minimal departure from the original 10m raster.

Introduction

Raster-based digital elevation models (DEMs) are an important data source used in a variety of spatial analyses. DEMs can be used to derive products such as slope (Weih and Mattson 2004), stream networks (O'Callaghan and Mark 1984), line-of-sight, hillshade, and irradiance maps (Burrough and McDonnel 1998), soil models (Thompson et al. 2001) and measures of terrain ruggedness (Sappington et al. 2007). The quality of these products depends on the quality and resolution of the source DEM. In a sensitivity analysis of stormwater runoff models, Cho and Lee (2001) found that runoff increased with increasing DEM resolution. Wu et al. (2008) found that mean topographic index values for a watershed increased with increasing grid cell size. Thompson et al. (2001) found the values of slope and several other DEM-derived data to depend on the horizontal and vertical resolution of the source DEM, as did the adjusted R^2 value for their soil model. Clearly, the resolution of raster DEMs has an impact on the quality of derived products (Figure 1).

DEMs are commonly available at 100m, 30m, and 10m horizontal resolutions, but some analyses may benefit from higher resolutions, especially when trying to model landforms that are smaller or narrower than the horizontal resolution of the raster DEM. The United States Geological Survey also has 3m data available in some areas as part of the national elevation dataset (NED), but most areas of the US do not have resolutions higher than 10m available (Gesch et al. 2009).

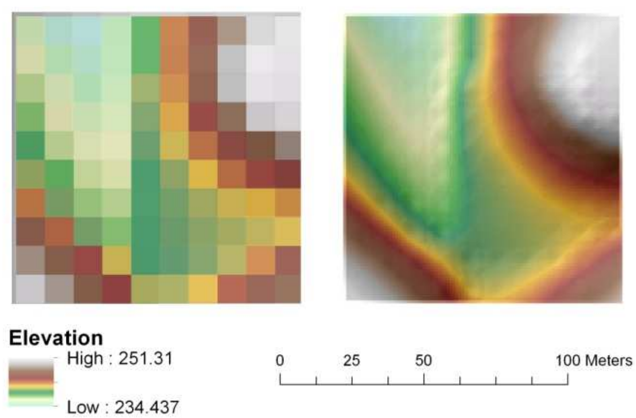

Figure 1. Comparison of original 10m DEM (left) and 1m DEM interpolated using a local polynomial function (right) Note the finer details that can be discerned in the interpolated DEM.

The ideal solution would be to collect data at a higher resolution, but this is typically limited by cost and manpower. An alternative approach is to use interpolation to estimate the elevation values at a higher resolution. Conceptually, the process is the same as "enhancing" the resolution of a digital photograph to produce a clearer image. There are many interpolation methods available (Burrough and McDonnel 1998, Yang and Hodler 2000, Aguilar et al. 2005), but the methods we reviewed fall into four general types: inverse distance weighted, radial basis functions, local polynomials, and kriging. There has been extensive literature published that compared the performance of these interpolators using scattered data (Weber and Englund 1992, 1994, DeClercq 1996, Yang and Hodler 2000, Aguilar et al. 2005). However, little attention has been paid to their performance on interpolating the regularly spaced data of an existing raster DEM to a finer resolution.

Inverse distance weighted (IDW) interpolation assumes that the values of a variable of interest, such as elevation, are more influenced by values at nearby locations than those at distant locations (Burrough and McDonnel 1998, Aguilar et al. 2005). It models the value Z at point x_0 using the expression:

$$Z(x_0) = \frac{\sum_{i=1}^{n} w(d_i) \cdot z(x_i)}{\sum_{i=1}^{n} w(d_i)} \qquad (1)$$

where $w(d_i)$ is a weight function, $z(x_i)$ is the elevation at known point x_i, and d_i is the distance between x_i and x_0. Since $w(d) \rightarrow \infty$ as $d \rightarrow 0$, the weight function can be expressed as $w = d^{-u}$. Increasing values of u decreases the weight of more distant points and creates a more localized interpolation.

