Chapter 12 – Query, measurement, and transformation

§ 1 Introduction

§ 2 Queries

§ 3 Transformations

Learning Objectives

- After reading this chapter you will be able to:
  - Definitions of geographic analysis and modeling, and tests to determine whether a method is geographic
  - The range of queries possible with a GIS
  - Methods for measuring length, area, shape, and other properties, and their caveats
  - Transformations that manipulate objects to create new ones, or to determine geometric relationships between objects

§ 1 Introduction

- Spatial analysis can reveal things that might otherwise be invisible – it can make what is implicit explicit
- Spatial analysis is the crux of GIS, the means of adding value to geographic data, and of turning data into useful information
- Effective spatial analysis requires an intelligent user, not just a powerful computer
- Spatial analysis helps us in situations when our eyes might otherwise deceive us.
- Spatial analysis is a set of methods whose results change when the locations of the objects being analyzed change
  - calculation of an average income vs calculation of the center of the US population

§ 1 Introduction

- **spatial analysis** The process of modeling, examining, and interpreting model results
- **Spatial analysis** is useful for evaluating suitability and capability, for estimating and predicting, and for interpreting and understanding

A conceptual model for solving spatial problems

Step 1: Stating the problem

What is your goal?

Step 2: Breaking the problem down

What are the objectives to reach your goal?

What are the phenomena and interactions (process models) necessary to model?

What datasets will be needed?

Step 3: Exploring input datasets

What is contained within your datasets?

What relationships can be identified?

Step 4: Performing analysis

Which GIS tools will you use to run the individual process models and build the overall model?

Step 5: Verifying the model's result

Do certain criteria in the overall model need changing?

If Yes—go back to step 4.

Step 6: Implementing the result

Description and Analysis

- Most GIS systems are acquired by large organizations for the purpose of *representing and describing* features of the real world which are relevant to that organization's mission
  - Spatial databases perform this function
  - This is what we have discussed so far
    - Coordinate systems as fundamental properties of spatial data
    - Points, lines, polygons concepts for representation
    - geographic file formats for storage
  - Most GIS system capabilities are focused here

Description and Analysis

- *Analysis* involves gaining an understanding of the "cause and effect" underlying the features which have been described in order to
  - Help the organization better carry out its mission
    - Make better decisions, for example
  - Understand the phenomena as a goal in itself
    - This is the role of science
  - GIS systems are less capable here, and often must be supplemented

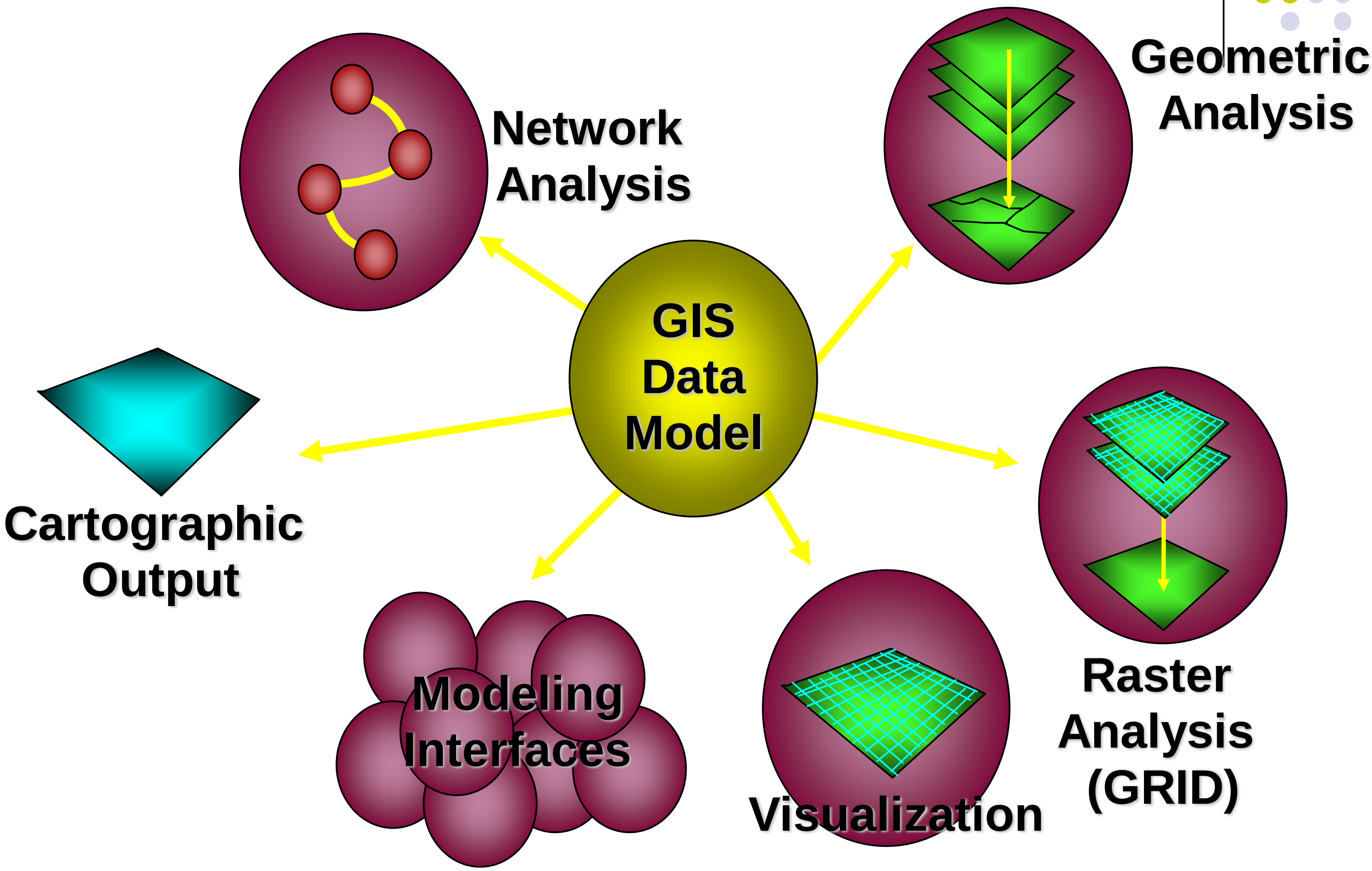
Process, Pattern and Analysis

- *Processes* operating in space produce *patterns*
- *Spatial Analysis* is aimed at:
  - Identifying and describing the pattern
  - Identifying and understanding the process

Spatial Analysis: ***successive levels of sophistication***

- Spatial data manipulation: classic GIS capabilities
  - Spatial queries & measurement, buffering, map layer overlay
- Spatial data analysis: descriptive and exploratory
  - Visualization through data manipulation and mapping
    - John Snow's maps of cholera in 1850s London
- Spatial statistical analysis: hypothesis testing
  - Are data "to be expected" or are they "unexpected" relative to some statistical model, usually of a random process
- Spatial modeling: prediction
  - Constructing models (of processes) to predict spatial outcomes (patterns)
  - *What if* analyses

Spatial Analysis



The Pitfalls of Spatial Analysis

- Spatial autocorrelation
  - Data from location near to each other are more likely to be similar than data from location remote from each other
  - Causes serious problems with traditional statistical models
    - ***Spatial*** statistical models are essential
- Modifiable areal unit problem (MAUP)
  - Results may depend on the areal unit used
    - Census tracts versus counties (scale issue)
    - Census tracts versus zip codes (not a scale issue)
- Ecological fallacy
  - Results obtained from aggregated data (e.g. census tracts) cannot be assumed to apply to individual people
  - A special case of the MAUP problem
  - Encountered in spatial and non-spatial analysis
- *Scale* affects representation and results
  - Cities may be points or polygons
  - MAUP may be viewed as a scale issue
- Nonuniformity of Space and Edge Issues
  - Phenomena is not distributed evenly in space
    - Bank robberies cluster ‘cos banks are clustered in space
  - Edges, beyond which there is no data, can significantly effect results

Fundamental Spatial Concepts

- Distance
  - The magnitude of spatial separation
  - Euclidean (straight line) distance often only an approximation
- Adjacency
  - Nominal or binary equivalent of distance
  - Levels of adjacency exist: 1^{st}, 2^{nd}, 3^{rd} nearest neighbor, etc..
- Interaction
  - The strength of the relationship between entities
  - An inverse function of distance
- Neighborhood
  - An association between one entity and those around it
  - May be based upon
    - Interaction: flows or connections (functional)
    - Similarity of attributes (formal)

§ 1 Introduction(cont.)

§ 1.1 Examples

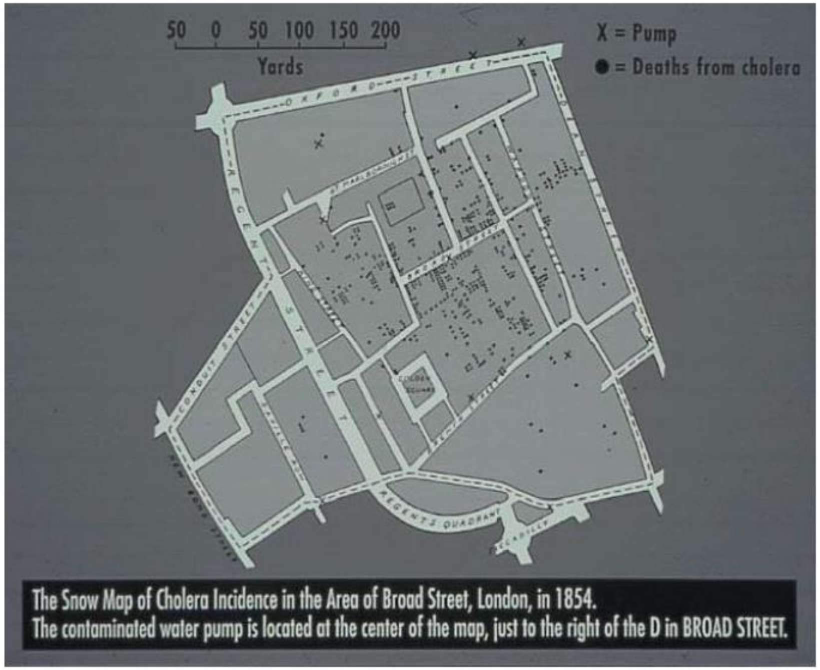


Figure 14.1 A redrafting of the map made by Dr John Snow in 1854 of the deaths that occurred in an outbreak of cholera in the Soho district of London. The existence of a public water pump in the center of the outbreak (the cross in Broad Street) convinced Snow that drinking water was the probable cause of the outbreak. Stronger evidence was obtained in support of this hypothesis when the water supply was cut off, and the outbreak subsided. (*Source*: Gilbert E. W. 1958 'Pioneer maps of health and disease in England.' *Geographical Journal*, **124**: 172–183) (Reproduced by permission of Blackwell Publishing Ltd.)

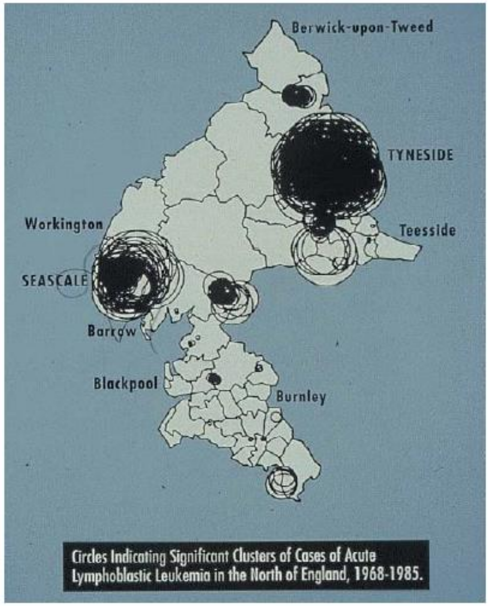


Figure 14.4 The map made by Openshaw and colleagues by applying their Geographical Analysis Machine to the incidence of childhood leukemia in northern England. A very large number of circles of random sizes is randomly placed on the map, and a circle is drawn if the number of cases it encloses substantially exceeds the number expected in that area given the size of its population at risk (*Source*: Openshaw S., Charlton M., Wymer C., and Craft A. 1987 'A Mark I geographical analysis machine for the automated analysis of point datasets.' *International Journal of Geographical Information Systems*, **1**: 335–358. www.tandf.co.uk/journals)

§ 1 Introduction(cont.)

§ 1.2 Types of spatial analysis and modeling

- *Queries* are the most basic of analysis operations, in which the GIS is used to answer simple questions posed by the user. No changes occur in the database, and no new data are produced. The perations vary from simple and well-defined queries like ‘how many houses are found within 1 km of this point’, to vaguer questions like ‘which is the closest city to Los Angeles going north’,where the response may depend on the system’s ability to understand what the user means by ‘going north’
- *Measurements* are simple numerical values that describe aspects of geographic data. They include measurement of simple properties of objects, like length, area, or shape, and of the relationships between pairs of objects, like distance or direction

§ 1 Introduction(cont.)

§ 1.2 Types of spatial analysis and modeling

- *Transformations* are simple methods of spatial analysis that change datasets, combining them or comparing them to obtain new datasets, and eventually new insights. Transformations use simple geometric, arithmetic, or logical rules, and they include operations that convert raster data into vector data, or vice versa.They may also create fields from collections of objects, or detect collections of objects in fields
- *Descriptive summaries* attempt to capture the essence of a dataset in one or two numbers. They are the spatial equivalent of the descriptive statistics commonly used in statistical analysis, including the mean and standard deviation

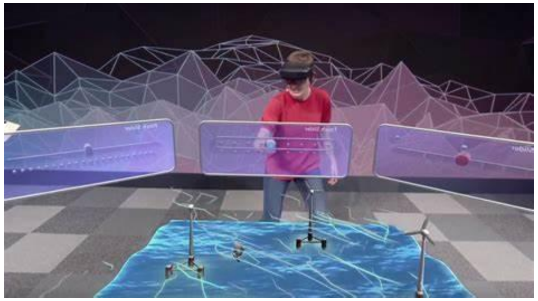
§ 1 Introduction(cont.)

