Chapter 6
Generalisation Operators

Lawrence V. Stanislawski, Barbara P. Buttenfield, Pia Bereuter, Sandro Savino and Cynthia A. Brewer

Abstract This chapter summarises cartographic generalisation operators used to generalise geospatial data. It includes a review of recent approaches that have been tested or implemented to generalise networks, points, or groups. Emphasis is placed on recent advances that permit additional flexibility to tailor generalisation processing in particular geographic contexts, and to permit more advanced types of reasoning about spatial conflicts, preservation of specific feature characteristics, and local variations in geometry, content and enriched attribution. Rather than an exhaustive review of generalisation operators, the chapter devotes more attention to operators associated with network generalisation, which illustrates well the logic behind map generalisation developments. Three case studies demonstrate the application of operators to road thinning, to river network and braid pruning, and to hierarchical point elimination. The chapter closes with some summary comments and future directions.

L. V. Stanislawski (✉)
Center of Excellence for Geospatial Information Science,
United States Geological Survey, Rolla MO, USA
e-mail: lstan@usgs.gov

B. P. Buttenfield
Department of Geography, University of Colorado, Boulder CO, USA
e-mail: babs@colorado.edu

P. Bereuter
Department of Geography, University of Zurich, Zurich, Switzerland
e-mail: pia.bereuter@geo.uzh.ch

S. Savino
Department of Information Engineering, University of Padova, Padova, Italy
e-mail: sandro.savino@dei.unipd.it

C. A. Brewer
Department of Geography, Pennsylvania State University, University Park PA, USA
e-mail: cbrewer@psu.edu

D. Burghardt et al. (eds.), *Abstracting Geographic Information in a Data Rich World*,
Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-319-00203-3_6,

6.1 Introduction

While there is a rich history of developments in manual generalisation discussed by numerous authors (for example McMaster 1983, 1989; McMaster and Shea 1992; Kilpeläinen 1997; Sarjakoski 2007), automated strategies for generalisation have been demonstrated to reduce manual workloads, and to minimise errors and inconsistencies. The challenge for automatic generalisation was, and continues to be, the development of rules and constraints which control the intensity of generalisation, the sequencing of operators, and conflict detection and resolution. Recent advances in computer science have introduced capabilities for artificial intelligence, amplified intelligence, pattern recognition, and automated spatial reasoning into generalisation software tools. Previous research mainly focused on orchestrating the logical sequence of operators, but models for some cases are still missing (e.g., for a broad range of geographic features, across a wide range of scales, or tailored to landscape differences). Practical application of these advanced techniques requires faster processing speeds as well as data enrichment, which enables context-sensitive generalisation that can be tailored to specific geographies (such as urban or rural areas, or dry and humid landscapes).

Regnauld and McMaster (2007) give a thorough overview of frameworks for generalisation operators. For the purposes of discussion throughout this chapter, a generalisation operator is defined as a generic descriptor for the type of spatial or attribute modification to be achieved on some set of geospatial data (Regnauld and McMaster 2007; Roth et al. 2011). An algorithm refers to a specific method by which one or more operators are implemented, and several algorithms usually exist for the same operator.

There are a myriad of ways in which generalisation may be implemented. Generalisation operators or algorithms may affect data at a micro or macro level, or any level in between. For instance, an operator may affect data at a micro (or atomic) level in a uniform way, such as resampling a digital terrain model, or reducing the number of vertices in linear features by simplification. A different operator may affect features at a subatomic level in a non-uniform way, such as simplification of linear features divided into parts based on the level of coalescence of each part. At the other end of the spectrum, 'super-operators' exist that may affect meso- or macro-level objects in uniform or non-uniform ways with regard to a local context (Ruas and Duchêne 2007). Examples include enrichment or thinning of road or hydrographic networks.

An exhaustive review of generalisation operators is not possible in the space of this chapter. The chapter devotes more attention to operators associated with network generalisation than to other operators. Network generalisation does illustrate the logic behind map generalisation developments. In the past, less has been written about network operations than other operators because the network graph data structure is complex to build and manipulate. However, conceptual and technological advances have enabled more network capabilities in recent years and additional research has been undertaken. Three case studies demonstrate the application of operators to road thinning, to river network and braid pruning, and to hierarchical point elimination. The chapter closes with some summary comments and future directions.

6.2 Generalisation Operators: Chronology of Typologies

A chronology of typologies developed for generalisation operators mirrors advances in processing power, in establishment of linked multi-representation databases, and in improved methods for automatic reasoning about spatial conflicts and geographic context. The shift from early typologies to subsequently published frameworks reflects a still accepted view that generalisation encompasses more than invocation of a list of isolated operators.

Early typologies emphasised paper map production. Ratajski's (1967) model distinguished between quantitative operations that reduced feature details for scale-changing and qualitative operations that transformed symbol designs. Ratajski's model was guided by the concept of a *generalisation point* identifying a critical scale at which representation methods begin to fail, mandating a change in geometry or cartographic symbols. Robinson and Sale's (1969) typology included four elements that were re-sequenced by Morrison (1974) into classification, simplification, symbolisation and induction, in order to focus specifically on abstraction and implementation of cartographic symbols. Nickerson and Freeman (1986) developed a typology similarly suited to map production, with feature deletion, simplification, merging, and reclassification as a first stage of processing, followed by symbol scaling and placement, scale reduction, and label placement. Throughout these developments, the challenge for automation is grounded in the need for 'intelligent' decisions driving the choice of algorithms and parameters. 'Intelligence' in this context relates to rules or constraints which are based upon reasoning on the spatial context.

Sarjakoski (2007, p. 19) reported that the earliest proposals for automatic reasoning and rule formation appeared in Britain, Germany, and the United States, with an operational prototype (GENEX) developed at Hannover. Buttenfield and Mark (1991) proposed an expert system for map design which included an inference engine to drive generalisation. Their typology for generalisation included operators to simplify geometric detail, classify attribute detail, and enhance detail with "… the purposeful and controlled introduction of information to augment or emphasize structures already present in the data" (Buttenfield and Mark 1991, p. 137). Their system was never implemented. Weibel (1991) argued that the lack of progress in operationalising an expert system was due in part to an incomplete understanding about generalisation operators and their interactions. He proposed amplifying the human acuity for "holistic reasoning, visual perception, [and] design" (Weibel 1991, p. 177) with machine acuity for handling repetitive or tedious tasks.

Various proposals for generalisation based upon constraints, rule formation, and amplified intelligence happened concurrently with adoption of object-oriented programming methods, and continued advances in processing speeds and data storage. Beard (1991) identified graphic, structural, application, and procedural constraints on generalisation operators, which respectively preserve display legibility, protect spatial relations, constrain generalisation according to map purpose, and ensure correct operator sequencing. The AGENT project (Ruas and Duchêne 2007; Ruas 1999) encapsulated strategies to eliminate spatial conflicts introduced by generalisation operators. Constraints guided the selection and trial of strategies within which the objects could self-modify. The AGENT project significantly advanced formalisation of automatic reasoning, and initiated examination about the impacts of varying geographical context upon generalisation.

Finally, Roth et al. (2011) propose a detailed typology of operators for multi-scale mapping that includes map design modifications on symbols and labels with modifications of feature geometry and attributes. Operators are categorised based on their impact on map content, geometry, symbol, and label. The typology has greater emphasis on map design than earlier approaches with the expectation that additional operators may be included as advancements are made in multi-scale and web-based mapping.

6.2.1 Vector-Based Operators

Figure 6.1 places the operators into three categories based on the intended function of the operator. Vector data may be pre-processed through enrichment or reclassification operators to enable subsequent operators that affect the amount of detail (visual quantity) or aesthetics (visual quality) of features retained for cartographic display. Arrows suggest the functional sequence of operators (or operations) typically implemented in workflows, with larger arrows representing more common approaches. Operators that reduce the quantity of content, or clutter on a visual display usually occur before retained features are massaged through operators that affect the visual clarity intended for the display, ergo visual quality. Less often, visual quality operators may require additional removal of content to achieve desired results, e.g. constraints for displacement or alignment may not be satisfied with

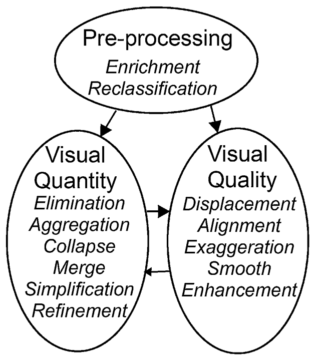

Fig. 6.1 Functional classification of primary vector generalisation operators. *Arrows* represent typical sequences of operators used in generalisation workflows

existing content mandating further feature elimination. This set of vector operators is presented here to provide a quick and simple overview. It follows Ratajski's (1967) concepts for map production, but with the addition of pre-processing functions. It closely resembles Morrison's (1974) typology, and the Roth et al. (2011) typology without the symbol and label operators except enhancement.

6.2.1.1 Pre-Processing Operators

Enrichment adds feature attribution to describe aspects of geometry, existing attributes, and/or local geography that are not explicitly stored in the original database. For instance, prominence of hydrographic network features may be derived by adding upstream length or drainage area, stream order, or hydrologic index (Verdin and Verdin 1999; Ai et al. 2006; Stanislawski 2009; Savino et al. 2011a; Wu et al. 2012). Information gathered from "structure recognition" or "structure analysis" (Steiniger and Weibel 2007) also provides enrichment information.

The *Reclassify* operator groups features based on existing attribution, including enriched data. Reclassification is a common form of pre-processing that facilitates the subsequent application of generalisation operations.

6.2.1.2 Operators Affecting Quantity of Visual Information

Elimination refers to removal of one or more features without replacement. It is generally used to reduce feature content for display at a reduced scale. It has also been referred to as *selection* (McMaster and Shea 1992; Jiang and Claramunt 2004; Gülgen and Gökgöz 2008; Touya 2010), *class selection* (Foerester et al. 2007), *extraction* (Wu et al. 2012), *thinning* (Punt and Watkins 2010; Briat et al. 2011; Stanislawski et al. 2012a; Brewer et al. 2013a, b), or *pruning* (Stanislawski 2009; Stanislawski and Savino 2011).

An *Aggregation* operation replaces many related features with a representative feature of increased dimension, such as replacing many point features with a single polygon feature. To clarify, point features may be considered zero dimensional, lines one dimensional, and polygons two dimensional. In slight contrast McMaster and Shea (1992) exclusively define aggregation as replacement of multiple points with a polygon.

Merge replaces more than one feature with a representation of equal dimension, such as creating a single envelope around multiple proximal polygons. In other typologies merge operates exclusively on linear features, while *amalgamate* operates exclusively on polygons. Some authors refer to aggregation only when combining isolated features, regardless of dimension, and refer to merging when combining connected or adjacent features. Here, as in Roth et al. (2011), we use change in dimension, or lack of, by an operator for a more concise distinction between operators where possible.

