Position Paper for Scale and Spatial Analytics (SSA):

A SPARC Workshop

Kevin Butler and Charlie Frye
Environmental Systems Research Institute (Esri), 380 New York St., Redlands, CA 92373 USA
kbutler@esri.com, cfrye@esri.com

Challenges and Opportunities in Addressing Issues of Scale in Spatial Analytics

Spatial analytical results are greatly impacted by the sampling design used to collect data, the type of areal unit in which they are enumerated, and most importantly, the extent to which data are aggregated – the well-known but often ignored modifiable areal unit problem in geography. Many spatial processes, such as agricultural yield or population, can only be interpreted in an areal unit. They cannot be meaningfully represented by a point. The choice of the geometric shape (or volume) and size of the areal unit used in a spatial analysis is often chosen arbitrarily or prescribed by a data provider. Arguably the most critical challenge is to identify a spatial unit of analysis, often an areal unit, appropriate to the exploration of the spatial problem under investigation. We contend that current practitioners often don't understand the implications of choosing an areal unit of analysis nor the implications of both intentional and unintentional spatial transformations of that areal unit as part of an analytical workflow. We present a series of commonly used practices related to scale in spatial analysis, illustrate potential challenges inherent in the practice, and highlight some recent advancements in GIS software that mitigate the challenges.

Using regular tessellations on a sphere

The fishnet grid is the predominant tessellation for spatial data representing a continuous potential surface. In its simplest form, this tessellation consists of a matrix of purportedly equal area cells (or pixels) organized into rows and columns (or a grid) where each cell contains a value. The square raster grid structure was adopted primarily for computational reasons related to integer indexing schemes and simplistic storage rather than a deep consideration of spatial analysis.

By definition a raster is planar and immutable. This presents significant challenges when using rasters for large-area analyses as the immutable grid cells cannot align to areas of a sphere. This problem has been exacerbated by the explosion of web mapping applications and portends even greater problems for web based spatial analyses. For example, the Web Mercator projection, (EPSG: 3857) popularized by Google has become the most commonly-used coordinate system for web mapping applications. It is currently used by Google Maps, Bing Maps, and Esri ArcGIS Online basemaps, among others. Again, this square projection became popular for computation reasons and not because it provided a good foundation for analysis.

The problem is so pervasive that the National Geospatial-Intelligence Agency (NGA) has released an advisory notice on the use of the Web Mercator coordinate system in military and intelligence applications. In the advisory, NGA warns about positional accuracy issues with Web Mercator, especially at higher latitudes, and "…reminds the community to use DoD approved World Geodetic System 1984 (WGS 84) applications for all mission critical activities."

Recommendation: Practitioners should be aware of the limitations of using a planimetric data structure to model spatial processes on a sphere or ellipsoid. GIS software should provide the ability for users to visualize and quantify the level of distortion introduced when using a regular tessellation on a sphere akin to Tissot's Indicatrix.

Transformations of regular tessellations on a sphere

One of the most important features of contemporary GIS is their ability to integrate disparate data sets potentially each with different coordinate systems. Commercial off the shelf (COTS) and open-source GIS must transform or translate coordinates from one geographic coordinate system to another for both display and analysis. Automatic on-the-fly projections can often create confusion as to which coordinate system is being used, and this confusion can lead to analytical problems further in the workflow.

In most GIS software, the cell size is converted from one geographic coordinate system to another using a simple linear unit conversion when projecting from a Projected Coordinate System (PCS) to another PCS or an angular unit conversion when going from Geographic Coordinate System (GCS) to another GCS. For example, projecting a 30-meter raster dataset from UTM zone 10N to WGS84 Web Mercator would continue to use 30 meters as the output (transformed) cell size. When projecting from GCS to PCS or vice versa, the common practice is to calculate an average cell size based on the ratios of diagonal lengths in source and destination projections. This is referred to as the *convert units* method, simply copying the cell size from input to output, changing units if necessary (Chatterjee and Tenbrink, 2018). This simple approach has some limitations; it does not account for distortion when projecting from one PCS to another PCS (1 projected meter in, say, a UTM zone at a given location may not cover the same ground distance as 1 projected meter at the same location in an Albers projection) and when computing the ratios of diagonal lengths, the behavior of the projection at only the four corner points is used. This may introduce excessive distortion, depending on the projection and the extent.

More robust methods of transforming regular tessellations have been recently introduced. The *preserve resolution* method ensures that the same number of square cells are preserved in the projected extent as are in the original extent. The output cell size is calculated based on the ratios of the areas of the projected extent to the original extent. The *center of extent* projects the center of the original extent to the output coordinate system. The output cell size is calculated by taking the average of the projected distances from the center point to its four adjacent points.

Recommendation: Practitioners should be aware that transformations (resampling and reprojection) can be introduced implicitly by analytical workflows and should specify the cell size projection method most appropriate to their problem.

