Map Generalization

Unit: III

Semester: II

Paper Code: GIS 08
Name of Paper: Computer Cartography

PG Diploma in RS & GIS

Dr. SHYAMA PRASAD MUKHERJEE UNIVERSITY, RANCHI

MAP GENERALISATION

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.

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 medication processes can be used to achieve generalization, each introducing increased abstraction to the data, its symbolization, or both.

The **International Cartographic Association** defines **Cartographic Generalisation** as "the selection and simplified representation of detail appropriate to the scale and/or the purpose of a map" (ICA 1967). More generally, the objective of generalisation is to supply information on a content and detail level corresponding to the necessary information for correct geographical reasoning.

Generalisation inputs are:

- ✓ The needs
- ✓ The geographical data: density, distribution, size, diversity etc.
- ✓ The readability rules
- ✓ The means: time, money, technique etc.

1.1. The Generalisation Necessity

On a map, the available space for all the cartographic representations of the objects and elements of a landscape model, is very small and decreases disproportionately from scale to scale. Therefore a limitation on the essential map elements and objects is necessary with the reduction of an image area from scale to scale.

1.2. Characteristics of generalisation are:

- ✓ An extreme reduction compared to reality. Example: with a 1:25 000 scale, the image area 625 million times smaller than reality.
- ✓ A pure photographic reduction of the original scale leads to an illegible map.
- ✓ Already in the 1:25 000 scale, many objects of the landscape cannot be represented any more.
- ✓ At smaller scales, a representation of all objects of the landscape is impossible.
- ✓ A complex and unclear structured reality must be simplified according to the scale of the map.

1.3. Aims of Generalisation

Cartographic generalisation is born of the necessity to communicate. As it is not possible to communicate map information at 1:1 scale, generalisation has many aims. The following aims can also be considered as generalisation rules.

1.3.1. Structure: The map content is well structured.

1. The estimation of map content priorities has to be adapted to the mapscale and to the intended purpose.
2. The objects have to be classified according to clear and reasonable criteria.
3. The grouping of objects has to be logical.

1.3.2 Legend: Expressive and associative symbols constitute the base for clear map communication.

1. The size and the form of the symbols are adapted to the other symbols and to the reality.

1.3.3. Generalisation level: The level of generalisation implies simplification and detailing.

1. A low level of generalisation signifies a high information density and a fine structured map.
2. A high level of generalisation signifies a low information density and a thick structured map.
3. The level of generalisation varies according to the purpose and to the map scale.
4. The level of generalisation is carefully defined.
5. The level of generalisation affects the legend and the symbols.

1.3.4. Selection of objects: The objects selection complies with the map purpose.

1. The objects selection complies with the map scale and with the intended purpose.
2. The objects that are visible in reality (e.g. houses) are completed with non-visible objects as borders or labelling.

1.3.5. Accuracy of objects: The optimal accuracy of the objects regarding position and form is reached. However, the visual placement of objects is more important than the geometrical accuracy.

1. Displacing objects is only needed for raising the legibility and for clarification.
2. The symbols of visible objects (in reality) have a high accuracy.
3. The symbols of non-visible objects (in reality) have a limited accuracy.
4. Object displacement is necessary, and the neighbouring objects are adapted.
5. The form accuracy is only limited by the good legibility and the respect proportions demand.
6. The contour lines are not treated as a single line, but are adjusted to the correct reproduction of the ground structure on each other.

1.3.6. Reality accuracy: Indeed, the reality is revised and changed, but is still, as far as possible, represented truthfully.

1. All objects present in the map really exist.
2. Appropriate legend symbols are assigned to the objects.
3. Labelling is correctly raised, written and assigned.

1.3.7. Legibility of the map elements: The map must be readable without auxiliary means (e.g. magnifying glass), and in bad conditions.

1. Good legibility is conditioned to the respect of the graphical minimal dimensions (sizes and distances) of the symbols.
2. Graphical minimal dimensions lead to an unscaled representation, i.e. to an enlargement of the dimensions scale.
3. Graphical readability rules support legibility.