Collapse replaces a feature with a representation of lower dimension, such as replacing a polygon with a point or a line. *Refinement* is used to reduce multiple features or sets of features to a more simple representation of fewer features. For instance, a network of streams and canals near the shoreline of a body of water may be converted to polygonal delta feature through refinement. McMaster and Shea (1992) consider refinement as a process to reduce clutter in a display, after the primary elimination process. When numerous features are replaced by fewer features (or symbols) of the same type, it is referred to as *typification* (Regnauld and McMaster 2007; Foerster et al. 2007; Roth et al. 2011).

Simplification reduces the number of points used to represent a line or polygon boundary. Algorithms that *filter* vertices from lines fit into this operator class.

6.2.1.3 Operators Affecting Quality of Visual Information

Displacement adjusts the location of a feature to avoid coalescence with nearby features while maintaining topological integrity with each feature. In some instances, displacement is affine, while other typologies may augment positional shifting with shape adjustment, such as for building generalisation.

Alignment rotates or adjusts the orientation of a feature to maintain or emphasize its relation to other, proximal or adjacent features. This operation may also be referred to as *rotation* or *squaring*, and is sometimes subsumed under displacement.

Smoothing removes small variations in the geometry of a feature to improve its appearance. Smoothing can insert additional coordinates to protect against abrupt changes in directionality of shape, or modify original coordinates as in the case of low- or high-pass filtering.

Exaggeration refers to amplification of a specific part of a feature to maintain the clarity or a particular aspect of the feature. *Caricature* or *enlargement* operators may be included in exaggeration.

Enhancement includes graphic embellishments around or within a symbolised feature to maintain or emphasise spatial relationships.

6.2.2 Raster Operators

We might argue that the compression of raster images is a form of map generalisation, but overall vector based approaches have come to dominate. However McMaster and Monmonier (1989) did distinguish four types of raster generalisation, and Jenny et al. (2011) applied multi-scale Laplacian pyramids to generalise terrain and used curvature coefficients, similar to Leonowicz et al. (2010), to preserve or accentuate edges or other important relief features. Numerical categorisation techniques (McMaster and Monmonier 1989) reduce content of raster data by image segmentation, cell-value slicing, classification, zonal statistics, and channel and ridge extraction. Applications of numerical categorisation include formation of morphologically similar terrain partitions (Chaudry and Mackaness 2008a), and construction of natural drainage density partitions for hydrologic generalisation (Stanislawski et al. 2012b).

Weibel (1992) focused specifically on terrain and identified three raster strategies including global filtering, iterative filtering, and a heuristic approach based upon structure lines. Schylberg (1993) introduced a set of raster ("area feature") operators that process contiguous pixels from the same category. Raster techniques see important application in altering image resolution through either resampling or aggregation. As an alternative to such global operators, neighbourhood operators can be applied in order to generalise portions of a raster based on a user-specified search radius, or a moving window (Fotheringham et al. 2000).

6.3 Operators in Commercial Software

The Euro SDR project (Stoter et al. 2010) was a major coordinated research effort by several European National Mapping agencies, academia, and software vendors to evaluate four state-of-the-art (at the time of testing) commercial generalisation software systems. The project identified strengths and weaknesses of the systems, which stimulated vendor enhancements; suggesting that further customisations are needed, particularly to detect and handle differing geographical contexts. Here we provide a brief description of five commercial geographic information system (GIS) software systems that offer tools or systems for cartographic generalisation. Table 6.1 provides a summary of vector operators that are available in each of these systems. Most of the systems were described in Regnauld and McMaster (2007) and evaluated in the Euro SDR project (Stoter et al. 2010).

Axpand™ ng is a product of Axes Systems in Switzerland. The system employs about 40 different algorithms in combinations or 'operators'. It is a constraint-based system founded on a multi-representation (MRDB) data model (Sarjakoski 2007; Bobzien et al. 2008) and includes automatic generalisation and incremental updating through selective re-generalisation of updates to the source data. Axpand ng is process-based and operates through workflow processing. Operators and constraints are invoked from within a workflow, which can contain sub-workflows for specialised generalisation of 'zones', or regions within the data with constraints specific to that region. This newer axpand ng system was not tested during EuroSDR research. Section 11.4 demonstrates axpand ng.

Esri's ArcGIS® software (version 10.1) has a cartography toolbox that includes tools for cartographic generalisation. It furnishes a geoprocessing environment with numerous functions that can be sequenced into algorithms and tools through the Python scripting language. The environment enables automated partitioning for processing large datasets.

GIPS is Intergraph® Corporation's Geospatial Intelligence Production Solution set of software products that is built upon the GeoMedia® family. GIPS provides a

Table 6.1 A summary of vector generalisation operators available in commercial GIS software systems: Axes axpandTM ng, Esri ArcGIS®, Intergraph GeoMedia® GIPS, Leibniz University of Hanover's CPT, and 1Spatial's Radius Clarity

Operators/algorithms	Axes expand ng	ArcGIS 10.1	Geomedia GIPS	Hanover CPT	1Spatial clarity
Line simplification					
Nth point	✔		✔		
Douglas	✔	✔	✔		✔
Lang			✔		
Reuman-witkam			✔		
House algorithms	✔	✔	✔	✔	✔
Line smoothing					
Brophy			✔		
Averaging	✔		✔		
McConalogue interpolation					✔
Bezier interpolation	✔	✔			
Akima interpolation					✔
Other cubic splines	✔	✔			✔
House algorithms	✔	✔		✔	
Exaggeration/Bend caricature					
Accordion	✔				✔
Min break	✔				✔
Max break	✔				✔
Bend removal	✔				✔
Exaggeration	✔			✔	✔
All-in-one line generalisation					
Plaster	✔				✔
Generalise-by-parts	✔				✔
Line merging					
Blend line	✔	✔	✔		
Area simplification					
Irregular shape	✔	✔			
Orthogonal shape	✔	✔	✔	✔	✔
Turn to rectangle	✔	✔	✔	✔	✔
Area enhancement					
Area extend	✔			✔	
Squaring	✔	✔	✔		✔
Merge/Aggregation					
Merging	✔	✔	✔		✔
Irregular amalgamation	✔	✔		✔	✔
Orthogonal amalgamation	✔	✔	✔	✔	
Point aggregation	✔	✔	✔		✔
Collapse					
Area > Point	✔	✔	✔		✔
Line > Point	✔	✔			✔

(continued)

Table 6.1 (continued)

Operators/algorithms	Axes expand ng	ArcGIS 10.1	Geomedia GIPS	Hanover CPT	1Spatial clarity
Points > Point	✔	✔			✔
Area > Line	✔	✔	✔		✔
Area > Edge	✔				
2 Lines > Line	✔	✔			✔
Refinement/Typification					
Points	✔			✔	✔
Network simplification					
Street network	✔	✔	✔		✔
River network	✔		✔		✔
Neighbourhood detection					
Hierarchical Blend line	✔		✔		
Exaggeration/Area enlargement					
Scaling	✔		✔	✔	✔
Enlarge to rectangle	✔		✔		✔
Enlarge bottleneck	✔				✔
Conflict detection	✔	✔			✔
Clustering	✔	✔			✔
Displacement					
Vertex	✔	✔		✔	✔
Holistic	✔	✔		✔	✔

rich set of data capture, review, and editing as well as product finishing capabilities for geospatial data. The GIPS Feature Cartographer product provides tools for using generalisation functionality in a production mapping environment.

Change, Push and Typify (CPT) provides parameterised batch tools which simplify, aggregate, displace, deform (i.e. object specific geometric operations, such as exaggerate), and typify (refine) objects. The system was developed by the Institute of Cartography and Geoinformatics, Leibniz University of Hanover. Change and Typify are designed for generalisation of building features, whereas Push performs holistic displacement of any feature type.

1Spatial's Radius Clarity (now rebranded as 1Generalise) is a constraint-based, object-oriented system, with a multi-agent framework designed to find optimum generalisation solutions in complex situations, especially where contextual relationships between geographic features are important. The operators iterate towards a solution which maximises the satisfaction of a set of constraints. A framework for operator sequencing is also available, as well as batch processing. The system requires configuration of numerous parameters and constraints.

6.4 Recent Advances in Operator Development

Since publication of the 2007 ICA book (Mackaness et al. 2007), much research and development has focused on advancing methods to automate generalisation over large areas, such as a country, or on large databases. Consequently work has focused on identifying and implementing rule-based constraints to control generalisation, enrichment, tailoring sequences of operations for contextual classes, and metrically assessing data and generalisation results to validate and refine processes.

6.4.1 *Enrichment*

Most developments in generalisation frameworks extend the premise that "place matters", developing generalisation strategies which vary depending on local spatial context or geographic conditions. These contexts are made explicit through the process of 'enrichment'—an essential pre-processing stage.

The most common reason for enriching data with regard to generalisation is to assign relative prominence estimates to data for use in subsequent elimination operations. In addition, data enrichment formalises spatial relations and adds data characteristics explicitly to the set of attributes (Bobzien et al. 2008). Detected and defined patterns in spatial data or local geography can guide generalisation to retain feature characteristics based on context, spatial distribution, and geographic conditions (Buttenfield et al. 2011; Touya et al. 2010). Steiniger and Weibel (2007) enrich cartographic data with vertical and horizontal relations to reduce subsequent processing. Neun et al. (2008) consider enrichment a labour-intensive yet mandatory operation for web-based generalisation services.

Enrichment can be applied to all forms of vector data (i.e., lines, networks, polygons, points) for generalisation purposes. Zhang et al. (2010) utilise enrichment to preserve shape and alignment in generalised building polygons. Steiniger et al. (2008) use supervised classification by way of discriminant analysis of geometric shape of buildings to assign urban structure characteristics for building generalisation. Use of geometric algorithms or statistical measures have been investigated for defining building patterns (Christophe and Ruas 2002; Zhang et al. 2010) or detecting island structures (Steiniger et al. 2006).

6.4.1.1 Partitioning

Partitioning is typically performed in order to subdivide data into manageable units. Partitioning data into manageable units is regularly performed for parallel processing of raster data (Wallace et al. 2010). With ArcGIS 10.1, Esri introduced partitioning to enable processing of large datasets for several algorithms, including road network thinning (Briat et al. 2011; Esri 2012).

In addition, partitioning may be used as an enrichment operator to form context-based geographic spaces that allow proper application of site-specific generalisation operations (Ruas 1995; Bobzien et al. 2008; Chaudhry and Mackaness 2008a; Stanislawski and Buttenfield 2011; Touya 2010; Touya et al. 2010). Partitioning for contextual classification may be completed on individual or multiple data themes, and the partitioning method depends on the generalisation strategy. For generalisation purposes, a partition is a logically-derived subset of a larger dataset which is to be separately processed with one or more generalisation operators or algorithms. Among others, partitions may be referred to as clusters, areas, regions, or groups.

For example, vector and raster algorithms were used to assign line-density partitions to hydrography and road network features to help retain density variations during generalisation (Stanislawski 2009; Stanislawski and Buttenfield 2011; Stanislawski et al. 2012a; Brewer et al. 2012).