§ 1.2 Types of spatial analysis and modeling

- *Optimization* techniques are normative in nature,designed to select ideal locations for objects given certain well-defined criteria. They are widely used in market research, in the package delivery industry, and in a host of other applications.
- *Hypothesis testing* focuses on the process of reasoning from the results of a limited sample to make generalizations about an entire population. It allows us, for example, to determine whether a pattern of points could have arisen by chance, based on the information from a sample. Hypothesis testing is the basis of inferential statistics and lies at the core of statistical analysis, but its use with spatial data is much more problematic

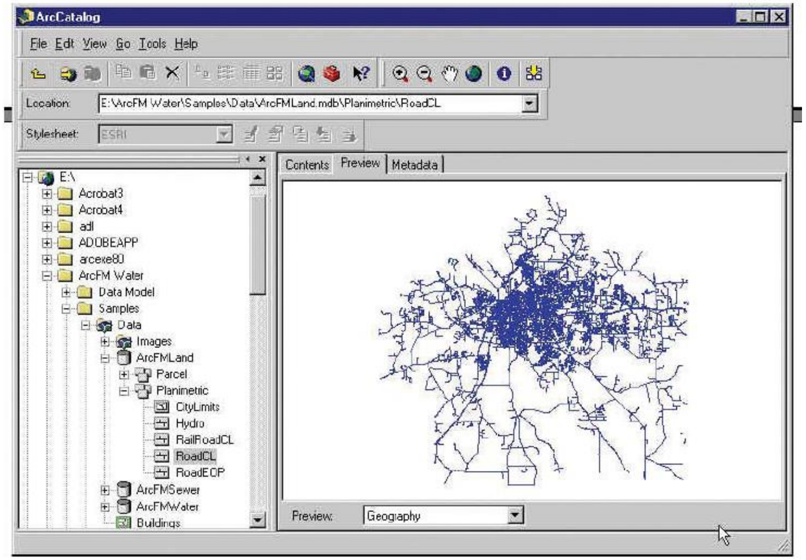
§ 2 Queries

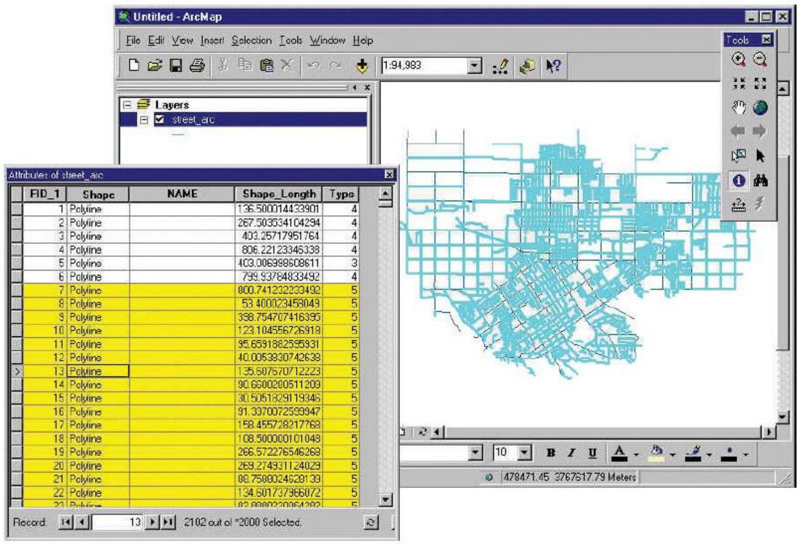
- In the ideal GIS it should be possible for the user to interrogate the system about any aspect of its contents, and obtain an immediate answer
- Today's user interfaces are very versatile, and have very nearly reached the point where it will be possible to interrogate the system by speaking to it



§ 2 Queries(cont.)

- Users query a GIS database by interacting with different views
  - Catalog, map, table
- SQL is a standard language for querying tables and relational databases





§ 2 Queries(cont.)

- Exploratory spatial data analysis allows its users to gain insight by interacting with dynamically linked views

Figure 14.8 Screen shot of GeoDa, a spatial analysis package that integrates easily with a GIS and features the ability to display multiple views (clockwise from the top left a map, a scatterplot, a map showing local indicators of spatial association (LISA; Box 15.3), and a histogram – see Section 14.2 for a discussion of histograms and scatterplots) and to link them dynamically. The tracts selected by the user in the upper left are automatically highlighted in the other windows. GeoDa is available at **www.csiss.org/clearinghouse/GeoDa**. (*Source*: courtesy of Luc Anselin, University of Illinois, Urbana-Champaign)

§ 3 Measurements

- Area

$$A = 0.5\sum_{i=1}^{N-1}\left[(x_2 - x_1)(y_1 + y_2)\right]$$

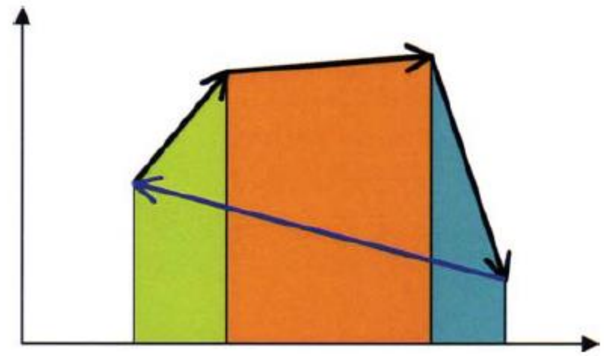


Figure 14.9 The algorithm for calculation of the area of a polygon given the coordinates of the polygon's vertices. The polygon consists of the three black arrows, plus the blue arrow forming the fourth side. Trapezia are dropped from each edge to the x axis and their areas are calculated as (difference in x) times average of y. The trapezia for the first three edges, shown in green, brown, and blue, are summed. When the fourth trapezium is formed from the blue arrow its area is negative because its start point has a larger x than its end point. When this area is subtracted from the total the result is the correct area of the polygon

§ 3 Measurements(cont.)

- Length or Distance

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

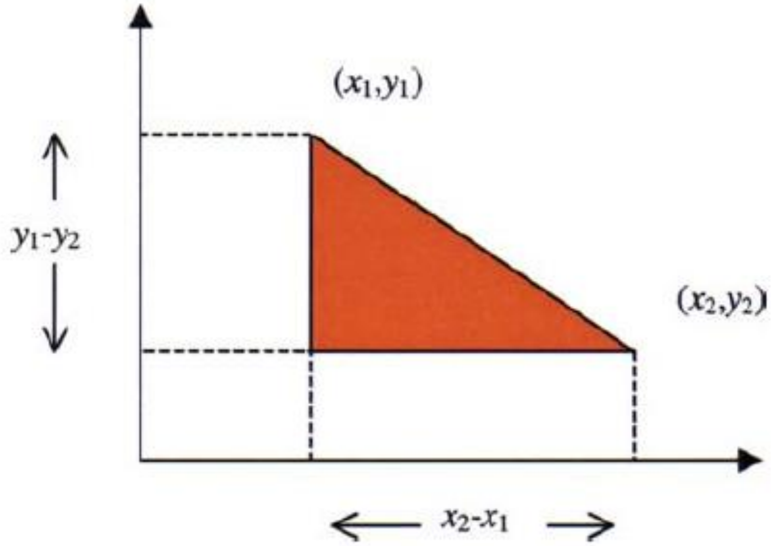
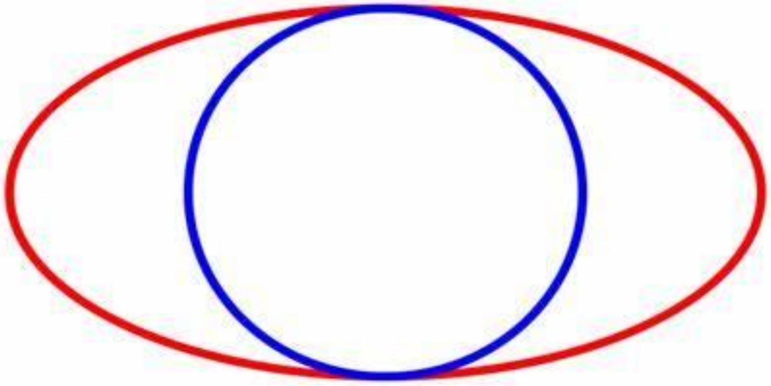


Figure 14.10 Pythagoras's Theorem and the straight-line distance between two points on a plane. The square of the length of the hypotenuse is equal to the sum of the squares of the lengths of the other two sides of the right-angled triangle

§ 3 Measurements(cont.)

- Shape

$$S = P/3.54\sqrt{A}$$



§ 3 Measurements(cont.)

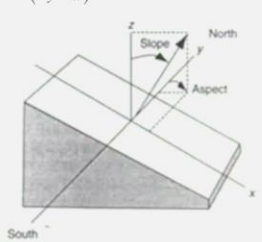
- Slope and aspect
  - The digital elevation model is the most useful representation of terrain in a GIS
  - The spatial resolution used to calculate slope and aspect should always be specified
  - Slope and aspect are the basis for many interesting and useful forms of analysis, especially in terrain analysis

Computing Algorithms for Slope and Aspect Using Grid

We can measure the slope and aspect for a cell in the elevation grid by the quantity and direction of the tilt of the cell's normal vector. Given a normal vector (n_x, n_y, n_z).

$$s = \left(n_x^2 + n_y^2{}^{0.5}\right)/n_z$$

$$A = \arctan\left(n_y/n_x\right)$$



The normal vector to the cell is the directed line perpendicular to the cell. The quantity and direction of tilt of the normal vector determine the slope and aspect of the cell.

Shows Normal Vector Determine the Slope and Aspect of the Cell

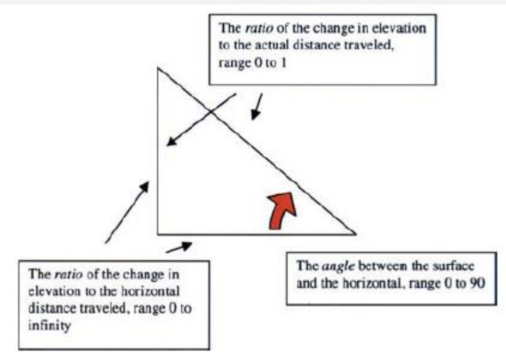


Figure 14.16 Three alternative definitions of slope. To avoid ambiguity we use the angle, which varies between 0 and 90 degrees

§ 4 Transformation

- methods that transform GIS objects and databases into more useful products, using simple rules
- These operations form the basis for many applications, because they are capable of revealing aspects that are not immediately visible or obvious

§ 4 Transformation(cont.)

§ 4.1 Buffer

- One method of proximity analysis
- Given any set of objects, which may include points, lines, or areas, a buffer operation builds a new object or objects by identifying all areas that are within a certain specified distance of the original objects

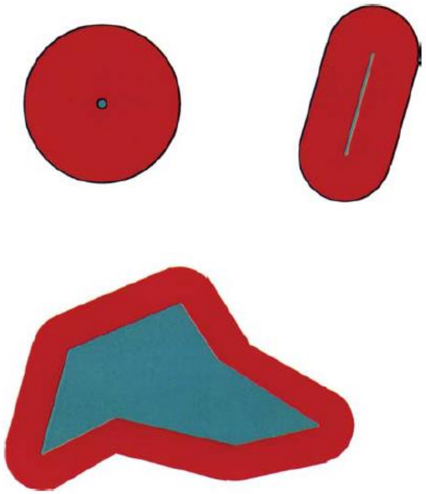
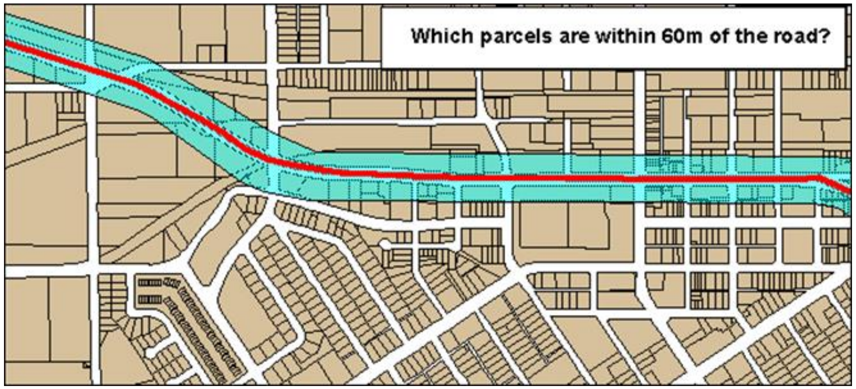


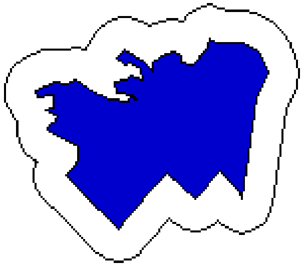
Figure 14.17 Buffers (dilations) of constant width drawn around a point, a polyline, and a polygon

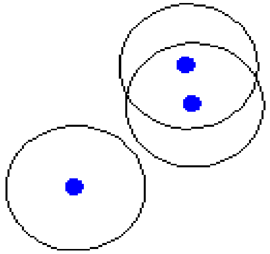


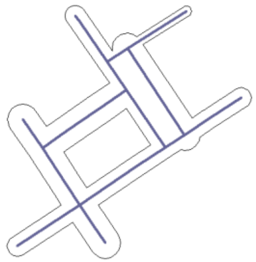
§ 4 Transformation(cont.)

§ 4.1 Buffer(cont.)

- A buffer is a map feature that represents a uniform distance around a feature
  - When creating a buffer, the user selects the feature to buffer from, as well as the distance to be buffered.
- The buffer operation creates a new polygon data set, where a specified distance is drawn around specific features within a layer.
  - The distances can either be constant or can vary depending upon attribute values
  - When features are close together, their buffers may overlap. The user can choose to preserve the overlaps or remove them
  - Answer questions based on proximity







§ 4 Transformation(cont.)

§ 4.1 Buffer(cont.)

- Examples:
  - permission to rebuild – the local planning authority could build a buffer around the parcel, in order to identify all homeowners who live within the legally mandated distance for notification of proposed redevelopments
  - A logging company wishes to clearcut an area, but is required to avoid cutting in areas within 100 m of streams – the company could build buffers 100 mwide around all streams to identify these rotected riparian areas
  - A retailer is considering developing a new store on a site of a type that is able to draw consumers from up to 4 km away from its stores – the retailer could build a buffer around the site to identify the number of consumers living within 4 km of the site, in order to estimate the new store's potential sales

Vector Distance Operation: Buffers & Setbacks

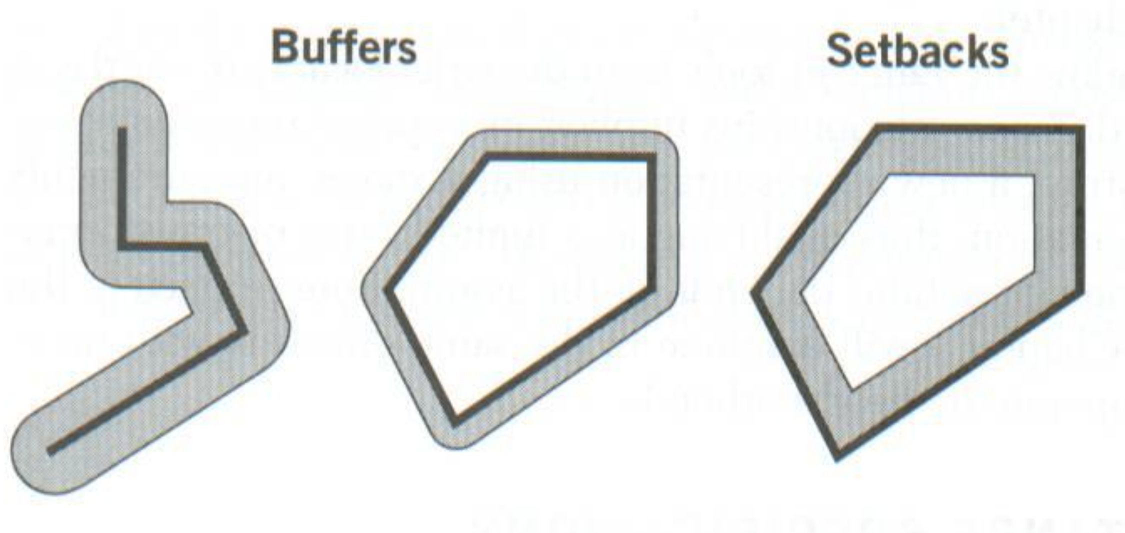
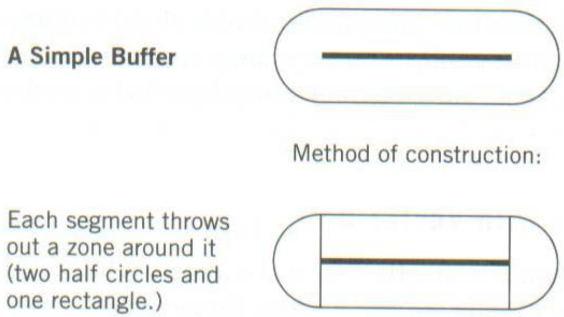


Diagram of simple buffers and a setback.
NOTE: buffers go outward from lines or areas; setbacks run inside of areas (not lines).

