ArcGIS™ Spatial Analyst: Advanced GIS Spatial Analysis Using Raster and Vector Data

An ESRI® White Paper • December 2001

ESRI 380 New York St., Redlands, CA 92373-8100, USA • TEL 909-793-2853 • FAX 909-793-5953 • E-MAIL info@esri.com • WEB www.esri.com

J-8747

ArcGIS Spatial Analyst: Advanced GIS Spatial Analysis Using Raster and Vector Data

An ESRI White Paper

Contents **Page**

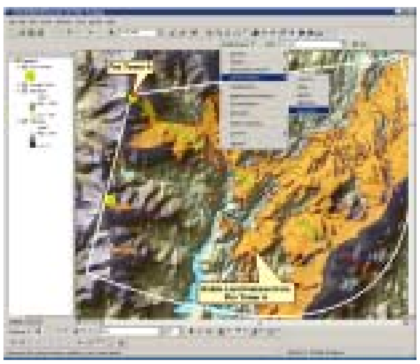
J-8747

Contents **Page**

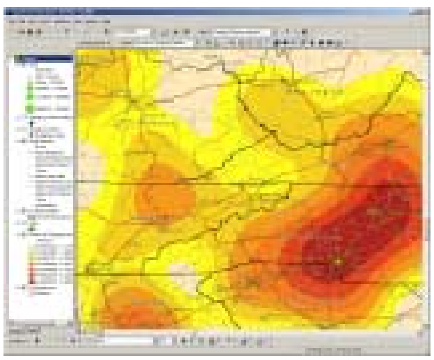
ArcGIS Spatial Analyst: Advanced GIS Spatial Analysis Using Raster and Vector Data

ArcGIS Spatial Analyst

ArcGIS™ Spatial Analyst provides a broad range of powerful spatial modeling and analysis features. Using ArcGIS Spatial Analyst, geographic information system (GIS) users can create, query, map, and analyze cell-based raster data; perform integrated raster/vector analysis; derive new information from existing data; query information across multiple data layers; and fully integrate cell-based raster data with traditional vector data sources. ArcGIS Spatial Analyst is integrated into the ArcGIS interface so the user can take advantage of all the advanced functionality in ArcGIS as well as work with other extensions such as ArcGIS Geostatistical Analyst and ArcGIS 3D Analyst™. Using ArcGIS Spatial Analyst, GIS users can derive information about geospatial data such as terrain analysis, spatial relationships, suitable locations, and the accumulated cost of traveling from one point to another. ArcGIS Spatial Analyst integrates real-world variables such as elevation into the geospatial environment to help solve complex problems. ArcGIS Spatial Analyst provides new functionality for advanced customization and interoperability. Using a common architecture and incorporating customization within any Component Object Model (COM)-compliant programming language, users can create more advanced raster models for their analysis.



Viewshed Analysis



Density Analysis

Who Uses ArcGIS Spatial Analyst?

Any GIS users who need a better spatial understanding of their vector or raster data can use ArcGIS Spatial Analyst.

From analyzing the optimal location for a new retail store to the location of the most sustainable area for a vegetation class, ArcGIS Spatial Analyst bridges the gap between a simple map on a computer and real-world analysis for deriving solutions to complex problems. Some of the various industries that utilize ArcGIS Spatial Analyst include agriculture production, exploration geology, meteorology, hydrology, archaeology, forestry, health care, mining, real estate, park services, city governments, retail chains, health care, and many others.

"The ArcGIS Spatial Analyst extension is essential to my day-to-day work analyzing water quality. I primarily use a landscape approach to modeling water quality in a watershed…I create hydrologically correct DEMs; weighted flow accumulation runoff grids; loading grids for expected mean concentration modeling; and model flow using various raster inputs such as precipitation, temperature, elevation, stream gradient, and drainage area from delineated watersheds. Multivariate statistical relationships are easily accomplished using various overlay analysis functions within ArcGIS Spatial Analyst… I could not imagine having ArcView without the ArcGIS Spatial Analyst extension."

