Quantifying digital elevation model (DEM) uncertainty introduced by interpolative gridding

Christopher J. Amante[1,2], Matthew R. Love[1,2], Barry W. Eakins[1,2]

[1]Cooperative Institute for Research in the Environmental Sciences, University of Colorado at Boulder
[2]NOAA, National Geophysical Data Center, Boulder, Colorado

Overview

Digital elevation models (DEMs) are representations of the Earth's solid surface and are the framework for modeling numerous oceanic processes, including tsunami propagation and ocean circulation. The accuracy of modeling such processes is in part dependent on the accuracy of the DEM. DEM errors, deviations from the actual seabed, originate from both the source depth measurements (e.g., multibeam sonar, lidar) and the interpolative gridding technique (e.g., spline, kriging, inverse distance weighting) used to estimate depths in areas with no source measurements. Previous research found that interpolation errors are as significant as the measurement errors and should be considered when generating and using DEMs (Guo et al., 2010). Numerous studies also indicate that interpolation errors are positively correlated with terrain complexity and distance to known measurements (e.g., Chaplot et al., 2006; Erdogan, 2009). The magnitude of interpolation errors is often unknown and the lack of knowledge about these errors represents the uncertainty introduced by the gridding process. We developed a computer program that utilizes a split-sample methodology to quantify interpolation errors introduced by gridding. A method for quantifying the DEM uncertainty introduced by interpolative gridding from the range of interpolation errors amongst the various techniques is described.

Study Area

Three study areas with different terrain characteristics were chosen to help quantify the uncertainty introduced by interpolative gridding as a function of terrain complexity (**Figure 1**). Kachemak Bay, AK has a steep slope offshore. The seafloor offshore of Crescent City, CA is rugged, while San Augustine, CA has a gradual slope with minimal variability offshore. All three study areas were recently surveyed with high-resolution multibeam swath sonar and are available for download from the NOS Hydrographic Survey database (http://www.ngdc.noaa.gov/mgg/bathymetry/hydro.html) in Bathymetric Attributed Grid (BAG) format. The surveys were downloaded as thoroughly evaluated, combined BAGs with either 4 or 5 meter horizontal resolution in UTM coordinates.

