

Making a Case for *Spatial/STEM*:

A Mathematical Structure for Teaching Fundamental Concepts in Spatial Reasoning, Map Analysis and Modeling

Joseph K. Berry

Adjunct Faculty in Geosciences, Department of Geography, University of Denver
Adjunct Faculty in Natural Resources, Colorado State University
Principal, Berry & Associates // Spatial Information Systems, Fort Collins, Colorado
Email: jkberry@du.edu — Website: www.innovativegis.com/basis/

Contents:

- [Simultaneously Trivializing and Complicating GIS](#) — describes a mathematical structure for spatial analysis operations
 - [Infusing Spatial Character into Statistics](#) — describes a statistical structure for spatial statistics operations
 - [Organizing Geographic Space for Effective Analysis](#) — an overview of data organization supporting grid-based map analysis
 - [To Boldly Go Where No Map Has Gone Before](#) — identifies Lat/Lon as a Universal Spatial Key for joining database tables
- ...www.innovativegis.com/Basis/Courses/SpatialSTEM/ for companion PowerPoint and additional materials
-

Simultaneously Trivializing and Complicating GIS *(GeoWorld, 4/12)*

Several things seem to be coalescing in my mind (or maybe colliding is a better word). GIS has moved up the technology adoption curve from *Innovators* in the 1970s to *Early Adopters* in the 80s, to *Early Majority* in the 90s, to *Late Majority* in the 00s and is poised to capture the *Laggards* this decade. Somewhere along this progression, however, the field seems to have bifurcated along technical and analytical lines.

The lion's share of the growth has been GIS's ever expanding capabilities as a "*technical tool*" for corralling vast amounts of spatial data and providing near instantaneous access to remote sensing images, GPS navigation, interactive maps, asset management records, geo-queries and awesome displays. In just forty years, GIS has morphed from boxes of cards passed to a megabuck mainframe, to today's sizzle of a 3D fly-by of terrain anywhere in the world with back-dropped imagery and semi-transparent map layers draped on top— all pushed from the cloud to a GPS enabled tablet or smart phone. What a ride!

However, GIS as an "*analytical tool*" hasn't experienced the same meteoric rise— in fact it might be argued that the analytic side of GIS has somewhat stalled over the last decade. I suspect that in large part this is due to the interests, backgrounds, education and excitement of the ever enlarging GIS tent. Several years ago (see figure 1 and author's note 1) I described the changes in breadth and depth of the community as flattening from the 1970s through the 2000s. By sheer numbers, the balance point has been shifting to the right toward general and public users with commercial systems responding to market demand for more technological advancements.

The 2010s will likely see billions of general and public users with the average depth of science and technology knowledge supporting GIS nearly "flatlining." Success stories in quantitative map analysis and modeling applications have been all but lost in the glitz n' flash of the technological

whirlwind. The vast potential of GIS to change how society perceives maps, mapped data and their use in spatial reasoning and problem solving seems relatively derailed.

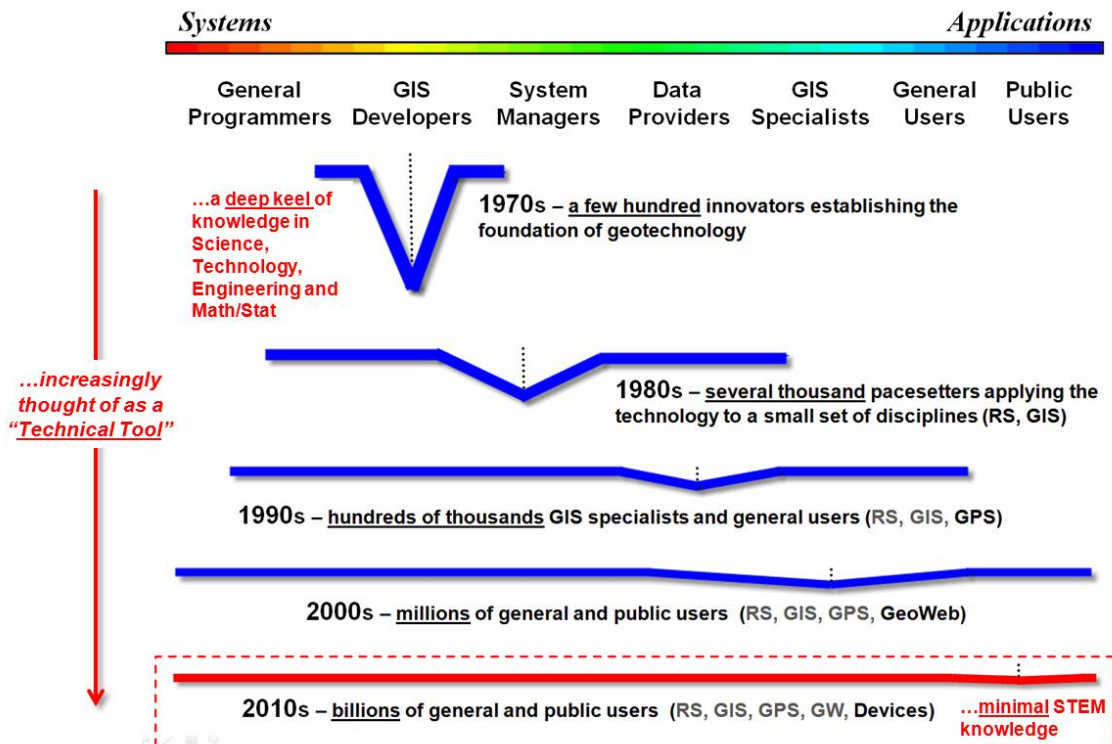


Figure 1. Changes in breadth and depth of the community.

In a recent editorial in *Science* entitled *Trivializing Science Education*, Editor-in-Chief Bruce Alberts laments that “Tragically, we have managed to simultaneously trivialize and complicate science education” (author’s note 2). A similar assessment might be made for GIS education. For most students and faculty on campus, GIS technology is simply a set of highly useful apps on their smart phone that can direct them to the cheapest gas for tomorrow’s ski trip and locate the nearest pizza pub when they arrive. Or it is a Google fly-by of the beaches around Cancun. Or a means to screen grab a map for a paper on community-based conservation of howler monkeys in Belize.

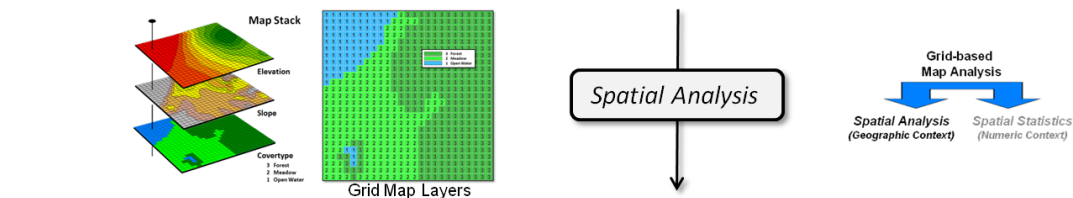
To a smaller contingent on campus, it is career path that requires mastery of the mechanics, procedures and buttons of extremely complex commercial software systems for acquiring, storage, processing, and display spatial information. Both perspectives are valid. However neither fully grasps the radical nature of the digital map and how it can drastically change how we perceive and infuse spatial information and reasoning into science, policy formation and decision-making— in essence, how we can “think with maps.”

