

Interpolating Surfaces in ArcGIS Spatial Analyst

By Colin Childs, ESRI Education Services

Editor's note: In addition to supplying tools for spatial analysis (i.e., for modeling suitability, distance, or hydrology), the ArcGIS 9 Spatial Analyst extension provides tools for spatial data analysis that apply statistical theory and techniques to the modeling of spatially referenced data.

Elevation, temperature, and contamination concentrations are types of data that can be represented by surfaces. Each raster cell represents a measurement such as a cell's relationship to a fixed point or specific concentration level. Because obtaining values for each cell in a raster is typically not practical, sample points are used to derive the intervening values using the interpolation tools in ArcGIS Spatial Analyst. This article provides an introduction to the interpolation methods used by these tools.

GIS is all about spatial data and the tools for managing, compiling, and analyzing that data. ArcGIS Spatial Analyst extension provides a toolset for analyzing and modeling spatial data. A set of sample points representing changes in landscape, population, or environment can be used to visualize the continuity and variability of observed data across a surface through the use of interpolation tools. These changes can be extrapolated across geographic space. The morphology and characteristics of these changes can be described. The ability to create surfaces from sample data makes interpolation both powerful and useful.

Understanding Surfaces

Before discussing different interpolation techniques, the differences in the methods used for surface representation need to be discussed. Each representation is useful for specific situations. This discussion concentrates on the creation of a surface in the native ArcGIS grid format.

A grid representation of a surface is considered to be a functional surface because for any given x,y location, it stores only a single Z value as opposed to multiple Z values. Functional surfaces are continuous because an x,y location has one and only one Z value regardless of the direction from which the x,y point is approached. Functional surfaces are 2.5-dimensional surfaces, not true three-dimensional surfaces.

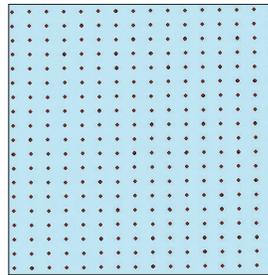
Functional surfaces can be used to represent terrestrial surfaces that depict the earth's surface, statistical surfaces that describe demographic and other types of data, and mathematical surfaces that are based on arithmetic expressions. Surface representation in its simplest form is done by storing x,y and Z values that define the location of a sample and the change characteristic represented by the Z value.

Contours or isolines are used to define a common characteristic along a line. Technically, contours join locations of equal value to each other. In the case of a contour line representing height, it is a line drawn on a map that connects points of equal elevation above a datum that usually represents mean sea level.

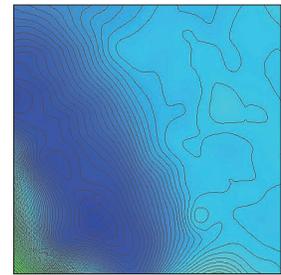
A triangulated irregular network (TIN) is a vector data structure used to store and display surface models. A TIN partitions geographic space using a set of irregularly spaced data points, each of which has x -, y -, and Z -values. These points are connected by edges that form contiguous, nonoverlapping triangles and create a continuous surface that represents the terrain.

A grid is a spatial data structure that defines space as an array of cells of equal size that are arranged in rows and columns. In the case of a grid that represents a surface, each cell contains an attribute value that represents a change in Z value. The location of the cell in geographic space is obtained from its position relative to the grid's origin.

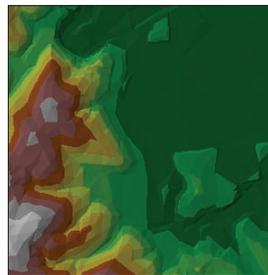
Methods of Surface Representation



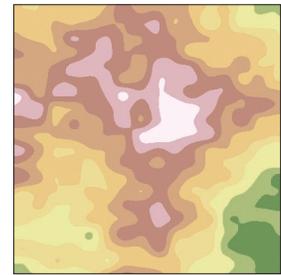
Points



Contours



TINs



Grids

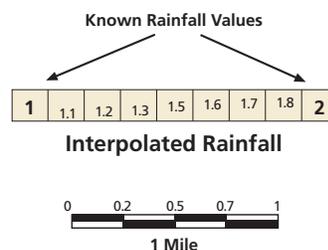
Interpolation Methods

To create a surface grid in ArcGIS, the Spatial Analyst extension employs one of several interpolation tools. Interpolation is a procedure used to predict the values of cells at locations that lack sampled points. It is based on the principle of spatial autocorrelation or spatial dependence, which measures degree of relationship/dependence between near and distant objects.