Michael P. Strager
Research Coordinator
Natural Resource Analysis Center
West Virginia University

Key Features of ArcGIS Spatial Analyst 8.1

ArcGIS Spatial Analyst 8.1 introduces many new features that add to the functionality of the existing ArcView® Spatial Analyst 3.x. With features such as an industry-standard customization environment and integration into the ArcGIS interface, users now have a variety of powerful tools for vector–raster analysis.

- Integration into ArcMap™ software for ArcView, ArcEditor™, and ArcInfo™ software.
- Performance of analysis on all raster formats.
- More analysis functions and options in the user interface.
- Map algebra in the raster calculator on all raster formats.
- Function dialogs allowing the user to browse for data; inputs do not need to be in the table of contents.
- Support for selection on raster inputs (on-the-fly masking).
- GRID functionality available through the COM objects provided.
- Increased data integration support.
- More robust selection environment.

- Analysis on string fields.
- Selection on raster and feature.

Raster Data

Cell-based raster data sets (images and grids) are especially suited to representing traditional geographic phenomena that vary continuously over space such as elevation, slope, and precipitation. They can also be used to represent less traditional types of information such as population density, consumer behavior, and other demographic characteristics. Rasters are also the ideal data representation for spatial modeling and analysis of flows and trends over data represented as continuous surfaces such as hydrologic modeling or the dynamics of population change over time.

Cell-Based Modeling

One of the strongest aspects of ArcGIS Spatial Analyst is its analytical capabilities. ArcGIS Spatial Analyst takes a locational perspective in which each cell represents a location and the value associated with each cell identifies the type of phenomenon that is at each location. Functions are operators that are utilized throughout the ArcGIS Spatial Analyst tools to perform spatial analysis at different cell levels. The functions associated with raster-cell cartographic modeling can be divided into five types.

- Those that work on single cells (local functions)
- Those that work on cells within a neighborhood (focal functions)
- Those that work on cells with zones (zonal functions)
- Those that work on all cells within the raster (global functions)
- When combined in a series, those that perform a specific application (application functions)