A large part of missing the mark on GIS’s full potential is our lack of reaching out to the larger science, technology, engineering and math/stat (STEM) communities on campus by insisting 1) that non-GIS students interested in understanding map analysis and modeling must be tracked into GIS courses that are designed for GIS specialists, and 2) that the material presented primarily focuses on commercial GIS software mechanics that GIS-specialists need to know to function in the workplace.

Much of the earlier efforts in structuring a framework for quantitative map analysis has focused on how the analytical operations work within the context of *Focal*, *Local* and *Zonal* classification by Tomlin, or even my own *Reclassify*, *Overlay*, *Distance* and *Neighbors* classification scheme (see top portion of figure 2 and author's note 3). The problem with these structuring approaches is that most STEM folks just want to understand and use the analytical operations properly— not to appreciate the theoretical geographic-related elegance, or to code an algorithm.

The bottom portion of figure 2 outlines restructuring of the basic spatial analysis operations to align with traditional mathematical concepts and operations (author's note 4). This provides a means for the STEM community to jump right into map analysis without learning a whole new lexicon or an alternative GIS-centric mindset.

GIS as “Technological Tool” (*Where is What*) vs. “Analytical Tool” (*Why, So What and What if*)



GIS Perspective:

Reclassify (*Position, Value, Size, Shape, Contiguity*)

Overlay (*Location-specific, Region-wide*)

Distance (*Distance, Proximity, Movement, Optimal Path, Visual Exposure*)

Neighbors (*Characterizing Surface Configuration, Summarizing Values*)

Map Analysis Toolbox



Mathematical Perspective:

Basic GridMath & Map Algebra (+ - * /)

Advanced GridMath (*Math, Trig, Logical Functions*)

Map Calculus (*Spatial Derivative, Spatial Integral*)

Map Geometry (*Euclidian Proximity, Narrowness, Effective Proximity*)

Plane Geometry Connectivity (*Optimal Path, Optimal Path Density*)

Solid Geometry Connectivity (*Viewshed, Visual Exposure*)

✓ **Unique Map Analytics** (*Contiguity, Size/Shape/Integrity, Masking, Profile*)

Figure 2. Alternative frameworks for quantitative map analysis involving mathematical operators.

For example, the GIS concept/operation of *Slope*= spatial “derivative”, *Zonal functions*= spatial “integral”, *Eucdistance*= extension of “planimetric distance” and the Pythagorean Theorem to proximity, *Costdistance*= extension of distance to effective proximity considering absolute and relative barriers to movement that is not possible in non-spatial mathematics, and *Viewshed*= “solid geometry connectivity”.

Figure 3 outlines the conceptual development of three of these operations. The upper-right graphic identifies the *Calculus Derivative* as a measure of how a mathematical function changes as its input changes by assessing the slope along a curve in 2-dimensional abstract space— calculated as the “slope of the tangent line” at any location along the curve.

In an equivalent manner, the *Spatial Derivative* creates a slope map depicting the rate of change of a continuous map variable in 3-dimensional geographic space—calculated as the slope of the “best fitted plane” at any location along the map surface.

Advanced Grid Math — Math, Trig, Logical Functions
Map Calculus — Spatial Derivative, Spatial Integral

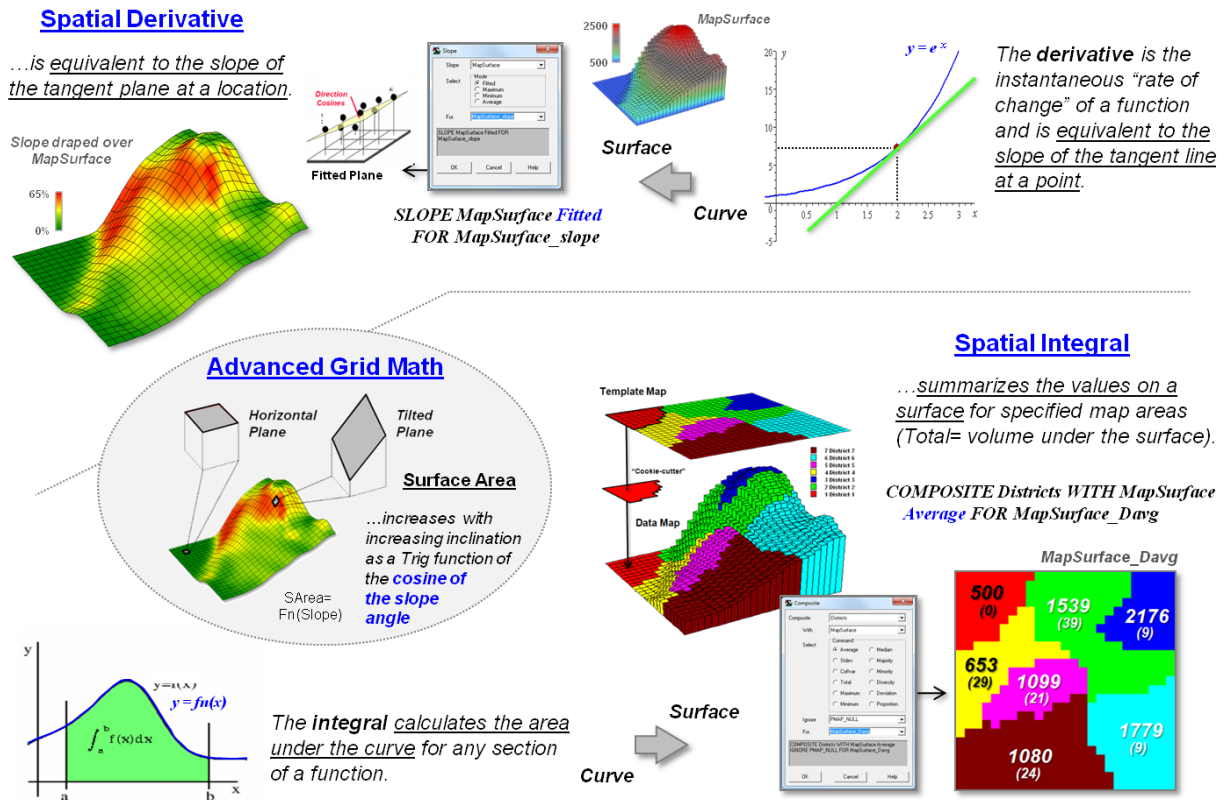


Figure 3. Conceptual extension of derivative, trigonometric functions and integral to mapped data and map analysis operations.

Similarly, *Advanced Grid Math* includes most of the buttons on a scientific calculator to include trigonometric functions. One can add, subtract, multiply, divide and even exponentiation modern maps because they “are numbers first and foremost, pictures later.” For example, calculating the “cosine of the slope values” along a terrain surface and then multiplying this value times the planimetric surface area of a grid cell solves for the increased real-world surface area of the “inclined plane” at each grid location (center inset).

In the lower-left portion of the figure the *Calculus Integral* is identified as the “area of a region under a curve” expressing a mathematical function. The *Spatial Integral* counterpart summarizes map surface values within specified geographic regions. The data summaries are not limited to just a total but can be extended to most statistical metrics. For example, the average map surface value can be calculated for each district in a project area. Similarly, the coefficient of variation ((Stdev / Average) * 100) can be calculated to assess data dispersion about the average for each of the regions.