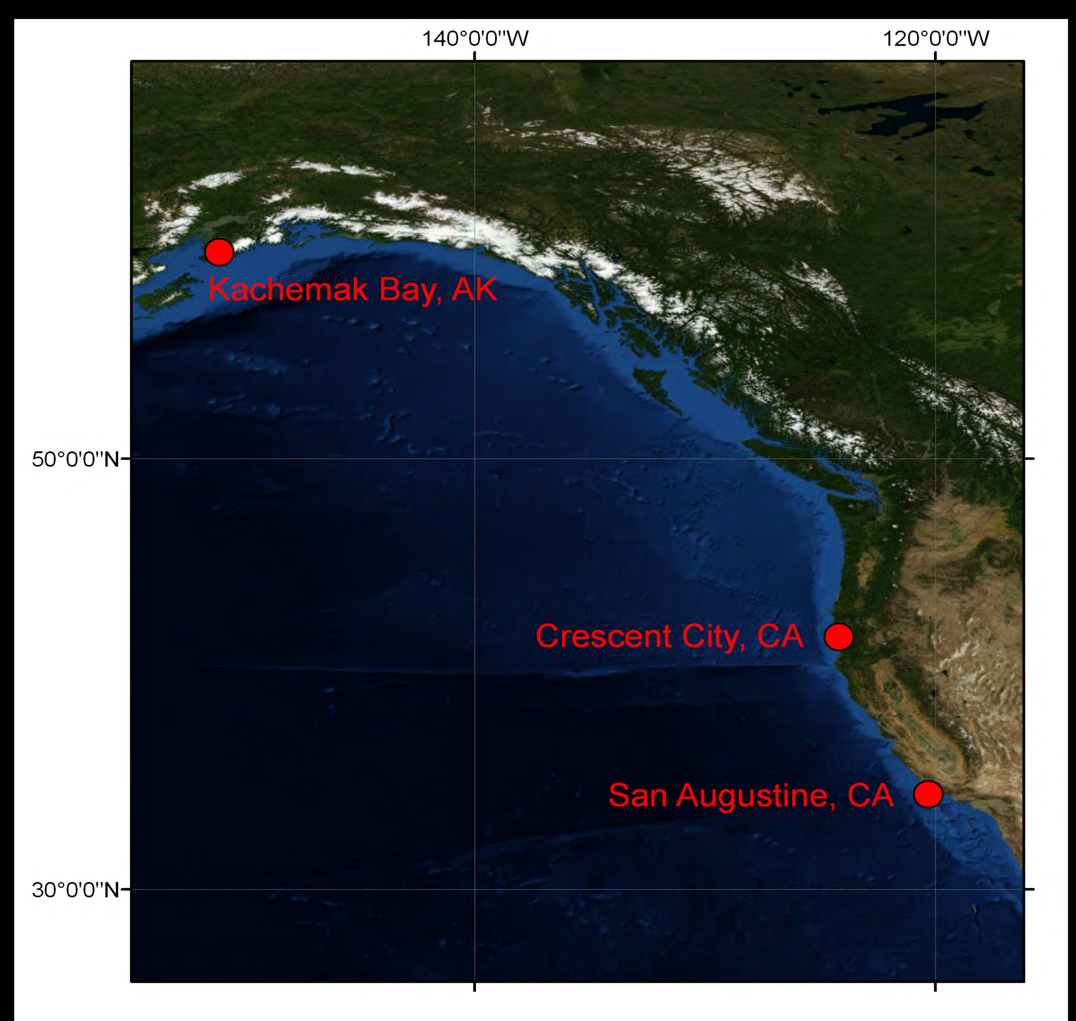

***Figure 1.** Locations of the three study areas. Background is ESRI World 2D imagery.*

Methodology

The survey BAGs were converted to XYZ, then median averaged using the Generic Mapping Tools (GMT) 'blockmedian' tool at 10-meter point spacing. This averaging filled small gaps in the surveys and ensured that every 10-m cell in each rectangular area had one constraining data point derived from depth measurements.

A split-sample approach was then used to quantify the errors of interpolated depths using the known blockmedianed data. Using this method, a percentage of the data is omitted, an interpolation method is applied to the data subset, and the interpolation errors are quantified as the differences between the interpolated depths and the original omitted depths (**Figure 2**). In order to quantify the errors of the interpolation method at every data point, this process is repeated at the same split percentage and the differences between the original omitted depths and the interpolated depths are aggregated. The split-sample method is often used to assess the stability of interpolation methods by omitting increasingly greater percentages of the original data and analyzing changes in the interpolation errors.

Each interpolation method has numerous parameters that can be adjusted to create different representations of Earth's surface (e.g., maximum neighbors for IDW; tension value for spline). The interpolation errors in each DEM developed from various combinations of parameters will be quantified and the uncertainty of the interpolation method can be defined as the maximum error for each DEM pixel from all possible combinations of parameters. If this process is repeated for other interpolation methods, the total uncertainty introduced by interpolative gridding is the maximum interpolation error for each pixel from all possible DEMs derived using multiple interpolation methods.