Image Source: Chrisman, Nicholas.(2002). 2nd Ed. *Exploring Geographic Information Systems*. p 154. fig. 6-1.

Buffer Creation: Illustrated



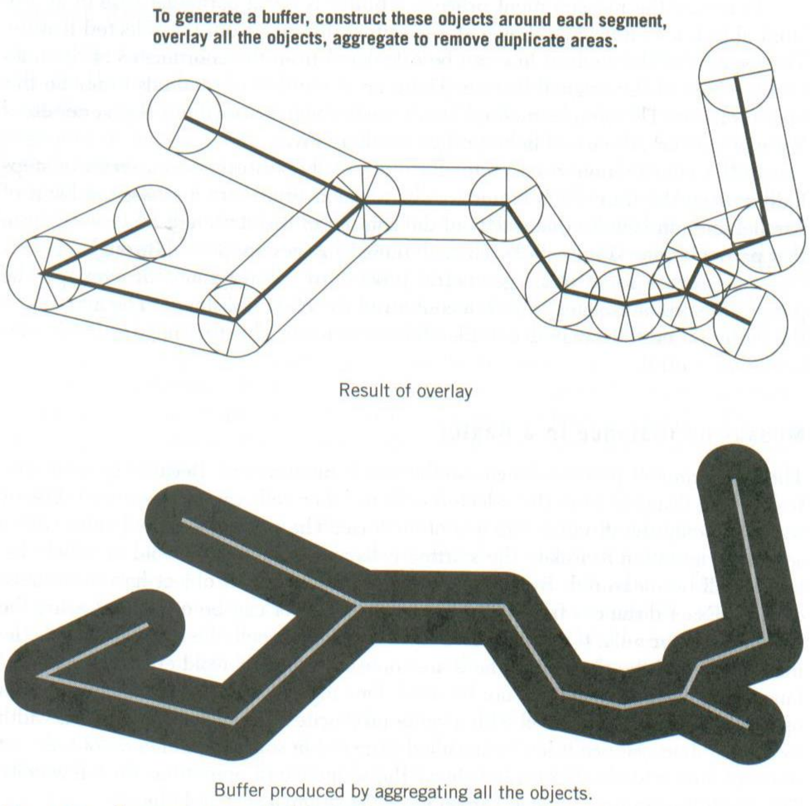
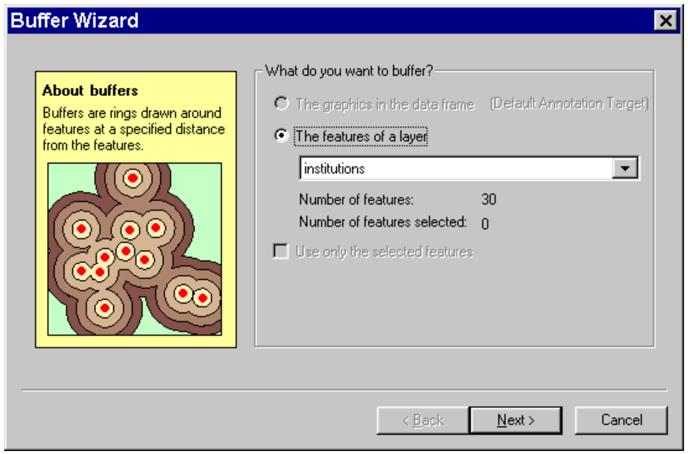
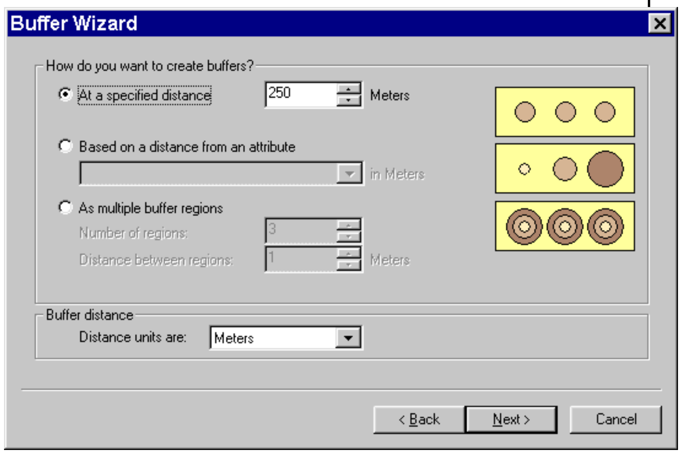


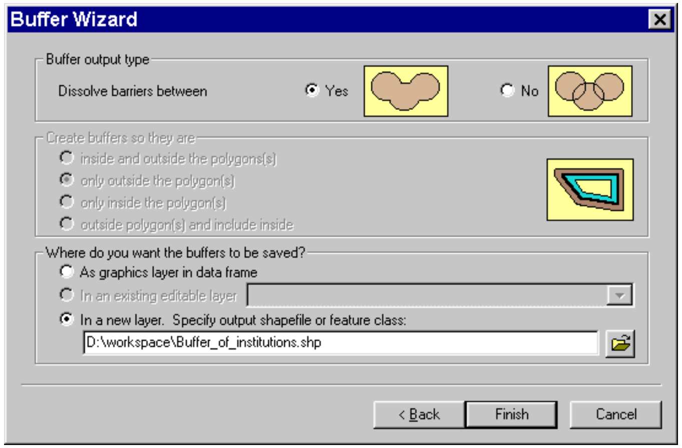
Figure 6-3 Construction method for buffers in a vector representation.

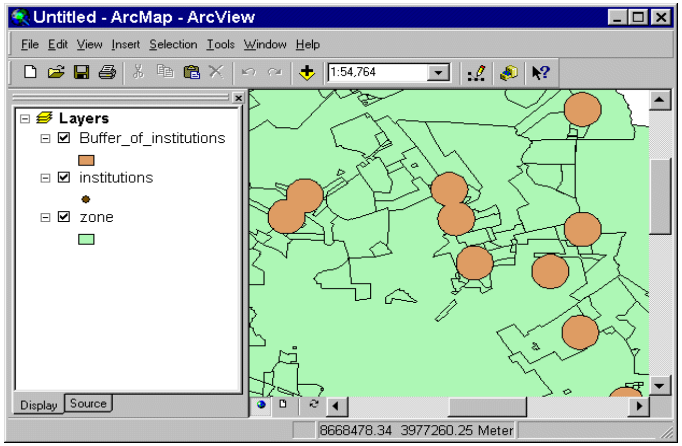
Image Source: Chrisman, Nicholas.(2002). 2nd Ed. *Exploring Geographic Information Systems*. p 60. fig. 6-3.

§ 4 Transformation(cont.)









§ 4 Transformation(cont.)

§ 4.2 Point in Polygon(PiP)

- In its simplest form, the point in polygon operation determines whether a given point lies inside or outside a given polygon
- In more elaborate forms, there may be many polygons, and many points, and the task is to assign points to polygons
- If the polygons overlap, it is possible that a given point lies in one, many, or no polygons, depending on its location

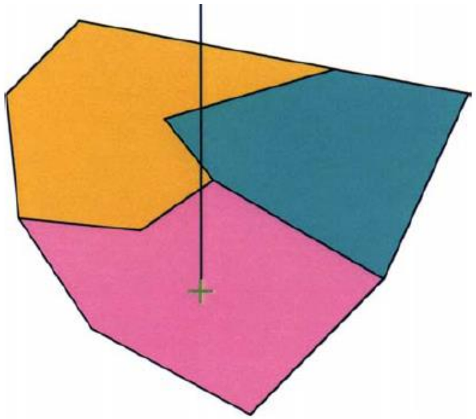
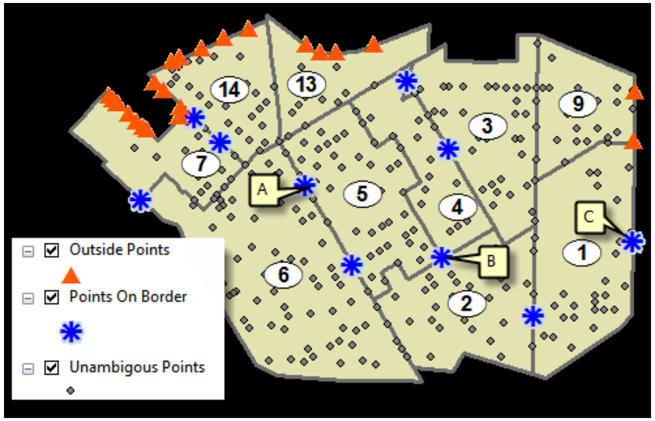


Figure 14.19 The point in polygon problem, shown in the continuous-field case (the point must by definition lie in exactly one polygon, or outside the project area). In only one instance (the pink polygon) is there an odd number of intersections between the polygon boundary and a line drawn vertically upward from the point

§ 4 Transformation(cont.)

§ 4.2 Point in Polygon(cont.)

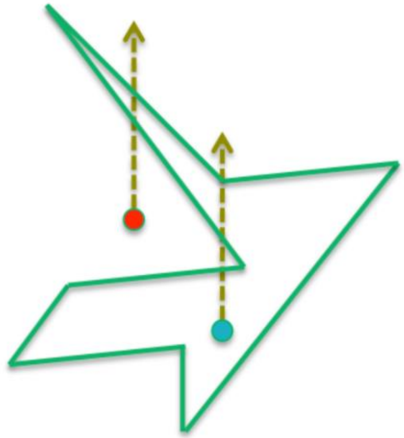
- Examples:
  - The points represent instances of a disease in a population, and the polygons represent reporting zones such as counties – the task is to determine how many instances of the disease occurred in each zone
  - The points represent the locations of transmission-line poles owned by a utility company, and the polygons are parcels of land – the task is to identify the owner of the land on which each pole lies, to verify that the company has the necessary easements and pays the necessary fees
  - The points represent the residential locations of voters, and the polygons represent voting districts – the task is to ensure that each voter receives the correct voting forms in the mail

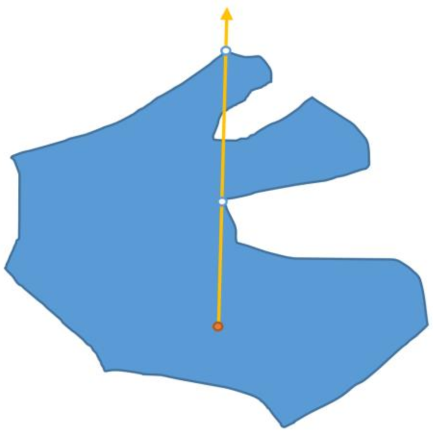


§ 4 Transformation(cont.)

§ 4.2 Point in Polygon(cont.)

- The standard algorithm for the point in polygon operation
  - it consists of drawing a line from the point to infinity, in this case parallel to the y axis
  - determining the number of intersections between the line and the polygon’s boundary
  - If the number is odd the point is inside the polygon and if it is even the point is outside
  - The algorithm must deal successfully with special cases – for example, if the point lies directly below a vertex (corner point) of the polygon





§ 4 Transformation(cont.)

§ 4.3 Polygon overlay

- One method of spatial join
- Polygon overlay is similar to point in polygon transformation in the sense that two sets of objects are involved, but in this case there is polygon involved
- It exists in two forms, depending on whether a continuous-field or discreteobjectperspective is taken
- The development of effective algorithms for polygon overlay was one of the most significant challenges of early GIS and the task remains one of the most complex and difficult to program

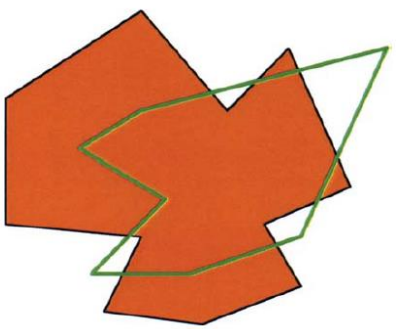


Figure 14.20 Polygon overlay, in the discrete-object case. Here, the overlay of two polygons produces nine distinct polygons. One has the properties of both polygons, four have the properties of the brown polygon but not the green polygon, and four are outside the brown polygon but inside the green polygon

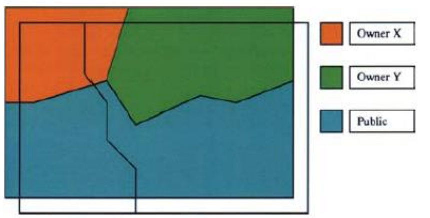


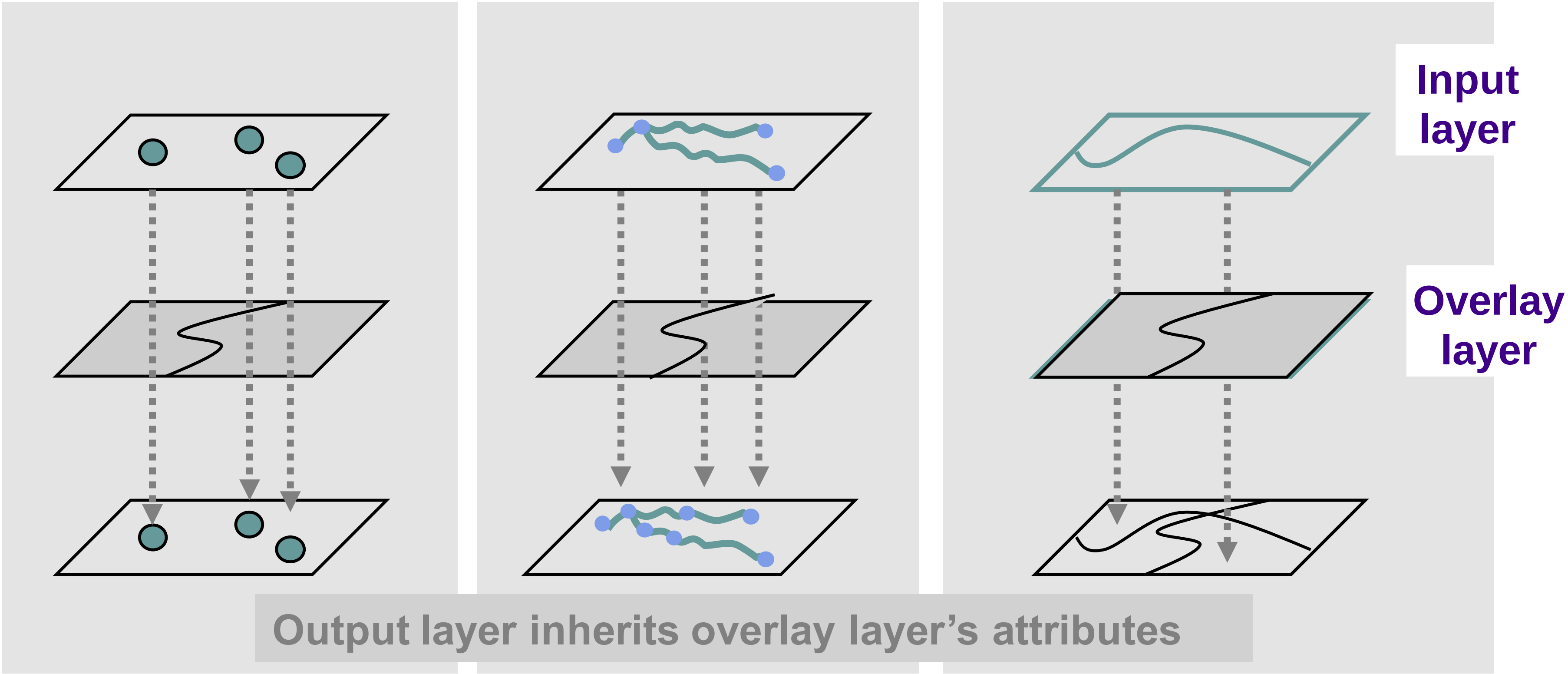
Figure 14.21 Polygon overlay in the continuous-field case. Here, a dataset representing two types of land cover (A on the left, B on the right) is overlaid on a dataset representing three types of ownership (the two datasets have been offset slightly for clarity). The result will be a single dataset in which every point is identified with one land cover type and one ownership type. It will have five polygons, since land cover A intersects with two ownership types and land cover B intersects with three

Overlay

- A common spatial analysis tool
- Overlays allow the user to identify areas where features in two layers overlap.
- A new data set is often created from these overlaps.
  - In a ***Union Overlay***, all features are included in the new data set but the features that overlap represent a new feature.
  - In an ***Intersect Overlay***, only the areas that overlap are contained in the new data set.

