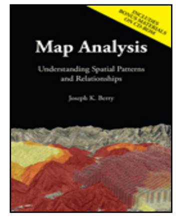
Beyond Mapping III

Topic 2 – *Fundamental Map Analysis Approaches (Further Reading)*



Map Analysis book

GIS Represents Spatial Patterns and Relationships — *discusses the important differences among discrete mapping, continuous map surfaces and map analysis* *(April 1999)*

<*Click here*> *for a printer-friendly version of this topic (.pdf).*

(***Back to the Table of Contents***)

GIS Represents Spatial Patterns and Relationships

(GeoWorld, April 1999)

*(**return to top of Topic**)*

Many of the topics in the *Beyond Mapping Compilation Series* of online books discuss the subtle (and often not so subtle) differences between mapping and map analysis. Traditionally, mapping identifies distinct map features, or spatial objects, linked to aggregated tables that are visually interpreted for spatial relationships. The thematic mapping and geo-query capabilities of this approach enable users to "see through" the complexity of spatial data and the barrage of associated tables.

Map analysis, on the other hand, slogs around in the complexity of geographic space, treating it as a continuum of varying responses and utilizing map-*ematical* computations to uncover spatial relationships. A major distinction between the two approaches lies in the extension of the traditional map elements of points, lines and areas to map surfaces—an old concept that has achieved practical reality with the advent of digital maps.

Zones and Surfaces

While much of the information in a GIS is discrete (e.g., the infrastructure of roads, buildings, and power lines), the focus of many applications extend to decision factors that widely vary throughout geographic space. As a result, surface modeling plays a dominant role in site-specific management of geographically diffuse conditions.

Map surfaces, also termed spatial gradients, often are characterized by grid-based data structures. In forming a surface, the traditional representation based on irregular polygons is replaced by a highly resolved matrix of grid cells superimposed over an area (top portion of figure 1).

