[CV-03-004] Scale and Generalization

Abstract

Scale and generalization are two fundamental, related concepts in geospatial data. Scale has multiple meanings depending on context, both within geographic information science and in other disciplines. Typically it refers to relative proportions between objects in the real world and their representations. Generalization is the act of modifying detail, usually reducing it, in geospatial data. It is often driven by a need to represent data at coarsened resolution, being typically a consequence of reducing representation scale. Multiple computations and graphical modi□cation processes can be used to achieve generalization, each introducing increased abstraction to the data, its symbolization, or both.

Keywords: abstraction, cartographic design, data modeling, granularity, map design fundamentals, representation, resolution, transformation, visualization

Author & citation

Raposo, P. (2017). Scale and Generalization. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed.). DOI: 10.22224/gistbok/2017.4.3

DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). Scale and generalization. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). Data abstraction: classification, selection, and generalization. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

Explanation

1. Definitions
2. Scale
3. Generalization
4. Types of and Reasons for Generalization
5. Operators & Algorithms
6. Generalization Across Data Layers and The State of the Art

1. Definitions

extent: The area or distance in real space over which some geographic entity exists. In cartography and GIS, the extent of a representation is the size of the real space being represented.

generalization: The processes of abstracting and transforming geospatial data in order to reduce their detail and generate versions that instead retain only their main, common, or

principal components or forms.

Modifiable Areal Unit Problem: An issue arising in the presentation or analysis of aggregate data wherein the same procedures on the same data differ in computed results when the data aggregation scheme is different. The MAUP is caused both by changing the size of aggregation units and by changing their boundaries; these are the scale and zone effects, respectively.

multiple representation: The practice of generating various representative symbols or data objects for any given mapped object, typically intending variants for differing mapping scales and resolutions.

operator: A single, atomic process of spatial data generalization, often meant to transform particular types of geometries such as lines or points. Operators are frequently combined in sequence or in parallel to effect generalization.

parameter: A variable, usually user-set, determining the degree to which an algorithm is applied. In generalization algorithms, parameters (also sometimes termed "tolerances," or, less often, "bandwidths") typically control, either directly or indirectly, how aggressively geographic features are abstracted.

precision: The repeatability of similar measures or representations; the granularity or frequency at which distinctions can be made by a sensor or display medium. See Mapping Uncertainty.

representative fraction: A ratio indicating the relative difference in size between real objects and their cartographic symbols. Given in the form 1:x (e.g., 1:24,000), meaning that 1 unit of distance on the map is equal to x units in the real world.

resolution: The degree of detail to which a phenomenon is detected or represented. Data are stored and rendered at some degree of representation resolution. In raster sensor arrays, resolution is defined by the dimensions of the individual sensors in terms of ground units (i.e., the width of one pixel in meters on the Earth). In vector geospatial data, resolution is defined both by the spatial precision to which vertices are defined and how densely vertices are posted, though both measures can vary greatly throughout a single dataset. Resolution is closely related to precision. Also called granularity.

scale: A measure of relative size, of objects or representations of them. Given constant display or data capture granularity, cartographic scale is correlated to resolution (i.e., "larger-scale" maps are generally of higher data and graphic resolution).

phenomenon scale: The relative spatiotemporal sizes or durations at which objects and processes occur in the natural world.

analysis scale: The relative spatiotemporal sizes or durations over which something is studied or simulated.

cartographic scale: The ratio between the size of an object and its representative symbol on a map.

2. Scale

Scale is a fundamental concept in virtually all sciences, but is especially important to geography, geographic information science, remote sensing, and cartography. The word has multiple meanings across contexts, even within single disciplines. Most definitions of the term refer to some kind of sizing, measure, or range in a certain space; the space in question can be temporal or thematic as well as spatial. Measures of scale can be any of ordinal, interval or ratio in nature; they are rarely nominal, since difierent scales usually involve some quantitative comparison (see Statistical Mapping). Measures are often referred to in the plural as "scales" (e.g., local, regional, and global scales).