Chaudhry and Mackaness (2008b) clustered buildings around road nodes to estimate settlement boundaries and develop partonomic relations for a multiple representation database and for generalisation. Werder et al. (2010) proposed an unsupervised classification to cluster buildings into settlement areas based on geometric and topological characteristics of the building outlines and road network data. Bildirici et al. (2011) combined buffers around building point features separated by road lines to build point clusters for typification algorithms.

Raposo et al. (2013) used a scale-dependent regular tessellation to subdivide topographic summit points for enrichment and thinning operations. Bereuter and Weibel (2013) apply a quadtree tessellation to subdivide point data for real-time generalisation (Sect. 6.7).

6.4.2 Transformations of Groups of Objects

It is often necessary to simultaneously process groups of objects (points, lines, polygons) through generalisation operators in order to retain relative geometric characteristics among the features through scale changes. This section describes some recent research or new approaches involving transformation of groups of objects.

Bildirici et al. (2011) proposed typification operators that use length and angle measures between the points (buildings) in each cluster to retain the geometric structure of each cluster through scale reduction. Also, an incremental approach for automatically displacing building points away from road lines and other points was proposed for generalisation (Aslan et al. 2012).

Working with point objects, Raposo et al. (2013) propose an automated way to separately thin each partition of points by summit prominence, which is enriched on the points through integration with elevation contours. Partitions are derived through a scale-dependent regular tessellation. The objective is to automatically select the most prominent set of summit points for legibly labeling summit points through multi-scale displays. Bereuter and Weibel (2013) partition point data using quadtrees and present operators for selection, simplification, aggregation, and displacement for on-the-fly generalisation (Sect. 6.7).

van Oosterom (2005) and van Oosterom and Meijers (2011) present a truly vario-scale structure for smooth zooming through a polygonal thematic map (Chap. 4).

6.4.3 Generalisation of Networks

Networks have a variety of geographical uses, which include mapping transportation and hydrography, routing materials and services, and hydrologic modelling. Increasingly, networks are being used as a reference framework for relating ancillary data or for managing supply chains (Long et al. 2012; Simley and Doumbouya 2012; Yager et al. 2012). Vector geospatial data for networks may be stored in a directed or undirected graph structure, which enable traversal and other network analysis functions. Because the spatial pattern and connectivity of a network affects interpretation and modelling, networks are typically generalised as a macro object through super-operators or processes, which combine several operators on a group of features.

Several methods to generalise hydrography or road networks have been developed, and research continues to refine these methods. Development of processes or algorithms for building a hierarchy of 'strokes' or paths of best continuation (Thomson and Richardson 1999; Thomson and Brooks 2007; Chaudhry and Mackaness 2005) is still a focus for network thinning (a vector elimination operator, Sect. 6.2.1) strategies. Prominence estimates help define and rank strokes in the network.

For river networks, some factors contributing to overall prominence are stream order, stream name, longest path, drainage area, straightness, and upstream branches (Ai et al. 2006; Touya 2007; Stanislawski 2009; Savino et al. 2011b; Gutman 2012). In the United States, each network feature in the National Hydrography Dataset (NHD) is assigned a unique permanent 'reach code' address that defines a continuous segment of surface water in the network, which often spans more than one feature. Instead of generating strokes to thin the NHD network, reach codes are assigned prominence estimates based on enriched upstream drainage area estimates. Local density is also assigned to NHD network features through a line density partitioning algorithm. A stratified pruning process then thins each partition to a prescribed target density, which ensures that natural density variations reflecting local geographic conditions are maintained (Stanislawski 2009; Buttenfield et al. 2011; Stanislawski and Savino 2011).

Recent research on network generalisation has focused on refining thinning strategies to account for local contextual variations related to local geography. Network thinning operations have been enabled or refined through enrichment of local line density (Savino et al. 2011a; Stanislawski et al. 2012a; Benz and Weibel 2013), pattern (Heinzle et al. 2005, 2007; Touya 2007; Savino et al. 2011b), road type (Balboa and Lopez 2008; Savino et al. 2010; Stanislawski et al. 2012a), and road network or block structure (Jiang and Claramunt 2004; Touya 2010; Gülgen and Gökgöz 2011). With regard to road networks, Zhou and Li (2012) evaluated several stroke-building strategies. Later, Li and Zhou (2012) combined stroke and mesh-density strategies (Chen et al. 2009) to thin road networks. A mesh is a region enclosed by network roads, and each mesh is assigned a density. Mesh-density thinning progressively removes edges of the highest density mesh until a minimum density is achieved. Benz and Weibel (2013) further researched combined stroke- and mesh-based road thinning, adding refinements to maintain density patterns of settlement areas. Stanislawski et al. (2012a), Brewer et al. (2013a, b) tested road network thinning stratified by density between rural and urban areas. Results indicate the process will support thinning of road and road labels for multi-scale display (Sect. 6.5).

Other generalisation research on road networks has focused on techniques or algorithms to remove excess detail (e.g., roundabouts, divided highway) at the appropriate level of detail, which can enhance network thinning and road labelling through scales (Brewer et al. 2013a, b; Weiss and Weibel 2013).

6.5 Case Study I: Generalisation of Road Networks

Lawrence V. Stanislawski and Cynthia A. Brewer

This section describes development of a workflow to implement a super-operator that thins (eliminates) road network data to multiple levels of detail for cartographic display and data delivery. A goal of the United States Geological Survey (USGS) is to implement automated generalisation to enable multi-scale display and delivery of transportation and other geospatial data available through the USGS *National Map* (Sugarbaker and Carswell 2011). *The National Map* transportation data displayed on 1:24,000 (24k) US Topo maps (Sect. 11.6) are currently compiled by TomTom® North America and the U.S. Forest Service and stored in the Best Practices (BP) Data Model using the Esri geodatabase format (USGS 2006). The goal is to thin the BP road network data, using commercially available tools, to levels of detail appropriate for scales ranging from about 1:20,000 (20k) to 1:1,000,000 (1M). This research is being conducted by the USGS Center of Excellence for Geospatial Information Science (CEGIS) in collaboration with University of Colorado-Boulder and Pennsylvania State University.

6.5.1 Context and Objective

Esri ArcGIS® version 10 and later includes the Thin Road Network tool that applies simulated annealing to automate road network thinning (Punt and Watkins 2010; Briat et al. 2011). Connectivity, general morphology of road patterns, and integrity of navigable routes are maintained in the thinned network. The thinning algorithm considers relative importance, significance, and density of input features. Importance is determined through the hierarchical road class (e.g., Interstate, State Route, local road, etc.) assigned as an attribute on the input data. Feature significance or functional relevance (Thomson and Richardson 1995) is determined as a feature's participation in long routes across the extent of the network. Features that are part of long network routes are deemed more significant than those required only for local travel. Density is computed similar to any density metric, as a ratio of the length of a street segment to its associated area.

The tool can be used to thin a topologically clean road network that has a hierarchy field. Therefore, use of the tool usually requires pre-processing to populate the hierarchy and importance fields and enhance the integrity of the input network features. The tool does not actually thin the network, but rather enriches the network with a binary "invisibility" field populated with one (invisible) or zero (visible) determined by the thinning algorithm and input parameters. The level of thinning from a single run of the tool is controlled by the importance field and "minimum length" parameters. The minimum length value is a tolerance that roughly corresponds to the shortest road segment that is visually sensible to include in the thinned network at the final scale (Esri 2011). The network can be thinned based on results stored in one or more invisibility fields populated by one or more runs of the tool.

The relation between minimum length and resulting network density is not consistent, and depends on conditions of the input road network, which vary with local geography. In addition, CEGIS researchers have noted that using the Thin Road Network tool with a uniform minimum length tolerance to thin an area of roads with substantial line density variations tends to homogenise density variations at smaller scales more than expected by the Radical Law (Töpfer and Pillewizer 1966). Consequently, a stratified network thinning strategy, similar to the methods applied to the NHD (Sect. 6.4.3), can be used to better maintain density variations, although this method requires multiple runs of the tool on the same dataset (Stanislawski et al. 2012a; Brewer et al. 2012).

6.5.2 Methods

To facilitate processing with the Thin Road Network tool, data partitions (Fig. 6.2, black lines) of up to 50,000 features were created for the BP road layer for the United States with Create Cartographic Partitions (Esri 2012). The cartographic

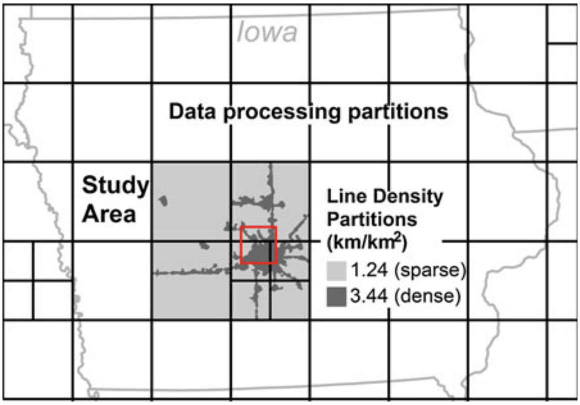

Fig. 6.2 Line-density partitions (*light and dark grey shaded areas*) generated for TomTom/USFS road features within the study area of seven data partitions (*bold black lines*) covering the Des Moines, Iowa area. *Redsquare* near middle of study area shows location of Fig. 6.3 examples

partitioning tool uses a quadtree-based approach to generate partitions, which seek to minimise the effect of partition shape on processing algorithms. A study area of seven data partitions (Fig. 6.2, light and dark grey shaded areas) covering the Des Moines, Iowa region were used to select a subset of BP road features for testing. This study area covers about 21,600 km^2 and includes over 104,000 BP road features.

The set of road features in the study area were prepared for the Thin Road Network tool to ensure that proper network topology and attribution exists. Pre-processing of feature geometry included projecting data from geographic coordinates to the North American Albers Equal Area Conic projection, removal of coincident (overlapping duplicate) features where necessary, conversion of multi-part to single-part features, and ensuring that features are split at all intersections. Feature attribution was pre-processed to transfer road names to retained features when removing coincident features, to populate the road class where needed based on feature names and types, and to assign importance values based on road class values. Custom tools were developed with Python and Esri geoprocessor functions to automate the pre-processing tasks.

Two line density partitions were generated for the test data using the raster-based partitioning process described by Stanislawski and Buttenfield (2011) with a density class break of 1.50 km/km^2 and a minimum polygon area tolerance of 45 km^2 (Fig. 6.2). Density partitioning assigns a density class to each road feature. The two partitions generated have overall road densities of 1.24 and 3.44 km/km^2 for the sparse and dense partitions, respectively.

Subsequently, the Thin Road Network tool was run in a ladder fashion on the test data, thinning with a series of minimum lengths incrementing from 500 m to 30 km. The resulting visibility settings were combined to produce a single 35-level attribute used to remove the least important roads in each partition through scale.