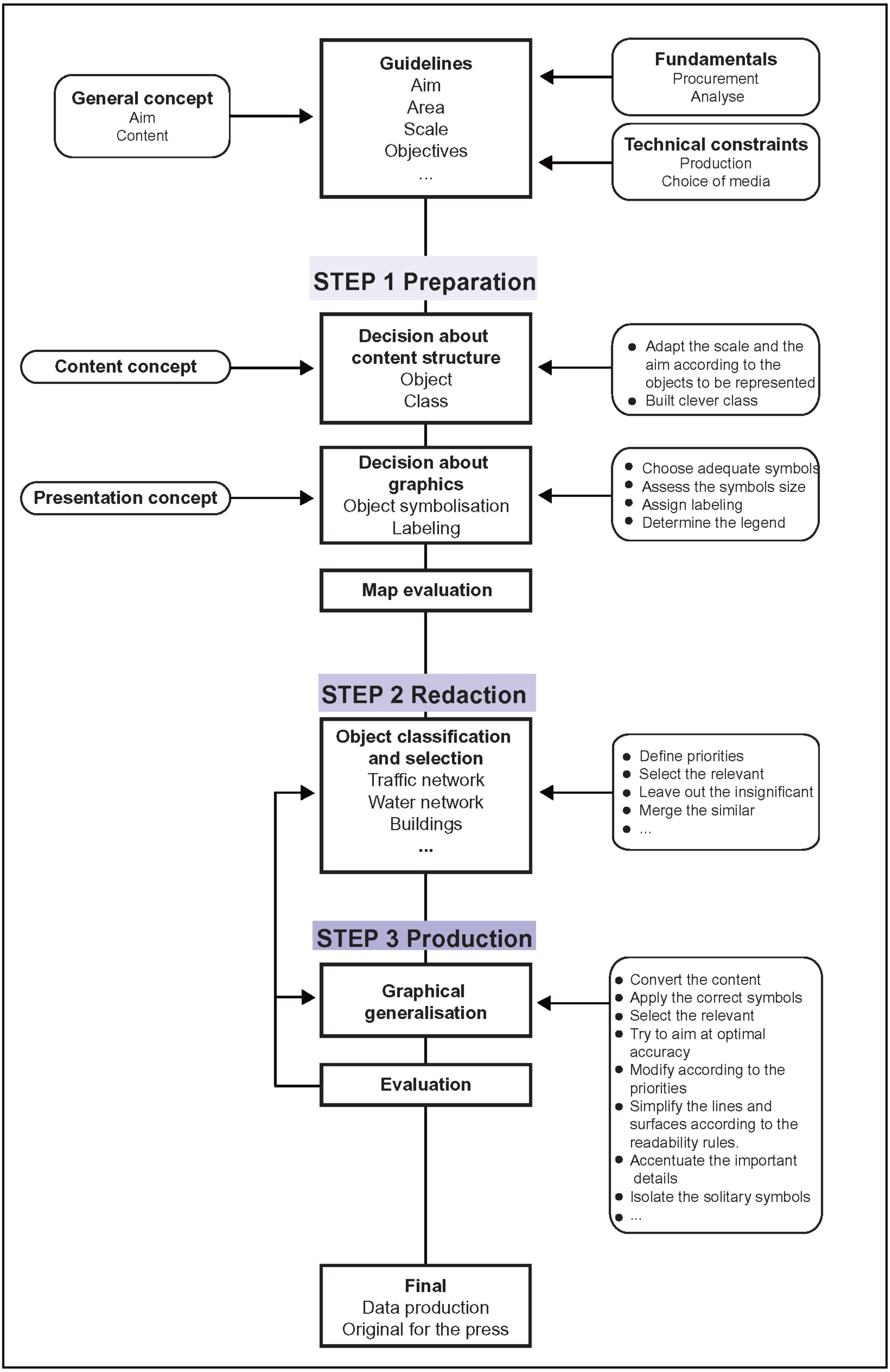
1.3.8. Graphical representation of the objects: Map content is adapted, legible and graphically convincing.

1. Legend is credible and exact.
2. The generalisation of the forms and line symbols respect the most exceptional forms and eliminates the small and fortuity ones.
3. The quantitative generalisation from strewn objects (e.g. houses) respects the density of objects in reality.
4. The relations and dependencies of objects in reality (e.g., streets, ways, waters, contour lines, etc.) are carefully considered.

1.4. Generalisation Workflow

In order to portray important aspects of reality, various manipulations of the data that represent information to be mapped are necessary. These manipulations can be divided in three steps: Preparation, Redaction, and Production.