Geographers have suggested various theoretical frameworks for scale, such as a hierarchical system of scales based on theories of human spatial cognition (Montello, 1993), a ranking of spatial extents based on human mobility and perception of the natural world (Granö, 1997), and a way of deciding what detail is and is not available at a given scale by simulating human vision at different viewing altitudes (Li & Openshaw, 1990). Three concepts of scale are especially relevant to GIS&T: phenomenon scale, analysis scale, and cartographic scale (Montello, 2015).

Phenomenon Scale - Geographers recognize that only at certain scales do objects and phenomena in the real world exist, or are observable at (e.g., wind and rain can be observed on a city street, but not the whole hurricane they are part of, which can be observed across several kilometers). Phenomenon scale refers to the spatiotemporal extent and resolution required to define, detect, or represent the given phenomenon meaningfully. Sometimes phenomena need to be considered at multiple scales, for example to determine how they are affected by larger or smaller systems of which they are a part or constituted of, respectively; examples include land use cases and climate and weather systems. Determining the scale of a phenomenon is central to analysis in geography and remote sensing, among other disciplines, and drives the choices over extent and resolution of data capture. "Large" and "small" phenomena scales refer to larger and smaller objects, respectively.

Analysis Scale - Analysis scale refers to the spatiotemporal extent and resolution at which any given phenomena is studied. In practice, it is often defined in part by the resolution of the data used, as well as any resolutions to which analysis steps are computed. Generally, analysis scale should reflect the scale of the phenomenon being studied; using an analysis scale that is too coarse or too fine for the analysis theme can obscure the phenomena in question.

Cartographic Scale - Cartographic scale unambiguously refers to the ratio of representation size to actual size for a given cartographic visualization or map. It is frequently expressed as a representative fraction, such as 1:25,000, meaning that one unit measured on the map represents 25,000 units in the real world (e.g., 1 cm on the map is 25,000 cm, or 250 m, in the real world). This relationship is sometimes given as a verbal scale (e.g., "one inch = 72 miles"). "Large" and "small" cartographic scales follow the mathematical definition of the ratio in question, with larger scales being ratios that compute to higher numbers than smaller scales (e.g., 1:25,000 is larger than 1:50,000). Cartographic scale can also be expressed graphically with a scale bar, being a line of a certain length drawn on the map and annotated to represent a certain real world length. Cartographic scale is sometimes referred to as "visualization" or "representation scale."

In addition to possessing some scale in each of the aspects mentioned above, digital geospatial data may be stored in a database at one scale and resolution, and rendered for display at another. In both cases, the geographer seeks to treat the data at phenomenon and analysis scales that correspond to the geographic theme in question, i.e., at which the concepts and objects in the landscape are observable and exist. Cartographic scales, in practice, vary more widely from any particular phenomena or analysis scale, as maps are drawn in many media and for many different audiences (e.g., smart phone screens and large wall maps, trained analysts and the general public).

Scale changes often drive the need to perform generalization on geospatial data. As cartographic scale decreases, so too does the representation resolution of the medium being drawn on (e.g., a pixel screen). This resolution is distinct from that at which the data are stored or analyzed at, and refers to the level of detail in its representation, usually in a cartographic form (Tobler, 1988). Coupled with the limits of what the human eye can resolve, this geometric fact often requires cartographers to deliberately remove detail from geospatial data when plotting them.

3. Generalization

Generalization is a process by which geospatial data undergo abstraction. Like projection, generalization constitutes a transformation of the data wherein certain geometric and topological properties are retained and others lost, though generalization can also lose, retain, or transform thematic or attribute properties as well. While there is inherent generalization in any data-gathering process (e.g., sensor resolution limits the detail that can be captured), generalization is usually considered when done deliberately. A common example is seen during mapmaking, when cartographers adjust levels of detail between various thematic data sources so that they correspond; this is typically done by removing detail from higher-resolution datasets until their levels of detail are commensurate with lower-resolution datasets. Undergoing generalization iteratively or continuously allows for multiple representations (Frank & Timpf, 1994) of the features in a given dataset, such as sets of renderings that continuously adapt to viewing scale and resolution. Research in automated digital generalization has been ongoing since the advent of GIS, often with the stated goal of identifying and computerizing human cartographic knowledge, heuristics, and techniques.