Spatial autocorrelation determines if values are interrelated. If values are interrelated, it determines if there is a spatial pattern. This correlation is used to measure

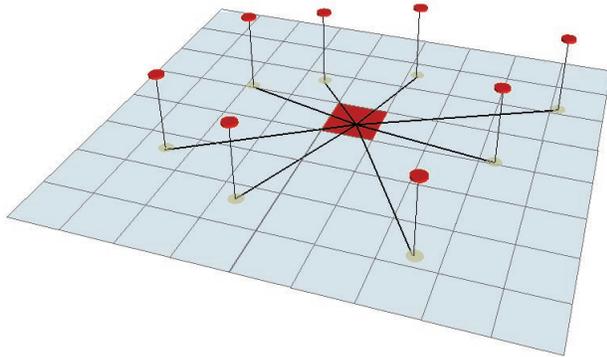
- Similarity of objects within an area
- The degree to which a spatial phenomenon is correlated to itself in space
- The level of interdependence between the variables
- Nature and strength of the interdependence

Different interpolation methods will almost always produce different results.

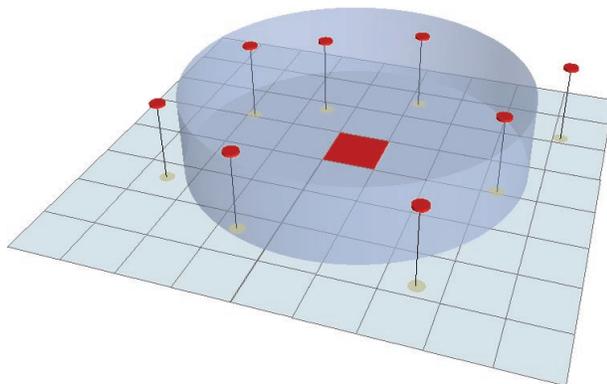


Interpolation is the procedure used to predict cell values for locations that lack sampled points.

Maximum Number



Fixed Radius

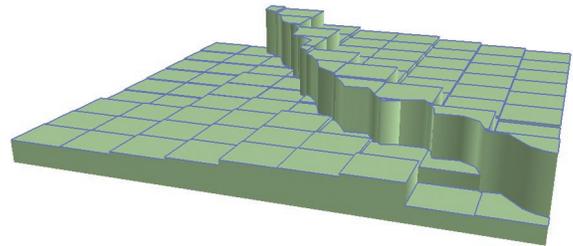


The characteristics of an interpolated surface can be controlled by limiting the input points used. Points can be limited by specifying a maximum number of points, which will select the nearest points until that maximum number of points is reached. Alternatively, a radius can be specified in map units.

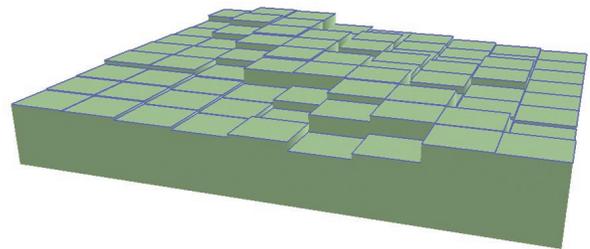
There are two categories of interpolation techniques: deterministic and geostatistical. Deterministic interpolation techniques create surfaces based on measured points or mathematical formulas. Methods such as Inverse Distance Weight (IDW) are based on the extent of similarity of cells while methods such as Trend fit a smooth surface defined by a mathematical function. Geostatistical interpolation techniques such as Kriging are based on statistics and are used for more advanced prediction surface modeling that also includes some measure of the certainty or accuracy of predictions.