6.5.3 Results and Discussion

Table 6.2 lists rounded scales calculated for each thinning level for the dense and sparse partitions, as well as for the study area as a whole. Target scales are calculated by inverting the modified Radical Law (Buttenfield et al. 2011) to compare total length of visible roads to the original length of all roads at the source scale of 24k (listed as 24 across the first row of Table 6.2). Looking across each row, scales become smaller from sparse to whole to dense areas (left to right). This trend shows that the denser areas are being thinned more aggressively than the sparse areas causing the whole study to become more homogeneous in density at a single thinning level. This pattern across rows lessens as scale decreases, and at a laddered minimum length of 23 km the target scale for both partitions is fairly similar. We are not yet able to direct the Thin Road Network tool to produce a specific density of features, but this inventory of lengths helps understand the results across dense and sparse areas through scale, with the goal of producing general guidance on use of this thinning algorithm.

Looking at Table 6.2 to find shared approximate scales offers direction on how to combine thinning levels to retain a relative sense of dense and sparse character in the road network that is suitable for a particular scale. Examples of similar scales in different rows in the Table 6.2 are highlighted by hue. For example, aiming at 100k, orange highlights in the table show 10 km thinning is suitable for approximately 106k scale mapping overall, but thinning with 11 and 4 km best approaches 100k scale for the sparse and dense partitions, respectively. Pink highlights show a similar pattern for about 400k, with less contrast in the thinning levels combined for the two density partitions.

Table 6.2 also shows that at longer minimum lengths (greater than 23 km) the thinning scales invert between sparse and dense partitions. It has been demonstrated in later research that this scale inversion can be eliminated by merging divided highways in the road network prior to the running the Thin Road Network Tool (Brewer et al. 2013b). Furthermore, elongated patterns in the line density partitions that follow divided highway (Fig. 6.2) may be eliminated by merging divided highways prior to generating the line density partitions.

Figure 6.3 demonstrates the example thinnings for 100k (b and c) and 400k (d), with all maps drawn at 500k for comparison (Fig. 6.3a showing all roads has a target scale of 24k). Figure 6.3c has more representative densities in the urban areas to the lower left of the sample area (in the purple, dense partition shown in Fig. 6.3a) with all roads adequately thinned for small-scale display. Short dangling roads near partition boundaries were trimmed.

6.5.4 Next Steps

Using similar methods to thin other BP road data, Stanislawski et al. (2012a) noted that target 100k thinning densities estimated through the Radical Law are

Table 6.2 Levels of thinning for the study area shown in Fig. 6.2. For one level of thinning (row), each column lists suitable scales calculated using the total length of road segments by density partition (e.g., a scale of 34 represents 1:34,000). Suitable scales are computed through the Radical Law based on length (Töpfer and Pillewizer 1966; Brewer et al. 2013a)

Thin Road Network (minimum length)	scale sparse partition	scale whole study area	scale dense partition	
all roads	24	24	24	
500m	26	28	34	
1000m	28	32	47	
1500m	29	35	57	
2000m	30	36	66	
2500m	30	38	74	
3000m	31	39	82	
3500m	31	40	88	
4000m	32	42	96	
5000m	34	45	111	
6000m	43	56	128	
7000m	54	69	143	
8000m	64	81	167	
9000m	71	90	185	
10km	85	106	208	
11km	100	122	225	
12km	119	142	240	
13km	135	160	259	
14km	155	185	311	
15km	178	210	338	
16km	208	241	366	
17km	243	275	387	
18km	271	303	409	
19km	307	335	424	
20km	352	376	445	
21km	391	409	459	
22km	440	453	490	
23km	507	504	498	inverts
24km	567	549	509	
25km	623	593	527	
26km	834	734	552	
27km	928	792	562	
28km	1,045	859	570	
29km	1,143	912	573	
30km	1,205	976	630	

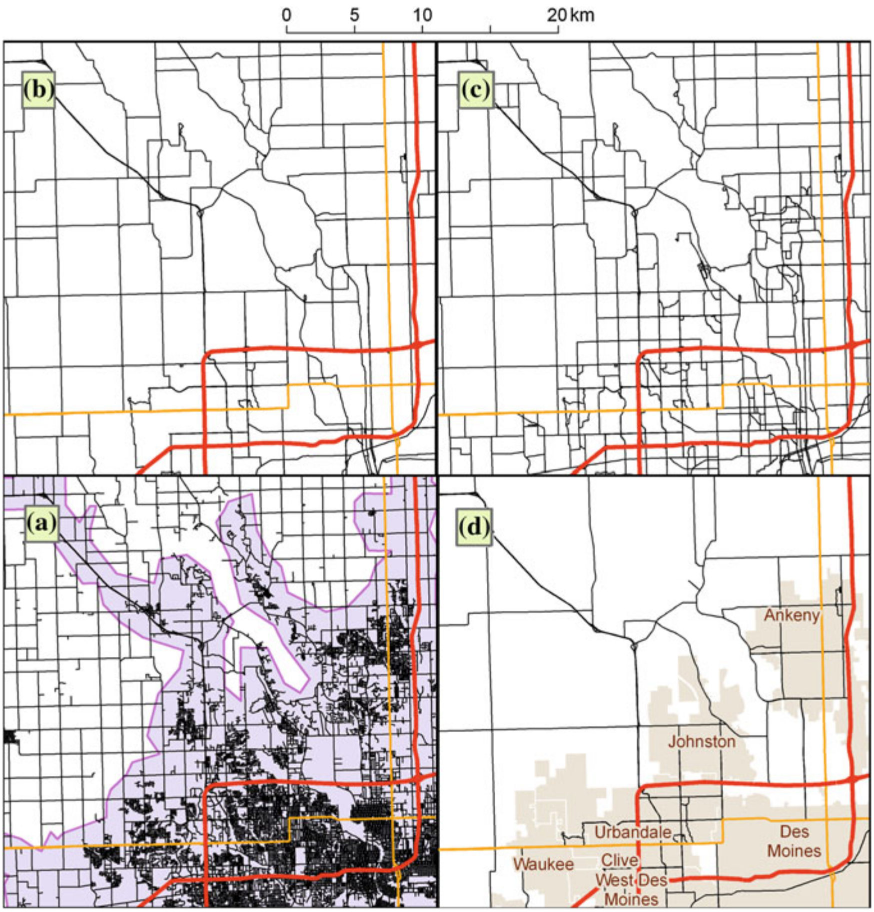

Fig. 6.3 Results of using Esri Thin Road Network tool on **a** original road data from USGS Best Practices data model, with line density partition in *purple*. Original roads thinned with minimum lengths of **b** 11 km, **c** 4 and 11 km in dense and sparse partitions, respectively, with trimming of short features at partition boundaries, and **d** 18 and 21 km in dense and sparse partition, respectively, with trimming of short features (incorporated places in beige, named for location reference). The four maps are each displayed at 500k

substantially less than benchmark densities from 100k USGS road data. USGS standards indicate that the vast majority of 24k road features are retained for 100k mapping (USGS 1985). The primary purpose of the map will govern the density and prioritisation of feature content. Scale-dependent densities governed by the Radical Law or other scaling rules (Jiang et al. 2013) may be adequate for interactive review of data through The National Map Viewer, but greater dominance of transportation features, as in historical USGS Topographic maps, may still be required for large and medium scale US Topo products. These decisions are yet to be determined, but enriching BP data with thinning levels as described in this section provides the flexibility to thin the data as needed for each purpose.

This approach is being tested on a variety of geographical locations across the United States to develop general thinning recommendations. Upon compiling results for other locations, an optimal implementation strategy will be designed to automate the process on the entire BP transportation theme of *The National Map*, including railroads. The design should minimise processing time (i.e. the number of runs of the tool), the number of density partitions, and the number levels of detail (LoD) datasets that must be generated and stored. Each LoD will be used for mapping a range of scales, which must be determined, and appropriate invisibility fields for each density must be joined to the attributes of each LoD. Among other details, research must identify appropriate simplification operators for each LoD and verify the LoD best suited for merging divided highways.

In other work on the same test data, Brewer et al. (2013a, b) identified a set of LoD scale ranges adequate for the target display scales of 20k to 1M and made use of thinning levels to prioritise road labels. Their system for road labelling will be further tested. Additional work with transportation data will examine how these thinning levels combine with other base information, such as hydrography, terrain, and populated place labels for complete mapping. Appropriate ranking of themes and features for displacement operators and integration (snapping) must be identified and available algorithms must be tested.

6.6 Case Study II: River Network Pruning by Enrichment and Density Analysis

Sandro Savino

This section illustrates an algorithm for the generalisation of river networks developed at the University of Padua, Italy. The network pruning algorithm enriches the network features with relevant information to assess the importance of each part of the network. Local and global density analyses are also used to further improve the selection process. The algorithm can generalise both natural and artificial streams, although an approach involving typification is probably best suited for generalising artificial streams that show regular patterns (e.g. irrigation systems) (Savino et al. 2011b). The algorithm was originally developed to generalise 1:5,000 scale data to 1:25,000 scale, but by tuning its parameters it was successfully adapted to data at different scales (Stanislawski and Savino 2011).

The workflow can be divided into four distinct operations: data enrichment, network pruning, braided sections generalisation, and density tuning. Where not otherwise stated, it is assumed that the network is composed only of linear elements. In cases where hydrographic drainage features include areal elements, their linear counterpart should be generated (e.g., by calculating the middle axis) in order to be processed (e.g., Regnauld and Mackaness 2006). In the following sections, a *feature* refers to a single linear element of the hydrographic network that may include one or more vertices between its endpoints. The term *river* or *stream* refers to an object spanning more than one feature, and the point where two or more features connect is called a *node*.

6.6.1 Data Enrichment

During enrichment, important data are extracted from the input hydrographic data and stored in a data structure that is used to drive the selection process. Extraction of important data components is based on input data attributes and channel morphological analysis.

The first step is to understand the relationship of each feature with the neighbouring ones. Given a feature R, all the features connected to it are classified either as 'fathers', 'children' or 'siblings': a father is a feature ending into R, a child is a feature beginning from R, and a sibling is a feature sharing a father or a child with R. The classification is based on the flow direction of each feature; in case this information is not explicitly available, it is computed by analyzing the values of the Z coordinate of each feature, when available.

Given a feature R, its two vertices with highest and lowest Z value pR_{zmax} and pR_{zmin} and their values R_{zmax} and R_{zmin}, and a feature C and its maximum and minimum Z values C_{zmax} and C_{zmin}:

- C is a father of R if $C_{zmax} > R_{zmax}$ and C is connected to R at pR_{zmax}
- C is a child of R if $C_{zmin} < R_{zmin}$ and C is connected to R at pR_{zmin}
- C is a sibling of R if $C_{zmax} > R_{zmin}$ and C is connected to R at pR_{zmin} or if $C_{zmin} < R_{zmax}$ and C is connected to R on pR_{zmax}

The "equal" case is handled as a special case in the algorithm (Savino et al. 2011a), though not discussed here for reasons of brevity.