4. Types of and Reasons for Generalization

Cartographers make a distinction between generalization performed on data objects for the purpose of efficient storage or analysis, being model generalization, and that performed to prepare objects for symbolization and visual presentation, being cartographic generalization (Grünreich, 1985; Brassel & Weibel, 1988). Model generalization is typically data-reducing, and motivated by a desire for economy in storage space or computational complexity. It can also reflect scale changes made to bring data to an appropriate resolution for some context-specific analysis. Cartographic generalization, which often follows model generalization, does not always reduce the volume of data, though it frequently does.Instead, the principle motivation is to derive geographic feature representations that are suitable (e.g.,graphically resolvable) for analysis or display in some

target cartographic context, such as cartometric analysis, or a zoom level in a digital interactive map display.

Both model and cartographic generalization are frequently driven by a reduction in map scale (i.e., a zooming-out), causing a commensurate reduction in graphic resolution (Tobler, 1988). Some procedures and algorithms for generalization have been developed with direct reference to a quantified change in scale and/or resolution (Perkal, 1956; Buttenfield, 1989; Li & Openshaw, 1990; Dutton, 1999), with the most famous of these (Töpfer & Pillewizer, 1966) being known as The Radical Law for its mathematical root-based definition of how many features should remain on a map after a measured scale change. Other commonly-used procedures are guided by heuristic or ad-hoc relationships to scale change or target scale.

In addition to scale-driven reasons, generalization may also be performed in order to use a dataset for some purpose other than that which it was compiled for (e.g., an expressway with two single-direction lines compiled for GPS navigation calculations is collapsed to a single line for map representation), or for graphic simplicity or aesthetic reasons (e.g., simplified and abstracted geometry in subway maps such as London's famous Tube map).

An important consideration, perhaps more frequently relevant in model generalization, is the effect generalization has on analysis. As a simple example, Figure 1 demonstrates how area calculations are affected by polygon simplification. The same effects are seen in generalized continuous data such as rasters, as demonstrated in Figure 2. Generalization can reduce both accuracy and precision (see Mapping Uncertainty), and analysts must decide whether or not the levels of either after generalization are appropriate to the task at hand. In analytical contexts, generalization often causes the Modifiable Areal Unit Problem (see Statistical Mapping) (Openshaw, 1984).



Figure 1. The area of Tennessee calculated by a GIS before and after polygon simplification. Both polygons are projected in the NAD 83 Tennessee State Plane coordinate system.