By recasting GIS concepts and operations of map analysis within the general scientific language of math/stat we can more easily educate tomorrow's movers and shakers in other fields in "spatial reasoning"—to think of maps as "mapped data" and express the wealth of quantitative analysis thinking they already understand on spatial variables.

Innovation and creativity in spatial problem solving is being held hostage to a trivial mindset of maps as pictures and a non-spatial mathematics that presuppose mapped data can be collapsed to a single central tendency value that ignores the spatial variability inherent in the data. Simultaneously, the "build it (GIS) and they will come (and take our existing courses)" educational paradigm is not working as it requires potential users to become a GIS'perts in complicated software systems. GIS must take an active leadership role in "leading" the STEM community to the similarities/differences and advantages/disadvantages in the quantitative analysis of mapped data—there is little hope that the STEM folks will make the move on their own. In the next installment we'll consider recasting spatial statistics concepts and operations into a traditional statistics framework.

Author's Notes: 1) see "A Multifaceted GIS Community" in Topic 27, *GIS Evolution and Future Trends in the online book Beyond Mapping III*, posted at www.innovativegis.com. 2) Bruce Alberts in *Science*, 20 January 2012., 335:6066 p. 263. 3) see "An Analytical Framework for GIS Modeling" posted at www.innovativegis.com/basis/Papers/Other/GISmodelingFramework/. 4) see "SpatialSTEM: Extending Traditional Mathematics and Statistics to Grid-based Map Analysis and Modeling" posted at www.innovativegis.com/basis/Papers/Other/SpatialSTEM/.

Infusing Spatial Character into Statistics (GeoWorld, 5/12)

The previous section discussed the assertion that we might be simultaneously trivializing and complicating GIS. At the root of the argument was the contention that "innovation and creativity in spatial problem solving is being held hostage to a trivial mindset of maps as pictures and a non-spatial mathematics that presupposes that mapped data can be collapsed into a single central-tendency value ignoring the spatial variability inherent in data.

The discussion described a mathematical framework that organizes the spatial analysis toolbox into commonly understood mathematical concepts and procedures. The following discussion does a similar translation to describe a statistical framework for organizing the spatial statistics toolbox into commonly understood statistical concepts and procedures. But first we need to clarify the differences between spatial analysis and spatial statistics. *Spatial analysis* can be thought of as an extension of traditional mathematics involving the "contextual" relationships within and among mapped data layers. It focuses on geographic associations and connections, such as relative positioning, configurations and patterns among map locations.

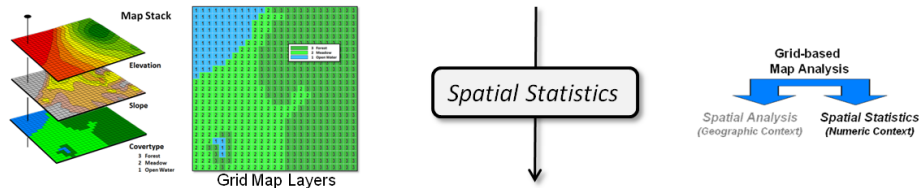
Spatial statistics, on the other hand, can be thought of as an extension of traditional statistics involving the "numerical" relationships within and among mapped data layers. It focuses on 1) mapping the variation inherent in a data set rather than characterizing its central tendency and 2) summarizing the coincidence and correlation of the spatial distributions.

The top portion of figure 1 identifies the two dominant GIS perspectives of spatial statistics—*Surface Modeling* that derives a continuous spatial distribution of a map variable from point sampled data and *Spatial Data Mining* that investigates numerical relationships of map variables.

The bottom portion of the figure outlines restructuring of the basic spatial statistic operations to align with traditional non-spatial statistical concepts and operations (see author's note). The first three

groupings are associated with general descriptive statistics, the middle two involve unique spatial statistics operations and the final two identify classification and predictive statistics.

GIS as “Technological Tool” (*Where is What*) vs. “Analytical Tool” (*Why, So What and What if*)



GIS Perspective:

Surface Modeling (*Density Analysis, Spatial Interpolation, Map Generalization*)

Spatial Data Mining (*Descriptive, Predictive, Prescriptive*)

Map Analysis Toolbox



Statistical Perspective:

Basic Descriptive Statistics (*Min, Max, Median, Mean, StDev, etc.*)

Basic Classification (*Reclassify, Binary/Ranking/Rating Suitability*)

Map Comparison (*Joint Coincidence, Statistical Tests*)

✓ **Unique Map Descriptive Statistics** (*Roving Window Summaries*)

✓ **Surface Modeling** (*Density Analysis, Spatial Interpolation*)

Advanced Classification (*Map Similarity, Maximum Likelihood, Clustering*)

Predictive Statistics (*Map Correlation/Regression, Data Mining Engines*)

Figure 1. Alternative frameworks for quantitative map analysis involving statistical operators.

Figure 2 depicts the non-spatial and spatial approaches for characterizing the distribution of mapped data and the direct link between the two representations. The left side of the figure illustrates non-spatial statistics analysis of an example set of data as fitting a standard normal curve in “data space” to assess the central tendency of the data as its average and standard deviation. In processing, the geographic coordinates are ignored and the typical value and its dispersion are assumed to be uniformly (or randomly) distributed in “geographic space.”

The top portion of figure 2 illustrates the derivation of a continuous map surface from geo-registered point data involving spatial autocorrelation. The discrete point map locates each sample point on the XY coordinate plane and extends these points to their relative values (higher values in the NE; lowest in the NW). A roving window is moved throughout the area that weight-averages the point data as an inverse function of distance— closer samples are more influential than distant samples. The effect is to fit a surface that represents the geographic distribution of the data in a manner that is analogous to fitting a SNV curve to characterize the data’s numeric distribution. Underlying this process is the nature of the sampled data which must be numerically quantitative (measurable as continuous numbers) and geographically isopleth (numbers form continuous gradients in space).

The lower-right portion of figure 2 shows the direct linkage between the numerical distribution and the geographic distribution views of the data. In geographic space, the “typical value” (average) forms a horizontal plane implying that the average is everywhere. In reality, the average is hardly anywhere and the geographic distribution denotes where values tend to be higher or lower than the average.

In data space, a histogram represents the relative occurrence of each map value. By clicking anywhere on the map, the corresponding histogram interval that location's value is highlighted; conversely, clicking anywhere on the histogram highlights all of the corresponding map values within the interval. By selecting all locations with values greater than + 1SD, areas of unusually high values are located— a technique requiring the direct linkage of both numerical and geographic distributions.