The radial basis functions are a family of interpolators that model unknown values by calculating a set of localized linear equations using basis functions (Yang and Hodler 2000, Aguilar et al. 2005). They are represented as:

$$Z(x) = \sum_{i=1}^{m} a_i f_i(x) + \sum_{j=1}^{n} b_j \Psi(d_j) \qquad (2)$$

where $\Psi(d_j)$ is the radial basis function of the distance d_j from each sample location and point x, $f_i(x)$ is a trend function, and a_i and b_j are coefficients calculated separately. Radial basis functions include multiquadric, inverse multiquadric, spline functions, and others.

Polynomial functions attempt to use polynomial equations to fit a surface through source data points. Each additional degree of polynomial allows for bends in an additional dimension, but also increases model instability. Global polynomials attempt to fit a polynomial through all of the source data, whereas local polynomials fit a series of polynomials, each considering only the data within an area that overlaps its neighbors. The localized approach allows the model to capture more detail than the globalized approach, while using a lower-order, more stable polynomial.

The last interpolation method we explored is kriging. Kriging is a robust geostatistical method that uses regionalized variable theory to determine the parameters of the interpolation (Oliver and Webster 1990, DeMers 1997, Burrough and McDonnell 1998). As with IDW, kriging assumes that points that are close are more similar than those that are far away. Kriging takes the extra step of quantifying the relationship between difference in values and distance between points. This is accomplished with the use of a variogram, where the semivariance, $\gamma(h)$, or one-half of the difference between values at two points, is plotted as a function of the distance, or lag, h, between these points. Using the semivariance function, the value of a variable Z at point x is considered as:

$$Z(x) = m(x) + \gamma(h) + \varepsilon'' \qquad (3)$$

where $m(x)$ is a function of the overall surface trend, $\gamma(h)$ accounts for the spatially correlated deviations from this trend, and ε'' is the "noise" or residual of the measurements. Taken together, these components account for all sources of variation in the values of Z between points.

There are several forms of kriging, including ordinary kriging, where no assumptions are made regarding the overall trend or mean and simple kriging, which assumes a known mean value for all points. Universal kriging assumes an overall directional trend in the data. It is also possible to use additional variables to predict the values of the primary variable, using a process called co-kriging (Oliver and Webster 1990, Burrough and McDonnell 1998). The method of kriging that we tested was ordinary kriging, as the other forms were not appropriate for our elevation data.

Interpolation by ordinary kriging takes the form of:

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \qquad (4)$$

where $\hat{z}(x_0)$ is the predicted value of z at point x_o, λ_i is a weight function for point x_i, with values of λ_i are chosen to eliminate bias in $\hat{z}(x_0)$ and so that $\sum_{i=1}^{n} \lambda_i = 1$. $Z(x_i)$ is a function of the known values at n sampled points, x_i.

Materials and Methods

The study site is an area of approximately 360km^2 in the eastern Ouachita Mountains in Garland, Perry, and Saline counties in Arkansas (Figure 2). Elevations range from 149m to 577m above sea level. The area was selected to encompass a set of surveyed control points collected by Weih (2010).

We downloaded 10m DEM data for the study area from the NED in December 2011 (Gesch 2007, Gesch et al. 2002), and reprojected the data from GCS NAD 83 to UTM Zone 15N projection, NAD 83 using ArcMap 10.0 (Environmental Systems Research

https://scholarworks.uark.edu/jaas/vol66/iss1/16
DOI: https://doi.org/10.54119/jaas.2012.6609

Table 1. Comparison of absolute mean difference (AMD) values and elevation ranges among the original 10m DEM, inverse distance weighted (IDW), inverse multiquadric (IMQF), ordinary kriging (OK), first and second order local polynomial, multiquadric (MQF), completely regularized spline (CRS), spline with tension (TNSP), and thin plate spline (TPS) interpolations at 1m, 2.5m, and 5m. Minimum and maximum elevations of each DEM are also reported.