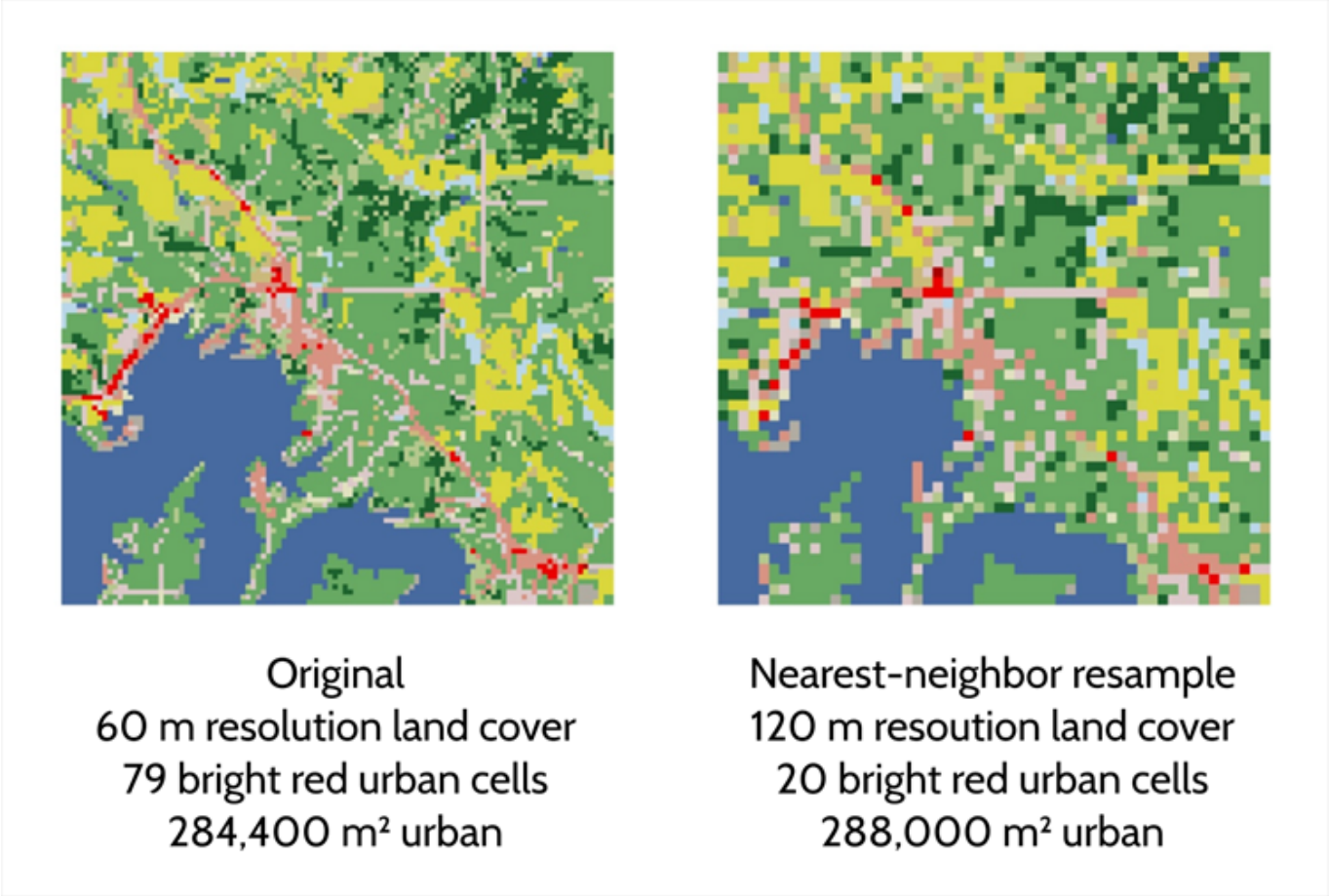


Figure 2. The area of a land cover class before and after a coarsening of resolution and nearest-neighbor resampling.

5. Operators & Algorithms

Particular atomic processes of abstraction applied to geospatial data in order to produce generalized versions are called operators. These typically are defined over a certain kind of input geometry (e.g., polygons) and produce a certain kind of output geometry. Any given operator can be effected using one of any number of algorithms. Various operators and algorithms have been heuristically classified as better or worse for certain kinds of geographic features (e.g., one line simplification algorithm may work well overall on human-made features but not on river lines). Often, particular algorithms afford the ability to calibrate their effects by allowing users to specify input parameter values; these values are sometimes commensurate with measurable generalization effects (Raposo, 2013), and other times are set by heuristic methods such as trial and error.

Several scholars have sought to define typologies of generalization operators (McMaster & Monmonier, 1989; Li, 2007; Roth, Brewer, & Stryker, 2011). Many operators exist, though their names and exact definitions are not universally agreed-upon. Figures 3 and 4 illustrate a few of these on vector and raster data, respectively, while Figure 5 demonstrates line simplification effected to various degrees using different user-set tolerance parameter values. Chains or workflows involving several operators are typically used to achieve desired generalization results. The operators illustrated in Figures 3 and 4 are defined below.

Simplification - The reduction in sinuosity or complexity of a linear or polygonal shape, usually involving a reduction in vertices along its constituent polylines.

Aggregation - The combination of polygon symbols into a smaller number, usually by filling space between the initial polygons to create a lesser number of contiguous polygons.

Smoothing - The replacement sharp angles in a polyline or polygon with curves so that the overall shape is softened.

Selection/Elimination - The retention of certain features and rejection of others.

Typification - The transformation of detailed polygonal features into canonical, usually simpler versions of the type of object being represented (e.g., complex buildings to simple rectangles).

Displacement - Moving features away from their planimetrically-accurate locations for legibility or to emphasize a spatial relationship (e.g., moving a building closer to or further away from a road).

Exaggeration - Adding visual emphasis, usually with increased symbol size, to an object.