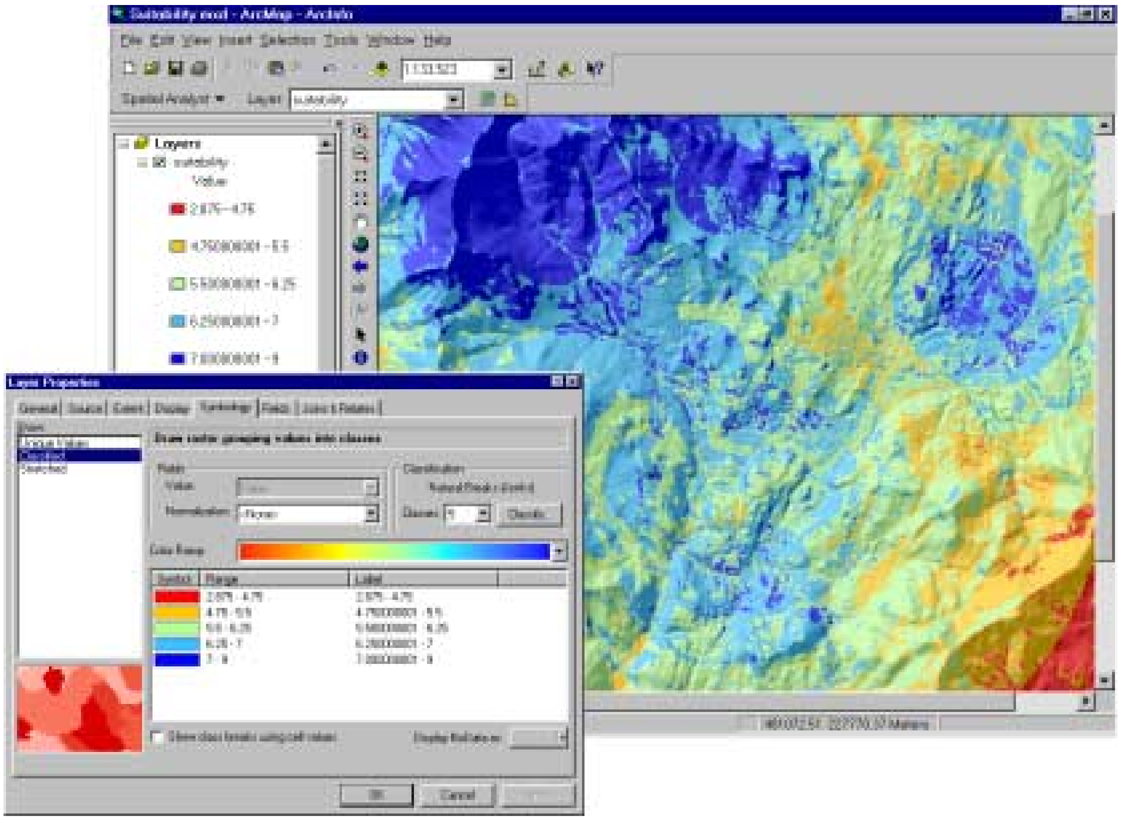
Data Integration

ArcGIS Spatial Analyst integrates the user's data enabling interaction between data of many different types. Images, elevation models, and other raster surfaces can be combined with computer-aided design (CAD) data, vector data, Internet data, and many other formats to provide an integrated analysis environment. For instance, regional or national maps showing the mean or the maximum precipitation for states or counties could be created by overlaying state or county boundary lines on a raster precipitation map.

ArcGIS Spatial Analyst can create raster data from any point, line, or polygon feature source such as ArcInfo coverages, shapefiles, geodatabases, CAD files, vector product format files, and ArcGIS themes created from tabular data. In addition, data in standard formats can be imported including TIFF, JPEG, BMP, USGS DEM, DTM, NIMA DTED, generic ASCII, MrSID™, and others.

Visualization

In addition to high-powered analysis and modeling, ArcGIS Spatial Analyst also allows analysts to visualize their data as never before. ArcGIS Spatial Analyst is integrated with ArcMap so that the user can create stunning visual displays with the powerful symbology and annotation options available.



Visualize the data with advanced symbology and annotation options.

Advanced Raster Calculations: Map Algebra

Map algebra is the analysis language for ArcGIS Spatial Analyst. In addition to the many functions that are available through the ArcGIS Spatial Analyst user interface, a wide variety of additional functions are available through map algebra such as hydrological modeling.

Not only does the algebra allow access to additional functions in the user interface, but it also allows the analyst to build more complex expressions and process them as a single command. For example, the user can calculate the mean between multiple rasters to assess overall changes of a geographical region.

Sophisticated Raster Data Analysis

ArcGIS Spatial Analyst provides a robust environment for advanced raster data analysis. This environment enables density mapping, distance analysis, surface analysis, advanced analysis with map algebra, grid statistics, spatial modeling, and surface creation. With so many features integrated, ArcGIS Spatial Analyst gives the analyst the ability to identify solutions to real-world problems in a dynamic mapping environment. Some of the various analytical tools include querying raster data, analyzing densities, deriving new raster data, and calculating the accumulated cost of travel.

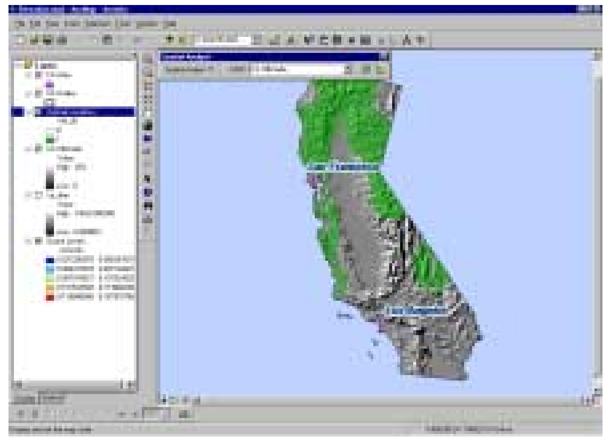
Query

A key component of ArcGIS Spatial Analyst is the ability to perform queries across different raster data sets in the raster calculator. This allows the analyst to ask questions that span multiple data types and levels of information (e.g., what areas are zoned for residential development and have a high water table on a steep slope greater than

J-8747

15 percent?). The query functionality gives the analyst the ability to leverage existing data and to make more informed decisions.