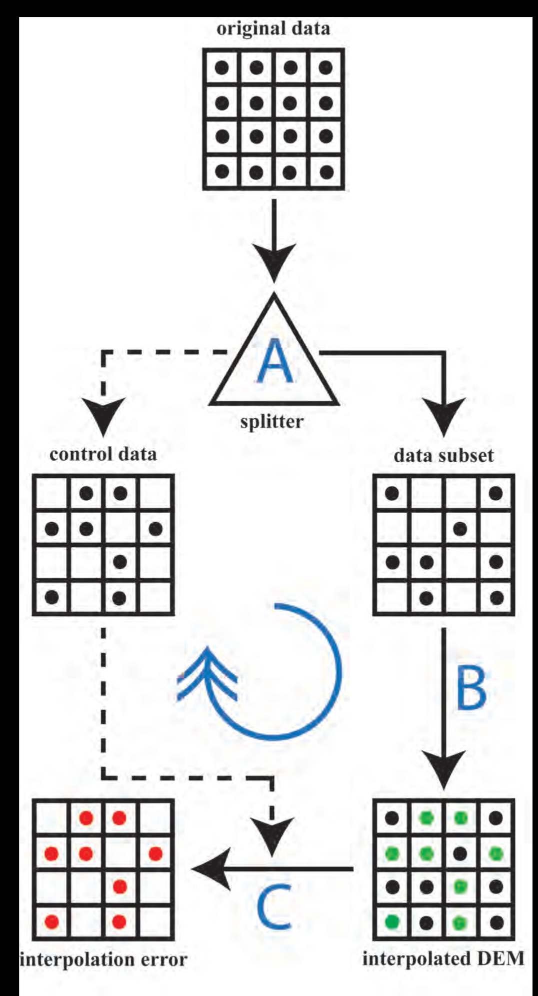

***Figure 2.** Flowchart depicting the split-sample methodology for quantifying interpolation errors. A) The original data are averaged to have exactly one depth value per grid cell. They are then randomly split by a fixed percentage (e.g., 50%) into control data and data subset. B) An interpolation method (e.g., spline, triangulation, IDW) is applied to the data subset to build an interpolated DEM. C) The interpolated DEM is compared to the control data to quantify the interpolation errors as a function of distance from data. Steps A to C are repeated at the same split percentage (randomness resulting in different control data and data subset) to determine interpolation error at every grid cell and account for bathymetric variability. The method is run iteratively using different split percentages to evaluate the stability (e.g., ability to reproduce the principal bathymetry) of the chosen interpolation method with various data densities.*