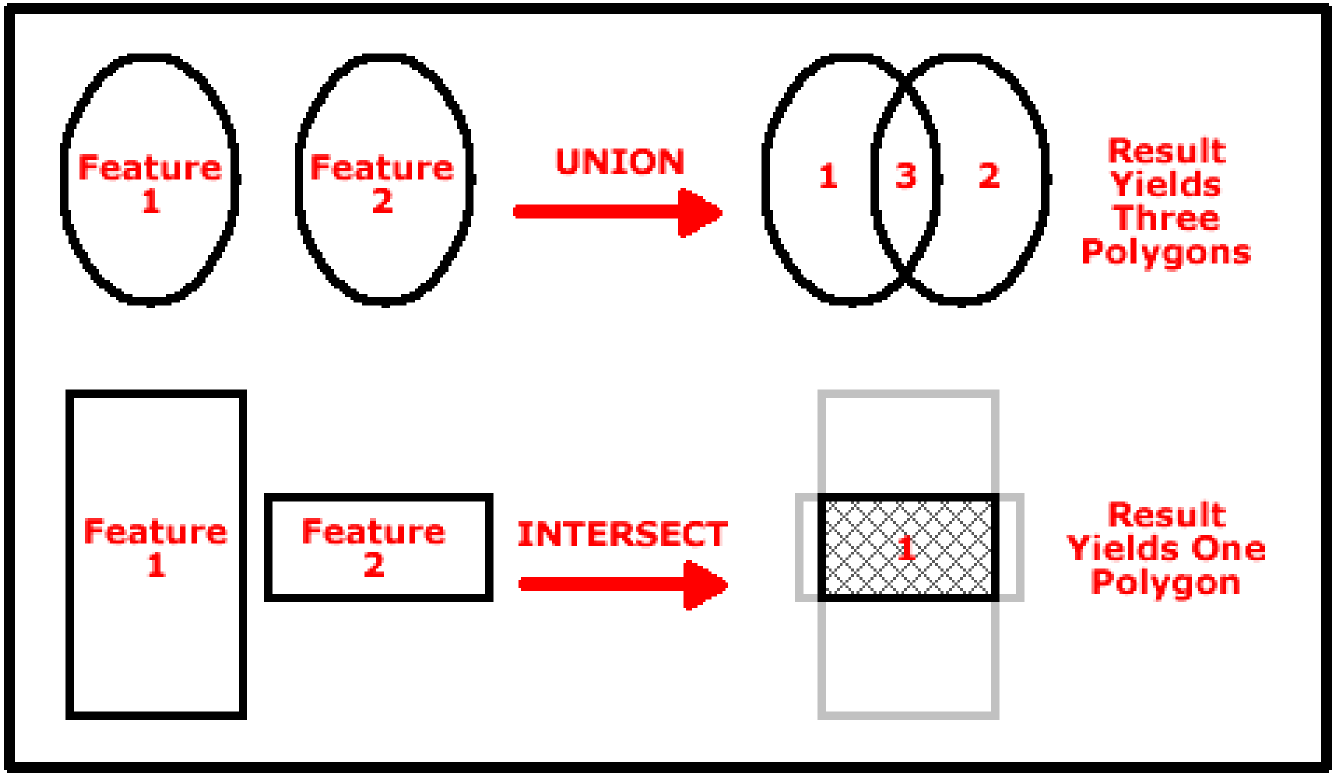
Overlay Analysis and Geoprocessing

Point-in-polygon Line-in-polygon Polygon-on-polygon



- Different from select by location

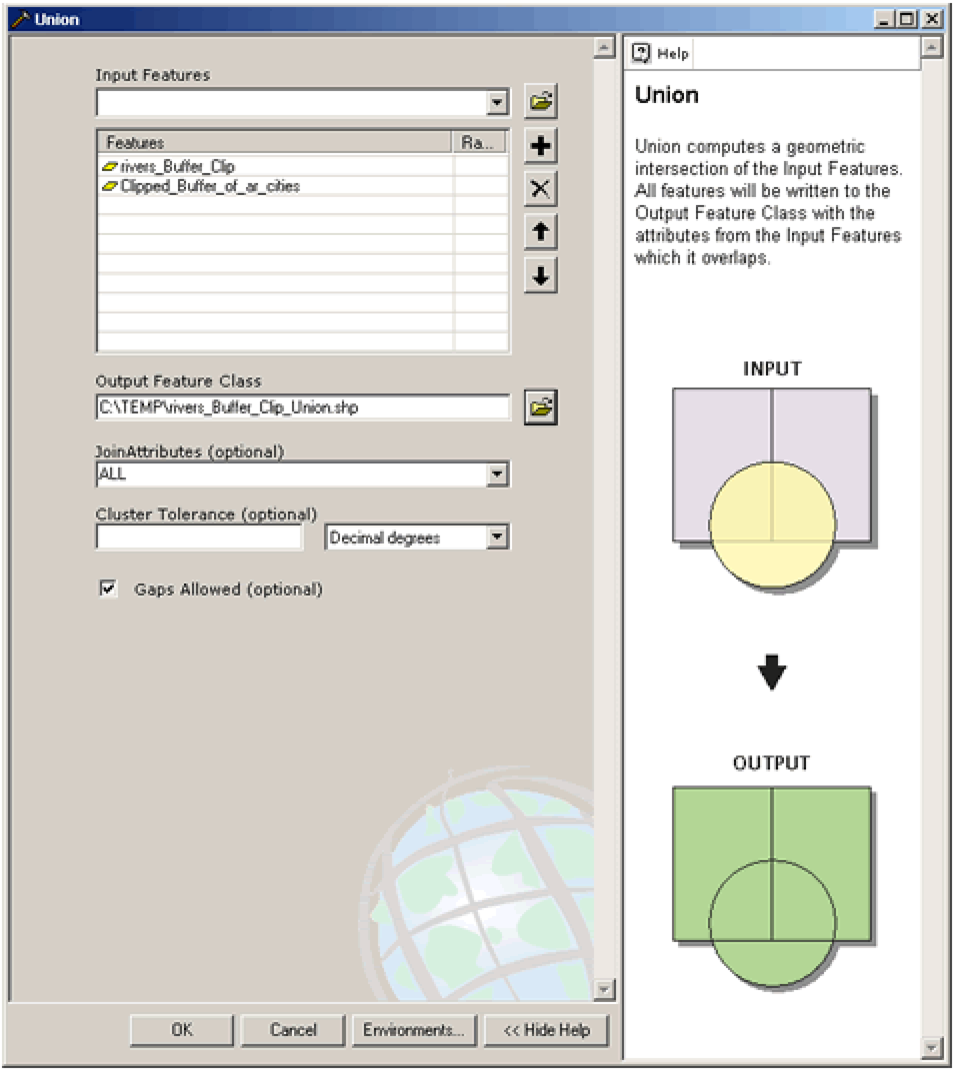
Overlay Example

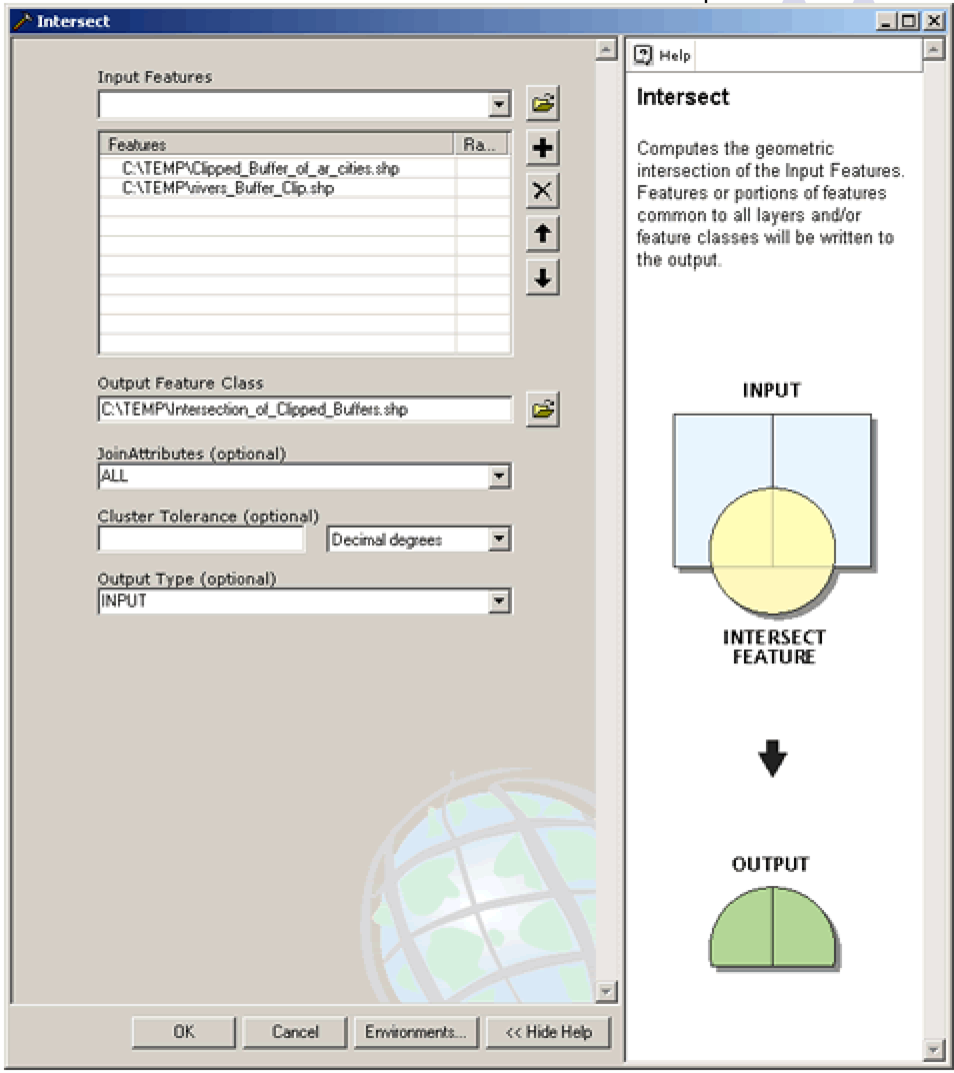


- Analysis Tools
  - select Overlay
    - Intersection tool

- Analysis Tools
  - select Overlay
    - Union tool

Overlay Example

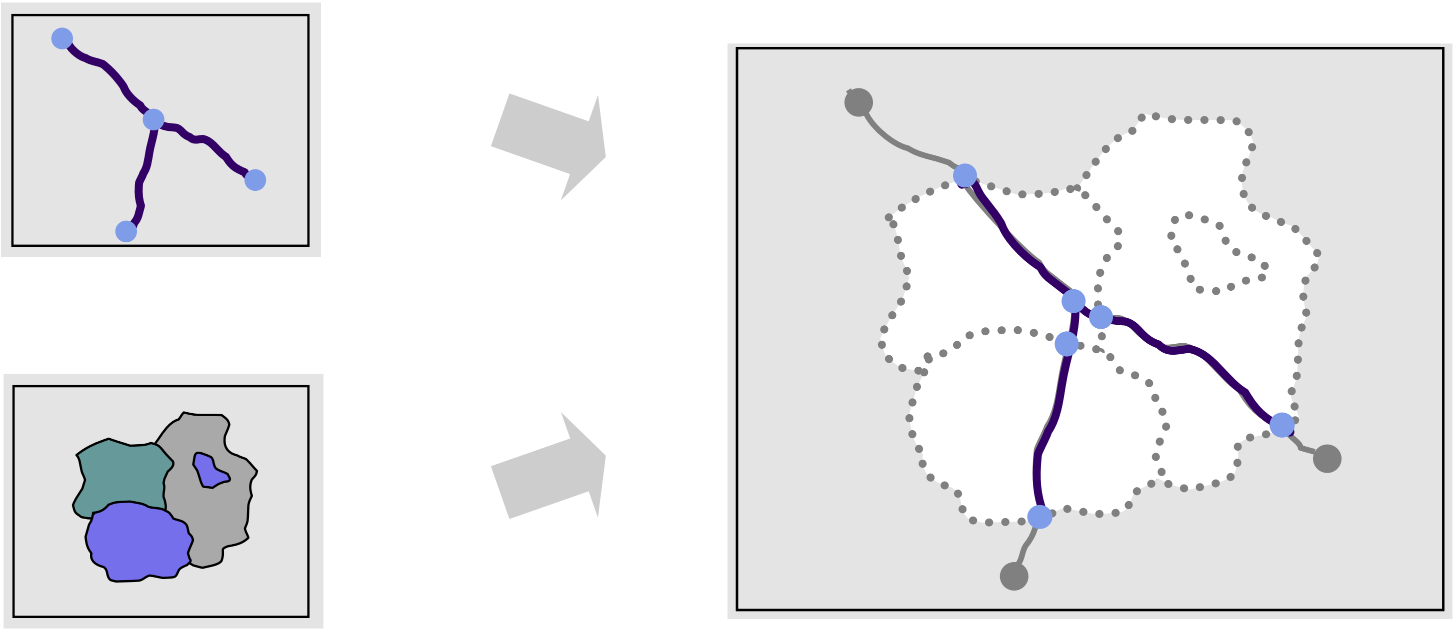




Intersect

Create new layer from common geometry between two input layers

Input points, lines, or polygons; overlay polygons



Analysis with Intersect

Attributes of Forest Stands

AREA	PERIMETER	STAND-ID	HS-INDEX	VALUE-PER-METER	OLDGROW#
107.582649	46.825623	291	1617.916748	64	1
321.611389	88.228783	251	735.220825	43	53
360.028687	92.255203	389	677.353882	75	1
43.471085	31.852316	413	822.655518	53	1
96.214729	50.517212	413	822.655518	53	1
361.912140	93.291679	0	0	0	48

Record: 0 Show: All Selected Records (0 out of 23140 Selected.) Options

Attributes of Road Buffer

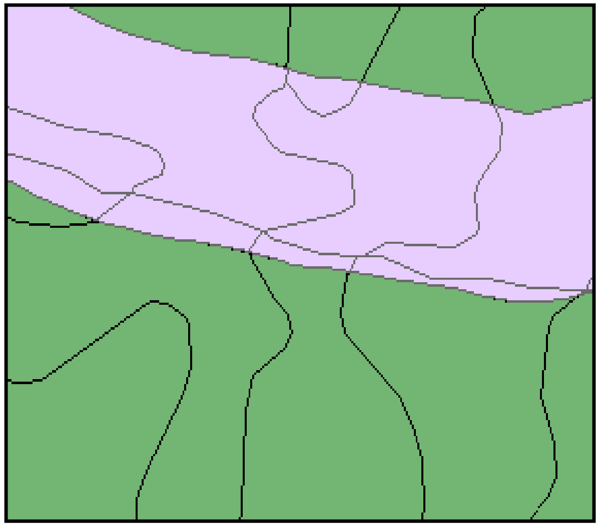
PERIMETER	ROADBUF#	ROADBUF-ID	IN-ROAD
84449.265625	2	1	1
2949.157471	3	2	1