Classification - Reducing the variety of measures in a dataset by binning similar measures together.

Trend Calculation - A relatively severe generalization of a surface into a mathematically-simple function approximating it, commonly defined by a lower-order polynomial.

Opening and Closing (Expand and Shrink) - The increasing or decreasing dilation, respectively, of the set of areas of a given class in a classified dataset. Often employed on classified raster regions, opening and closing tends to produce simplified region boundary geometries. The two operations are not commutative.

Resampling - Changing the unit of aggregate data by recollecting source data in differently-sized units (e.g., changing the resolution of a raster dataset).

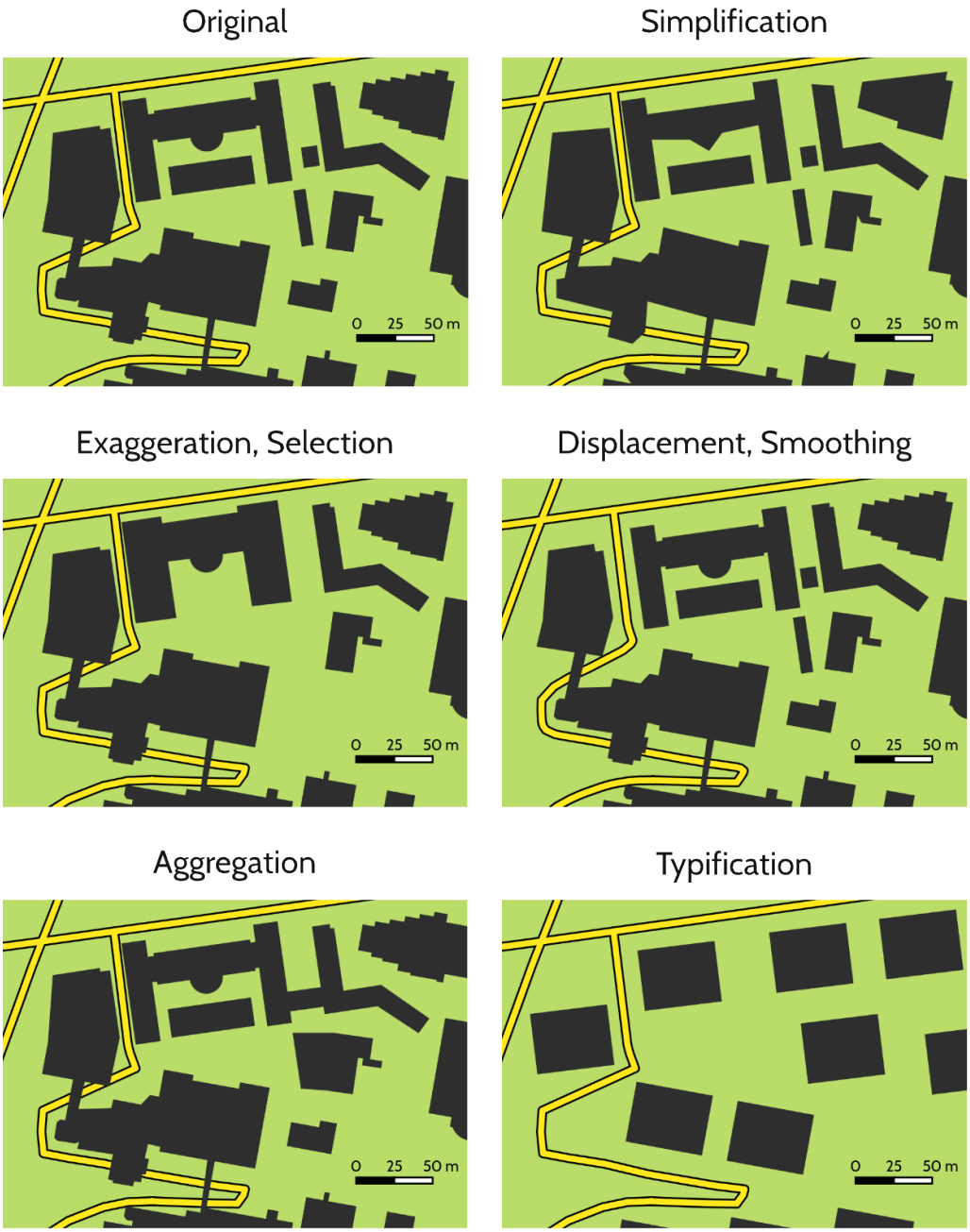


Figure 3. Various vector generalization operators illustrated over buildings and roads.