The following image shows you the most important elements of these.



General concept
Aim
Content
Guidelines
Aim
Area
Scale
Objectives
...
Fundamentals
Procurement
Analyse
Technical constraints
Production
Choice of media
STEP 1 Preparation
Content concept
Decision about content structure
Object
Class
Adapt the scale and the aim according to the objects to be represented
Built clever class
Presentation concept
Decision about graphics
Object symbolisation
Labeling
Choose adequate symbols
Assess the symbols size
Assign labeling
Determine the legend
Map evaluation
STEP 2 Redaction
Object classification and selection
Traffic network
Water network
Buildings
...
Define priorities
Select the relevant
Leave out the insignificant
Merge the similar
...
STEP 3 Production
Graphical generalisation
Evaluation
Convert the content
Apply the correct symbols
Select the relevant
Try to aim at optimal accuracy
Modify according to the priorities
Simplify the lines and surfaces according to the readability rules.
Accentuate the important details
Isolate the solitary symbols
...
Final
Data production
Original for the press

1.5. 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.

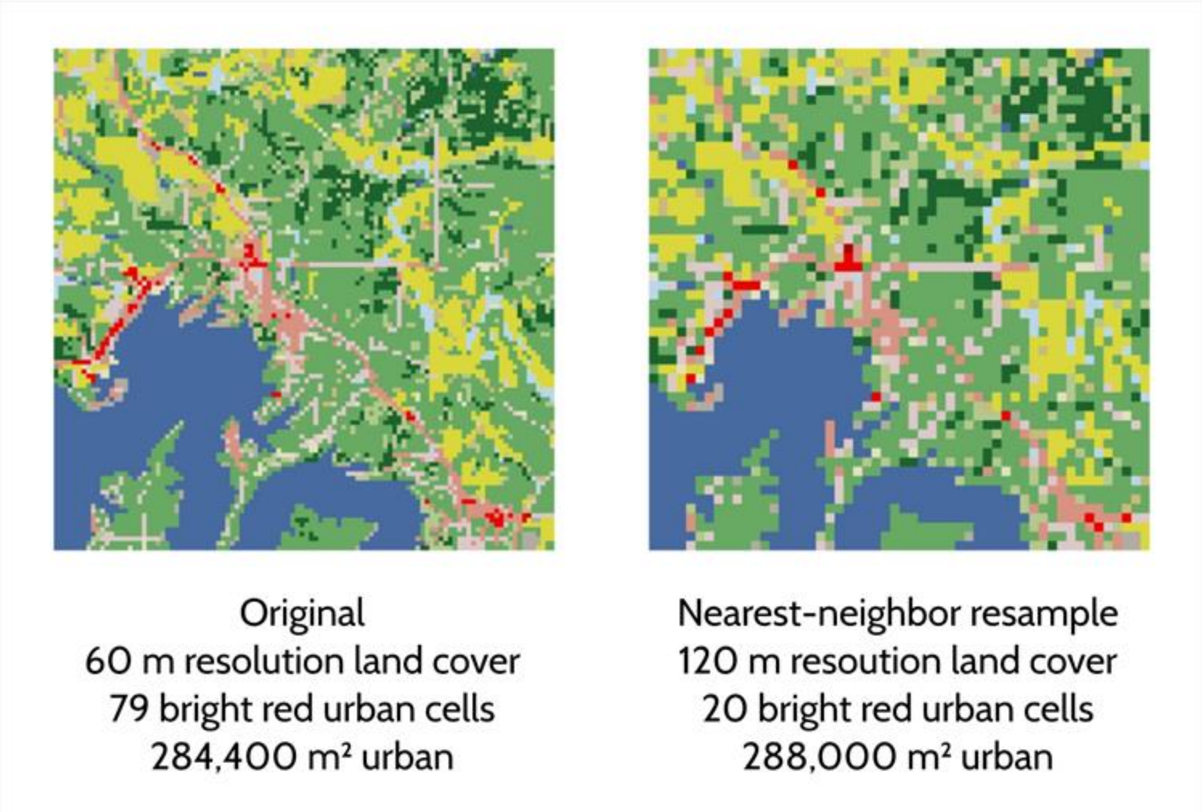


Figure 2. The area of a land cover class before and after a coarsening of resolution and nearest-neighbour resampling.

1.6. 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 affected 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).