Record: 1 Show: All Selected Records (0 out of 2 Se

Attributes of Intersection_Output

AREA	PERIMETER	ROADBUF_	ROADBUF_ID	IN_ROAD	OGRW_STND	OGRW_STND	ST.
8466290	84449.296875	2	1	1	23102	23101	
8466290	84449.296875	2	1	1	23128	23127	
8466290	84449.296875	2	1	1	21301	21300	
8466290	84449.296875	2	1	1	22688	22687	
8466290	84449.296875	2	1	1	22691	22690	
8466290	84449.296875	2	1	1	22645	22644	
8466290	84449.296875	2	1	1	22621	22620	

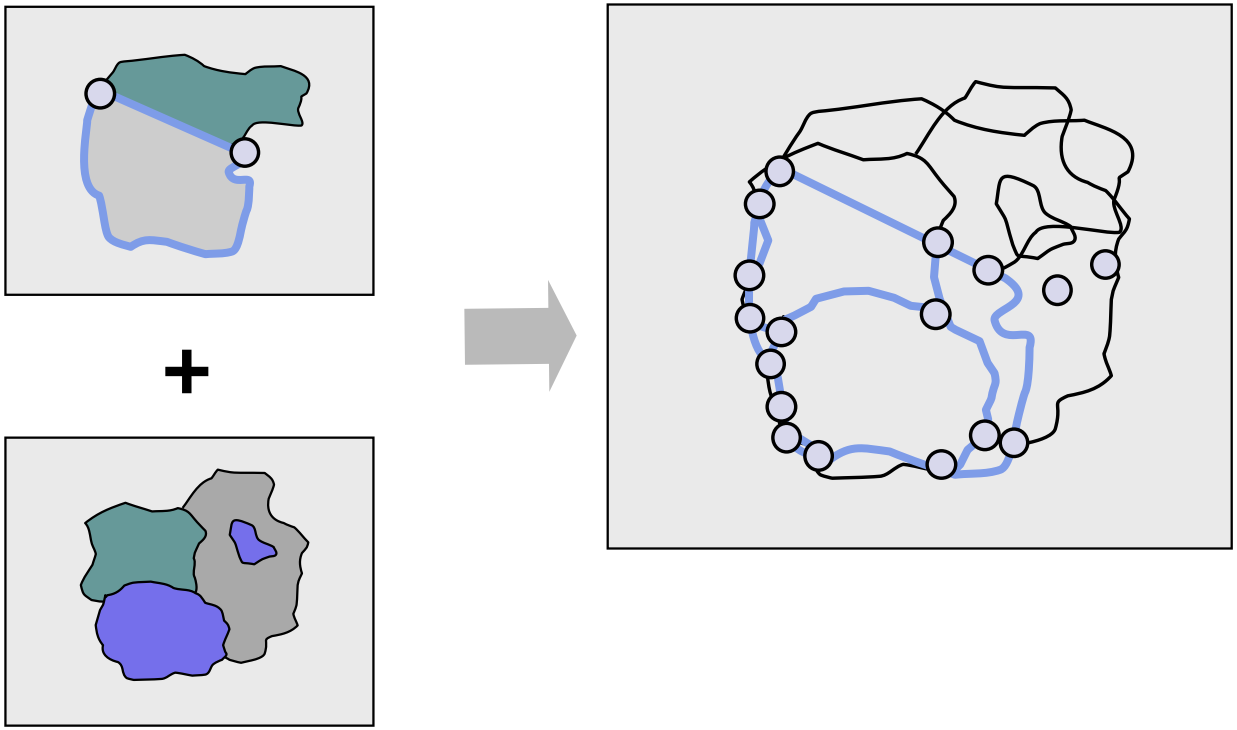
Record: 1 Show: All Selected Records (0 out of 1091 Selected.) Options



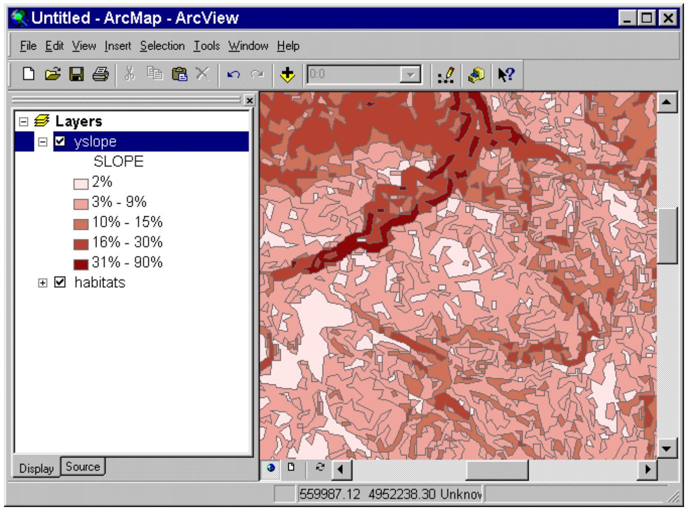
Union

Create a new layer from combined geometry of two input layers

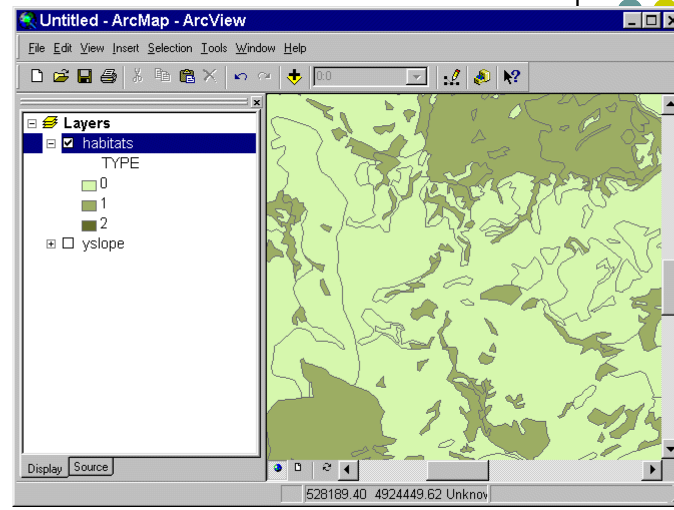
Both input and overlay layers must be polygons



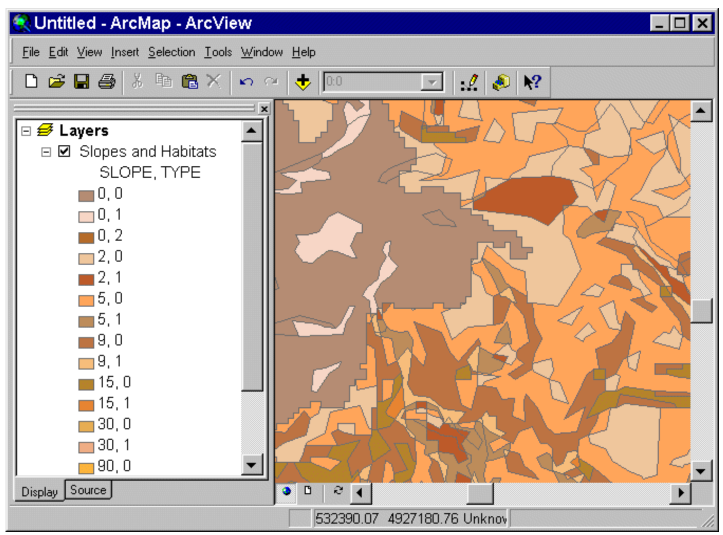
Analysis with Union



Percent slope

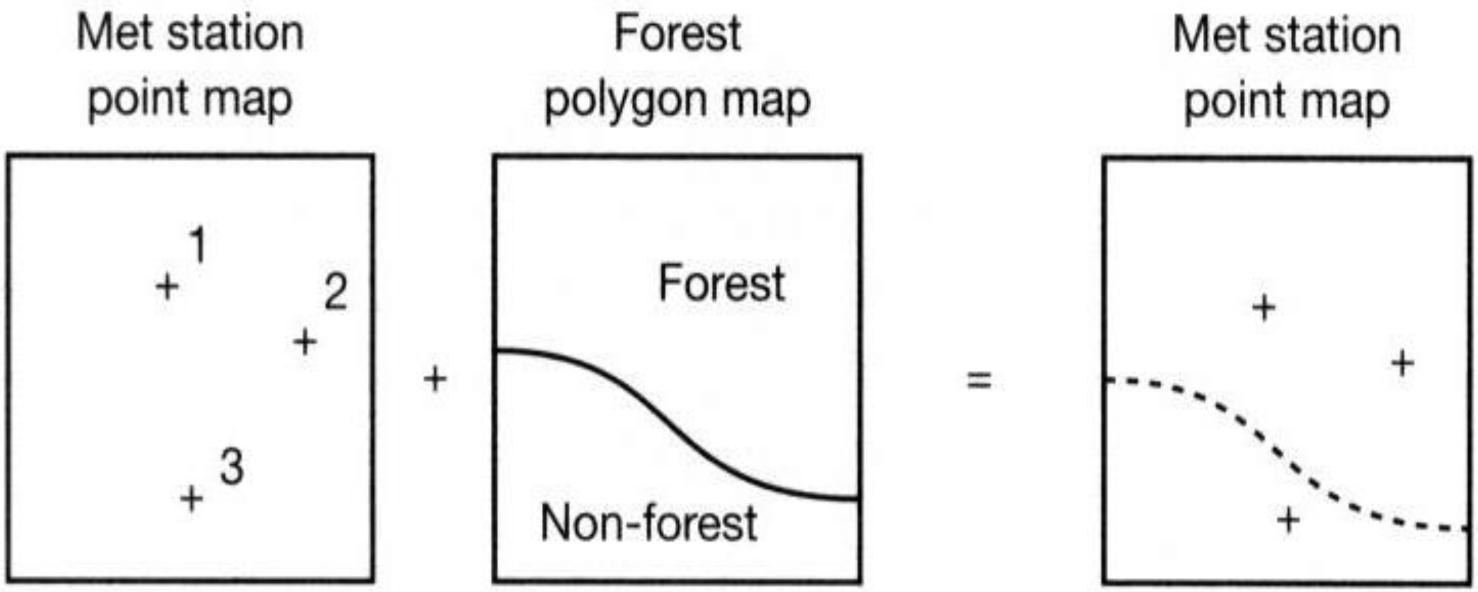


Habitat type



Query, display, and analyze new shapes

Vector based overlay

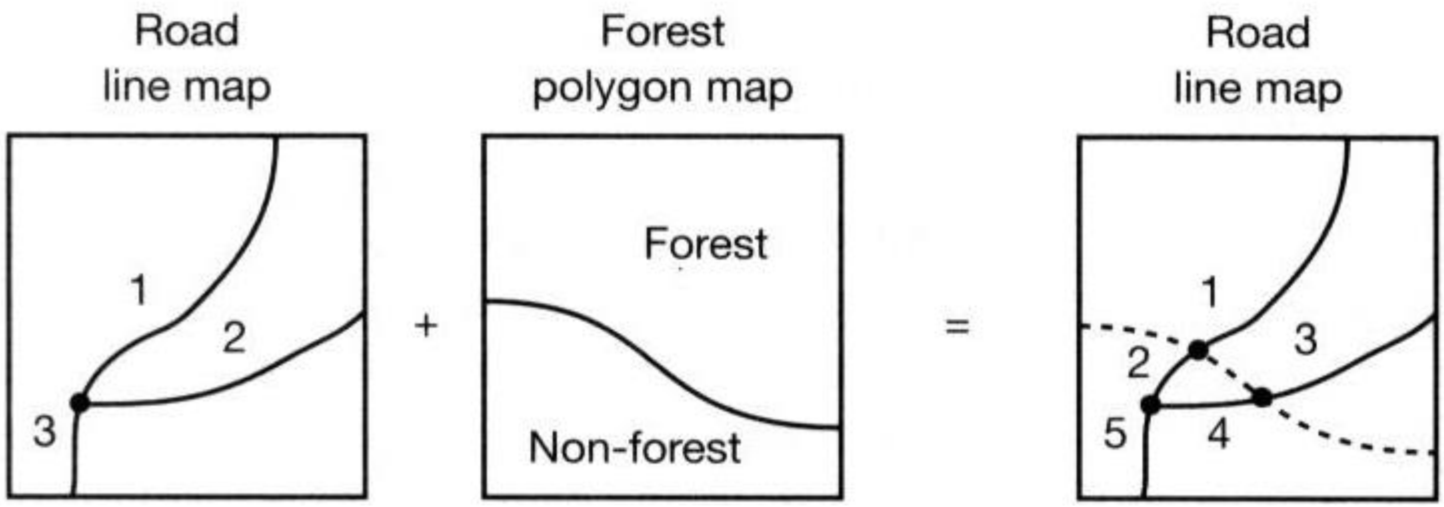


Met station
attribute table

Point ID	Land use
1	Forest
2	Forest
3	Non-forest

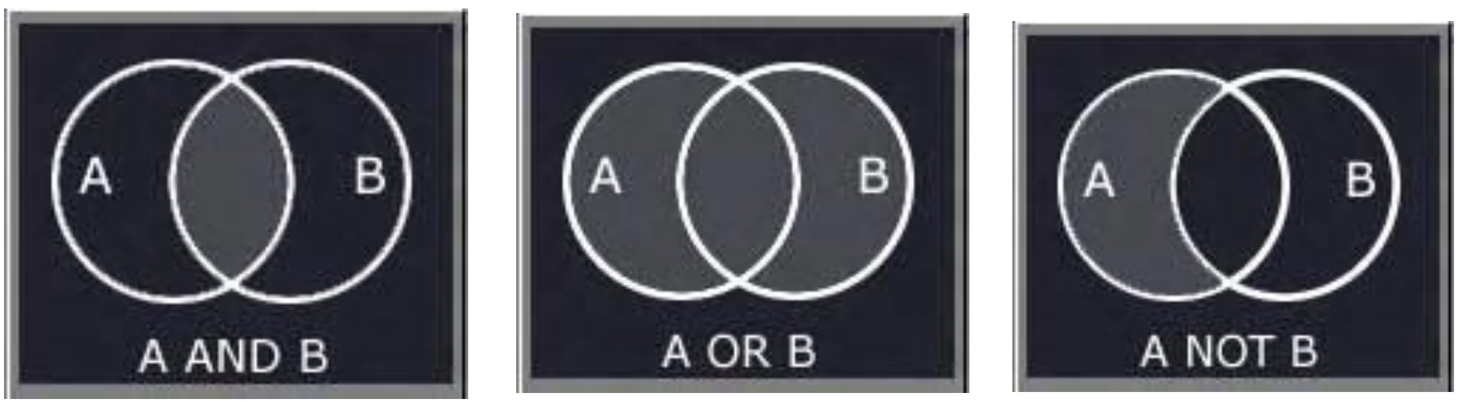
point-in-polygon example

Vector based overlay



line-in-polygon example

Vector based overlay



polygon-in-polygon example

Raster Based Overlay: Simple Addition

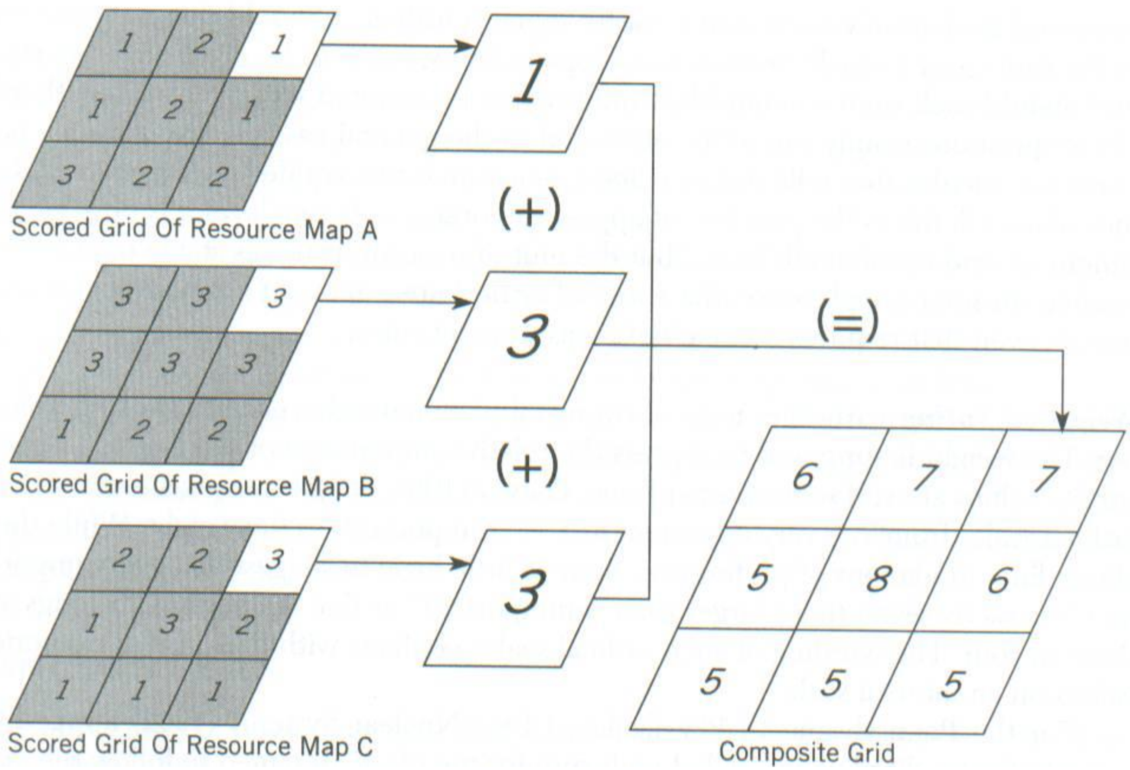


Image Source: Chrisman, Nicholas.(2002). 2nd Ed. *Exploring Geographic Information Systems.* p 144. fig. 5-12.