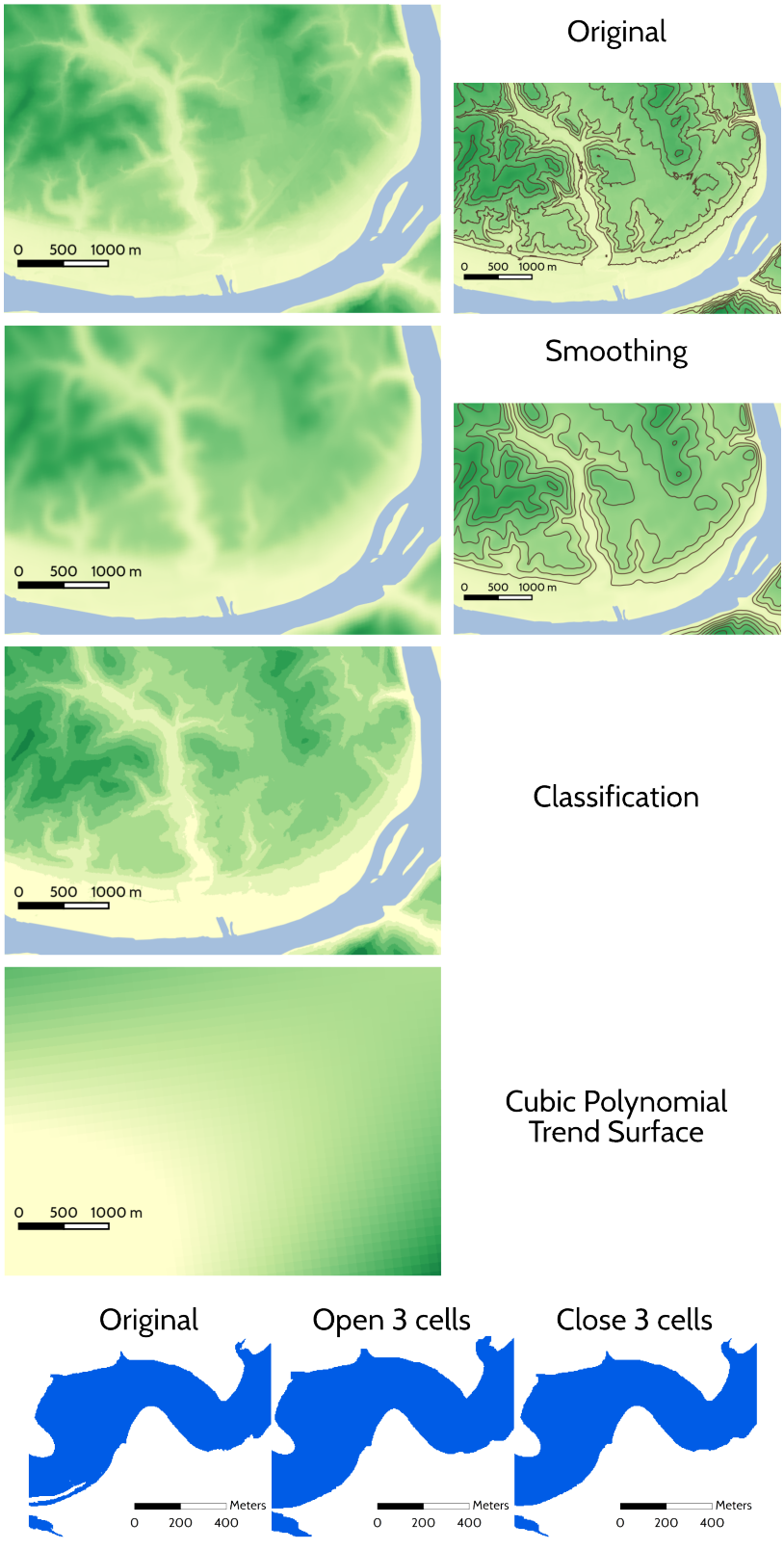


Figure 4. Various raster generalization operators illustrated over a digital elevation model (top, in greens), and over a single classed raster region (below, in blue). Greens are higher elevations while yellows are lower.