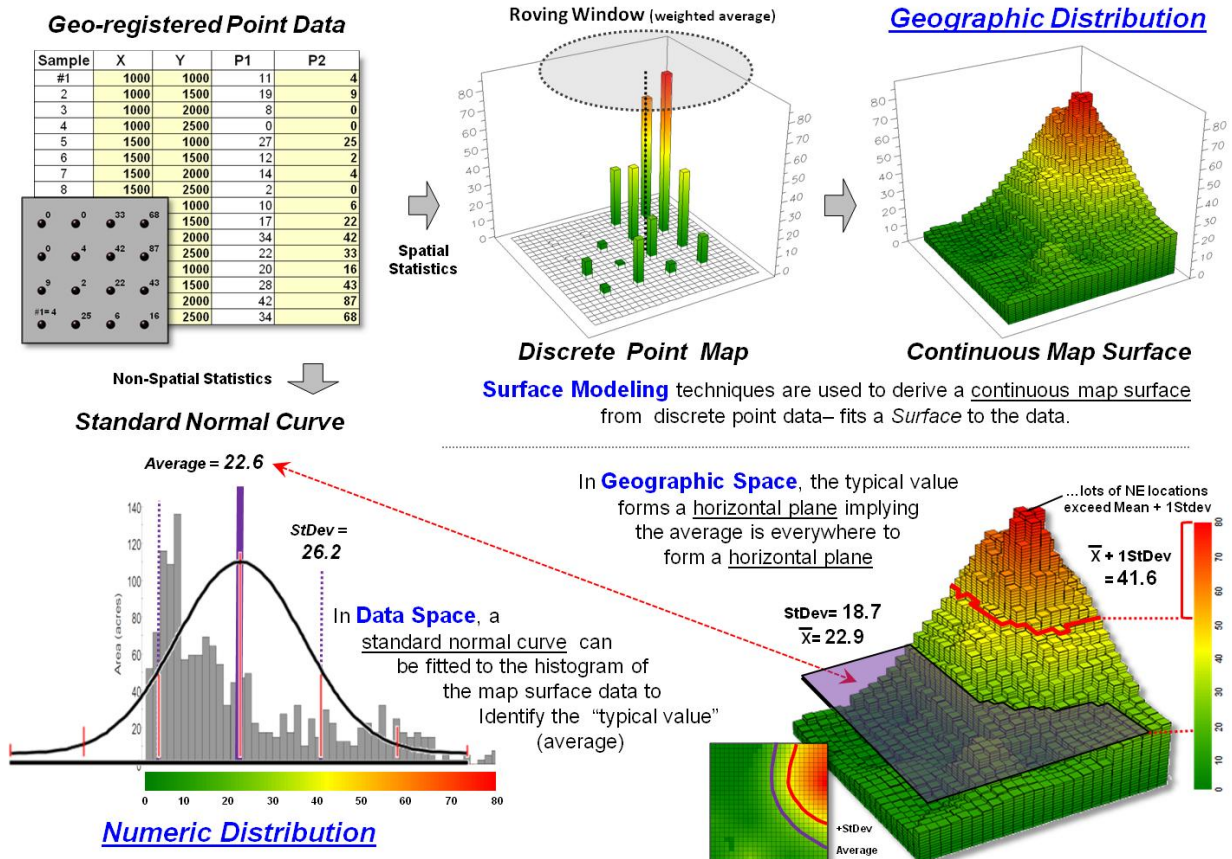


Figure 2. Comparison and linkage between spatial and non-spatial statistics

Figure 3 shows two examples of the advanced spatial statistics operations involving spatial relationships among two or more map layers. The top portion of the figure uses *map clustering* to identify the location of inherent groupings of elevation and slope data by assigning pairs of values into groups (called clusters) so that the value pairs in the same cluster are more similar to each other than to those in other clusters.

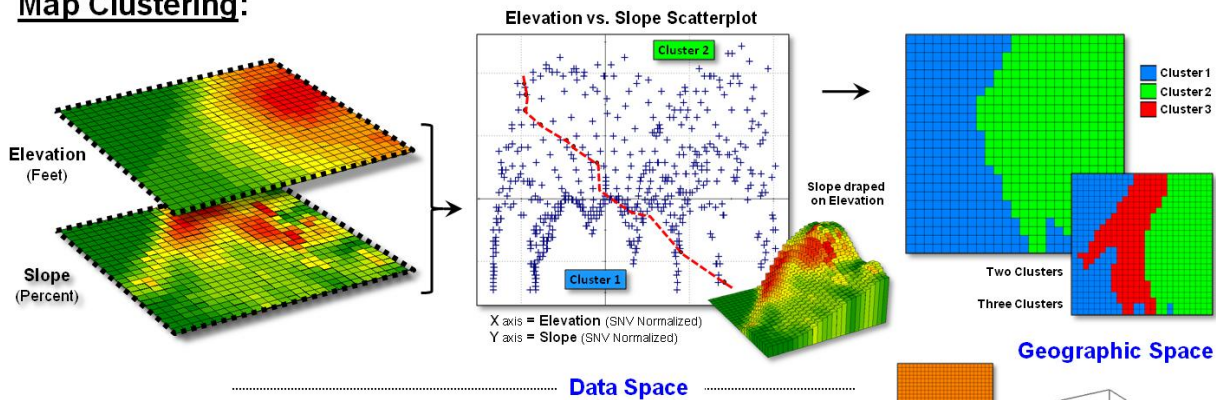
The bottom portion of the figure assesses *map correlation* by calculating the degree of dependency among the same maps of elevation and slope. Spatially "aggregated" correlation involves solving the standard correlation equation for the entire set of paired values to represent the overall relationship as a single metric. Like the statistical average, this value is assumed to be uniformly (or randomly) distributed in "geographic space" forming a horizontal plane.

"Localized" correlation, on the other hand, maps the degree of dependency between the two map variables by successively solving the standard correlation equation within a roving window to generate a continuous map surface. The result is a map representing the geographic distribution of

the spatial dependency throughout a project area indicating where the two map variables are highly correlated (both positively, red tones, and negatively, green tones) and where they have minimal correlation (yellow tones).

With the exception of unique Map Descriptive Statistics and Surface Modeling classes of operations, the grid-based map analysis/modeling system simply acts as a mechanism to spatially organize the data. The alignment of the geo-registered grid cells is used to partition and arrange the map values into a format amenable for executing commonly used statistical equations. The critical difference is that the answer often is in map form indicating where a statistical relationship is more or less than typical.

Map Clustering:



Map Correlation:

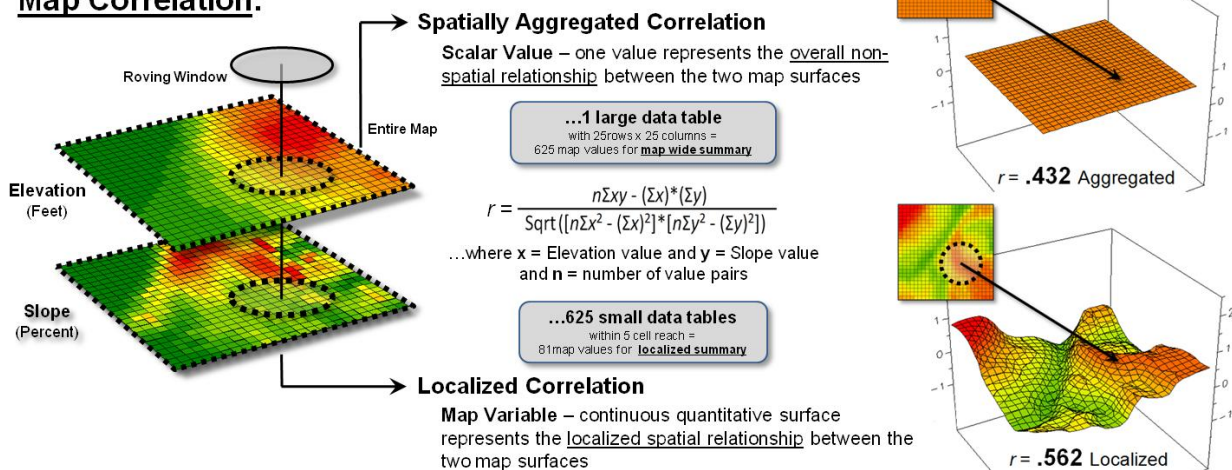


Figure 3. Conceptual extension of clustering and correlation to mapped data and analysis.