We also observe that:

- if R has no fathers, it is a source,
- if R has no children, it is a sink,
- the flow direction of R is from pR_{zmax} to pR_{zmin}.

If the flow direction is known, the procedure above can be applied without comparing Z values, assuming for each feature R flowing from its vertex R_a to R_b, $pR_{zmax} = R_a$ and $pR_{zmin} = R_b$.

Once the relationship of each feature with its neighbours is known, the algorithm computes and stores the following measures for each feature R:

- the Strahler order S_r (Strahler 1952)
- the total distance to the furthest source uphill L_r
- the total number of branching points uphill B_r (where multiple features converge)

These values are computed as follows, where A and B are fathers of R:

$S_r = S_a + 1$ if $S_a = S_b,$
otherwise $S_r = \max(S_a, S_b)$
$L_r = \max(L_a, L_b) + \mathrm{length}(R)$
$B_r = B_a + B_b + 1$

The process starts from a randomly chosen source and continues downhill on one of the children of R (each source has values $S_r = 1$ and $B_r = 0$); if R has no children the algorithm picks another source and starts the process from there; if one of the fathers of R has not been visited yet, R is not processed and the algorithm picks another source to process. The whole procedure ends when all the edges have been processed.

As a last step, the algorithm uses the information gathered to detect the rivers in the network. A river is a sequence of connected features, starting at a source and ending either in a sink or into another river. The river is conceptually similar to the idea of a "stroke" (Thomson and Richardson 1999; Thomson and Brooks 2002) and it is the basic unit on which the pruning process is applied.

The process is bottom-up and starts from one of the sinks: for each feature R the algorithm decides which father of R is the best continuation of the river; the choice is performed by giving a score to each father F of R based on the enriched data.

The score for F will be increased if:

- F has the highest value L_F
- F has the highest value B_F
- F and R have a high collinear (straight) alignment

and, if the data contains this information

- F has the same name of R,
- F belongs to the same hydrographic class of R,
- F has the largest width

On the other hand the score for F will be decreased if:

- F has a different name of R
- F has a lower Strahler order than R

The process ends when each feature has been associated to one river (there might be rivers spanning only one feature); each river inherits the S, L, B values of its most downhill feature (this is similar to Horton 1945).

6.6.2 Network Pruning

Pruning selects rivers deemed important enough for the generalised network, and removes the less important ones. The importance of a river is relative to the target scale and is modeled based on user-defined thresholds for S, L and B. When

Table 6.3 Parameters used to generalise data at different scales

	1:25,000	1:50,000	1:100,000
Minimum river length (m)	250	600	1600
Buffer size for density pruning (m)	120	400	1200
P_{max} percentage of buffer overlap (%)	50	50	50
S strahler order	3	3	3
L length to furthest source (m)	1000	1500	3200
B number of branches uphill	4	8	16

deleting a river, all the features composing the river or converging with it (i.e. its father features) are removed from the network.

Selection happens in two steps, the first deletes all the rivers shorter than a minimum length threshold (ideally, too short to be represented at the target scale), while the second prunes the network in the areas where it is too dense, improving the legibility of the output. The first selection deletes all the rivers having L smaller than a minimum value. Removal of full rivers, instead of single features, maintains the connectivity of the network.

The second selection requires assignment of local density to each river, which is performed by drawing a buffer around each river and calculating the percentage of its area that overlaps the buffers drawn around neighbouring rivers. The higher this percentage P, the closer the river is to other rivers. The algorithm sorts the rivers by decreasing values of P and, starting from the river with highest P, analyzes one river at a time, removing only the less important. The importance of a river is assessed comparing its S, L and B values with the user defined thresholds: the river is removed only if all the three values are below the thresholds; upon removal, the values of P of the neighbouring rivers are updated and the list sorted accordingly. The process continues until the highest value of P is below a threshold P_{max} or every river having P bigger than the threshold is deemed too important to be removed. Table 6.3 shows the threshold values used to generalise at different scales; these parameters have been found empirically; the process and results are shown in Fig. 6.4.

6.6.3 Braided Sections Generalisation

Where a river flows in a flat region its stream may split, forming many branches flowing downhill, which may merge and split again. This is referred to as a braided section of a river. Braided sections are characterised by the presence of braid bars, the islands sitting among the braids.

To deal with braids, the algorithm changes focus to generalise the islands instead of the river network (see also Touya 2007). The algorithm targets the islands whose size falls below a user defined threshold; these are either amalgamated with nearby islands or enlarged if isolated. Generalising a braided section,

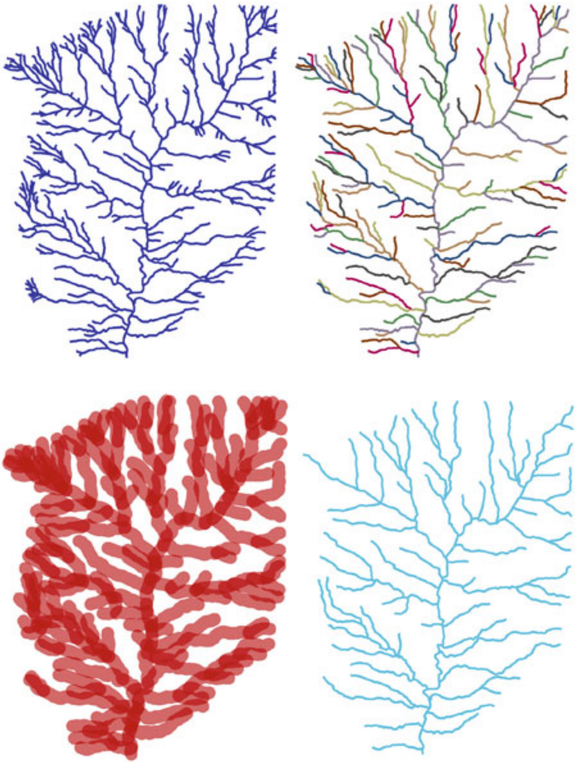

Fig. 6.4 Different steps in the selection process: (*top left*) input network, (*top right*) detected rivers, (*bottom left*) buffers used to compute density, and (*bottom right*) pruned network

the algorithm mainly pursues two objectives: (1) create compact shapes to avoid narrow parts in the generalised islands that could lead to legibility problems, and (2) preserve the streamlined pattern of the network, that is, avoid hard bends in the generalised network.

Braided sections are detected by finding sets of one or more islands where the network lines form polygons. The size of each island is estimated as the area of the associated polygon less the area within the user-defined buffer around the network lines that represents the width of the river (or the actual area of the river polygons where available).

The algorithm processes (Fig. 6.5) each island, starting from the smallest one, looking for candidates for amalgamation among its neighbours. Islands with no neighbours are either deleted or enlarged, depending on their size; enlargement is performed by applying a scaling filter to the geometry that displaces the vertices based upon their closeness to the center of the polygon: vertices closer to the center are displaced further, thus producing shapes that are more compact than results from a buffer operation; the same filter is applied to the braids surrounding the enlarged island.

The algorithm evaluates each amalgamation candidate with respect to the compactness of the resulting amalgamated geometry (compact shapes are

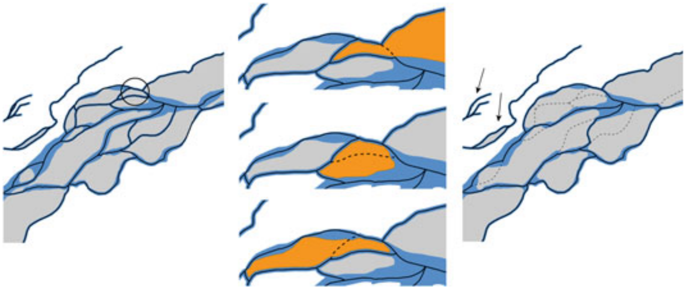

Fig. 6.5 Example amalgamation process of island polygons: (*left*) river polygons and islands of a braided section of the network; (*centre*) selection of the best amalgamation candidate (*orange*) for the island *circled at left*: middle solution is chosen because it produces an island that is more compact than that the top or bottom solution; (*right*) generalised braided section and the former network (*dashed lines*), with *arrows* pointing to two generalised isolated islands, one deleted and one enlarged

favoured, Fig. 6.5 middle center), the size of the braid between the island and the neighbour (narrow braids are favoured), and the angles in the resulting river network (hard bends are not allowed, Fig. 6.5 top center); the best candidate is selected through a scoring mechanism and upon amalgamation, the network and the list of islands are updated. The process ends when all the small islands have been generalised.

6.6.4 Selection Tuning by Density Analysis

In large datasets that span areas with different morphological traits, generalising the hydrographic network with fixed thresholds can produce an output too homogeneous, where local characteristics are lost (Stanislawski and Savino 2011). To handle relative density variations among different areas of the network, the last step of the generalisation process employs a technique that refines the selection process by comparing the local density in both the input and the generalised data.

Density is computed using a regular grid dividing the dataset in cells; the density is defined as the sum of the lengths of the geometries contained in each cell divided by the area of the cell; cells on the boundary of the dataset are not considered to avoid biasing the density calculation. Cells completely covered by braided sections are excluded from the process because braided sections are handled separately as previously described.

The algorithm computes the average cell density, D_{avg}, on the input and generalised datasets and then, for each cell, the difference as a percentage, between the cell density $D_{c,}$ and the average D_{avg}: this difference, $dD_{c,}$ can have a positive or negative value and marks the local variation of the density. By comparing the dD_c value of corresponding cells in the input and in the generalised dataset, it is possible to detect whether local variations have been lost: if dD_c is bigger in the

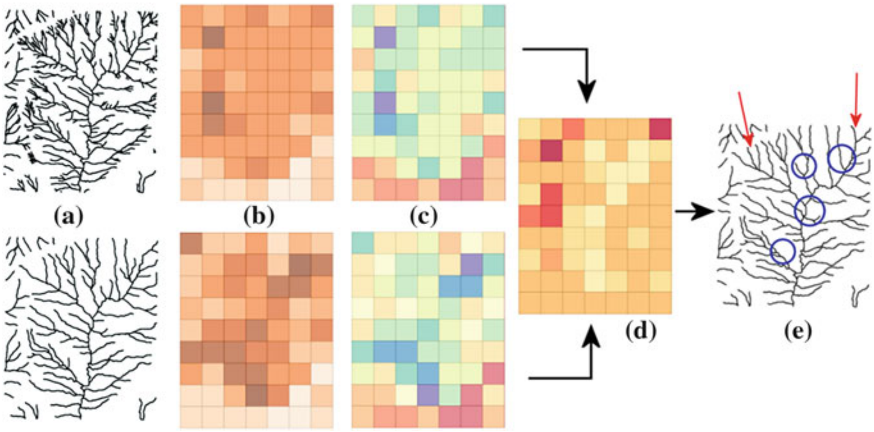

Fig. 6.6 The density tuning process: **a** the networks (*top* input data, *below* generalised data); **b** the density maps (bigger values have *darker colour*); **c** the density difference maps (the darker the colour, the bigger the difference); **d** the density difference comparison map (*dark colours* for over-generalised areas, *light colours* for under-generalised areas) and **e** the tuned network (removed rivers are *circled* while *arrows point* to added rivers)

generalised data, data are locally under generalised, if it is smaller, data are locally over generalised.