Raster Overlay: Boolean Combine

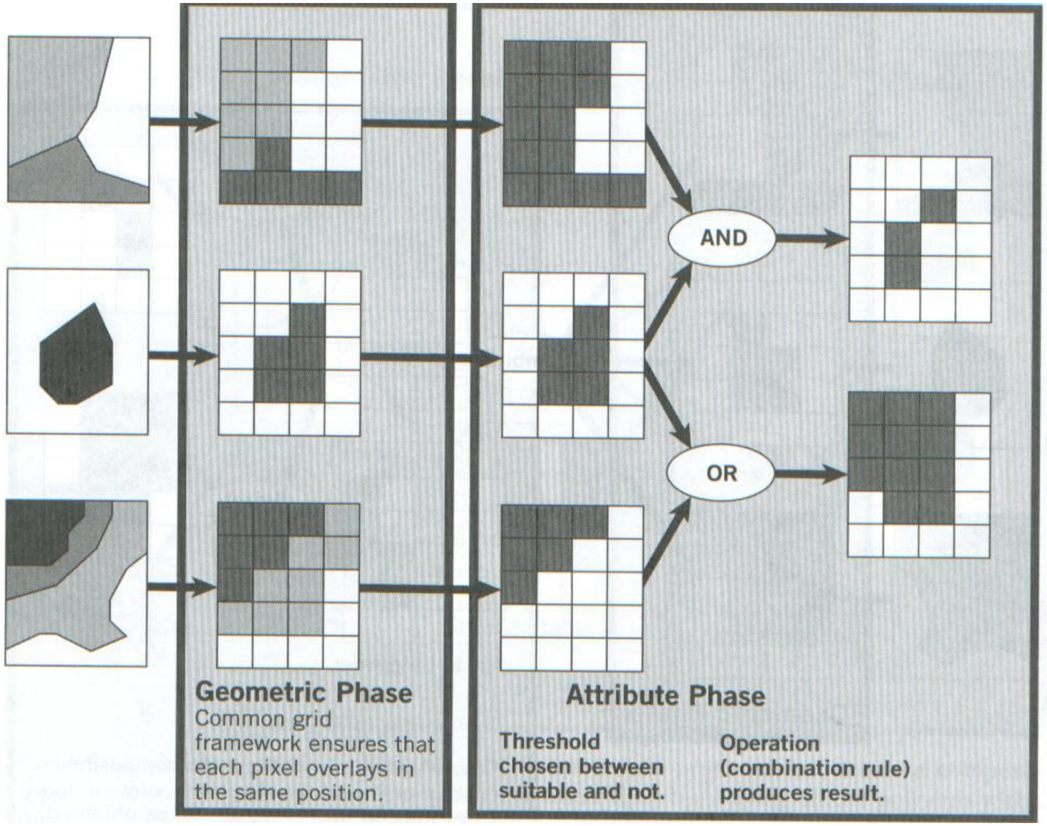
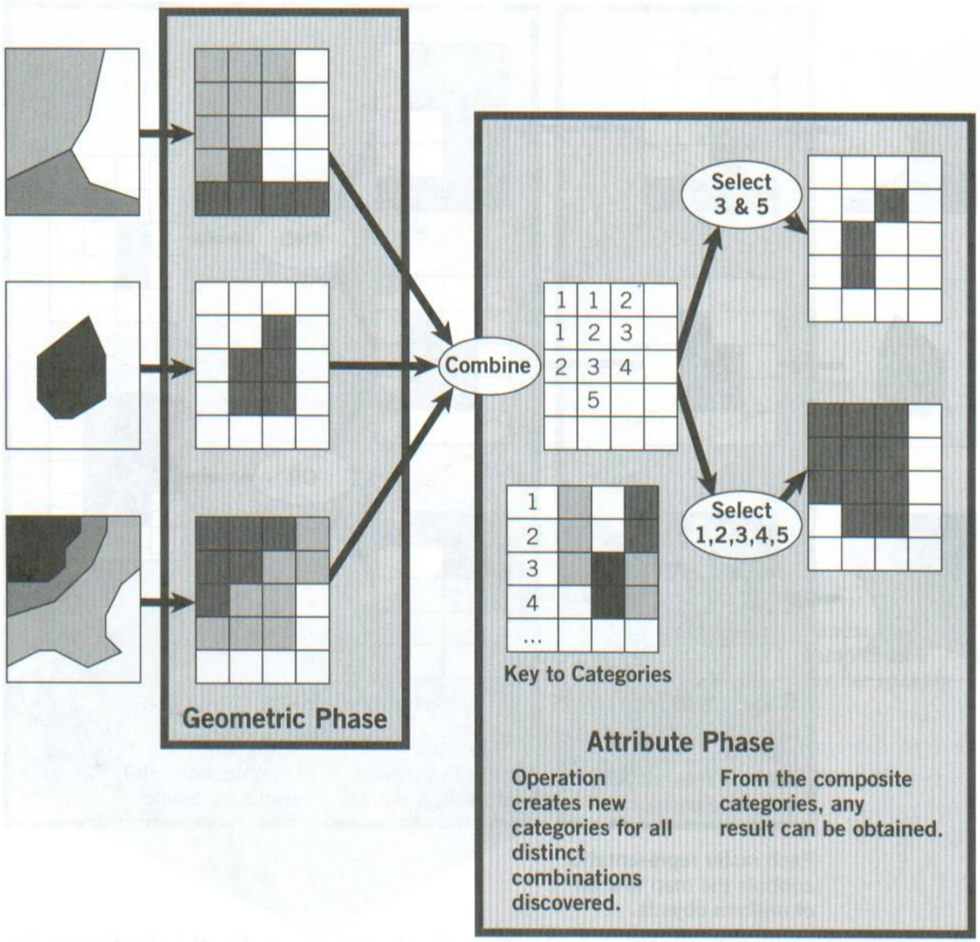


Image Source: Chrisman, Nicholas.(2002). 2nd Ed. *Exploring Geographic Information Systems*. p 125. fig. 5-3.

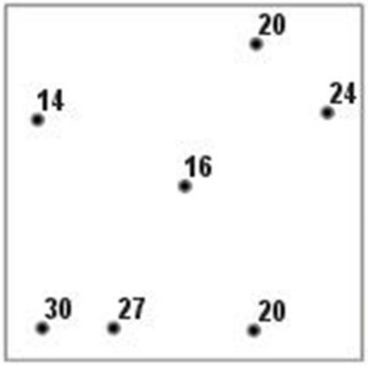
Raster Overlay: Composite Combine



§ 4 Transformation(cont.)

§ 4.4 Spatial interpolation

- Spatial interpolation is a pervasive operation in GIS
- Although it is often used explicitly in analysis, it is also used implicitly, in various operations such as the preparation of a contour map display, where spatial interpolation is invoked without the user's direct involvement
- Spatial interpolation is a process of intelligent guesswork, in which the investigator (and the GIS) attempt to make a reasonable estimate of the value of a continuous fieldat places where the field has not actually been measured
- Spatial interpolation is an operation that makes sense only from the continuous-field perspective
- the one principle that underlies all spatial interpolation is the Tobler Law

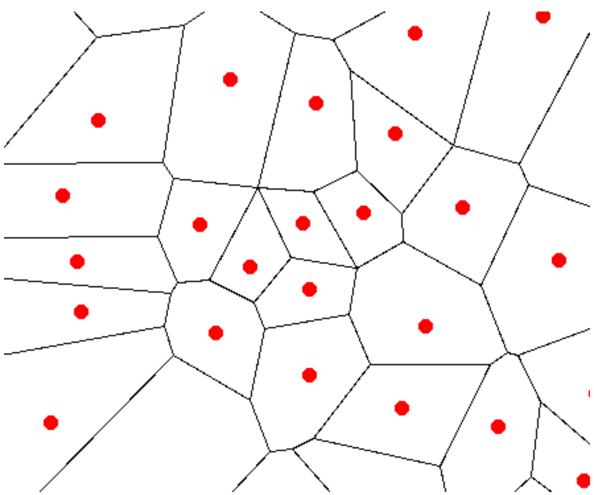


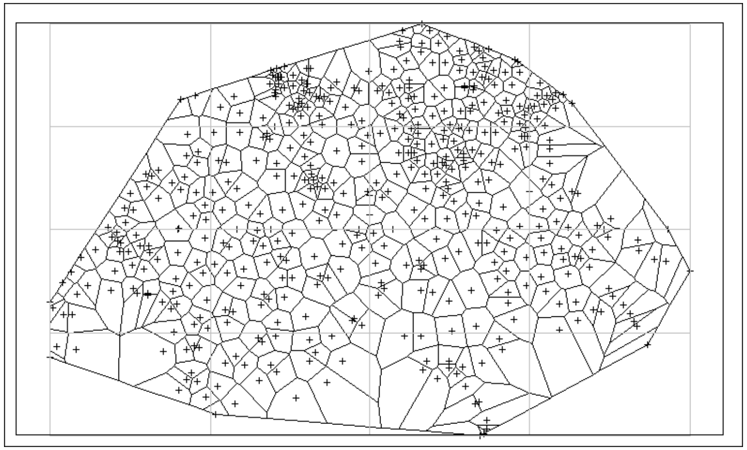
13	14	16	20	23
14	14	16	19	24
18	16	16	18	22
24	22	19	19	21
30	27	23	20	20

§ 4 Transformation(cont.)

§ 4.4.1 Thiessen Polygon(Voronoi Diagram)

- Thiessen polygons are generated from a set of sample points such that each polygon defines an area of influence around its sample point, so that any location inside the polygon is closer to that point than any of the other sample points

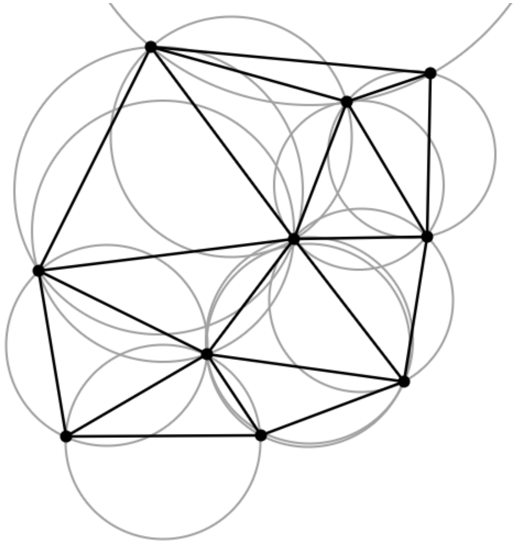




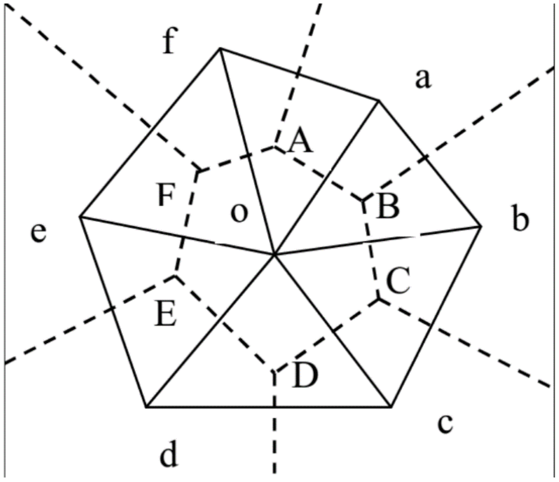
§ 4 Transformation(cont.)

§ 4.4.1 Thiessen Polygon(cont.)