While the technological applications of GIS have soared over the last decade, the analytical applications seem to have flat-lined. The seduction of near instantaneous geo-queries and awesome graphics seem to be masking the underlying character of mapped data— that maps are numbers first, pictures later. However, grid-based map analysis and modeling involving Spatial Analysis and Spatial Statistics is, for the larger part, simply extensions of traditional mathematics and statistics. The recognition by the GIS community that quantitative analysis of maps is a reality and the recognition by the STEM community that spatial relationships exist and are quantifiable should be the glue that binds the two perspectives. That reminds me of a very wise observation about technology evolution—

“Once a new technology rolls over you, if you're not part of the steamroller, you're part of the road.”
-- Stewart Brand, editor of the Whole Earth Catalog

Author's Notes: For a more detailed discussion, see “SpatialSTEM: Extending Traditional Mathematics and Statistics to Grid-based Map Analysis and Modeling” posted at www.innovativegis.com/basis/Papers/Other/SpatialSTEM/ containing a comprehensive appendix of URL links to over 125 additional readings on the grid-based map analysis/modeling concepts, terminology, considerations and procedures described in this presentation and royalty-free teaching materials.

Organizing Geographic Space for Effective Analysis (GeoWorld, 9/12)

A basic familiarity of the two fundamental data types supporting geotechnology—vector and raster—is important for understanding map analysis procedures and capabilities (see author's note). Vector data is closest to our manual mapping heritage and is familiar to most users as it characterizes geographic space as *collection of discrete spatial objects* (points, lines and polygons) that are easy to draw. Raster data, on the other hand, describes geographic space as a *continuum of grid cell values* (surfaces) that while easy to conceptualize, requires a computer to implement.

Generally speaking, vector data is best for traditional map display and geo-query—“*where is what*,” applications that identify existing conditions and characteristics, such as “where are the existing gas pipelines in Colorado” (a descriptive query of existing information). Raster data is best for advanced graphics and map analysis—“*why, so what and what if*” applications that analyze spatial relationships and patterns, such as “where is the best location for a new pipeline” (a prescriptive model deriving new information).

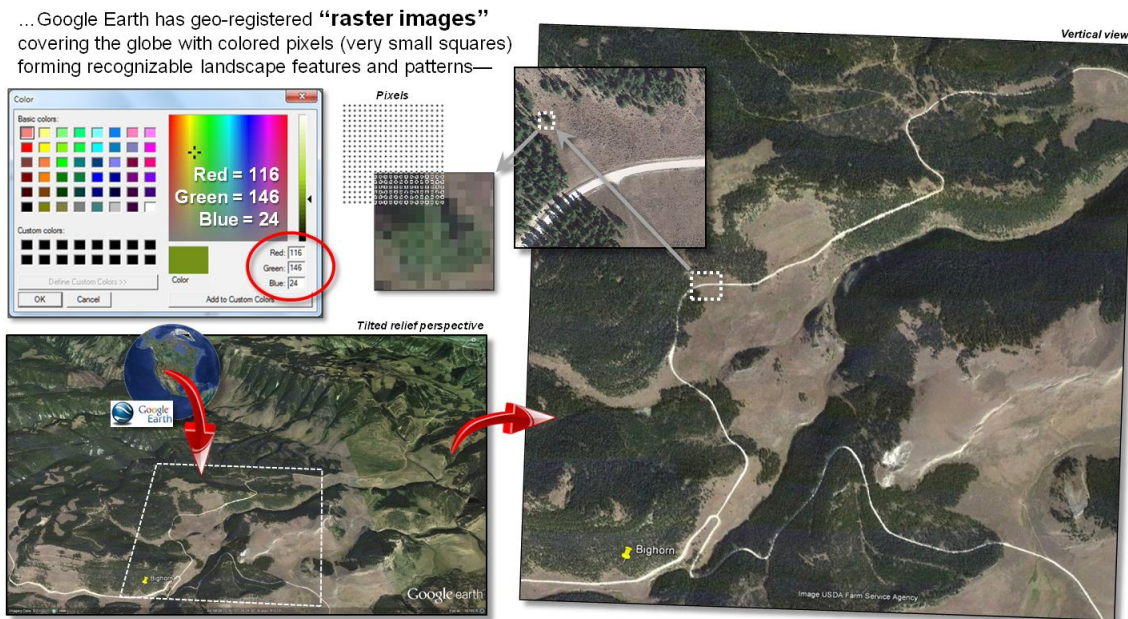


Figure 1. A raster image is composed of thousands of numbers identifying different colors for the “pixel” locations in a rectangular matrix supporting visual interpretation.

Most vector applications involve the extension of manual mapping and inventory procedures that take advantage of modern computers’ storage, speed and Internet capabilities (better ways to do things). Raster applications, however, tend to involve entirely new paradigms and procedures for

visualizing and analyzing mapped data that advances innovative science (entirely new ways to do things).

On the advanced graphics front, the lower-left portion of figure 1 depicts an interactive Google Earth display of an area in northern Wyoming's Bighorn Mountains showing local roads superimposed on an aerial image draped over a 3D terrain perspective. The roads are stored in vector format as an interconnecting set of line features (vector). The aerial image and elevation relief are stored as numbers in geo-referenced matrices (raster).

The positions in a raster image matrix are referred to as "pixels," short for picture elements. The value stored at each pixel corresponds to a displayed color as a combination of red, green and blue hues. For example, the green tone for some of the pixels portraying the individual tree in the figure is coded as red= 116, green= 146 and blue= 24. Your eye detects a greenish tone with more green than red and blue. In the tree's shadow toward the northwest the red, green and blue levels are fairly equally low (dark grey). In a *raster image* the objective is to generate a visual graphic of a landscape for visual interpretation.

A *raster grid* is a different type of raster format where the values indicate characteristics or conditions at each location in the matrix designed for quantitative map analysis (spatial analysis and statistics). The elevation surface used to construct a tilted relief perspective in a Google Earth display is composed of thousands of matrix values indicating the undulating terrain gradient.