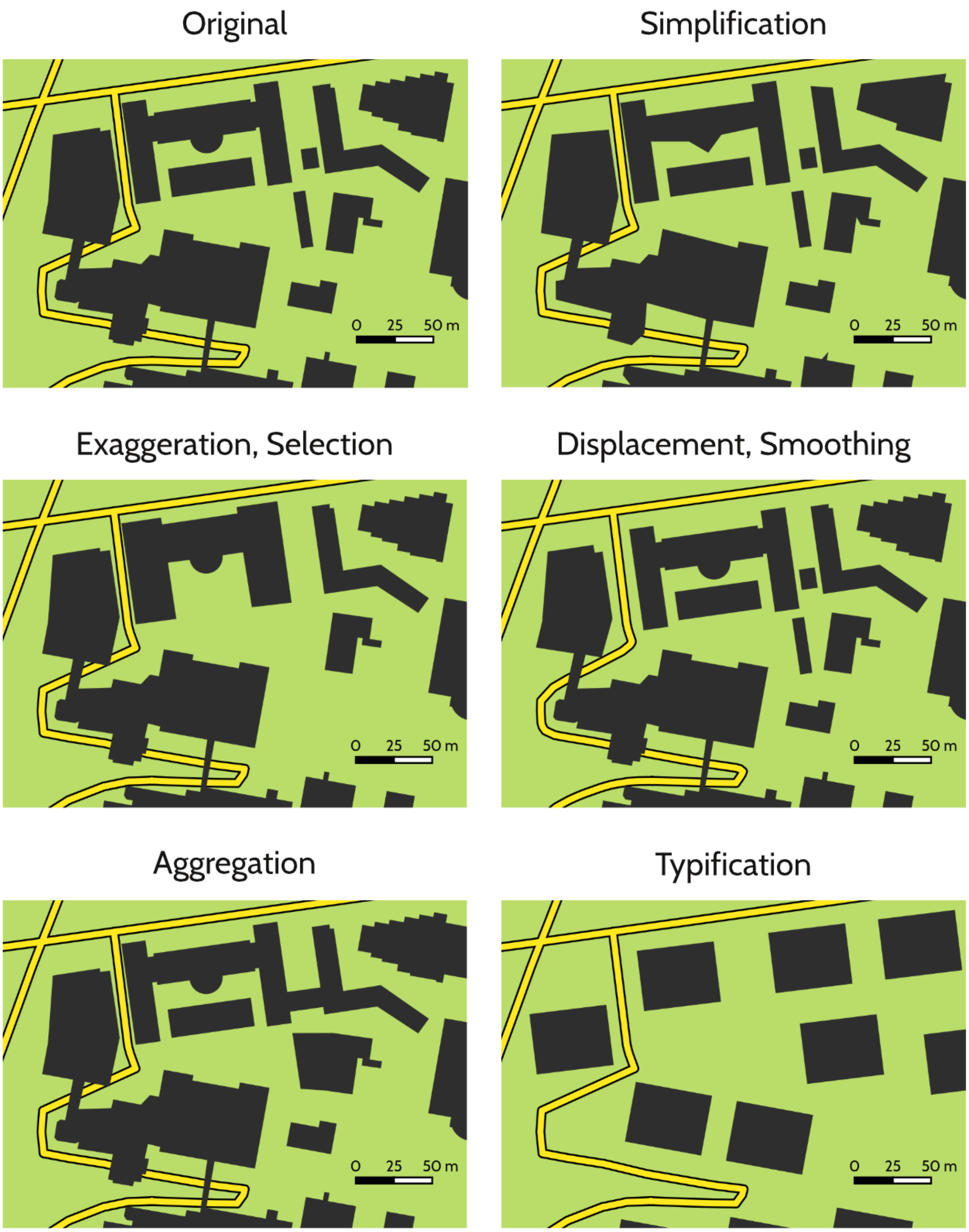


Figure 3. Various vector generalization operators illustrated over buildings and roads.