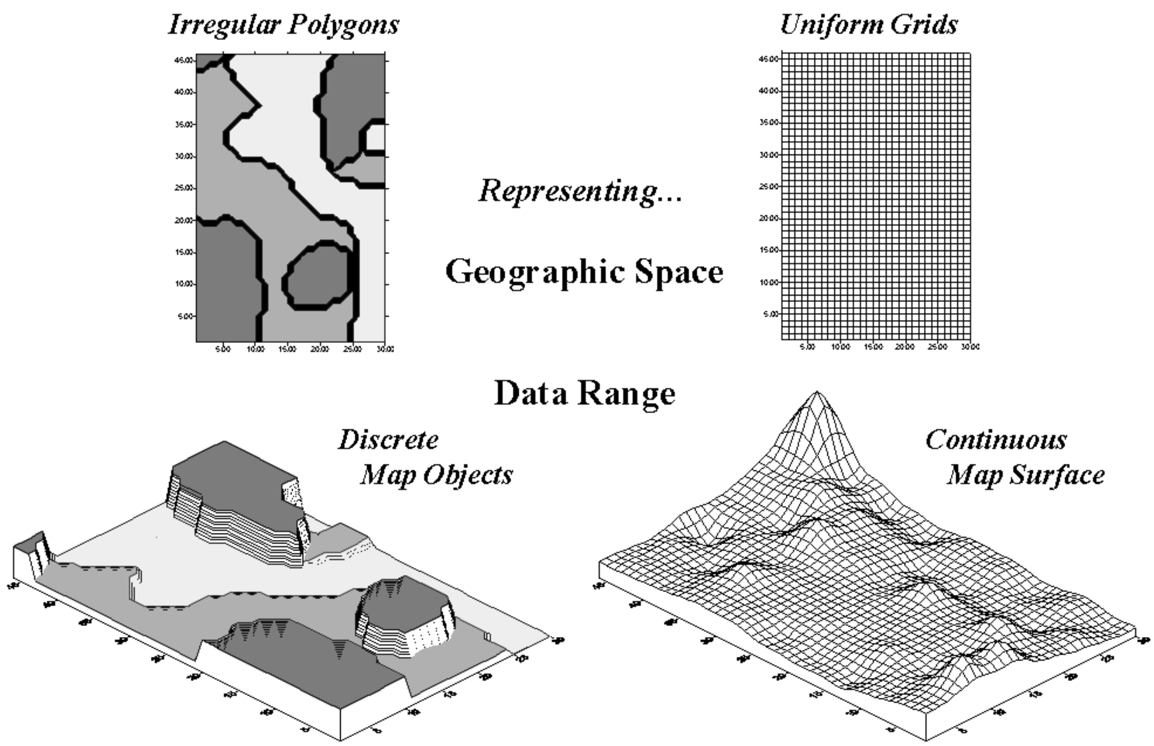


Figure 1. Comparison of zone (polygon) and surface (grid) representations for a continuous variable.

The representation of the data range for the two approaches is radically different. Consider the alternatives for characterizing phosphorous levels throughout a farmer's field. One approach, termed zone management, uses air photos and a farmer's knowledge to subdivide the field into similar areas (gray levels depicted on the left side of figure 1). Soil samples are randomly collected in the areas and the average phosphorous level is assigned to each zone. These, plus other soil data gathered for each of the zones are used to develop a fertilization program for the field.

An alternative approach, termed site-specific management, systematically samples the field, then interpolates these data for a continuous map surface of phosphorous levels (right side of figure 1). First, note the similarities between the two representations—the generalized levels (data range) for the zones correspond fairly well with the map surface levels (the darkest zone tends to align with the highest levels, while the lightest zone contains the lowest levels).

Now consider the differences between the two representations. Note that the zone approach assumes a constant level (horizontal plane) of phosphorous within each zone (Zone#1(dark gray)= 55, Zone#2= 46 and Zone#3(light gray)= 42, whereas the surface shows a gradient of change across the field varying from 22 to 140). Two important pieces of information are lost in the zone approach—the extreme high/low values and the geographic distribution of the variation. This "missing" information severely limits the potential for spatial analysis of the zone data.*

Ok, you're not a farmer, so what's to worry? If you think about it, the bulk of GIS data is zonal—average income by counties, dominant species by forest parcel, and total sales within market area are common examples. However, in many instances the parsing into irregular polygons has minimal relationship to the actual geographic distribution of the data. For example, knowing an average lead concentration in wells for a county is of minimal value if there are a lot of wells above the average and they are all concentrated in one area. Or, two adjoining market areas might both seem mediocre, but all of their high sales happen along their shared boundary. In both cases the averaging within zones severely reduces the spatial information contained in the "raw" data.

Surface Modeling

There are two ways to establish map surfaces—continuous sampling and spatial analysis of a dispersed set samples. By far the best way is to continuously sample and directly assign an actual measurement(s) to each grid cell. Remote sensing data with a measurement for each pixel is good example of this data type.