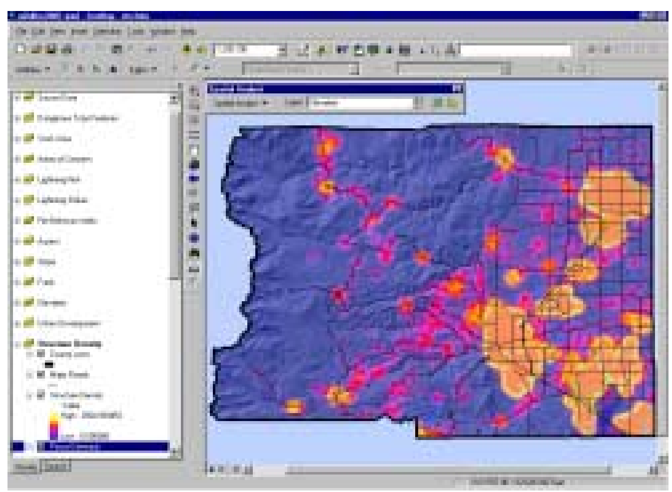
Query Raster Data Sets



Ozone Levels Less Than .094 ppm and Elevation Less Than 500 Feet

Density Analysis

By calculating density, the user can measure the quantity of an input feature data set (line or point) distributed throughout a landscape. A density value is calculated for each cell in the output raster. Density surfaces are good for showing where point or line features are concentrated. For example, an environmental planner may want to display building density to assess encroachment on areas designated as open space.



Building Density for Boulder County, Colorado

Deriving Raster Data Sets from Existing Maps

The user can convert any point, line, or polygon data set into a raster for advanced spatial analysis. Associated attributes can also be assigned to each individual cell for raster calculations and reclassifications. Once a vector data set, such as land use, is converted to a raster, the user can combine this data set with other raster data sets for a more accurate depiction of the spatial problem. For example, an analyst can combine a converted land use raster with an existing elevation raster to represent optimal suitability based on the type of land use and the elevation in the area.

Calculating Accumulated Travel Cost

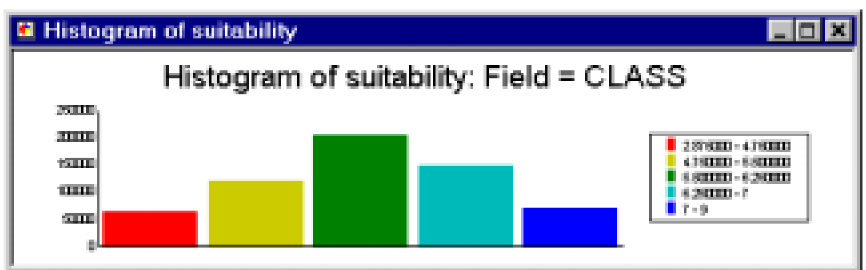
Calculating the accumulated cost of traveling can provide the user with a rich set of information from which to make decisions. For example, mobile defense units can calculate the most efficient path for deployment based on variables such as slope, elevation, vegetation density, and water bodies. Once all of these variables have been considered, the units can be deployed faster, utilize less fuel, and have a more timely response to foreign threats.

Cell Distributions and Statistics

ArcGIS Spatial Analyst provides the user with a sophisticated exploration environment for analyzing raster cell distributions and calculating cell statistics over time. By analyzing these components of the data, the user can accurately assess patterns that may be occurring throughout a landscape. These statistics and distributions can be utilized to foresee where possible problems may occur. For example, by using the cell-based statistics the user can visualize where a desert may be encroaching over a 10-year period. From this information a planner may choose an alternate zoning for that area.