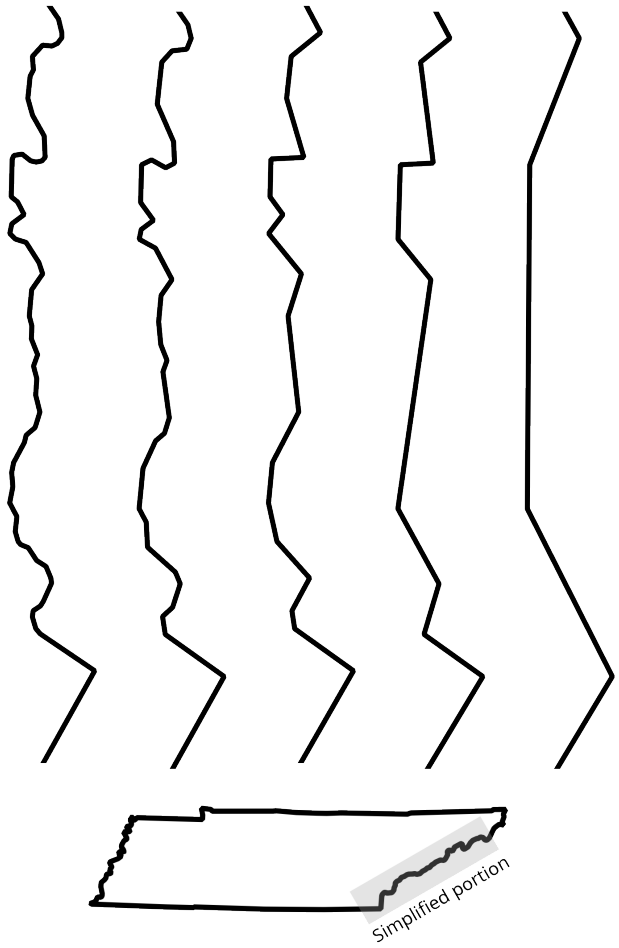


Figure 5. A line representing the eastern border of Tennessee, simpli□ed using the Douglas-Peucker (1973) algorithm to multiple levels of detail using multiple input tolerance values.

6. Generalization Across Data Layers and The State of the Art

Most GIS projects consist of a set of several or many data layers. In such sets, generalization (i.e., transformation of geometry and/or thematic attributes) in one layer must be propagated throughout the others, so that all layers correspond and vertically register correctly. For example, given a polyline representing a river and an adjacent polygon representing a city on its shore, simplifying the river may cause it to run through or deviate from the city; if the river simplification is to be accepted, the city polygon needs to be displaced such that it lies on the correct shore. The complexity of such inter-layer relationships in generalization makes the overall process necessarily holistic and highly contextual (Müller, 1991).

The majority of generalization operators have thus far been formulated to transform single data themes or layers, and are effectively oblivious of any others. The present state of the art reflects this: propagating generalization through multiple layers is usually done by error-correcting post-processing routines after having generalized individual layers. Such post-processing continues until no further artifacts or errors are detected. Production cartographic generalization work usually still involves some amount human inspection and

editing, but research continues on fully-automated methods that resolve clearly-defined cartographic design constraints (Harrie & Weibel, 2007). There has been some success in more comprehensive approaches to the generalization of multiple layers using hierarchical graphs (Frank & Timpf, 1994), agent-based models (Ruas, 2002; Duchêne, Ruas, & Cambier, 2012), continuous optimization approaches (Harrie & Sarjakoski, 2002), and combinatorial approaches (Ware & Jones, 1998). Also, several European national topographic mapping agencies already make use of multi-representation databases to produce map series.

References

Brassel, K. E., & Weibel, R. (1988). A Review and Conceptual Framework of Automated Map Generalization. International Journal of Geographical Information Science, 2(3), 229-244.