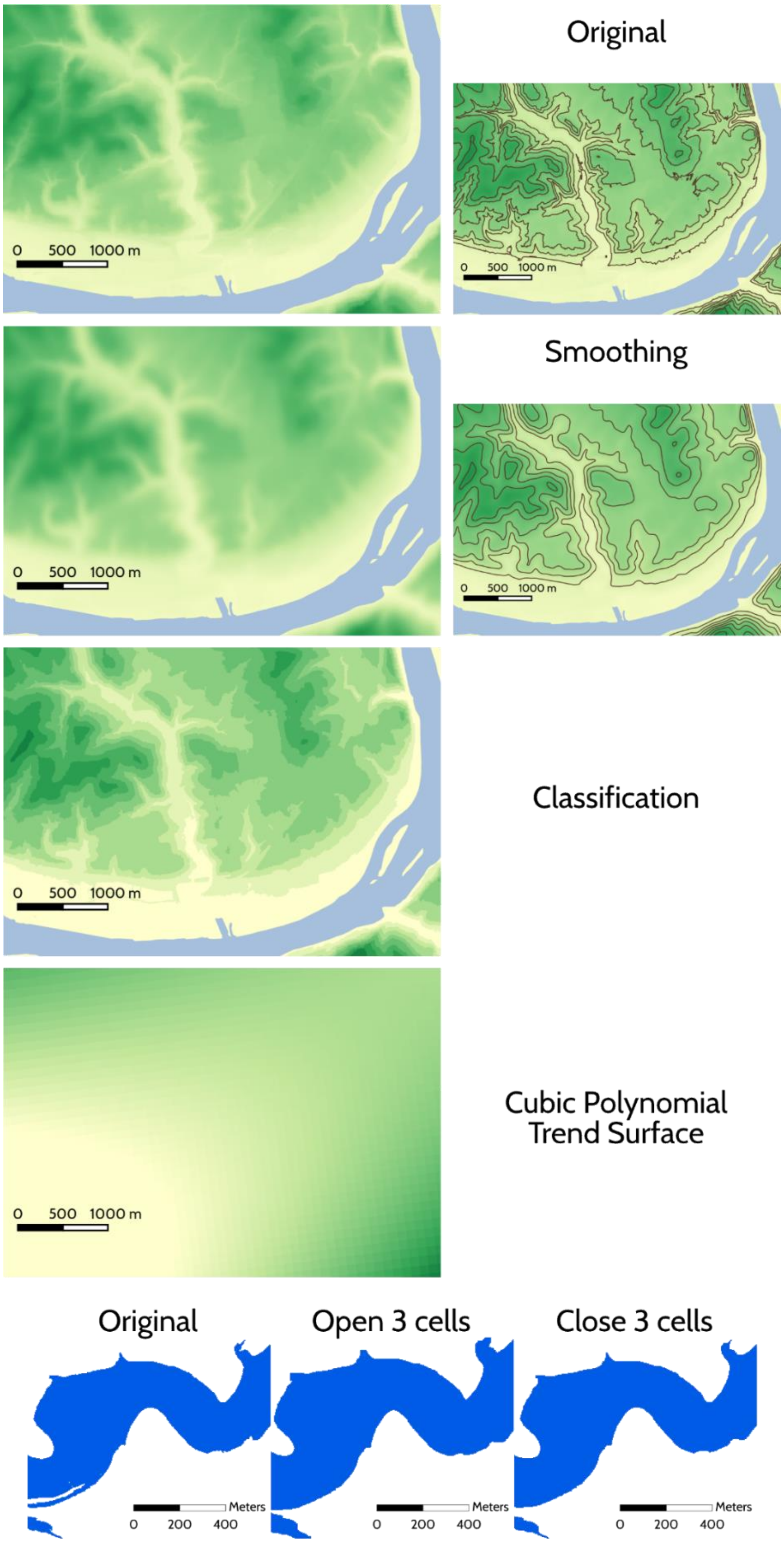


Figure 4. Various raster generalization operators illustrated over a digital elevation model (top, in greens), and over a single classed raster region (below, in blue). Greens are higher elevations while yellows are lower.