Histograms

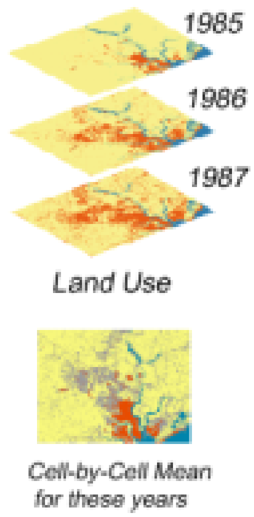
The distribution of information and the pattern within this information are often very important. Histograms have long been used to evaluate data and patterns. ArcGIS Spatial Analyst allows the user to create histograms from the raster grids either from selected features or from interactively defined graphic shapes to accurately identify dominant distributions throughout the data.



The histogram displays the number of cells for each value in the data.

Cell Statistics

It can be very important to visualize temporal changes over time to identify changes in crop yields, environmental contamination levels, changing landscape due to encroaching deserts, and many other factors that are essential for accurate analysis. All of these variables that play a role in the area in question can be displayed and analyzed over time utilizing cell statistics.



Cell-based statistics help identify temporal trends.

By computing cell statistics, the user can compute a statistic (such as majority, mean, maximum) for each cell in an output raster that is based on the values of each cell of multiple input rasters. The output raster can provide the user with information such as temporal trends that may be occurring in the data.

Neighborhood Statistics

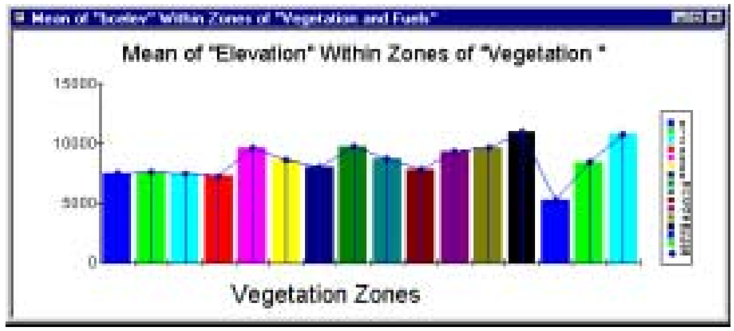
Neighborhood statistics are a focal function that computes an output raster in which the value at each location is a function of the input cells in some specified neighborhood of the location. Calculating neighborhood statistics is useful for obtaining a value for each cell based on a specified neighborhood. For example, when examining ecosystem stability, it might be useful to obtain the variety of species for each neighborhood to identify the locations that are lacking a variety of species.

Zonal Statistics

Zonal statistics calculate statistics for each zone of a zone data set, based on values from another data set. A zone is all the cells in a raster that have the same value, regardless of whether or not they are contiguous. Raster and feature data sets can be used as the "zone data set." So, for example, residential is a zone of a land use raster data set, or a roads feature data set can be the zone for an accident data set.

Zonal statistical functions perform operations on a per-zone basis; a single output value is computed for every zone in the input zone data set.

J-8747



Vegetation Types Within Zones of Elevation

Zonal statistics are used to calculate statistics such as the mean elevation for each vegetation zone. Alternatively, the user might want to know how many different types of vegetation there are in each elevation zone (variety).

Global Statistics

Global, or per-raster, functions compute an output raster data set in which the output value at each cell location is potentially a function of all the cells in the input raster data sets. There are two groups of global functions: Euclidean distance and weighted distance.

- Euclidean distance—Euclidean distance assigns its distance from the closest source cell (a source may be the location from which to start a new road) to each cell in the output raster data. The direction of the closest source cell can also be assigned as the value of each cell location in an additional output raster data set.

- Weighted Distance—By applying a global function to a weighted (cost) surface, the analyst can determine the cost of moving from a destination cell (the location where the user wishes to end the road) to the nearest source cell. The shortest path over a cost surface can be calculated over a nonnetworked surface from a source cell to a destination cell. In all the global calculations, knowledge of the entire surface is necessary to return the solution.