- Defines area of influence around a point that polygon boundaries are equidistant from neighboring points
  - Join points to nearest neighbor in a triangular tessellation
  - Compute perpendicular bisector
  - Remove triangles



Delaunay triangulation



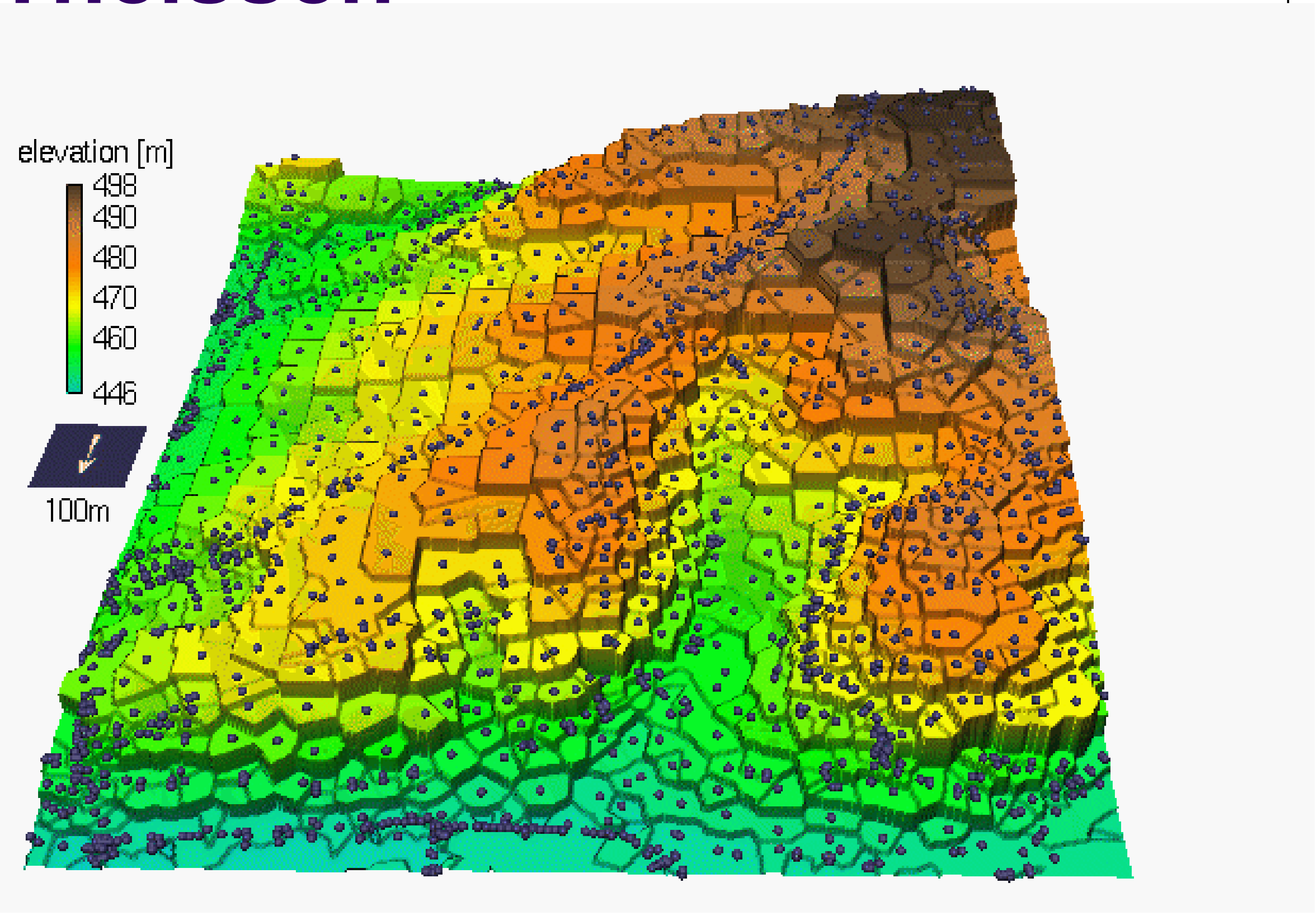
Thiessen polygons

§ 4 Transformation(cont.)

§ 4.4.1 Thiessen Polygon(cont.)

- Problems
  - Size and shape of area depends on point distribution
  - Value of property of interest in each cell is estimated by a sample of one
  - Computation of a value at an unsampled point is simply a point in polygon problem, and not related to nearest neighbor (hard boundaries)

Theissen



§ 4 Transformation(cont.)

§ 4.4.2 Inverse-distance weighting

- Denote the point of interest as x, and the points where measurements were taken as x_i, where i runs from 1 to n, if there are n data points
- Denote the unknown value as z(x) and the known measurements as z_i
- Give each of these points a weight d_i , which will be evaluated based on the distance from x_i to x.
- Then the weighted average computed at x is

$$z(\mathbf{x}) = \sum_i w_i z_i / \sum_i w_i$$

- where

$$w_i = 1/d_i^2$$

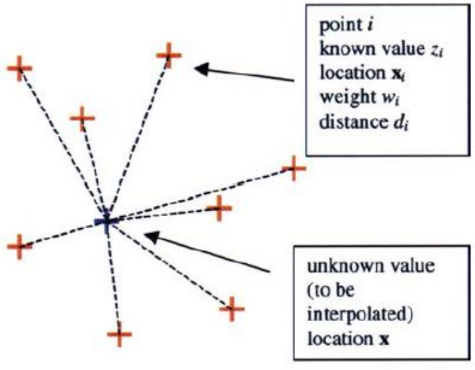
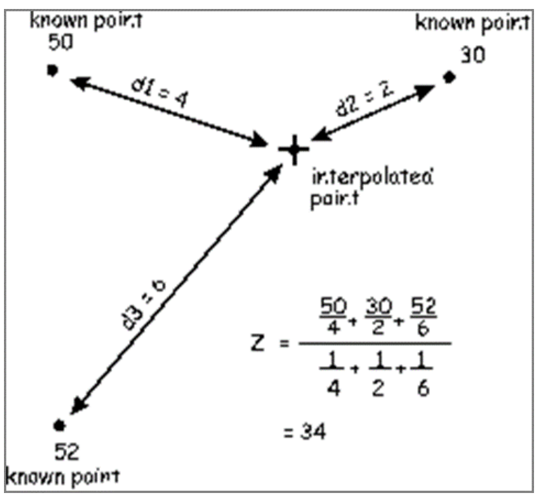


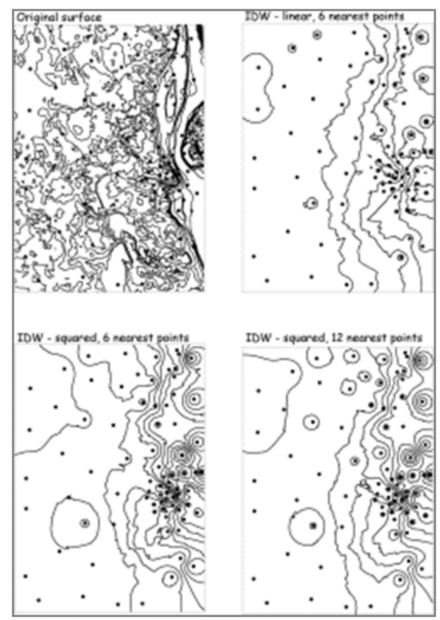
Figure 14.25 Notation used in the equations defining spatial interpolation

§ 4 Transformation(cont.)

§ 4.4.2 Inverse-distance weighting

- Inverse Distance Weighted and radial basis functions are exact interpolators
- The influence of an input point on an interpolated value is isotropic. Since the influence of an input point on an interpolated value is distance related, IDW is not "ridge preserving"
- The best results from IDW are obtained when sampling is sufficiently dense with regard to the local variation you are attempting to simulate

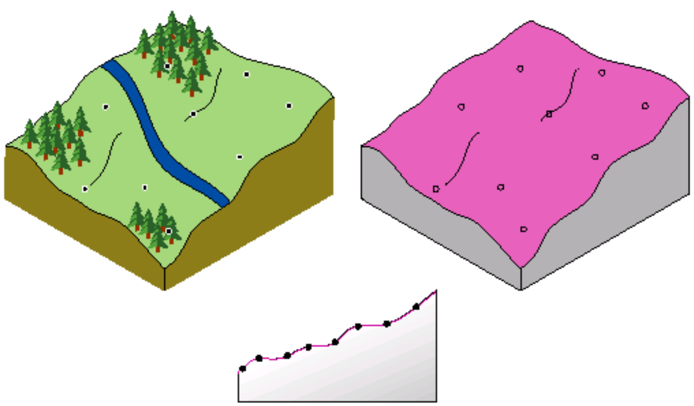




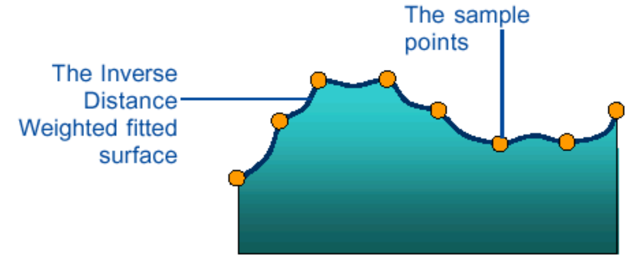
How radial basis functions work

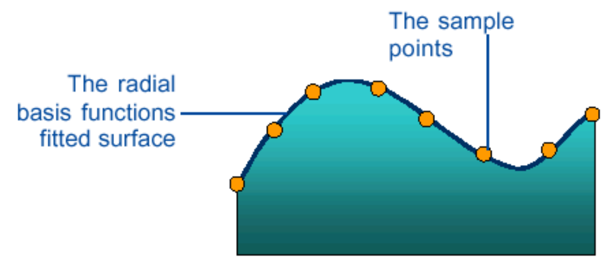
Radial basis functions (RBF) methods are a series of exact interpolation techniques, that is, the surface must go through each measured sample value. There are five different basis functions: thin-plate spline, spline with tension, completely regularized spline, multiquadric function, and inverse multiquadric spline. Each basis function has a different shape and results in a different interpolation surface. RBF methods are a form of artificial neural networks.

RBFs are conceptually similar to fitting a rubber membrane through the measured sample values while minimizing the total curvature of the surface. The selected basis function determines how the rubber membrane will fit between the values. The diagram below demonstrates conceptually how an RBF surface fits through a series of elevation sample values. Notice in the cross section that the surface passes through the data values.

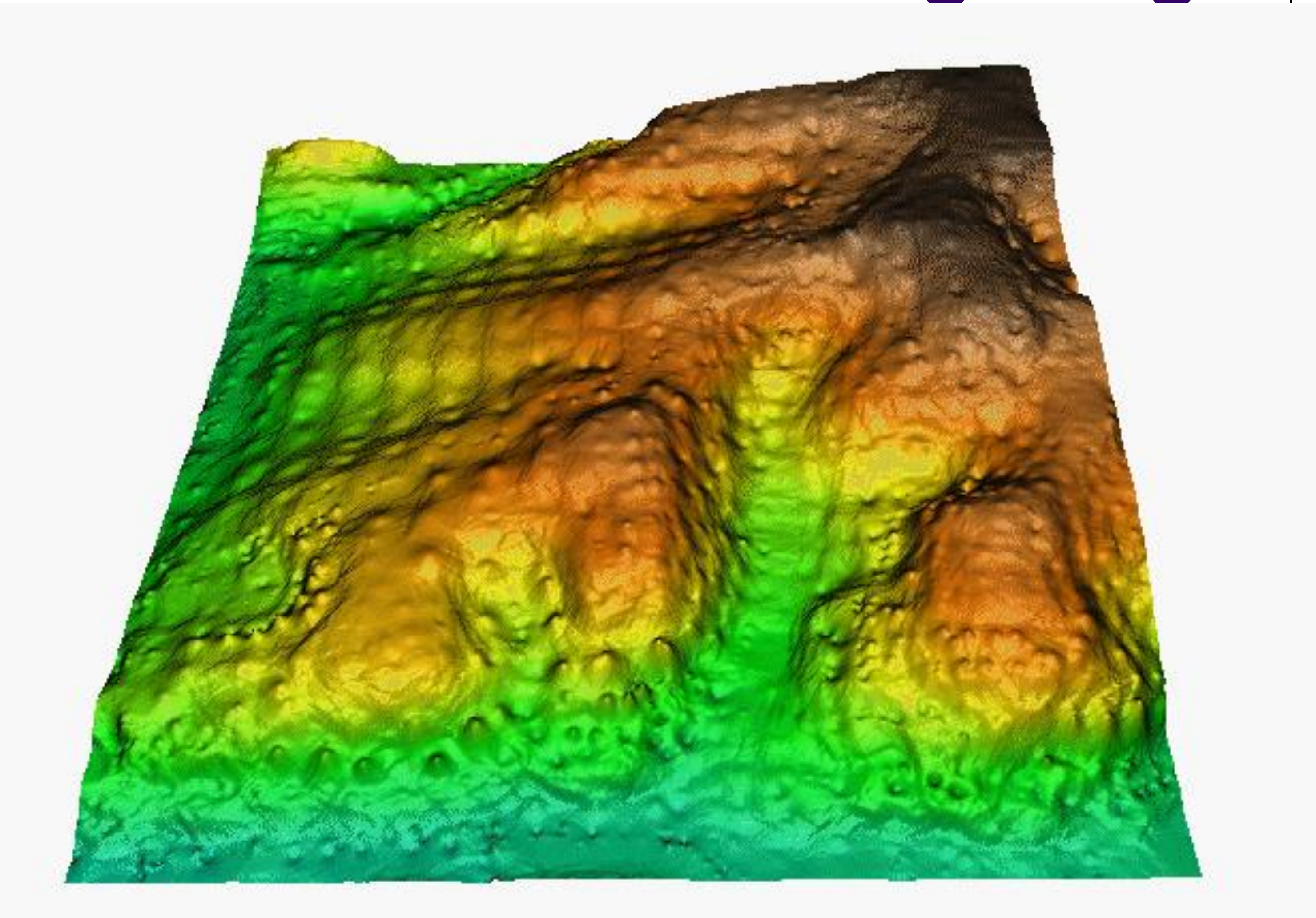


Being exact interpolators, RBF methods differ from the global and local polynomial interpolators, which are both inexact interpolators that do not require the surface to pass through the measured points. When comparing RBF to the IDW method, another exact interpolator, IDW will never predict values above the maximum measured value or below the minimum measured value, as you can see in the cross section of a transect of sample data below.





Inverse Distance Weighting



§ 4 Transformation(cont.)

§ 4.4.3 Kriging

- Of the common methods of spatial interpolation, Kriging makes the most convincing claim to be grounded in good theoretical principles
- The basic idea is to discover something about the general properties of the surface, as revealed by the measured values, and then to apply these properties in estimating the missing parts of the surface
- Smoothness is the most important property (note the inherent conflict between this and the properties of fractals), and it is operationalized in Kriging in a statistically meaningful way

§ 4 Transformation(cont.)

§ 4.4.3 Kriging(cont.)

- an advanced interpolation procedure that generates an estimated surface from a scattered set of points with z values.
- Assumes spatial correlation among data points
- Kriging is based on the regionalized variable theory that assumes that the spatial variation in the phenomenon represented by the z values is statistically homogeneous throughout the surface; that is, the same pattern of variation can be observed at all locations on the surface.
- Requires a strong background in spatial statistics

Geostatistical solutions

So far, the techniques that we have discussed are referred to as deterministic interpolation methods because they are directly based on the surrounding measured values or on specified mathematical formulas that determine the smoothness of the resulting surface. A second family of interpolation methods consists of geostatistical methods that are based on statistical models that include autocorrelation (statistical relationships among the measured points). Not only do these techniques have the capability of producing a prediction surface, but they can also provide some measure of the certainty or accuracy of the predictions.

The following example will guide you through the basic steps of geostatistics using ordinary kriging.

Kriging is similar to IDW in that it weights the surrounding measured values to derive a prediction for each location. However, the weights are based not only on the distance between the measured points and the prediction location but also on the overall spatial arrangement among the measured points. To use the spatial arrangement in the weights, the spatial autocorrelation must be quantified.

To solve the geostatistical example, you will walk you through a series of steps.

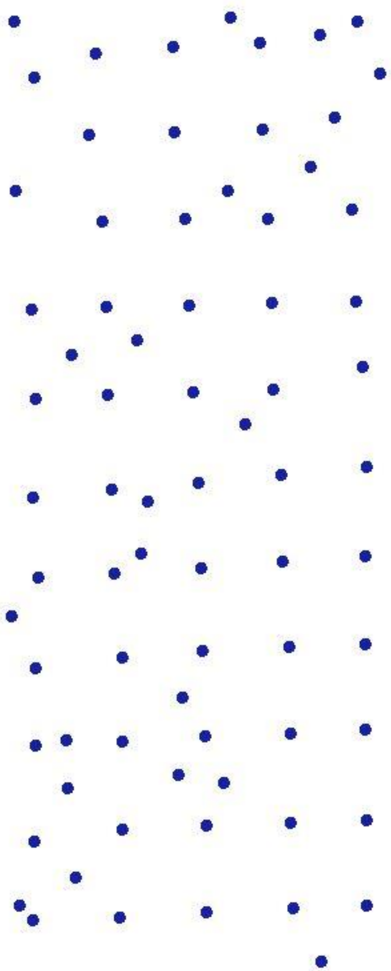
Calculate the empirical semivariogram—kriging, like most interpolation techniques, is built on the assumption that things that are close to one another are more alike than those farther away (quantified here as spatial autocorrelation). The empirical semivariogram is a means to explore this relationship. Pairs that are close in distance should have a smaller measurement difference than those farther away from one another. The extent that this assumption is true can be examined in the empirical semivariogram.