The algorithm compensates the density variation by removing or adding rivers in the cell until the target density difference is met. To perform this operation the scoring technique described before is used to choose the most important river to add, among those previously deleted, or to choose the least important river to delete. The algorithm keeps adding (or removing) rivers until the cell density is within a threshold around the desired value. The whole process is illustrated in Fig. 6.6.

6.7 Case Study III: Algorithms for On-the-Fly Generalisation of Point Data Using Quadtrees

Pia Bereuter

On-the-fly (or real-time) generalisation and adaptation to user interaction and content are essential for the dynamic use of web and mobile mapping. Typical applications, such as mashups or location-based services (LBS), usually encompass a thematic foreground layer predominantly in the form of points of interest (POIs) or large point collections (e.g. animal observations or twitter counts), against a spatial reference formed by background data such as a topographic map. Background data are typically rendered by a pre-generalised tile service to ensure seamless map interaction. On the other hand, the content of the foreground data, depending on user requests, requires dynamic adaption and therefore calls for cartographic generalisation in real-time (Weibel and Burghardt 2008). Bereuter and Weibel (2013) and Bereuter et al. (2012) provide a review of the relevant literature and propose several algorithms for real-time point data generalisation based on a quadtree data structure. Here, we present the generalisation of point groups as a Case study, focusing on results and ignoring the technical details. Below a short overview on the generalisation operators is provided, followed by a quantitative and qualitative cartographic analysis on data and conflict reduction, data enhancement, displacement measures and preservation of spatial patterns.

6.7.1 Overview of Generalisation Operators

The algorithms described by Bereuter and Weibel (2013) and Bereuter et al. (2012) provide implementations of the major generalisation operators that can be applied to point data (Sect. 6.2.1). Several algorithms are available for each of the generalisation operators summarised in Table 6.4. The basic idea of the quadtree-based generalisation approach is to apply generalisation operations to quadtree nodes according to the target level of detail (LOD), mapped to the depth of the quadtree and the selected point symbol size. Target LOD translates to the width of the quadnode side, which denotes the smallest required distance to resolve spatial conflicts. The LOD in this Case study is mapped to the zoom level of the background web map tile services (WMTS) and the zoom levels are named accordingly. Bereuter and Weibel (2013) also present a performance analysis of the algorithms.

Table 6.4 Point generalisation operators based on the quadtree (Bereuter and Weibel 2013)

Operator (see Sect. 6.2.1)	Description of implemented algorithms
Pre-processing operators affecting quantity of visual information	
Selection, elimination	Based solely on feature attributes, applying various selection functions per quadnode, such as rank, frequency or feature category distribution
Simplification	Returns one point feature per quadnode, governed by geometric criteria such as centrality, or weighted centrality
Aggregation	Reduces the number of points per quadnode by grouping together semantically similar or spatially close points, replacing the original points by a new placeholder feature, such as midpoint or based on clustering or collocation criteria
Typification (Refinement)	Replaces numerous points by fewer points of the same type within a quadnode
Operator affecting quality of visual information	
Displacement	Locally reconfigures as many point symbols as geometrically possible per zoom level to resolve spatial conflicts by moving points apart from each other using the quadtree for neighbour search

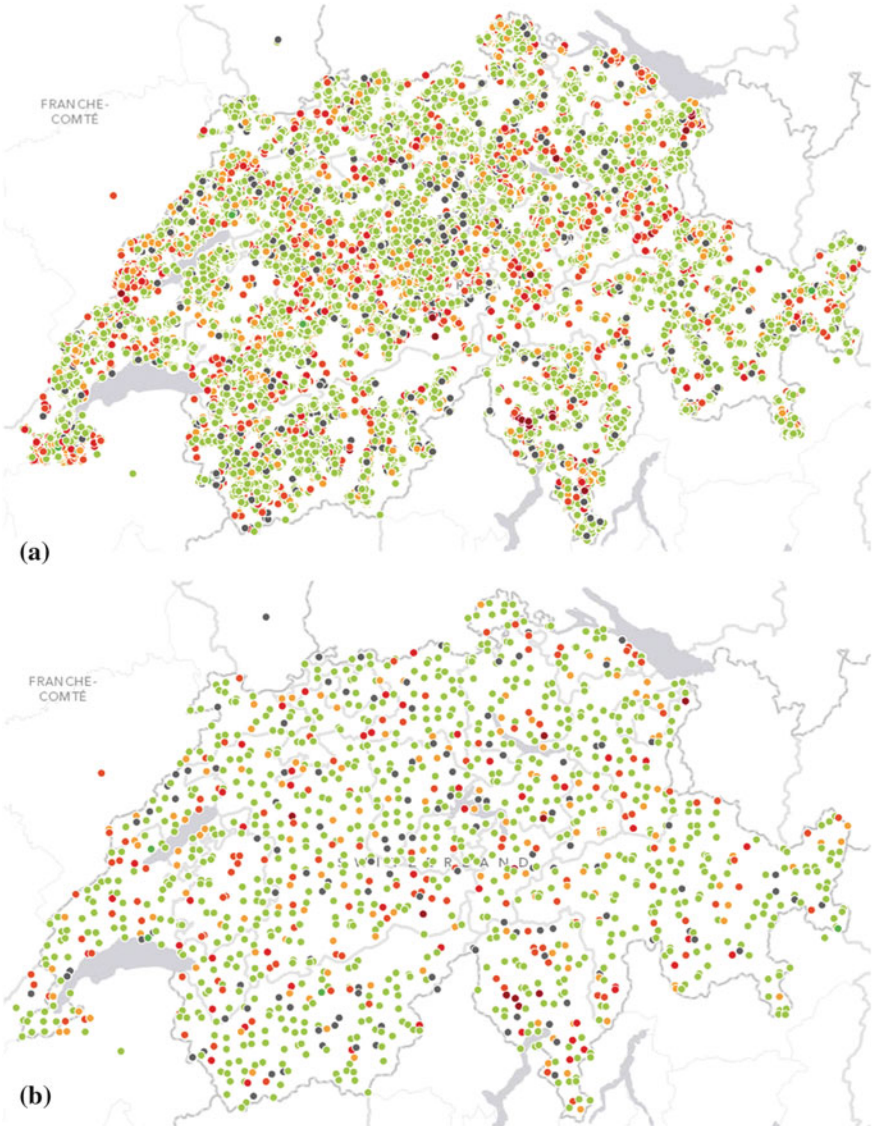

Fig. 6.7 Lichens observation in Switzerland at zoom level 8, color coded from least endangered (*green*) to most endangered (*red*) and no information (*grey*). (*Data* Stofer et al. 2012) *top* **a** raw data, *bottom* **b** centrality-based simplification (1,288 data points)

The following Case study presents several aspects of the introduced generalisation algorithms, implemented in a prototype generalisation platform based on Java and Processing (www.processing.org). The point collection used in this Case study originates from SwissLichens (Stofer et al. 2012), which is a database maintaining past and present population distributions of more than 500 different lichen species at over 86,000 locations within Switzerland.

An overview of the complete dataset is given in Fig. 6.7a, for the area of Switzerland. The map shows all observations of lichens in Switzerland and their red list status. Figure 6.7b shows a generalised view of the same data with the centrality-based simplification applied (Bereuter and Weibel 2013).

The following sections describes a quantitative analysis of the results obtained using the described generalisation algorithms. The analysis focuses on the main cartographic requirements for map generalisation by showing the data and conflict reduction rate, conservation of important point attributes, displacement measures, and the maintenance of spatial patterns.

6.7.2 Data and Conflict Reduction

The data reduction curves for the quadtree-based generalisation algorithms between zoom level, as well as the Radical Law (Töpfer and Pillewizer 1966) with its initial scale at zoom level 20, are shown in Fig. 6.8. The scale for the zoom levels in Fig. 6.8 spans from a small scale of ~1:500,000,000 for zoom level 0 to a large scale of ~1:500 for zoom level 20, with a scale change of factor two between each consecutive zoom level. It shows that quadtree-based operators retain more points than the Radical Law would suggest, mainly due the selected symbol size and the fact that the Radical Law ignores the spatial configuration of the input data. Quadtree-based generalisation operators account for proximity and density of point symbols and therefore they remove points only where conflicts arise. Once the average distance between data points reaches the size of the quad

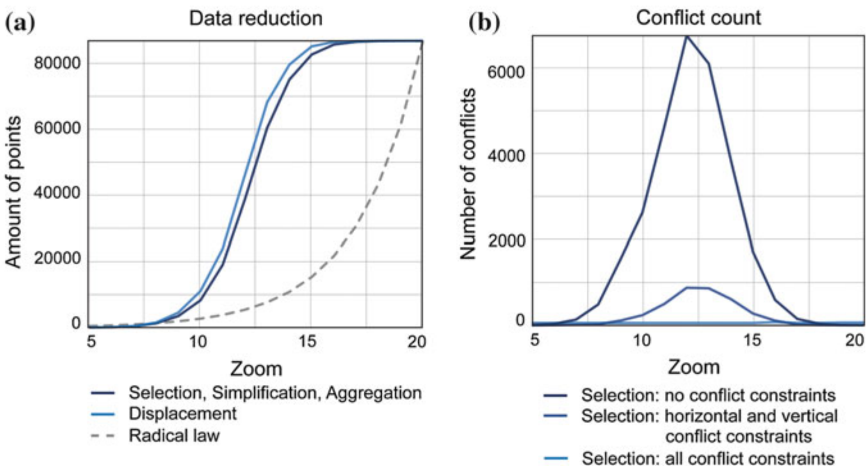

Fig. 6.8 **a** Global data reduction per zoom level, for point reduction algorithms and displacement. **b** Global conflict count per zoom level with different conflict constraints applied