Buttenfield, B. P. (1989). Scale-Dependence and Self-Similarity in Cartographic Lines. Cartographica, 26(1), 79-100.

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. Cartographica: The International Journal for Geographic Information and Geovisualization, 10(2), 112-122.

Duchêne, C., Ruas, A., & Cambier, C. (2012). The CartACom Model: Transforming Cartographic Features Into Communicating Agents for Cartographic Generalization. International Journal of Geographical Information Science(advance online publication), 1-30.

Dutton, G. (1999). Scale, Sinuosity, and Point Selection in Digital Line Generalization. Cartography and Geographic Information Science, 26(1), 33-53.

Frank, A. U., & Timpf, S. (1994). Multiple representations for cartographic objects in a multi-scale tree - an intelligent graphical zoom. Computers & Geosciences, 18(6), 823-829.

Granö, J. G. (1997). Pure Geography. In O. Granö & A. Paasi (Eds.), Pure Geography. Baltimore, MD: Johns Hopkins University Press. (Originally published in German, 1929.)

Grünreich, D. (1985). Computer-Assisted Generalization. In Papers CERCO-Cartography Course. Frankfurt a.M.: Institut für Angewandte Geodäsie.

Harrie, L. & Weibel, R. (2007). Modelling the Overall Process of Generalisation. In W. A. Mackaness, A. Ruas, & L. T. Sarjakoski (Eds.), Generalisation of Geographic Information: Cartographic Modelling and Applications (pp. 67-87). Elsevier.

Li, Z. (2007). Algorithmic Foundation of Multi-Scale Spatial Representation. Boca Raton, London, New York: CRC Press.

Li, Z., & Openshaw, S. (1990). A natural principle of objective generalization of digital map

data and other spatial data (Tech. Rep.). Newcastle upon Tyne: CURDS, University of Newcastle upon Tyne.

McMaster, R. B. & Monmonier, M. (1989). A Conceptual Framework for Quantitative and Qualitative Raster-Mode Generalization. In Proceedings of the Annual GIS/LIS Conference. Orlando, Florida.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), Spatial information theory: A theoretical basis for GIS. Proceedings of COSIT '93. Lecture Notes in Computer Science (Vol. 716, pp. 312-321). Berlin: Springer-Verlag.

Montello, D. R. (2015). Scale in Geography. In N. J. Smelser & P. B. Baltes (Eds.), International Encyclopedia of the Social & Behavioral Sciences (2nd ed.). Oxford, England: Pergamon Press.

Müller, J. C. (1991). Generalization of Spatial Databases. In: D. J. Maguire, M. Goodchild, & D. Rhind (Eds.), Geographical Information Systems. 1st Edition. London, U.K.: Longman Scientific.

Openshaw, S. (1984). The Modifiable Areal Unit Problem. Concepts and Techniques in Modern Geography No. 38. Norwich, England: Geo Books, Regency House.

Perkal, J. (1956). On epsilon length. Bulletin de l'Academie Polonaise des Sciences, 4, 399-403.

Raposo, P. (2013). Scale-Specific Automated Line Simplification by Vertex Clustering on a Hexagonal Tessellation. Cartography and Geographic Information Science, 40(5), 427-443.

Roth, R. E., Brewer, C. A., & Stryker, M. S. (2011). A typology of operators for maintaining legible map designs at multiple scales. Cartographic Perspectives, 2011(68), 29-64.

Ruas, A. (2002). Les problématiques de l'automatisation de la généralisation. In A. Ruas (Ed.), Généralisation et représentation multiple (pp. 75-90). Hermès.

Tobler, W. R. (1988). Resolution, resampling, and all that. In H. Mounsey & R. F. Tomlinson (Eds.), Building Databases for Global Science: the proceedings of the First meeting of the International Geographical Union Global Database Planning Project (pp. 129-137). Hampshire, U.K.: Taylor and Francis.

Töpfer, F., & Pillewizer, W. (1966). The Principles of Selection. The Cartographic Journal, 3(1), 10-16.

Ware, J. M., & Jones, C. B. (1998). Conflict Reduction in Map Generalization Using Iterative Improvement. GeoInformatica, 2(4), 383-407.