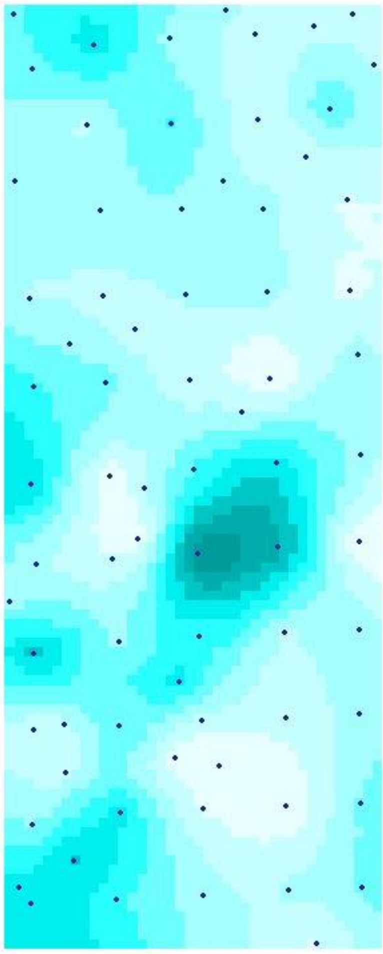
Fit a model—this is done by defining a line that provides the best fit through the points in the empirical semivariogram cloud graph. That is, you need to find a line such that the (weighted) squared difference between each point and the line is as small as possible. This is referred to as the (weighted) least-squares fit. This line is considered a model that quantifies the spatial autocorrelation in your data.

Create the matrices—the equations for ordinary kriging are contained in matrices and vectors that depend on the spatial autocorrelation among the measured sample locations and prediction location. The autocorrelation values come from the semivariogram model described above. The matrices and vectors determine the kriging weights that are assigned to each measured value.

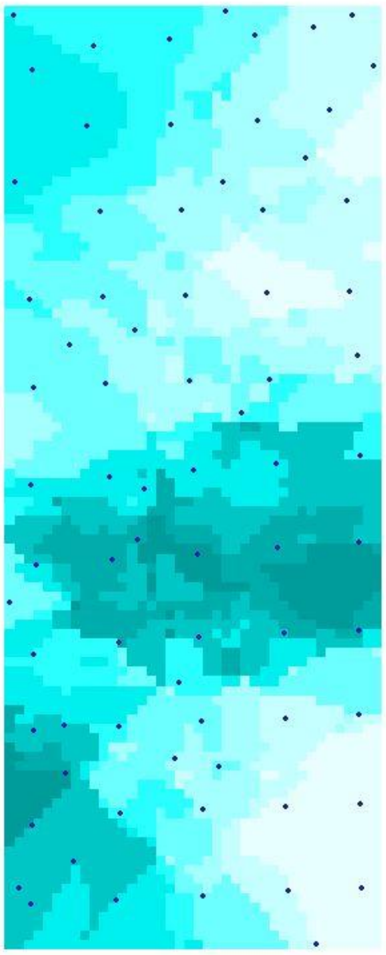
Make a prediction—from the kriging weights for the measured values, you can calculate a prediction for the location with the unknown value.

Kriging

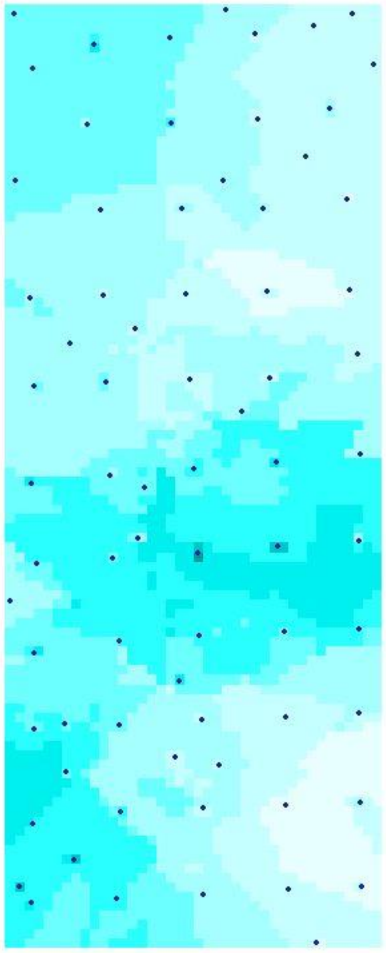




Universal

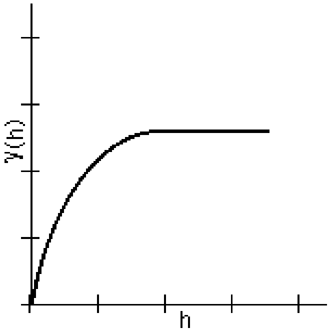


Exponential



Circular

SPHERICAL

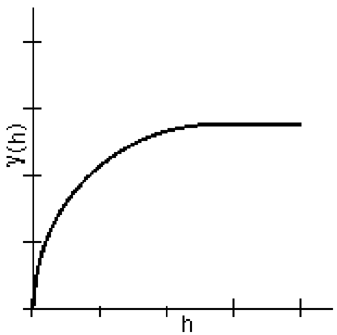


$$\gamma(\mathbf{h}) = c_0 + c\left(\frac{3h}{2a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right) \qquad 0 < h \le a$$

$$\gamma(\mathbf{h}) = c_0 + c \qquad h > a$$

$$\gamma(0) = 0$$

CIRCULAR



$$\gamma(\mathbf{h}) = c_0 + c\left(1 - \frac{2}{\pi}\cos^{-1}\left(\frac{h}{a}\right) + \sqrt{1 - \frac{h^2}{a^2}}\right)$$

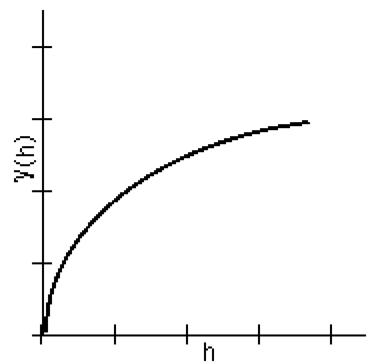
$$0 < h \le a$$

$$\gamma(\mathbf{h}) = c_0 + c \qquad h > a$$

$$\gamma(0) = 0$$

© ESRI, ArcView Help

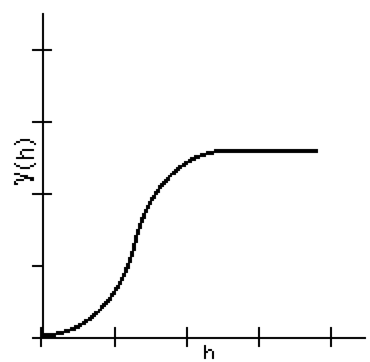
EXPONENTIAL



$$\gamma(\mathbf{h}) = c_0 + c\left(1 - exp\left(\frac{-h}{r}\right)\right) \qquad h > 0$$

$$\gamma(0) = 0$$

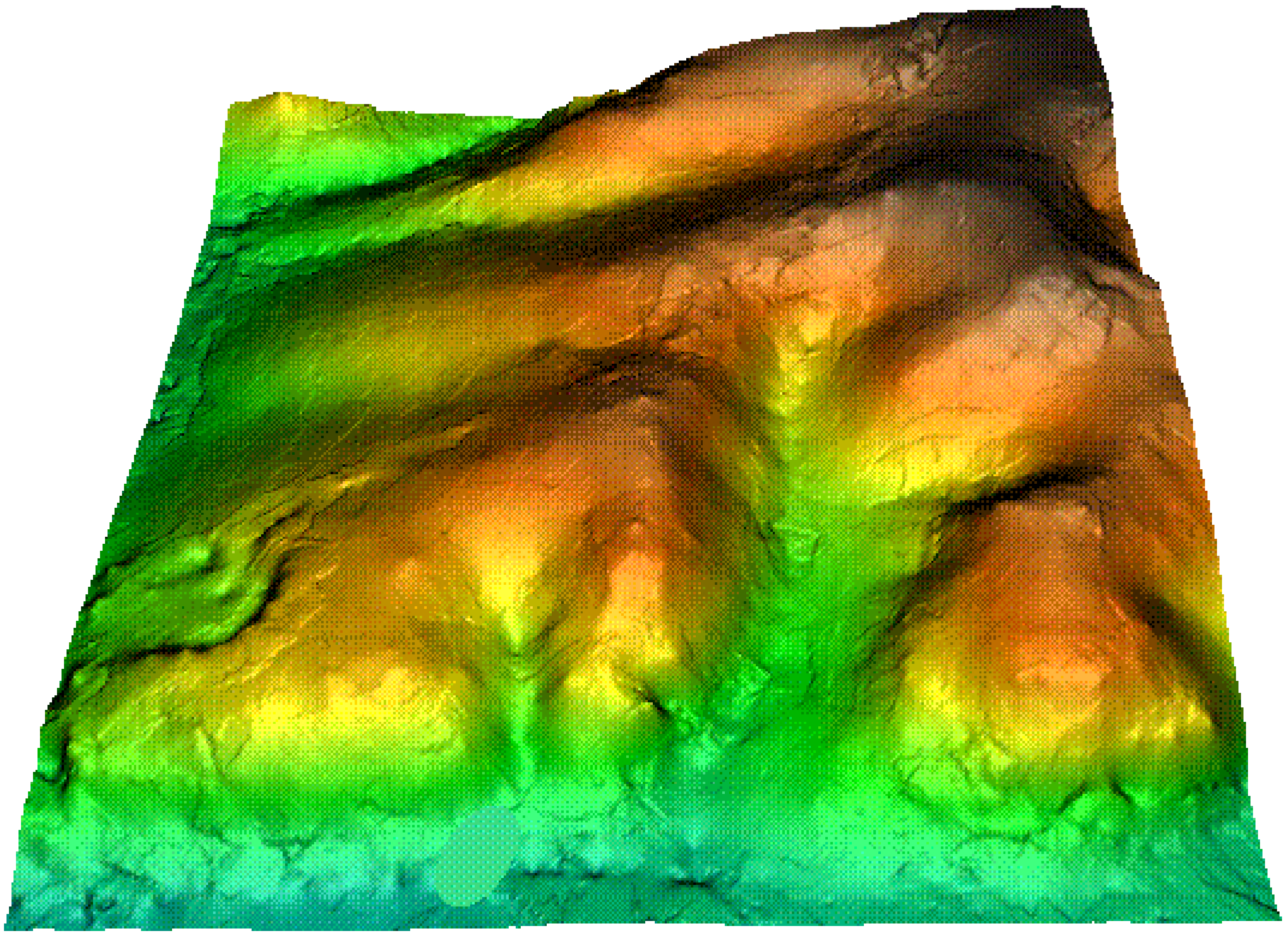
GAUSSIAN



$$\gamma(\mathbf{h}) = c_0 + c\left(1 - exp\left(-\frac{h^2}{r^2}\right)\right) \qquad h > 0$$

$$\gamma(0) = 0$$

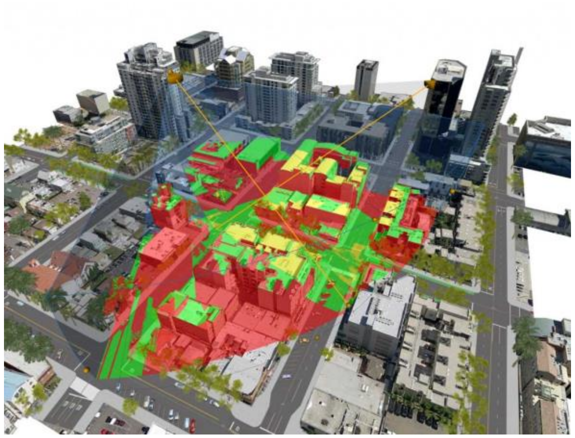
Kriging

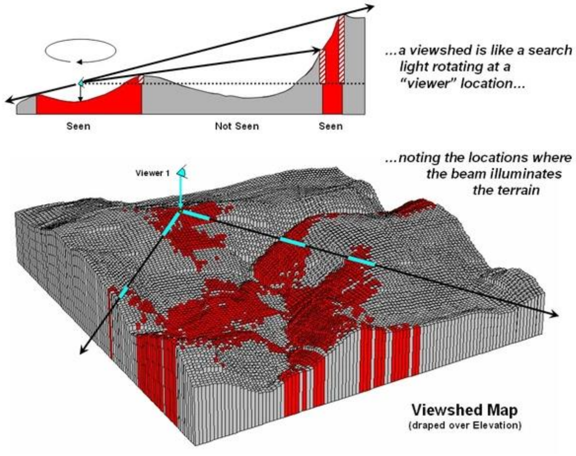


§ 4 Transformation(cont.)

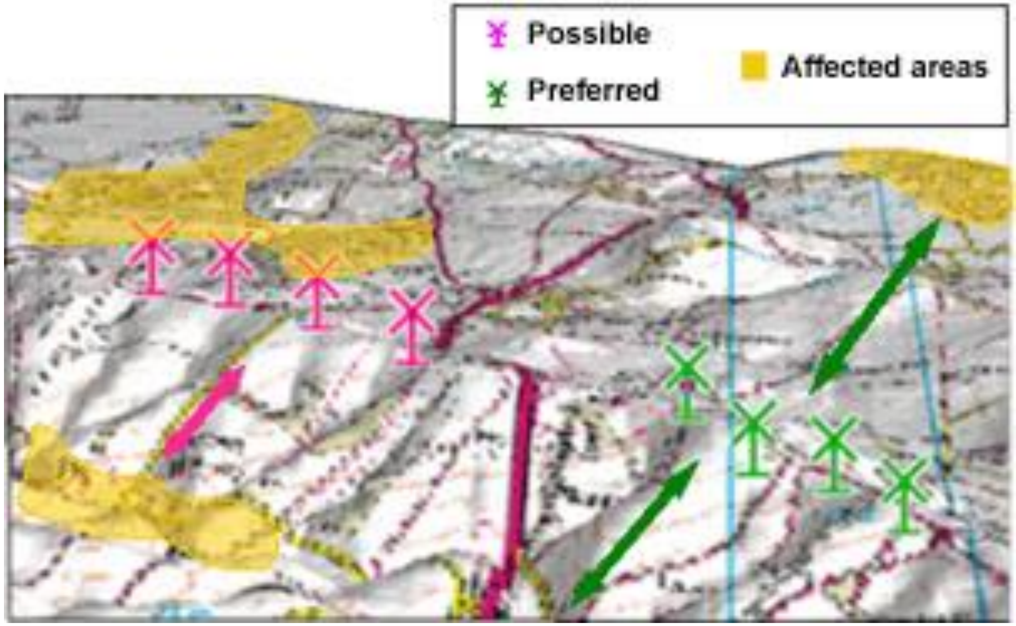
§ 4.5 Terrain Analysis

- Visibility Analysis Operations
  - identification of areas of terrain that can be seen from a particular point on the surface
- Viewshed Operation
  - uses digital elevation model data (DEMs) or*.....*
  - digital terrain model data (DTMs) or*......*
  - triangulated irregular network data (TINs)





§ 4 Transformation(cont.)



3D landscape model impact on natural beauty

§ 4 Transformation(cont.)

§ 4.5 Terrain Analysis(Cont.)

- Surface functions
  - density, contour, interpolation functions
  - aspect, slope, hillshade, etc.
  - watershed analysis and modeling (flow direction, flow accumulation, flow length, watershed delineation, stream ordering)
  - visibility modeling/mapping
    - determine the area that can be "seen" from the target location