Terrain Analysis

With ArcGIS Spatial Analyst anyone can derive useful information such as a hillshade, contour slope, viewshed, or aspect map. These topographic surfaces give the user the power to relate his or her data to real-world elevations and analyze how these varied surfaces might affect the data in question. By combining the terrain maps with vector data, a more realistic depiction of the area is presented. The user can utilize the transparency feature in ArcMap to mold these data sets into one effective and highly visual map of the area. The user can then see where elevation and other terrain fluctuations may play a role in the spatial problem at hand. These terrain maps are outstanding for visualizing the area in question; however, the user can also combine these raster images with other variables for more advanced spatial analysis.

- Slope—Slope identifies the slope, or maximum rate of change, from each cell to its neighbors. An output slope raster data set can be calculated as either a percentage of slope (for example, 10% slope) or a degree of slope (for example, 45-degree slope).

- Aspect—Aspect identifies the steepest downslope direction from each cell to its neighbors. The value of the output raster data set represents the compass direction of the aspect: "0" is true north, a 90-degree aspect is to the east, and so forth.

- Hillshade—Hillshade is used to determine the hypothetical illumination of a surface for either analysis or graphical display. For analysis, hillshade can be used to determine the length of time and intensity of the sun in a given location. For graphical display, hillshade can greatly enhance the relief of a surface.

- Viewshed—Viewshed identifies either how many of the observation points specified on the input observation raster data set can be seen from each cell or which cell locations can be seen from each observation point.

- Contour—The contour feature produces an output polyline data set. The value of each line represents all contiguous locations with the same height, magnitude, or concentration of whatever the values on the input data set represent. The function does not connect cell centers; it interpolates a line that represents locations with the same magnitude. Creating contours is an effective way to identify which locations have the same value. Contours are also useful for surface representation because they allow the analyst to simultaneously visualize flat and steep areas for analyzing the distance between contours. The user can also identify ridges and valleys.

- Curvature (directional flow of angles)—Curvature measures the slope of the surface at each cell. It calculates the second derivative of the input-surface raster data set—the slope of the slope. The result of the curvature function can be used to describe the physical characteristics of a surface such as the erosion and runoff processes within a landscape. The slope identifies the overall rate of downward movement, and aspect defines the direction of flow. The profile curvature is the shape of the surface in the direction of the slope. The platform curvature defines the shape of the surface perpendicular to the direction of the slope.

Spatial Relationships

Spatial Modeling

ArcGIS Spatial Analyst provides the ability to create sophisticated spatial models for many different geospatial problems. ArcGIS Spatial Analyst utilizes process models to attempt to describe the interaction of the objects that are modeled in the representation model. The relationships are modeled using spatial analysis tools. Since there are many different types of interactions between objects, ArcGIS and ArcGIS Spatial Analyst provide a large suite of tools to describe interactions.

Representation modeling is sometimes referred to as data models, which are considered descriptive models. Process modeling is sometimes referred to as cartographic modeling. Process models can be used to describe processes, but they are often used to predict what will happen if some action occurs. Some process models include

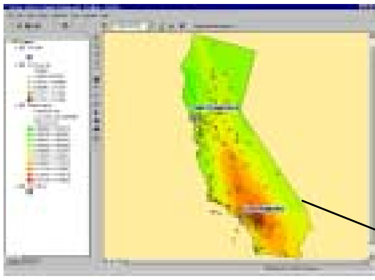
- Suitability modeling—Most spatial models involve finding optimum locations such as finding the best location to build a new school, landfill, or resettlement site.

- Distance modeling—What is the shortest flight distance from Los Angeles to San Francisco?

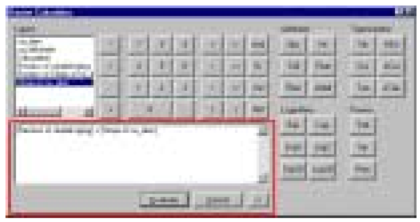
- Hydrologic modeling—Where will the water flow to?

- Surface modeling—What is the ozone pollution level for various locations in a county?