The characteristics of an interpolated surface can be controlled by limiting the input points used in the calculation of output cell values. This can be done by limiting the number of points sampled or the area from which sampled points are taken. Specifying the maximum number of points to be sampled will return the points closest to the output cell location until the maximum number is reached. Alternatively, specifying a fixed radius in map units will select only input points within the radius distance from the center of the output cell unless there are not enough points within that radius.

With Barrier



Without Barrier



Barriers, which reflect the presence of fault lines, cliffs, streams, and other features that create linear discontinuity in surfaces, also control how surfaces are generated. IDW and Kriging support the use of barriers.

Many interpolation tools incorporate barriers that define and control surface behavior in terms of smoothness and continuity. Barriers are needed because sometimes interpolation operations should not be performed across features, such as fault lines, levees, cliffs, and streams, that create a linear discontinuity in the surface. By using barriers, changes in the behavior of the surface can be described and enforced.

The characteristics of an interpolated surface can be controlled by limiting the input points used in the calculation of output cell values.

Available Interpolation Options

ArcGIS Spatial Analyst at version 9 offers several interpolation tools for generating surface grids from point data. PointInterp, Natural Neighbors, and Trend methods and the Topo to Raster command have been added to the IDW, Spline, and Kriging interpolation methods that were available in ArcGIS 8.3 Spatial Analyst. Natural Neighbor and Trend methods were available in ArcGIS 8.3 3D Analyst, although Trend was only accessible programmatically. Each method uses a different approach for determining the output cell values. The most appropriate method will depend on the distribution of sample points and the phenomenon being studied.

Continued on page 34

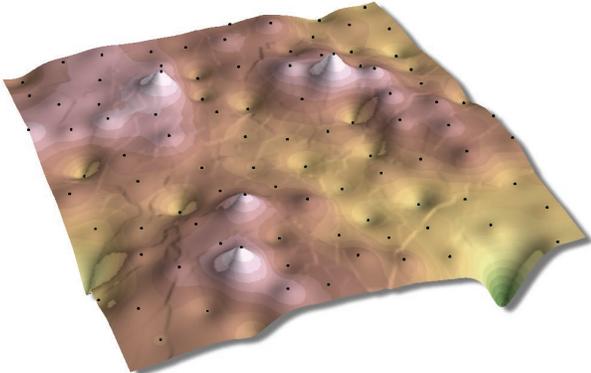
Interpolating Surfaces in ArcGIS Spatial Analyst

Continued from page 33

IDW

The IDW function should be used when the set of points is dense enough to capture the extent of local surface variation needed for analysis. IDW determines cell values using a linear-weighted combination set of sample points. The weight assigned is a function of the distance of an input point from the output cell location. The greater the distance, the less influence the cell has on the output value.

Inverse Distance Weighted

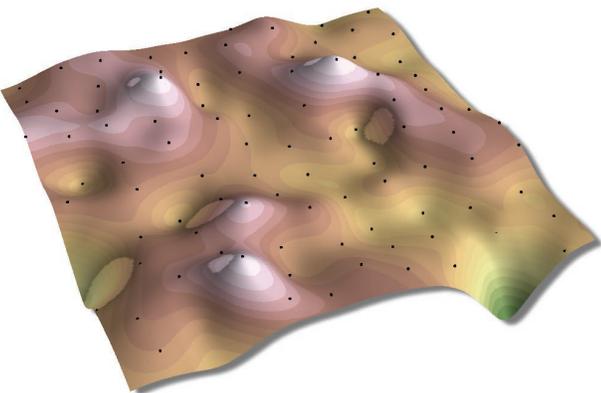


Spline

Spline estimates values using a mathematical function that minimizes overall surface curvature. This results in a smooth surface that passes exactly through the input points. Conceptually, it is like bending a sheet of rubber so that it passes through the points while minimizing the total curvature of the surface. It can predict ridges and valleys in the data and is the best method for representing the smoothly varying surfaces of phenomena such as temperature.