Method	**Resolution**	**n**	**Sum of Difference**	**AMD**	**Minimum Elevation**	**Maximum Elevation**
Original	10	105	286.239	2.726	149.01	576.39
IDW	1	105	286.239	2.726	149.01	576.39
IDW	2.5	105	286.239	2.726	149.01	576.39
IDW	5	105	287.925	2.742	149.01	576.39
IMQF	1	105	286.181	2.726	148.87	576.45
IMQF	2.5	105	285.752	2.721	148.87	576.45
IMQF	5	105	288.107	2.744	148.88	576.45
OK	1	105	286.529	2.729	148.90	576.43
OK	2.5	105	286.084	2.725	148.90	576.43
OK	5	105	285.552	2.720	148.92	576.43
1st Order Local Polynomial	1	105	286.644	2.730	149.05	576.36
1st Order Local Polynomial	2.5	105	286.062	2.724	149.05	576.35
1st Order Local Polynomial	5	105	287.555	2.739	149.06	576.36
2nd Order Local Polynomial	1	105	286.721	2.731	148.94	576.40
2nd Order Local Polynomial	2.5	105	286.234	2.726	148.94	576.40
2nd Order Local Polynomial	5	105	285.852	2.722	148.96	576.39
MQF	1	105	286.188	2.726	148.87	576.45
MQF	2.5	105	285.766	2.722	148.88	576.44
MQF	5	105	288.154	2.744	148.89	576.44
CRS	1	105	286.192	2.726	148.87	576.46
CRS	2.5	105	285.753	2.721	148.87	576.45
CRS	5	105	288.110	2.744	148.88	576.45
TNSP	1	105	286.616	2.730	148.92	576.42
TNSP	2.5	105	285.449	2.719	148.93	576.41
TNSP	5	105	288.374	2.746	148.93	576.40
TPS	1	105	286.616	2.730	148.92	576.42
TPS	2.5	105	286.088	2.725	148.92	576.41
TPS	5	105	288.374	2.746	148.93	576.40

Institute, Redlands, CA). Next, we converted the digital raster data to points, so that there was one point with an elevation value for each cell in the elevation grid. This point file was used as the basis for the interpolations.

All interpolations were performed using the Geostatistical Wizard function of ArcMap, which allows the user to preview the interpolation while adjusting the parameters, and offers an automated optimization algorithm. The variogram for kriging was fitted visually in the Wizard using a stable variogram model. Through a combination of manual adjustments and allowing the Wizard to optimize the model parameters, we generated interpolation surfaces using nine different interpolators: inverse distance weighted, multiquadric, inverse multiquadric, thin plate spline, spline with tension, completely regularized spline, ordinary kriging, and first and second order local polynomials. Other interpolators were tested, but were rejected due to instability evident in the preview.

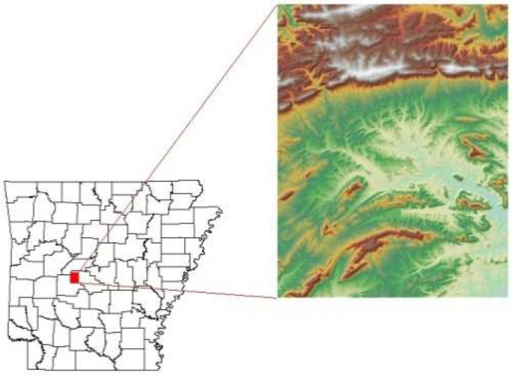

Figure 2. Study area showing original 10m DEM.

We extracted the values of the interpolated surfaces generated by the Geostatistical Wizard to rasters at 5m, 2.5m, and 1m resolutions. The original NED DEM and each interpolated raster were overlaid with the surveyed control points (Weih 2010) and raster elevation values extracted at each point. The raster elevations were evaluated using the absolute mean difference (AMD), which is calculated as:

$$AMD = \frac{\sum_{i=1}^{n}|z_i-\hat{z_i}|}{n} \qquad (5)$$

where z_i is the measured elevation at point i,$\hat{z_i}$ is the predicted elevation at point i, and n is the total number of points sampled. The resulting AMD values were then compared between interpolators and resolutions. We also compared the maximum and minimum elevations for each interpolated surface with the original DEM.

Results and Discussion

Overall, we found little variation in the statistical performance of each interpolator, with a range of less than 0.03m between all AMD values (Table 1). This is not too surprising, given the high density and regular spacing of source data. IDW interpolation showed the most similarity to the original DEM in terms of AMD and elevation range. The lowest AMD value was for spline with tension at 2.5m resolution, followed by ordinary Kriging at 5m resolution. Both of these elevation surfaces had better AMD values than the original DEM.