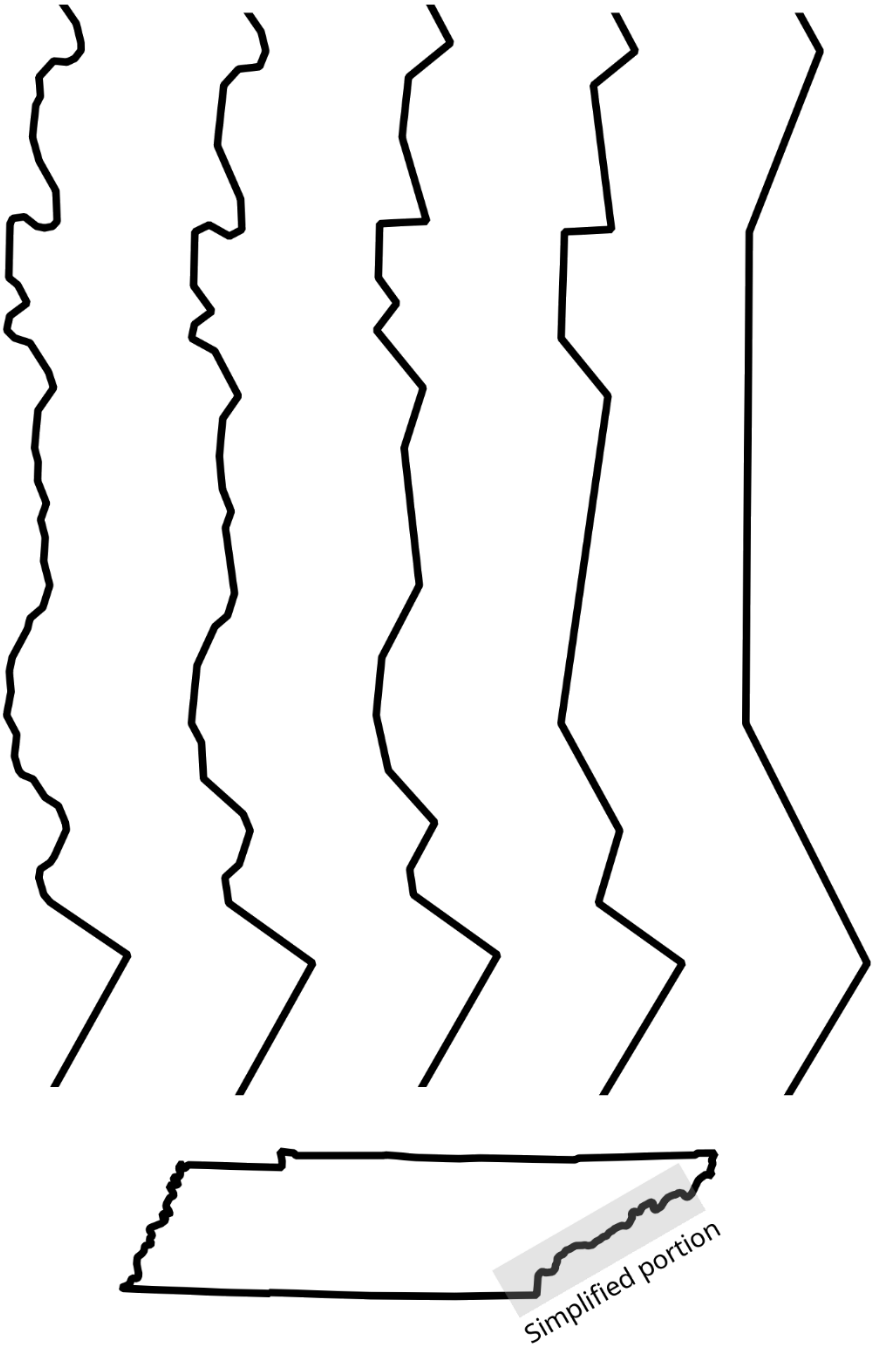


Figure 5. A line representing the eastern border of Tennessee, simplied using the Douglas-Peucker (1973) algorithm to multiple levels of detail using multiple input tolerance values.

1.7. 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. DOI: 10.1080/13658816.2011.639302(link is external).
- 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., & Sarjakoski, T. (2002). Simultaneous Graphic Generalization of Vector Data Sets. *GeoInformatica*, *6*(3), 233-261.
- 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. DOI: 10.1201/9781420008432.ch7(link is external)
- 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*. London, U.K.: Longman Scientific.

- Openshaw, S. (1984). The Modiable Areal Unit Problem. 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. DOI: 10.1080/15230406.2013.803707(link is external)
- 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. DOI: 10.14714/CP68.7(link is external)
- 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., & Jones, C. (1998). Conflict Reduction in Map Generalization Using Iterative Improvement. *GeoInformatica*, *2*(4), 383-407.