There are two variations of spline—regularized and tension. A regularized spline incorporates the first derivative (slope), second derivative (rate of change in slope), and third derivative (rate of change in the second derivative) into its minimization calculations. Although a tension spline uses only first and second derivatives, it includes more points in the spline calculations, which usually creates smoother surfaces but increases computation time.

Spline



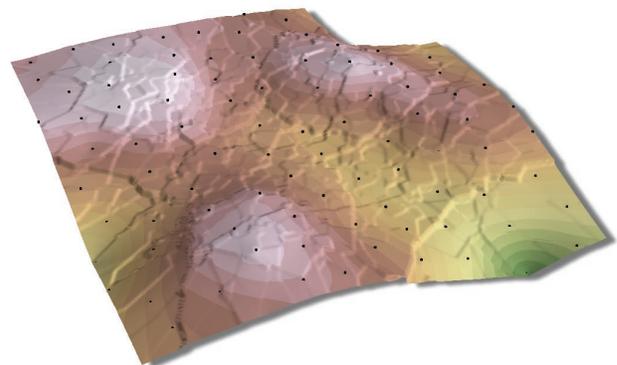
Kriging

A powerful statistical interpolation method used for diverse applications such as health sciences, geochemistry, and pollution modeling, Kriging assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variation in the surface. It fits a function to a specified number of points or all points within a specified radius to determine the output value for each location. Kriging is most appropriate when a spatially correlated distance or directional bias in the data is known and is often used for applications in soil science and geology.

The predicted values are derived from the measure of relationship in samples using sophisticated weighted average technique. It uses a search radius that can be fixed or variable. The generated cell values can exceed value range of samples, and the surface does not pass through samples.

There are several types of Kriging. Ordinary Kriging, the most common method, assumes that there is no constant mean for the data over an area mean (i.e., no trend). Universal Kriging does assume that an overriding trend exists in the data and that it can be modeled.

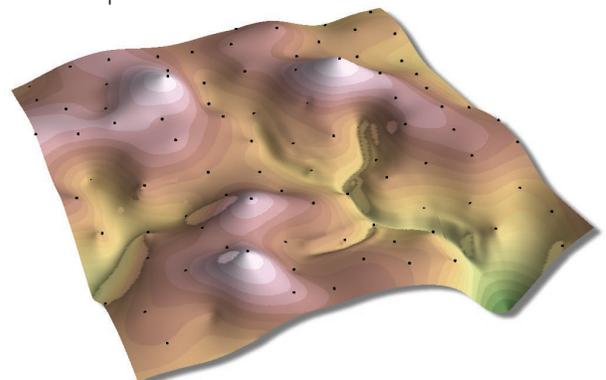
Kriging



PointInterp

A method that is similar to IDW, the PointInterp function allows more control over the sampling neighborhood. The influence of a particular sample on the interpolated grid cell value depends on whether the sample point is in the cell's neighborhood and how far from the cell being interpolated it is located. Points outside the neighborhood have no influence.

PointInterp

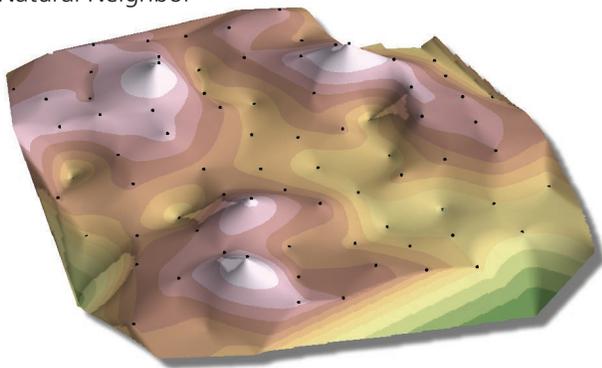


The weighted value of points inside the neighborhood is calculated using an inverse distance weighted interpolation or inverse exponential distance interpolation. This method interpolates a raster using point features but allows for different types of neighborhoods. Neighborhoods can have shapes such as circles, rectangles, irregular polygons, annuluses, or wedges.