Given the similarity in performance between interpolators, processing time may be a more important factor in deciding which method to use. We noted that IDW and local polynomial interpolations were the fastest, multiquadric, inverse multiquadric, regularized spline, and ordinary kriging were intermediate, and spline with tension and thin plate spline took the longest. Also, the time to export the interpolation surface to a raster format increases exponentially with decreasing cell size. This is because the number of cells in a raster increases proportionally to the square of the inverse in the decrease in cell size. For example, a 1m raster would have 100 times the number of cells that a 10m raster of the same extent would. As the number of cells increases, so does the processing time. It should also be noted that data storage requirements also increase with decreasing cell size.

Acknowledgements

Funding for this project was provided by the USDA National Forest Service and the Arkansas Forest Resources Center at the University of Arkansas at Monticello. We would also like to thank Dr. Robert Kissel for providing feedback on the interpolation process and Emma Rearick and Dr. Hal Liechty for critiquing this manuscript.

Literature Cited

Aguilar FJ, **F Agüera**, **MA Aguilar** and **F Carvajal**. 2005. Effects of terrain morphology, sampling density, and interpolation methods on grid DEM Accuracy. Photogrammetric Engineering & Remote Sensing 71(7):805-16.

Burrough PA and **RA McDonnell**. 1998. Principles of Geographic Information Systems. New York: Oxford University Press. 333p.

Cho S and **M Lee**. 2001. Sensitivity Considerations when modeling hydrologic processes with digital elevation model. Journal of the American Water Resources Association 37(4):931-4.

DeClercq FAN. 1996. Interpolation methods for scattered sample data: accuracy, spatial patterns, processing time. Cartography and Geographic Information Systems 23(3):128-44.

DeMers MN. 1997. Fundamentals of Geographic Information Systems. New York: J. Wiley & Sons. 486 p.

Gesch DB. 2007. The national elevation dataset. *In* Maume, D, editor. Digital elevation model technologies and applications: the DEM users manual, 2nd edition. American Society for Photogrammetry and Remote Sensing (Bethesda, Maryland). p 99-118.

https://scholarworks.uark.edu/jaas/vol66/iss1/16

DOI: https://doi.org/10.54119/jaas.2012.6609

Gesch D, **G Evans**, **J Mauck**, **J Hutchinson** and **WJ Carswell**, **Jr**. 2009. *The National Map*—Elevation: U.S. Geological Survey Fact Sheet 2009-3053: 4 p. Available at: http://ned.usgs.gov/downloads/documents/NED_Accuracy.pdf. Accessed 3 April 2012.

Gesch D, **M Oimoen**, **S Greenlee**, **C Nelson**, **M Steuck** and **D Tyler**. 2002. The national elevation dataset. Photogrammetric engineering and remote sensing 68(1):5-11.

O'Callaghan JF and **DM Mark**. The extraction of drainage networks from digital elevation data. Computer Vision, Graphics, and Image Processing 28:323-44.

Oliver MA and **R Webster**. 1990. Kriging: a method of interpolation for geographical information systems. International Journal of Geographical Information Systems 4(3):313-32.

Sappington JM, **KM Longshore** and **DB Thompson**. 2007. Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. Journal of Wildlife Management 71(15):1419-1426.

Thompson JA, **JC Bell** and **CA Butler**. 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma 100:67-89.

Weber D and **E Englund**. 1992. Evaluation and comparison of spatial interpolators. Mathematical Geology 24(4):381-91.

Weber D and **E Englund**. 1994. Evaluation and comparison of spatial interpolators II. Mathematical Geology 26(5):589-603.

Weih RC. 2010. Assessing the vertical accuracy of Arkansas five-meter digital elevation model for different physiographic regions. Journal of the Arkansas Academy of Science 64:123-8.

Weih RC and **TL Mattson**. 2004. Modeling slope in a geographic information system. Journal of the Arkansas Academy of Science 58:100-8.

Wu S, **J Li**, and **H Huang**. 2008. Characterization and evaluation of elevation data uncertainty in water resources modeling with GIS. Water Resource Management 22:959-972.

Yang X and **T Hodler**. 2000. Visual and statistical comparisons of surface modeling techniques for point-based environmental data. Cartography and Geographic Information Science 27(2):165-75.