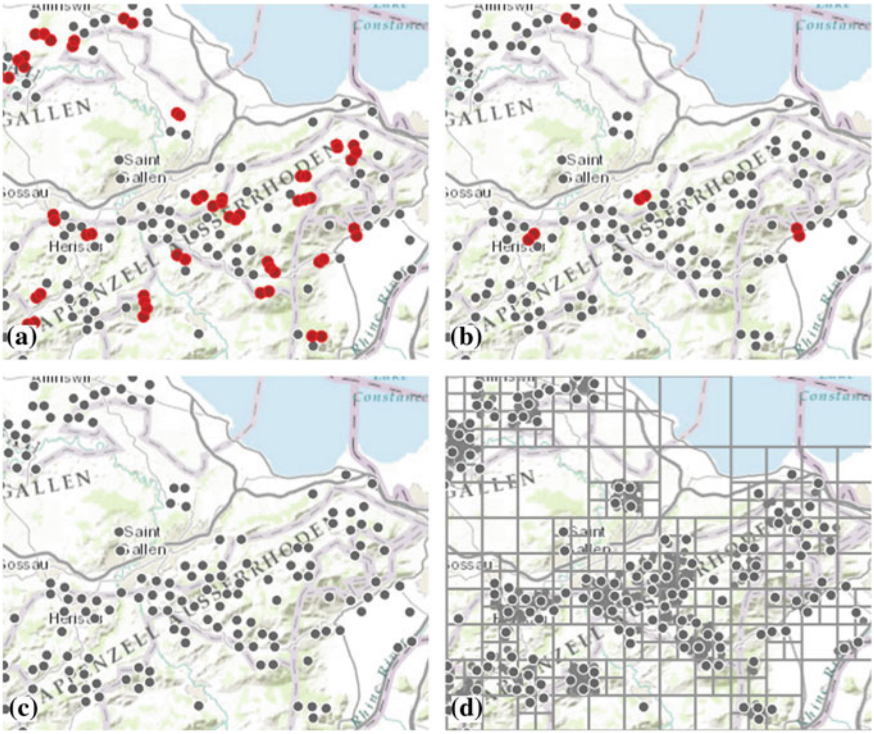

Fig. 6.9 **a**–**c** Cartographic conflicts (*red dots*) for quadtree-based selection with: **a** no conflict constraints, **b** *horizontal* and *vertical* conflict constraints, **c** including diagonal conflicts constraints, and **d** debug view with the underlying quadtree data structure

cells corresponding to the zoom level, the point reduction rate rapidly increases. Not surprisingly, a comparison between point reduction operators (selection, simplification and aggregation) and displacement illustrates that for those zoom levels where most conflicts arise, a process combining selection and displacement retains more points than with an elimination algorithm only.

Cartographic conflict—overlapping symbols—are not entirely removed by solely retaining one point per quadnode at the target LOD, as they do not consider *per se* potential overlaps from generalisation results residing in neighbouring quadnodes (Fig. 6.8b dark blue curve). Two points may lay across the border between two neighbouring quadnodes, separated by a distance less than the symbol size. This can be alleviated by checking for collisions in the adjacent quadnode neighbours by performing a further elimination or by constraining all points inside the quadnode not allowing for any overlap. Figure 6.8b shows the evolution and reduction of cartographic conflicts over the different zoom levels for the selection algorithm, with the different variations of collision checks applied.

Figure 6.9 shows conflict reduction applying the different variations of conflict checks in the case of value-based selection on lichens data for the Eastern part of

Switzerland. While pure selection still shows some cartographic conflicts (red dots in Fig. 6.9a), the application of conflict constraints in the algorithm reduces them significantly (Fig. 6.9b, c).

6.7.3 Data Enhancement

A variant of value-based selection (Fig. 6.10) illustrates how a particular attribute is retained and enhanced through a set of scales. It can be applied if comparatively rare point features need to be retained and are not evened out over the course of scales. Figure 6.10 highlights how the most endangered species are retained throughout the different zoom levels, rather than maintaining the overall distribution of categories.

6.7.4 Displacement Measures

To further resolve spatial conflicts and retain more elements than solely with point reduction operators, a displacement algorithm can be applied. The displacement algorithm tries to accommodate as many points as possible keeping at most one point per quadnode, and displacing points if the directly neighbouring quadnodes provide sufficient holding capacity for displacement. Remaining overlaps can be removed by further resolving boundary constraints as illustrated in Fig. 6.9. A comparison between the two algorithms (Fig. 6.11a, b) illustrates that displacement retains more points for the displayed zoom level. On the other hand, it shows that displacement has the effect of homogenising dense clusters and thus affecting the overall distribution pattern.

The characteristics of an applied displacement operator can be highlighted by plotting cumulated displacement vectors (in pixels) for each angle of displacement. Figure 6.11c shows nicely that the algorithm (considering in this example only horizontal and vertical neighbours) did not displace points to diagonal neighbours with the majority of displacement angles in horizontal and vertical direction.

6.7.5 Preservation of Spatial Patterns

How well a generalisation algorithm preserves the underlying spatial pattern can be investigated by visually comparing the kernel density estimation (KDE) of a point pattern or by calculating the difference between two kernel density estimations. The kernel density map in Fig. 6.12a for zoom level 9 and the KDE density difference map in Fig. 6.12b between zoom level 9 and 10, show the density distribution of the point pattern and where it changes most, respectively. It shows

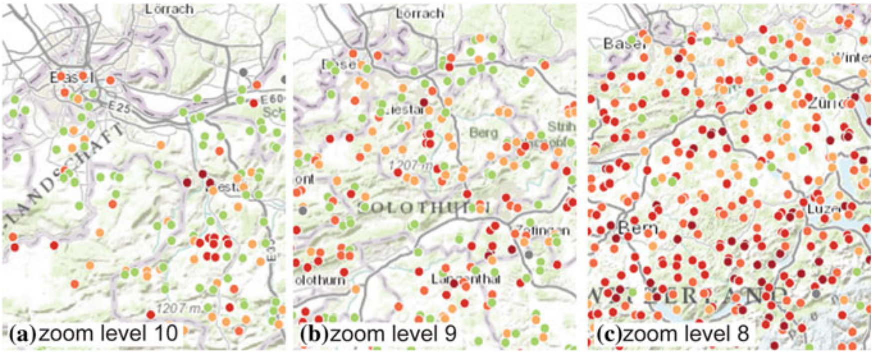

Fig. 6.10 Value-based selection, retaining most endangered lichens in the region of Basel with decreasing zoom level. Color codes range from least endangered (*green*) to most endangered (*red*)

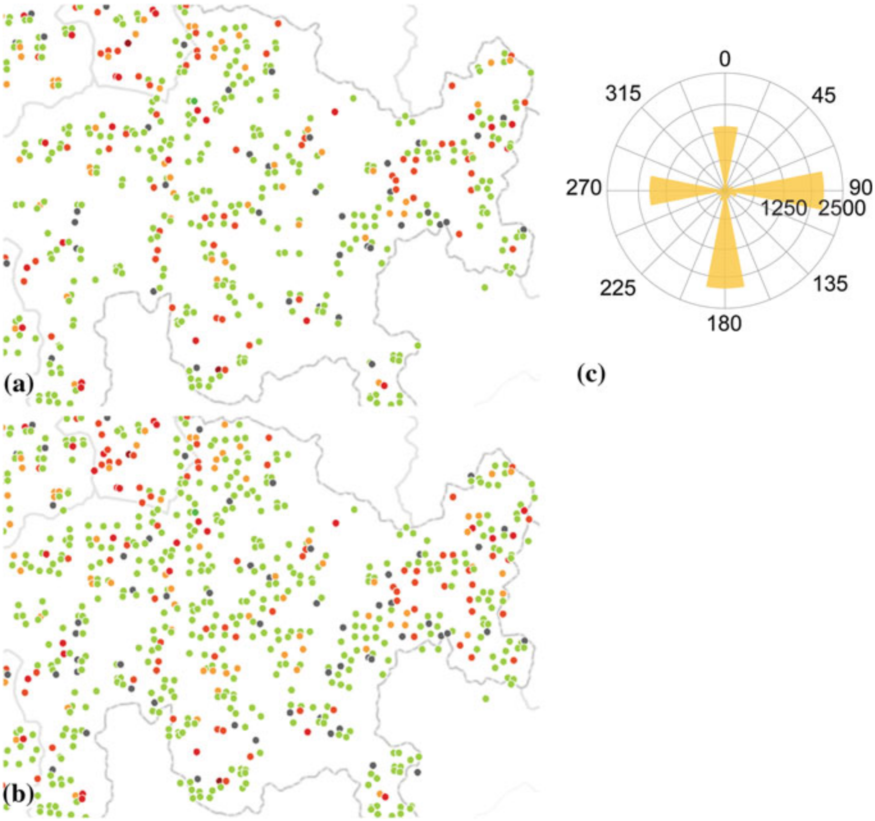

Fig. 6.11 **a** Centrality-based simplification (724 point), **b** displacement applied after centrality-based simplification (974 points) **c** cumulated displacement vectors for zoom level 9

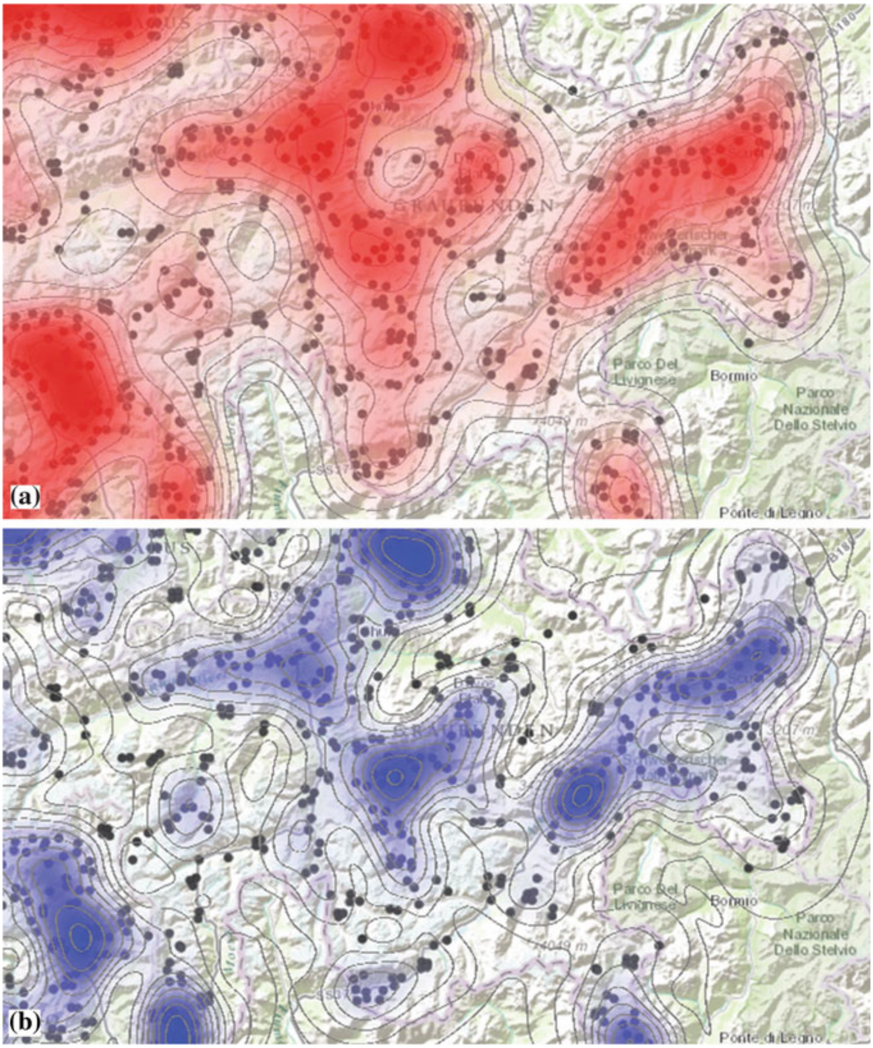

Fig. 6.12 **a** Kernel density estimation for centrality-based simplification at zoom level 9, **b** Kernel density difference between zoom level 9 and 10

that the applied algorithm reduces the point density most at local peaks, where the highest densities are located, and how much the density is decreased around local density peaks. The decrease around density peaks is however less pronounced than with the displacement operator (Fig. 6.11a, b).