Regular tessellations in 3D

As spatial analysis moves to three dimensions, the challenges of scale are exacerbated. Since full volumetric analysis is not available in desktop GIS software, the common practice is to construct a regular tessellation on the surface of the Earth and extend it above or below the surface. Kelly and Savric (2019) describe the process where a "rectangular mesh element is defined by perpendicularly raising a base face of a projected longitude-latitude surface grid above or below the projected surface such that opposite faces of the element are equal and parallel." This approach fails to account for the convergence of Earth radii and results in biased volumetric calculations as depth below or height above the surface increases. Mesh elements calculated on a sphere or ellipsoid have the potential to improve volumetric calculations.

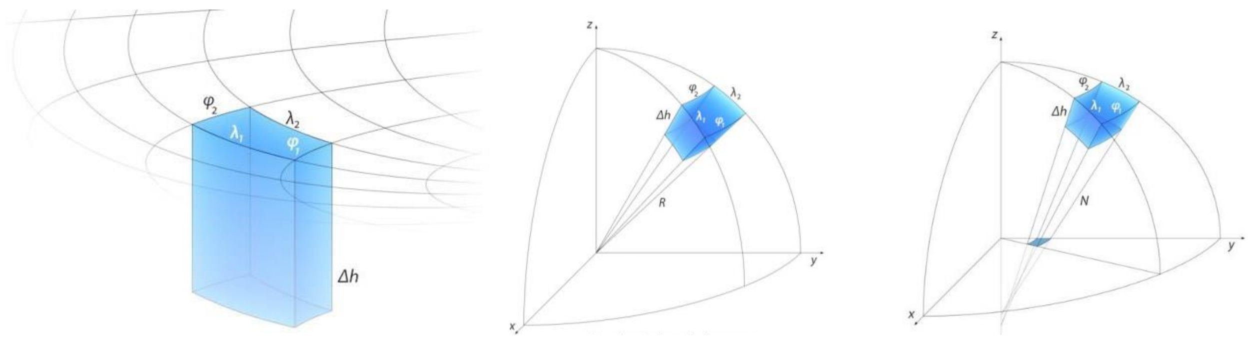

Figure 1 -- Volumetric units constructed on the surface of a plane, sphere and ellipsoid and extended below the surface. Source: Kelly and Šavrič (2019)

Alternative tessellations

Though the square (fishnet) grid is the predominantly used shape type in GIS analysis and thematic mapping, tessellations based on hexagons have well documented advantages (Birch et al., 2007). Hexagons reduce sampling bias due to edge effects of the grid shape, this is related to the low perimeter-to-area ratio of the shape of the hexagon. Hexagons are preferable when your analysis includes aspects of connectivity, movement paths, or local neighborhood analysis due to each hexagon having six equidistant neighbors. In addition, fishnet grids can draw the eye to the straight, unbroken, parallel lines which may inhibit the perception of the underlying patterns in the data. Hexagons tend to break up the lines and allow any curvature of the patterns in the data to be seen more clearly and easily. This breakup of artificial linear patterns also diminishes any orientation bias that can be perceived in fishnet grids. When analyzing a large area, a hexagon grid will suffer less distortion due to the curvature of the earth than the shape of a fishnet grid. Hexagon based tessellations on a global scale are available in the form of Geodesic Discrete Global Grid Systems (GDGGS) (Sahr et al. 2003) and have been adopted by the U.S. EPA for global sampling problems (White et al. 1992) and Uber for efficiently optimizing ride pricing and dispatch (Brodsky, 2019).

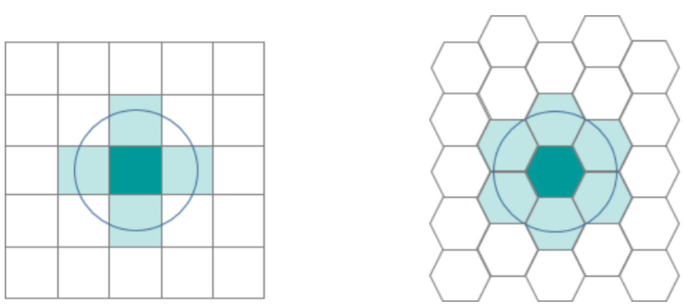

Figure 2 -- Since the edge or length of contact is the same on each side of the hexagon, the centroid of each neighbor is equidistant.

Change of support

The spatial support is the volume, shape, size, and orientation associated with each spatial measurement. A particular challenge is that spatial analysts often have little choice in the areal unit of their data; the unit is often determined by the sampling design of the data. However, COTS and open-source GIS do provide methods for transforming data from one form of spatial support to another (Table 1). However, great caution is required when changing the spatial support as it has significant impacts on statistical inference.