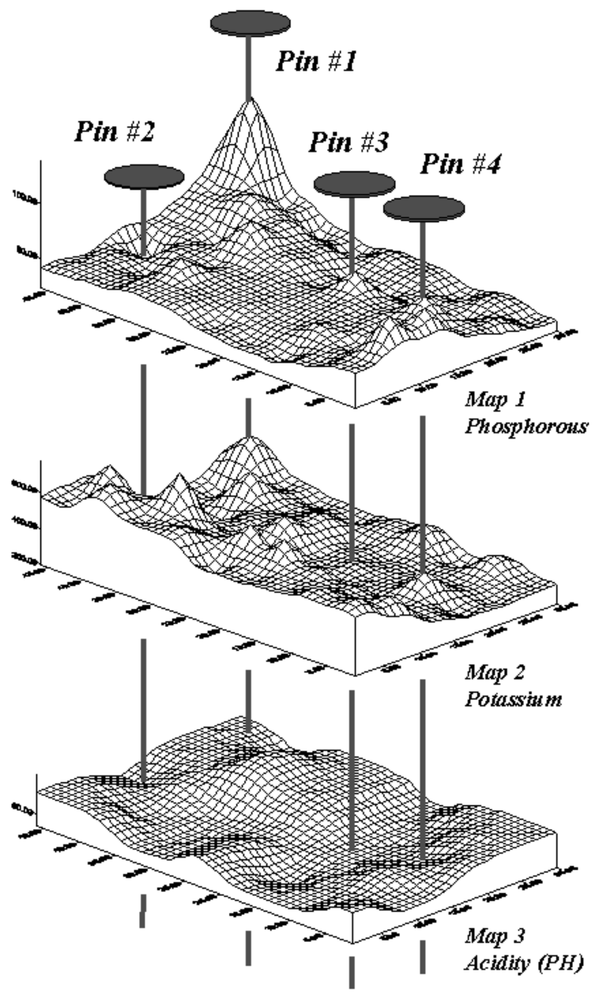


Figure 2. Geo-referenced map surfaces provide information about the unique combinations of data values occurring throughout an area.

However, map surfaces often are derived by statistically estimating a value for each grid cell based on a set of scattered measurements. For example, locations of a bank's home equity loan accounts can be geo-coded by their street address. Like "push-pins" stuck into a map on the wall, clusters of accounts form discernible patterns. An account density surface is easily generated by successively stopping at each grid cell and counting the number of accounts within $1/8^{th}$ of a mile. In a similar fashion, an account value surface is generated by summing the account values within the radius. These surfaces show the actual "pockets" (peaks and pits) of the bank's customers. By contrast, a zonal approach would simply assign the average number/value of accounts "falling into" predefined city neighborhoods, whether they actually matched the spatial patterns in the data or not.

Surface Analysis

The loss in spatial specificity for a map variable by generalizing into zones can be significant. However, the real kicker comes when you attempt to analyze the coincidence among maps. Figure 2 shows three geo-referenced surfaces for the farmer's field—phosphorous, potassium and acidity (PH). The pins depict four of the 1380 possible combinations of data for the field. By contrast, the zonal representation has only three possible combinations since it has just three distinct zones with averages attached.

The assumption of the zone approach is that the coincidence of the averages is consistent throughout the entire map area. If there is a lot of spatial dependency among the variables and the zones happen to align with actual patterns in the data, this assumption isn't bad. However in reality this is rarely the case.

Table 1. Comparison of zone and surface data for selected locations.

		Map Variable 1		Map Variable 2		Map Variable 3	
Pin#	Zone#	*Phos*_**Zone**	*Phos*_**Surf**	*Pot*_**Zone**	*Pot*_**Surf**	*PH*_**Zone**	*PH*_**Surf**
1	Z1	**55**	**140**	**457**	**584**	**6.4**	**6.9**
2	Z1	**55**	**66**	**457**	**446**	**6.4**	**6.3**
3	Z2	**46**	**64**	**419**	**357**	**6.5**	**6.1**
4	Z3	**42**	**22**	**384**	**194**	**6.8**	**6.3**

Consider the "shish kebab" of data values for the four pins shown in Table 1. The first two pins are in Zone #1 so the assumption is that the levels of phosphorous= 55, potassium= 457 and PH= 6.4 are the same for both locations (as they are for everywhere within Zone #1). But the surface data for Pin #1 indicates a sizable difference from the averages—150% ([[140-55]/55]*100) for phosphorous, 28% for potassium and 8% for PH. The differences are less for Pin #2 with 20%, 2% and –2%, respectively. Pins #3 and #4 are in different zones, but similar deviations from the averages are noted, with the greatest differences in phosphorous levels and the least in PH levels.

While zones might be sufficient for general description and viewing of spatial data, surfaces are needed in most applications to discover spatial relationships. As GIS technology evolves, the traditional thematic mapping and geo-query capabilities will be extended to data mining and knowledge discovery. In this expanded role, map variables are used to relate subtle changes in one variable to those in others. As a result, maps become an active ingredient in enlarging our understanding of things, not just a mechanism to "paint" existing non-spatial heuristics and science. The first step along the path to tomorrow's GIS is incorporating map surfaces and spatial analysis into our geographic psyche and science.

Author's Note: *a more detailed discussion of zones and surfaces is available online at www.innovativegis.com/basis, select Column Supplements.*

(return to top of Topic)

(Back to the Table of Contents)