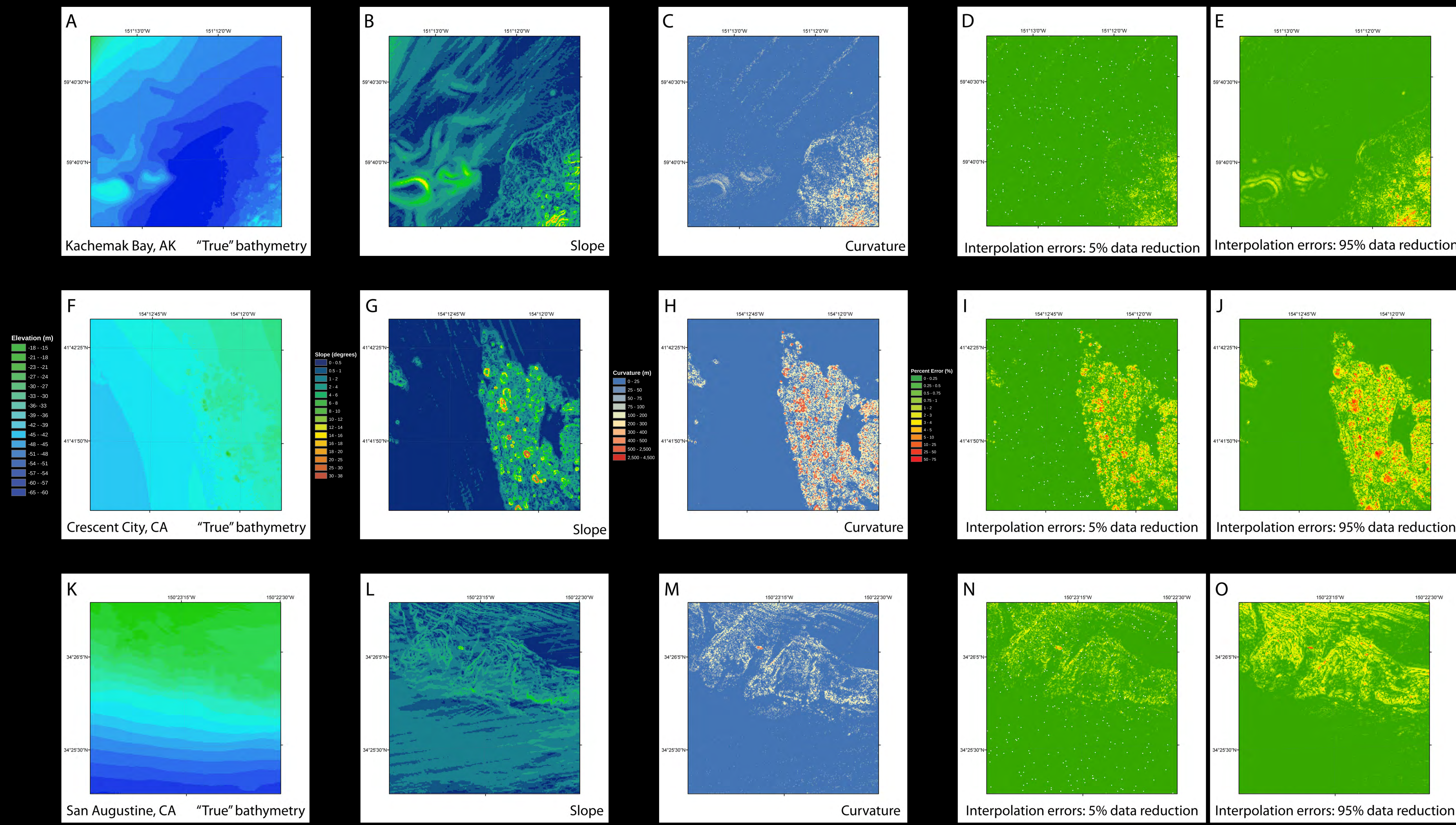

***Figure 3.** Interpolation errors from GMT 'surface' spline interpolation are greatest in areas of complex terrain (large slope and curvature) and increase with greater percentages of data removed.*

A – E. "True" bathymetry, slope, curvature, and interpolation error grids for 5%, and 95% of data points removed in Kachemak Bay, AK.

F – J. "True" bathymetry, slope, curvature, and interpolation error grids for 5%, and 95% of data points removed offshore of Crescent City, CA.

K – O. "True" bathymetry, slope, curvature, and interpolation error grids for 5%, and 95% of data points removed offshore of San Augustine, CA.

Preliminary Results

All of the original survey data are median-averaged at a given cell size and represent the "true" bathymetry. Slope and curvature grids are derived from the "true" bathymetry. The split-sample program quantifies interpolation errors as the differences between interpolated depths and the "true" bathymetry. The interpolation errors are then quantified as a function of terrain characteristics (slope and curvature) and distance to known measurements. Interpolation errors are greatest in areas of complex terrain (large slope and curvature) and increase with greater percentages of data removed (**Figure 3**).

Future Work

Chris Amante's Master's thesis project at the University of Colorado at Boulder is investigating the variations in DEM surfaces created by various gridding techniques (e.g., spline, inverse distance weighting, kriging, triangulation, nearest-neighbor) and the impact of these variations on the modeling of tsunami inundation at Crescent City, California. The University of Alaska, Fairbanks tsunami modeling code will be used to model inundation from the 1964 Alaska tsunami on each surface, and compare with inundation from the actual event.

References

Chaplot V., Darboux F., Bourennane H., Leguedois S., Silvera N. & Phachomphon K. (2006) Accuracy of interpolation techniques for the derivation of the digital elevation models in relation to landform types and data density, *Geomorphology*, 77, 126– 141.

Erdogan S. (2009) A comparison of interpolation methods for producing digital elevation models at the field scale. *Earth Surface Processes and Landforms*, 34, pp. 366–376.

Guo, Q., Li, W., Yu, H., Alvarez, O. (2010) Effects of topographic variability and lidar sampling density on several DEM interpolation methods. *Photogrammetric Engineering & Remote Sensing*, 76, 1-12.

Website and Contact Information

NGDC NOS Hydrographic Survey Database
http://www.ngdc.noaa.gov/mgg/bathymetry/hydro.html

Generic Mapping Tools (GMT)
http://gmt.soest.hawaii.edu/

Contact:
Christopher.Amante@colorado.edu
Work phone: (303) 497-4338