Table 1 -- Methods for Change of Spatial Support (Adapted from Gotway and Young, 2002)

Support of observed data	Desired support for analysis	GIS method
Point	Point	Kriging
Point	Line	Contouring (upscaling)
Point	Area	Interpolation and Zonal Summary
Point	Surface	Interpolation
Area	Point	Ecological inference
Area	Area	Areal interpolation Apportionment

Temporal Scale

Time is increasingly being integrated into spatial analyses. The great need to unit spatial and temporal analysis even led Waldo Tobler to propose a modification to the first law of geography: "… everything is related to everything else, but near and **recent** things are more related than distant things" (Tobler, 2014). Temporal scale is important because the geographical processes and events being analyzed operate at a variety of spatial and temporal scales. Determining a temporal scale of analysis is additionally challenging given that time can be viewed as linear or cyclical. Linear time has a distinct beginning and end and can be expressed using discrete or continuous intervals. Cyclical time captures events that occur in a sequence over and over.

We must also resist the impulse to model space and time independently and separately, which will require even more careful consideration when choosing a spatio-temporal scale (see Lee and Li, 2017 for an informative discussion).

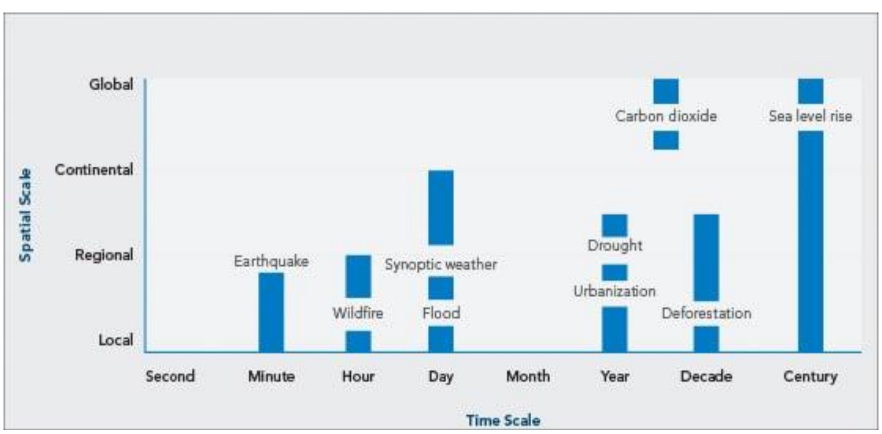

Figure 3 -- Geographical processes and events operate at various spatial and temporal scales.

Conclusion

Many of the challenges related to scale and spatial analysis exist because of implementation decisions made in a computationally deprived era. These choices were necessary to make spatial algorithms tractable, but these simplifications are no longer justified in an era of massive and extensible storage and computational power. COTS and open-source GIS recognized this reality and its shortcomings and are evolving to provide frameworks for true modern spatial analysis.

Birch, C. P., Oom, S. P., & Beecham, J. A. (2007). Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecological modelling, 206(3-4), 347-359.

Brodsky, I. (2019). H3: Uber's Hexagonal Hierarchical Spatial Index. https://eng.uber.com/h3/.

Gotway, C. A., & Young, L. J. (2002). Combining incompatible spatial data. Journal of the American Statistical Association, 97(458), 632-648.

Chatterjee, S., & Tenbrink, J. (2019, May 6). Retrieved from https://www.esri.com/arcgis-blog/products/spatial-analyst/analytics/introducing-the-new-resolution-preserving-cell-size-projection-method/

Kelly, K. M., & Šavrič B., (2019). Area and volume computation of longitude-latitude grids and three-dimensional meshes. Manuscript submitted for publication.

Lee, J., & Li, S. (2017). Extending Moran's index for measuring spatiotemporal clustering of geographic events. Geographical Analysis, 49(1), 36-57.

NGA Geomatics Office. (2007). *NGA of Advisory Notice on "Web Mercator"* https://earth-info.nga.mil/GandG/wgs84/web_mercator/

NGA Geomatics Office. (2014). NGA.STND.0036.1.0.0WGS84 https://earth-info.nga.mil/GandG/publications/NGA_STND_0036_1_0_0_WGS84/NGA.STND.0036_1.0.0_WGS84.pdf

Openshaw, S. (1984). "The Modifiable Areal Unit Problem". GeoBooks, Norwich, England.

Sahr, K., White, D., & Kimerling, A. J. (2003). Geodesic discrete global grid systems. Cartography and Geographic Information Science, 30(2), 121-134.

Tobler, W. (2014, December 10). Personal Communication. Esri/University of Redlands Colloquium. Redlands, California.

White, D. 2000. Global grids from recursive diamond subdivisions of the surface of an octahedron or icosahedron. Environmental Monitoring and Assessment 64(1): 93-103.