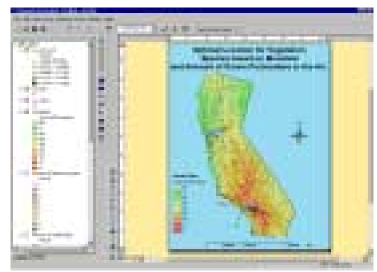
One of the most basic ArcGIS Spatial Analyst operations is adding two rasters together to create a more efficient analysis of the area in question. For example, by combining a land use grid and a slope grid and reclassifying these rasters with levels of importance, the analyst can effectively analyze areas that will have the greatest suitability for site location.



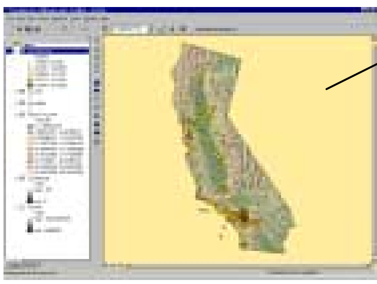
Sophisticated Surface



Raster Calculator



Suitability Analysis



Topographic Data

Suitability Modeling

Calculate optimal site locations by identifying possible influential factors; creating new data sets from existing data (e.g., slope, aspect); reclassifying the data to identify areas with high suitability; and, finally, aggregating these into one logical assessment of optimal suitability. This optimal suitability map may provide a project manager with new insight into the ideal areas where a new site should be located.

Distance Modeling

Determine the least expensive method for a new road, flight pattern, shipping route, or any factor that is affected by time and cost.

By mapping distance, the analyst can find out information such as the distance to the nearest hospital from certain areas for an emergency helicopter or all fire hydrants within 500 meters of a burning building. Alternatively, the analyst can find the shortest (or least-cost) path from one location to another based on some cost factor.

Distance Mapping Functions

The distance mapping functions are *global functions*. They compute an output raster data set where the output value at each location is potentially a function of all the cells in the input raster data sets.

There are several distance mapping tools for measuring both straight-line (Euclidean) distance and distance measured in terms of other factors such as the cost to travel over the landscape. The outputs from the Straight Line Distance functions are normally used directly, while the outputs from the Cost Weighted Distance functions are most commonly used to compute shortest (or least-cost) paths.

Straight Line Distance Functions

The *Straight Line Distance* function measures the straight line distance from each cell to the closest source (the source identifies the objects of interest such as a well, road, or school). The distance is measured from cell center to cell center.

The *Straight Line Allocation* function assigns each cell the value of the source to which it is closest. The nearest source is determined by the Straight Line Distance.

The *Straight Line Direction* function computes the direction to the nearest source, measured in degrees.

Cost Weighted Distance Functions

The *Cost Weighted Distance* function modifies the Straight Line Distance by some other factor that is the cost to travel through any given cell. For example, it may be shorter to climb over the mountain to the destination, but it is faster to walk around it.

The *Cost Weighted Allocation* function identifies the nearest source cell based on accumulated travel cost.

The *Cost Weighted Direction* function provides a road map identifying the route to take from any cell along the least-cost path and back to the nearest source.

The distance and direction raster data sets are normally created to serve as inputs to the pathfinding function, the *shortest* (or least-cost) *path*.

Shortest Path (Path Analysis)

The Shortest Path function determines the path from a destination point to a source. Once the user has performed the Cost Weighted Distance function and created distance and direction rasters, he/she can then compute the least-cost (or shortest) path from a chosen destination to the source point. The shortest path function can be used for analyses such as finding the best route for a new road based on construction costs.

Hydrologic Modeling

The shape of a surface determines how water will flow across it. The hydrologic modeling functions provide methods for describing the hydrologic characteristics of a surface. Using an elevation raster data set as input, it is possible to model where water will flow, create watersheds and stream networks, and derive other hydrologic characteristics.

Surface Modeling

Typically, it is not possible or economically feasible to collect data points for every value within the area of interest. Therefore, an accurate continuous surface creation is a must for predicting these values. ArcGIS Spatial Analyst introduces a set of spatial interpolation functions, allowing the user to generate results for areas of missing data. For example, the analyst can use global positioning system points to interpolate an elevation surface.