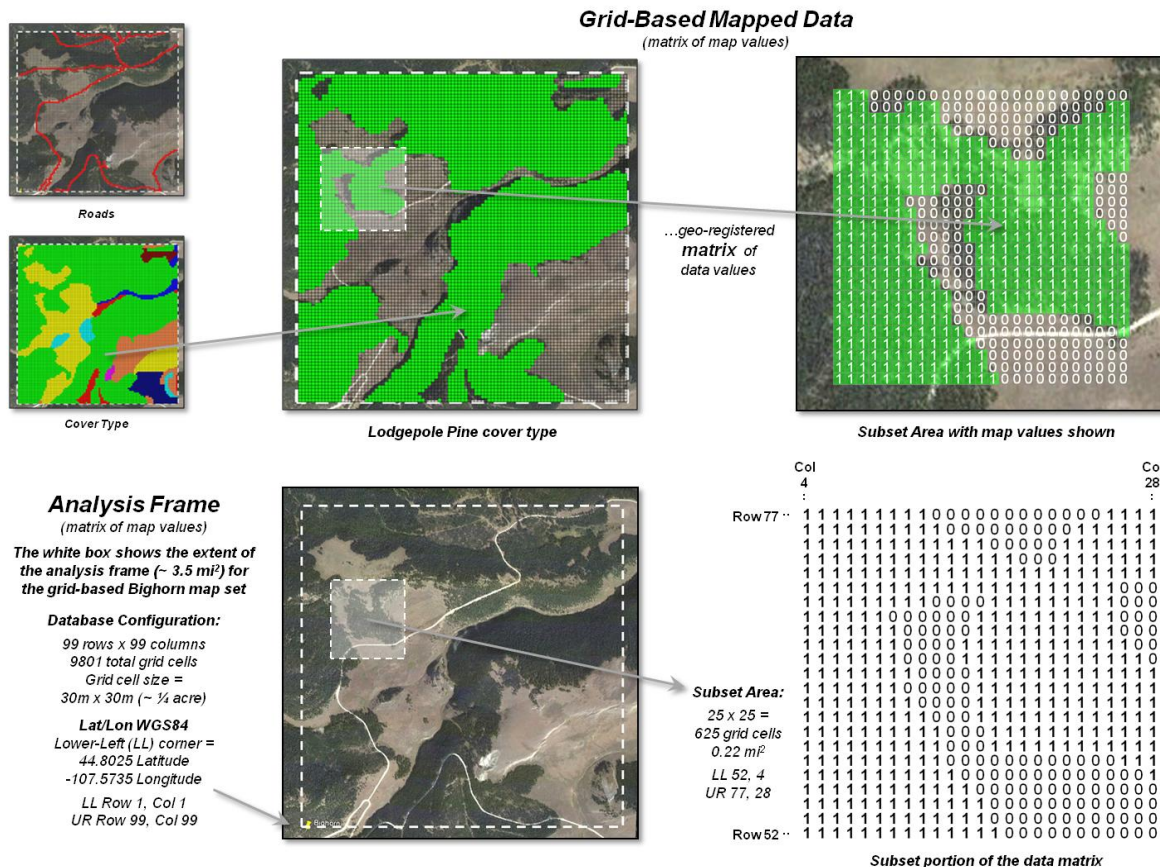


Figure 2. A raster grid contains a map values for each "grid cell" identifying the characteristic/condition at that location supporting quantitative analysis.

Figure 2 depicts a raster grid of the vegetation in the Bighorn area by assigning unique classification values to each of the cover types. The upper portion of the figure depicts isolating just the Lodgepole Pine cover type by assigning 0 to all of the other cover types and displaying the stored matrix values for a small portion of the project area. While you see the assigned color in the grid map display (green in this example), keep in mind that the computer “sees” the stored matrix of map values.

The lower portion of the figure 2 identifies the underlying organizational structure of geo-registered map data. An “analysis frame” delineates the geographic extent of the area of interest and in the case of raster data the size of each pixel/grid element. In the example, the image *pixel size* for the visual backdrop is less than a foot comprising well over four million values and the *grid cell size* for analysis is 30 meters stored as a matrix with 99 columns and 99 rows totally nearly 10,000 individual cell locations.

For geo-referencing, the lower-left grid cell is identified as the matrix’s origin (column 1, row1) and is stored in decimal degrees of latitude and longitude along with other configuration parameters as a few header lines in the file containing the matrix of numbers. In most instances, the huge matrix of numbers is compressed to minimize storage but uncompressed on-the-fly for display and analytical processing.

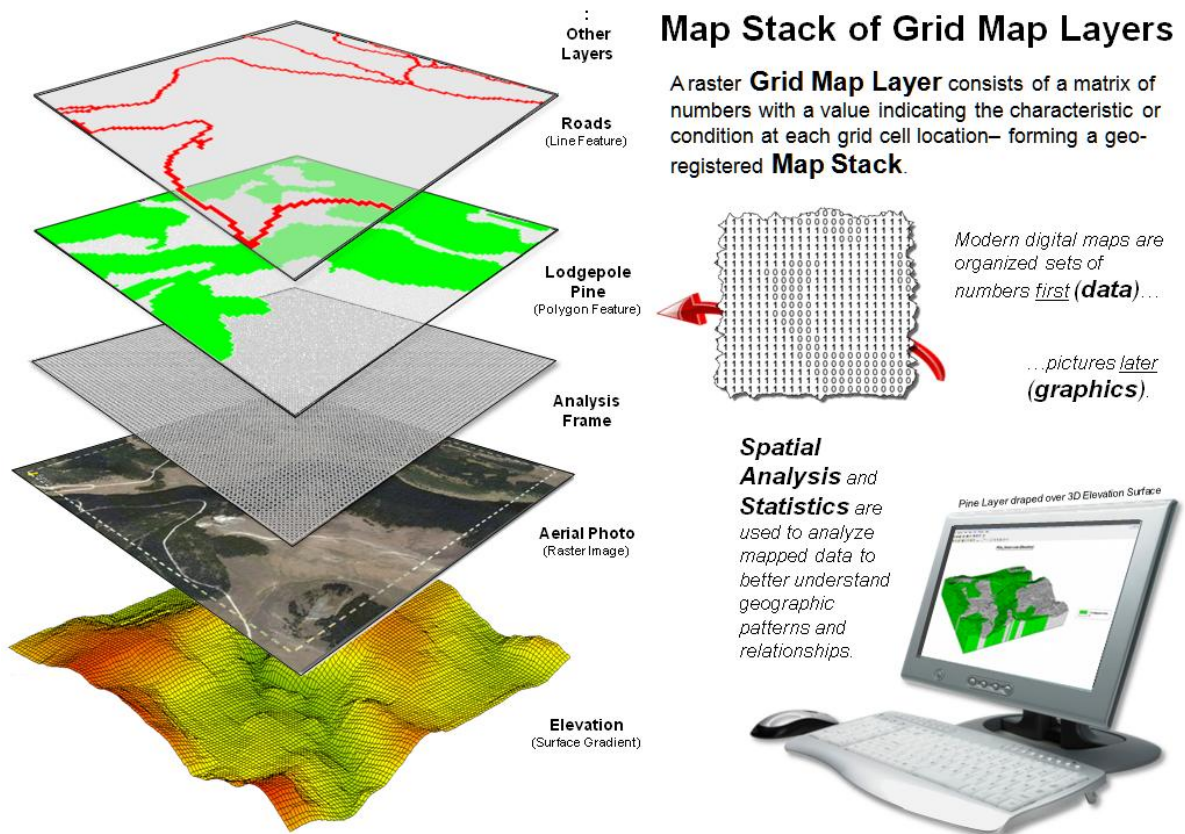


Figure 3. A set of geo-registered map layers forms a “map stack” organized as thousands upon thousands of numbers within a common “analysis frame.”

Figure 3 illustrates a broader level of organization for grid-based data. Within this construct, each grid map layer in a geographically registered analysis frame forms a separate theme, such as roads, cover type, image and elevation. Each point, line and polygon *map feature* is identified as a grid cell grouping having a unique value stored in implied matrix characterizing a discrete spatial variable. A *surface gradient*, on the other hand, is composed of fluctuating values that track the uninterrupted increases/decreases of a continuous spatial variable.

The entire set of grid layers available in a database is termed a *map stack*. In map analysis, the appropriate grid layers are retrieved, their values map-atically processed and the resulting matrix stored in the stack as a new layer—in the same manner as one solves an algebraic equation, except that the variables are entire grid maps composed of thousands upon thousands of geographically organized numbers.