Natural Neighbor

Natural neighbor interpolation has many positive features, can be used for both interpolation and extrapolation, and generally works well with clustered scatter points. Another weighted-average method, the basic equation used in natural neighbor interpolation is identical to the one used in IDW interpolation. This method can efficiently handle large input point datasets. When using the Natural Neighbor method, local coordinates define the amount of influence any scatter point will have on output cells.

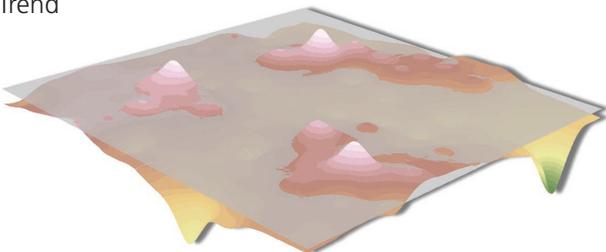
Natural Neighbor



Trend

Trend is a statistical method that finds the surface that fits the sample points using a least-squares regression fit. It fits one polynomial equation to the entire surface. This results in a surface that minimizes surface variance in relation to the input values. The surface is constructed so that for every input point, the total of the differences between the actual values and the estimated values (i.e., the variance) will be as small as possible. It is an inexact interpolator, and the resulting surface rarely passes through the input points. However, this method detects trends in the sample data and is similar to natural phenomena that typically vary smoothly.

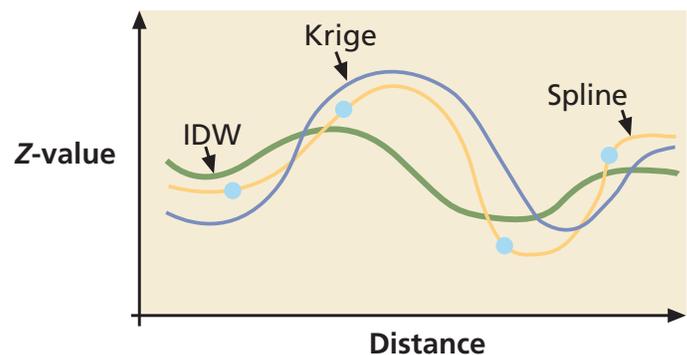
Trend



Each of the interpolation methods available in the ArcGIS Spatial Analyst extension uses a different approach for determining output cells. Choose a method based on the distribution of sample points and the phenomenon being studied.

Topo to Raster

By interpolating elevation values for a raster, the Topo to Raster method imposes constraints that ensure a hydrologically correct digital elevation model that contains a connected drainage structure and correctly represents ridges and streams from input contour data. It uses an iterative finite difference interpolation technique that optimizes the computational efficiency of local interpolation without losing the surface continuity of global interpolation. It was specifically designed to work intelligently with contour inputs.



IDW and Spline are two deterministic methods that create surfaces from samples based on the extent of similarity or degree of smoothing. However, while a spline surface passes exactly through each sample point, an IDW will pass through none of the points. Kriging is a geostatistical method that uses a powerful statistical technique for predicting values derived from the measure of relationship in samples and employs sophisticated weighted average techniques.

Conclusion

This brief introduction supplies a framework for exploring the interpolation tools available in ArcGIS 9 Spatial Analyst. A more robust and comprehensive arsenal of tools for data exploration and spatial data analysis is available from the ArcGIS 9 Geostatistical Analyst extension. More information on how to perform spatial data analysis and applications for the tools in ArcGIS 9 Spatial Analyst may be found in the product manual, *Using ArcGIS Spatial Analyst*. Papers that provide background information on interpolation methods and their application in various industries and disciplines can be found at www.esri.com/geostatisticalanalyst.