ArcGIS Spatial Analyst includes the following interpolation methods:

- Spline—Spline estimates values using a mathematical function that minimizes overall surface curvature, resulting in a smooth surface that passes exactly through the input points.

- Inverse Distance Weighted—Inverse Distance Weighted estimates cell values by averaging the values of sample data points in the vicinity of each cell. The closer a point is to the center of the cell being estimated, the more influence, or weight, it has in the averaging process. This method assumes that the variable being mapped decreases in influence with distance from its sampled location.

- Basic Kriging (Ordinary, Universal)—Kriging is based on statistical models that include autocorrelation (the statistical relationship among the measured points). The two methods included can be used to produce a prediction surface as well as identify the certainty or accuracy of the predictions.

ArcGIS Spatial Analyst also provides the ability to generate a density map across an area where the value of each cell is the result of a units-per-specified-area calculation. This could be population density per square mile or grasshopper infestations per square kilometer. Density maps can be used as weights for modeling, such as business models or pesticide models, to best make use of limited resources.

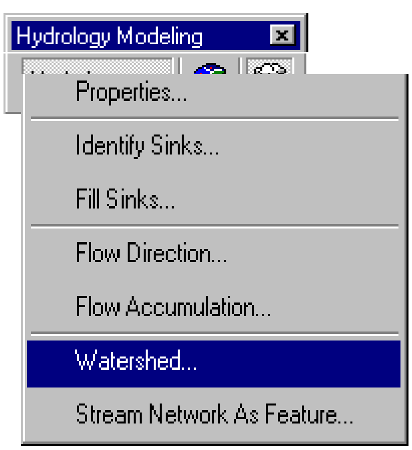
Customization

ArcGIS Spatial Analyst gives the analyst the ability to create custom tools for spatial modeling and incorporate these tools directly into the ArcGIS interface with any COM-compliant development environment. With the advanced customization tools, the user can create advanced spatial models for his/her specific spatial analysis.

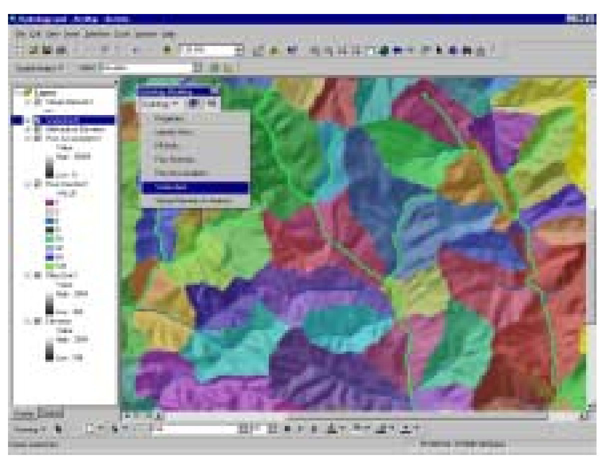
New customization never before seen in ArcGIS Spatial Analyst is as follows:

- Customize the interface with the drag-and-drop and menu-driven tools.
- Create custom models and user interfaces.
- Add custom analysis functions.
- Use custom .dll or .exe files.
- Support new formats.

For more technical information, downloads, and developer scripts, please visit ArcObjects™ online at http://arconline.esri.com/arcobjectsonline.



Add custom tools such as hydrology modeling.



J-8747

Conclusion

ArcGIS Spatial Analyst provides users with the freedom to analyze various spatial problems within their specific industry. ArcGIS Spatial Analyst is integrated with the ArcGIS Desktop products (ArcInfo, ArcEditor, ArcView) for interoperability and advanced symbology and mapping capabilities. ArcGIS Spatial Analyst gives the analyst the ability to derive new information from existing data, query information across multiple data layers, and fully integrate cell-based raster data with traditional vector data sources. From identifying areas of suitability for a new subdivision to creating a continuous surface from ozone pollution data measurements, ArcGIS Spatial Analyst has a variety of analytical modeling tools for many spatial problems.