The major advantages of grid-based maps are their inherently uncomplicated data structure and consistent parsing within a holistic characterization of geographic space—just the way computers and math/stat mindsets like it. No sets of irregular spatial objects scattered about an area that are assumed to be completely uniform within their interiors... rather, continuously defined spatial features and gradients that better align with geographic reality and, for the most part, with our traditional math/stat legacy.

The next section's discussion builds on this point by extending grid maps and map analysis to “a universal key” for unlocking spatial relationships and patterns within standard database and quantitative analysis approaches and procedures.

Author's Notes: For a more detailed discussion of vector and raster data types and important considerations, see Topic 18, “Understanding Grid-based Data” in the online book *Beyond Mapping III* posted at www.innovativegis.com/basis/MapAnalysis/.

To Boldly Go Where No Map Has Gone Before (GeoWorld, 10/12)

The previous sections have described a mathematical framework— dare I say a “map-ematical” framework— for quantitative analysis of mapped data (see Author's Notes). Recall that *Spatial Analysis* operations investigate the “contextual” relationships within and among maps, such as variable-width buffers that account for intervening conditions. *Spatial Statistics* operations, on the other hand, examine the “numerical” relationships, such as map clustering to uncover inherent geographic patterns in the data.

The cornerstone of these capabilities lies in the grid-based nature of the data that treats geographic space as continuous map surfaces composed of thousands upon thousands of cells with each containing data values that identify the characteristics/conditions occurring at each location. This simple matrix structure provides a detailed account of the unique spatial distribution of each map variable and a geo-registered stack of map layers provides the foothold to quantitatively explore their spatial patterns and relationships.

The most fundamental and ubiquitous grid form is the Latitude/Longitude coordinate system that enables every location on the Earth to be specified by a pair of numbers. The upper portion of figure 1, depicts a 2.5⁰ Lat/Lon grid forming a matrix of 73 rows by 144 columns= 10,512 cells in total with each cell having an area of about 18,735mi².

The lower portion of the figure shows that the data could be stored in Excel with each spreadsheet cell directly corresponding to a geographic grid cell. In turn, additional map layers could be stored as separate spreadsheet pages to form a map stack for analysis.

Of course this resolution is far too coarse for most map analysis applications, but it doesn't have to be. Using the standard single precision floating point storage of Lat/Long coordinates expressed in decimal degrees, the precision tightens to less than half a foot anywhere in the world ($365214 \text{ ft/degree} * 0.000001 = .365214 \text{ ft} * 12 = 4.38257 \text{ inches}$ or 0.11132 meters). However, current grid-based technology limits the practical resolution to about 1m (e.g., Ikonos satellite images) to 10m (e.g., Google Earth) due to the massive amounts of data storage required.

For example, to store a 10m grid for the state of Colorado it would take over two and half billion grid cells ($26,960\text{km}^2 = 269,601,000,000\text{m}^2 / 100\text{m}^2 \text{ per cell} = 2,696,010,000 \text{ cells}$). To store the entire earth surface it would take nearly a trillion and a half cells ($148,300,000\text{km}^2 = 148,000,000,000,000\text{m}^2 / 100\text{m}^2 \text{ per cell} = 1,483,000,000,000 \text{ cells}$).

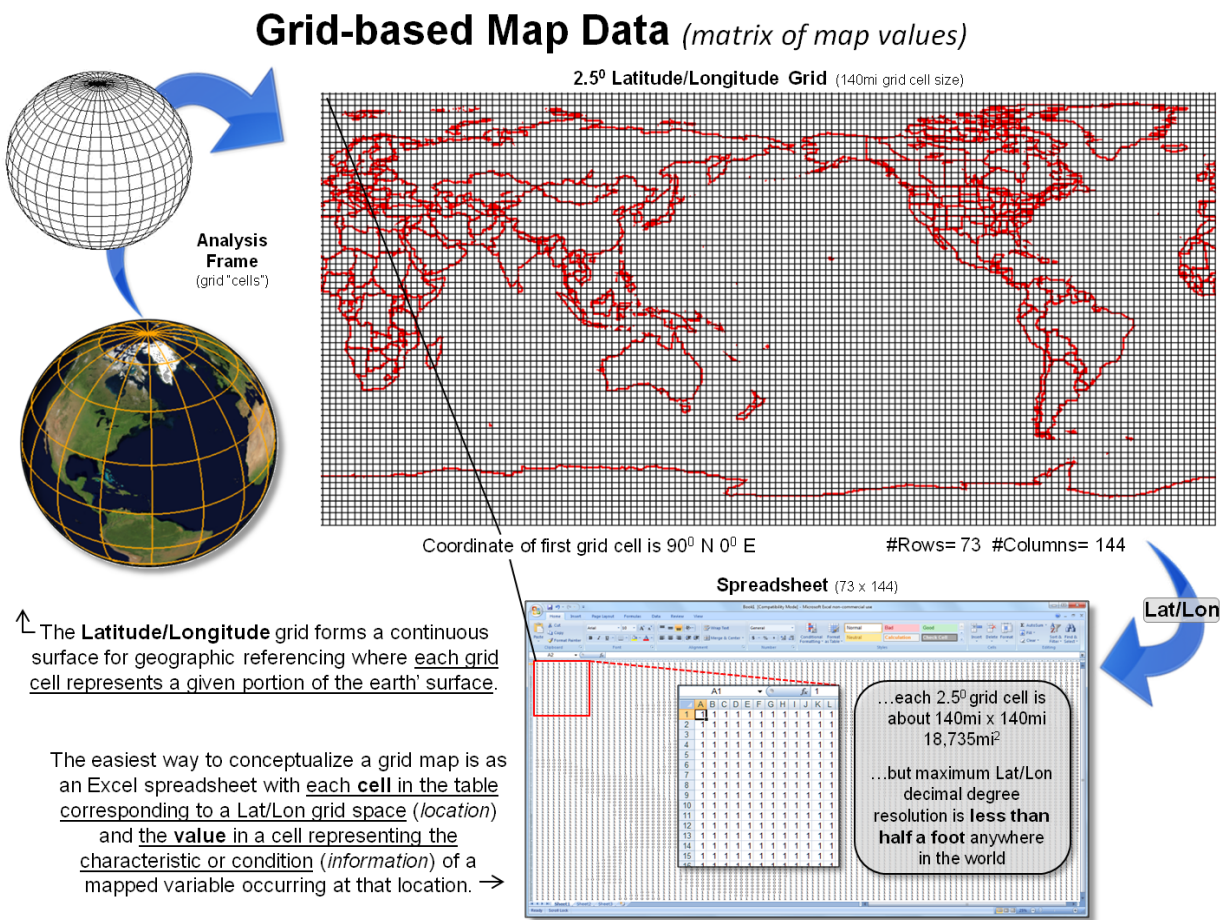


Figure 1. Latitude and Longitude coordinates provide a framework for parsing the earth's surface into a standardized set of grid cells.

At first these storage loads seem outrageous but with distributed cloud computing the massive grid can be "easily" broken into manageable mouthfuls. A user selects an area of interest and data for that area is downloaded and stitched together. For example, Google Earth responds to your screen interactions to nearly instantaneously download millions of pixels, allowing you to pan/zoom and

turn on/off map layers that are just a drop in the bucket of the trillions upon trillions of pixels and grid data available in the cloud.

Figure 2 identifies another, more practical mechanism for storage using a relational database. In essence, each of the conceptual grid map spreadsheets can be converted to an interlaced format with a long string of numbers forming the columns (data fields); the rows (records) identify the information available each of the individual grid cells that form the reference grid.

For fairly small areas of up to a million or so cells this is an excellent way to store grid maps as their spatial coincidence is inherent in the organization and the robust standard set of database queries and processing operations is available. Larger grids use more advanced, specialized mechanisms of storage to facilitate data compression and virtual paging of fully configured grid layers.

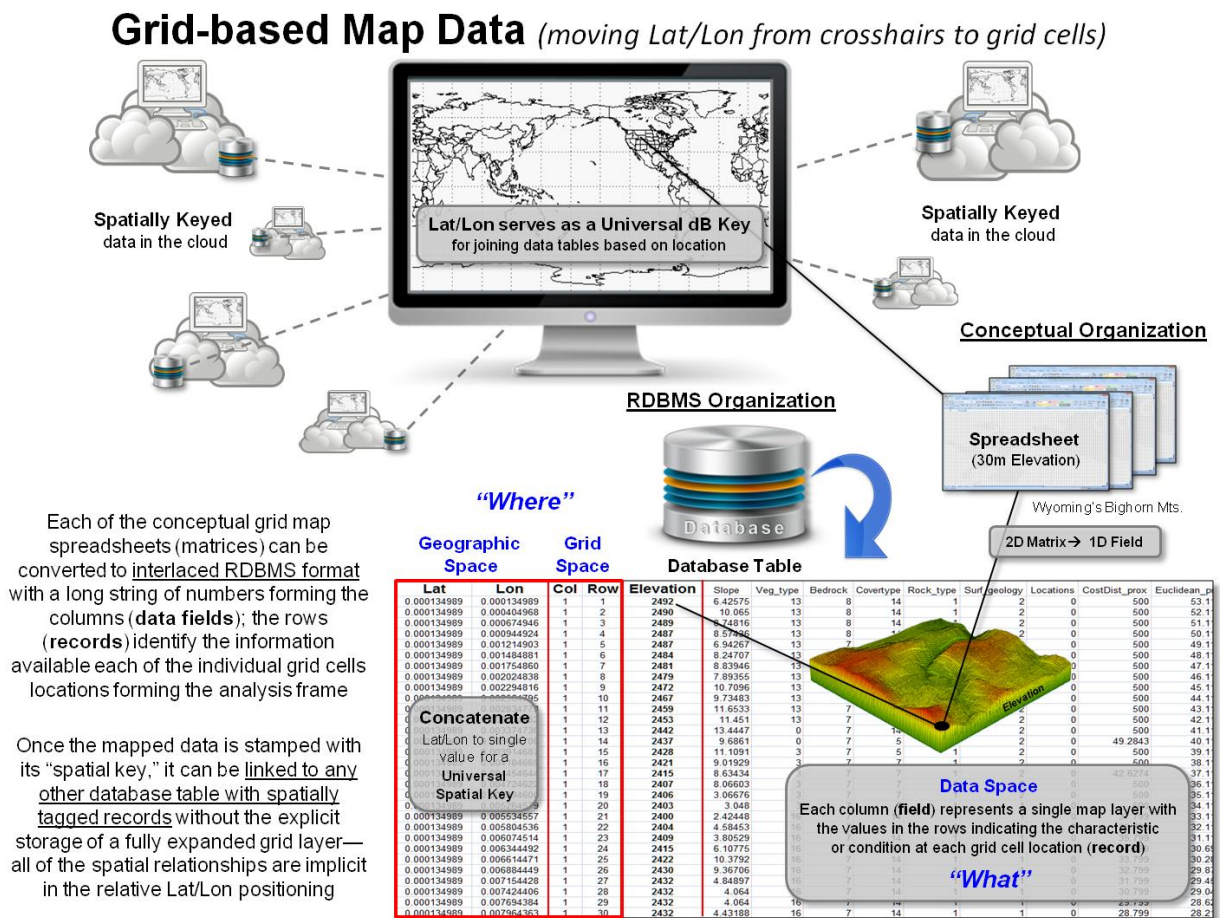


Figure 2. Within a relational database, Lat/Lon forms a Universal DBMS Key for joining tables.

But the move to a relational database structure is far more important than simply corralling megagulps of map values. It provides a “Universal DBMS Key” that can link seemingly otherwise disparate database tables. The process is similar to a date/time stamp, except the “where information” provides a spatial context for joining data sets. Demographic records can be linked to resource records that in turn can be linked to business records, health records, etc—all sharing a common Lat/Lon address.

All that is necessary is to tag your data with its Lat/Lon coordinates (“where” it was collected) just as you do with the date/time (“when” it was collected) ...not a problem with the ubiquitous availability and increasing precision of GPS that puts a real-time tool for handling detailed spatial data right in your pocket. In today’s technology, most GPS-enabled smart phones are accurate to a few meters and specialized data collection devices precise to a few centimeters.

Once your data is stamped with its “spatial key,” it can be linked to any other database table with spatially tagged records without the explicit storage of a fully expanded grid layer. All of the spatial relationships are implicit in the relative positioning of the Lat/Lon coordinates.

For example, a selection operation might be to identify of all health records jointly occurring within half a kilometer of locations that have high lead concentrations in the top soil. Or, all of the customer records within five miles of my store; better yet, within a ten-minute drive from a store.

Geotechnology is truly a mega-technology that will forever change how we perceive and process spatial information. Gone are the days of manual measurements and specialized data formats that have driven our mapping legacy. Lat/Lon coordinates move from cross-hairs for precise navigation (intersecting lines) to a continuous matrix of spaces covering the globe for consistent data storage (grid cells). The recognition of a universal spatial key coupled with spatial analysis/statistics procedures and GPS/RS technologies provides a firm foothold “to boldly go where no map has gone before.”

Author’s Notes: *Beyond Mapping* columns for December 2004 – March 2005, April – May 2012 and September – October 2012 have been compiled into Topic 24, “Overview of Spatial Analysis and Statistics” in the online book *Beyond Mapping III* posted at www.innovativegis.com/basis/MapAnalysis/. Also see Topic 28, “Spatial Data Mining in Geo-Business,” section on *The Universal Key for Unlocking GIS’s Full Potential* (October 2011 column).

This paper is based on earlier writings that present a combined conceptual framework supporting a SpatialSTEM approach for teaching map analysis and modeling fundamentals within a mathematical/statistical context that resonates with science, technology, engineering and math/stat communities. The papers were originally published in the monthly “Beyond Mapping” column in GeoWorld, April, May and October 2012. Other columns in the Beyond Mapping series since 1989 are posted at www.innovativegis.com/basis/MapAnalysis/ChronList/ChronologicalListing.htm. Additional papers and instructional materials on Geotechnology concepts, procedures, practices and issues are posted at www.innovativegis.com/basis/.

This paper is posted at www.innovativegis.com/basis/Papers/Other/SpatialSTEM/SpatialSTEM_case.pdf. A supporting PowerPoint slide set is posted at www.innovativegis.com/basis/Papers/Other/SpatialSTEM/SpatialSTEM_case.ppt.

For a more detailed discussion, see “SpatialSTEM: Extending Traditional Mathematics and Statistics to Grid-based Map Analysis and Modeling” posted at www.innovativegis.com/basis/Papers/Other/SpatialSTEM/ containing a comprehensive appendix of URL links to over 125 additional readings on the grid-based map analysis/modeling concepts, terminology, considerations and procedures described in this presentation and royalty-free teaching materials.