6.8 Conclusions

This chapter reviewed generalisation operators, defined as a generic descriptor for the type of spatial or attribute modification to be achieved on some set of geospatial data. Examples of operator typologies demonstrate the wide variety of methods to organise types of operators for paper mapping, digital mapping, and for multi-scale data modelling. The development of methods by which to formalise differing requirements for topographic and thematic mapping across a range of scales has caused a shift from application of individual operators in isolation, to generalisation strategies encompassing integrated sequences of operators.

Recent advances in the design of generalisation strategies increase processing requirements and in some cases mandate data enrichment, which supports generalisation that can be tailored to specific landscape and settlement contexts. Improved methods for reasoning about relative feature priorities permit advanced processing of road networks, stream networks, and very large point data sets, as documented in the three case studies. Additional research on networks should assess whether algorithms are suitable for more complex cyclic, anthropogenic networks (i.e., road and rail networks, or stream and canal networks) or whether refinements or a different set of algorithms are required. Much work in the realm of generalisation operators has been devoted to methods that preserve local density variations of feature distributions. Feature density variations are often related to changing anthropogenic or geophysical conditions, and the case studies all show how maintenance of spatial patterns for these variations enhances map displays with more realistic context.

In the future, one can expect to see continued work to formally characterise feature geometry, spatial context, and to identify and resolve spatial conflicts arising from generalisation for mapping at reduced scales. The extent to which multi-scale characterisation can be articulated will continue to expand the extent to which generalisation can be fully automated; and the foundations for such advances will be initiated in development and refinement of generalisation operators.

Acknowledgments Work for this chapter was funded by several federal agencies in several countries. The work of Dr. Brewer is funded by the U.S. Geological Survey's Center of Excellence for Geospatial Information Science (USGS-CEGIS) grant #06HQAG0131. The work of Dr. Buttenfield is supported by USGS-CEGIS grant #04121HS029. The research reported in the third Case study represents part of the PhD project of Pia Bereuter, funded by the Swiss National Science Foundation through project GenW2 + (SNF no. 200020-138109). Particular thanks go to Silvia Stofer of the National Data and Information Center for Swiss lichens and all the contributors for collecting and providing the data of the SwissLichens dataset.

References

Ai T, Liu Y, Chen J (2006) The hierarchical watershed partitioning and data simplification of river network. Prog Spat Data Handling Part 11:617–632. doi:10.1007/3-540-35589-8_39

Aslan S, Bildirici IO, Simav O, Cetinkaya B (2012) An incremental displacement approach applied to building objects in topographic mapping. In: 13th ICA workshop on generalisation and multiple representation, Istanbul

Balboa JLG, Lopez FJA (2008) Generalization-oriented road line classification by means of an artificial neural network. Geoinformatica 12:289–312

Beard K (1991) Constraints on rule formation. In: Buttenfield BP, McMaster RB (eds) Map generalisation: making rules for knowledge representation. Longman Group, Harlow, pp 121–135

Benz SA, Weibel R (2013) Road network selection using an extended Stroke-Mesh combination algorithm. In: 16th ICA workshop on generalisation and multiple representation, Dresden, Germany, 23–24 Aug 2013

Bereuter P, Weibel R, Burghardt D (2012) Content zooming and exploration for mobile maps. In: Proceedings of the 15th AGILE international conference on geographic information science, pp 74–80

Bereuter P, Weibel R (2013) Real-time generalisation of point data in mobile and web mapping using quadtrees. Cartogr Geogr Inf Sci 40(4):1–11

Bildirici IO, Aslan S, Simav O, Cobankaya ON (2011) A generic approach to building typification. In: 14th ICA workshop on generalisation and multiple representation, Paris, 30 June–1 July

Bobzien M, Burghardt D, Petzold I, Neun M, Weibel R (2008) Multi-representation databases with explicitly modeled horizontal, vertical, and update relations. Cartography Geogr Inf Sci 35(1):3–16

Brewer CA, Stanislawski LV, Buttenfield BP, Raposo P, Sparks KA, Howard MA (2012) Multiscale design for the national map of the United States: road thinning for topographic mapping. In: AutoCarto 2012 Proceedings, Columbus, Ohio, 16–18 Sept 2012

Brewer CA, Stanislawski LV, Buttenfield BP, Sparks KA, McGilloway J, Howard MA (2013a) Automated thinning of road networks and road labels for multiscale design of the national map of the United States. Cartogr Geogr Inf Sci

Brewer CA, Guidero EM, Stanislawski LV, Buttenfield BP, Raposo P (2013b) Labeling through scale using hierarchies of thinned road networks for design of the national map of the United States. In: Proceedings 26th international cartographic conference, Dresden, Germany, Aug 25–30

Briat M-O, Monnot J-L, Punt EM (2011) Scalability of contextual generalisation processing using partitioning and parallelization. In: 14th ICA/ISPRS workshop on generalisation and multiple representation, Paris, 30 June–1 July

Buttenfield BP, Mark DM (1991) Expert systems in cartographic design Chapter 7. In: Taylor DRF (ed) Geographic information systems. New York: Pergamon: pp. 129–150

Buttenfield BP, Stanislawski LV, Brewer CA (2011) Adapting generalisation tools to physiographic diversity for the United States national hydrography dataset. Cartography Geogr Inf Sci 38(3):289–301

Chaudhry O, Mackaness WA (2005) Rural and urban road network generalisation deriving 1:250,000 from OS mastermap. In: Proceedings of the 22nd international cartographic conference, A Coruña, Spain, 11–16 July

Chaudhry O, Mackaness WA (2008a) Partitioning to make manageable the generalisation of national spatial datasets. In: 11th ICA workshop on generalisation and multiple representation, Montpellier, France, 20–21 June 2008

Chaudhry O, Mackaness WA (2008b) Automatic identification of urban settlement boundaries for multiple representation databases. Comput Environ Urban Syst 32(2):95–109
Chen J, Hu Y, Li Z, Zhao R, Meng L (2009) Selective omission of road features based on mesh density for automatic map generalisation. Int J Geogr Inf Sci 23(8):1013–1032
Christophe S, Ruas A (2002) Detecting building alignments for generalisation purposes, advances in spatial data handling. Springer, Berlin, pp 419–432
Esri (2011) ArcGIS desktop 10 help documentation: thin road network (cartography). Environ Syst Res Inst, 11 Sept 2011 http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#//007000000014000000. Accessed 6 Feb 2013
Esri (2012) ArcGIS 10.1 help: cartographic partitions (environment setting). Environ Syst Res Ins, http://resources.arcgis.com/en/help/main/10.1/index.html#/Welcome_to_the_ArcGIS_Help_Library/00qn0000001p000000/. Accessed 18 Feb 2013
Foerster TJ, Stoter J, Kobben B (2007) Towards a formal classification of generalisation operators. In: Proceedings of the 23rd international cartographic conference. Moscow, Russia
Fotheringham AS, Brunsdon C, Charlton M (2000) Quantitative geography: perspectives on spatial data analysis. Sage, London
Gülgen F, Gökgöz T (2008) Selection of roads for cartographic generalization. the international archives of the photogrammetry. Remote Sens Spat Inf Sci 37(B4):615–620
Gülgen F, Gökgöz T (2011) A block-based selection method for road network generalisation. Int J Digital Earth 4(2):133–153
Gutman N (2012) Continuous and adaptive cartographic generalization of river networks. Doctoral dissertation. School of Computer Science, University of Oklahoma
Heinzle F, Anders KH, Sester M (2005) Graph based approaches for recognition of patterns and implicit information in road networks. In: Proceedings of the 22nd international cartographic conference, A Coruna, Spain, 9–16 July
Heinzle F, Anders K-H, Sester M (2007) Automatic detection of patterns in road networks methods and evaluation. In: Proceedings of joint workshop visualization and exploration of geospatial data 26(4)
Horton RE (1945) Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology. Geol Soc Am Bull 56(3):275–370
Jenny B, Jenny H, Hurni L (2011) Terrain generalisation with multi-scale pyramids constrained by curvature. Cartography Geogr Inf Sci 38(1):110–116
Jiang B, Claramunt C (2004) A structrual approach to the model generalisation of an urban street network. GeoInformatica 8(2):157–171
Jiang B, Liu X, Jia T (2013) Scaling of geographic space as a universal rule for map generalization. Ann Assoc Am Geogr 103(4):844–855
Kilpeläinen T (1997) Multiple representation and generalisation of geo-databases for topographic maps, Doctoral dissertation. Publications of the Finnish Geodetic Institute, Kirkonummi, (124)
Leonowicz AM, Jenny B, Hurni L (2010) Terrain sculptor: generalizing terrain models for relief shading. Cartographic Perspect 67:51–60
Li Z, Zhou Q (2012) Integration of linear and areal hierarchies for continuous multi-scale representation of road networks. Int J Geogr Info Sci 26(5):855–880
Long S, Shoberg T, Corns S, Carlo H, Ramachandran V (2012) Integrating geospatial data from the national map with supply chain networks to improve resiliency following large-scale urban disasters. American association of geographers national conference, New York
Mackaness WA, Ruas A, Sarjakoski LT (2007) Generalisation of geographic information: cartographic modelling and applications, Elsevier, Amsterdam pp 370

McMaster RB (1983) A quantitative analysis of mathematical measures in linear simplification. Unpublished Ph.D. dissertation, Department of Geography-Meteorology, University of Kansas

McMaster RB (1989) Numerical generalisation in cartography. In: McMaster, RB (ed) Monograph 40, numerical generalisaton in cartography. Cartographica 26(1):1–6

McMaster RB, Monmonier M (1989) A conceptual framework for quantitative and qualitative raster-mode generalisation. In: Proceedings of GIS/LIS'89, Orlando. American society for photogrammetry and remote sensing, Maryland, pp 390–403

McMaster RB and Shea KS (1992) Generalisation in digital cartography. In: Association of american geographers resource publications in college geography, Washington, DC, pp 134

Morrison JL (1974) A theoretical framework for cartographic generalisation with emphasis on the process of symbolization. Int Yearb Cartography 14:115–127

Neun M, Burghardt D, Weibel R (2008) Web service approaches for providing enriched data structures to generalisation. Int J Geogr Inf Sci 22(2):133–165

Nickerson BG, Freeman HR (1986) Development of a rule-based system for automatic map generalisation. In: Proceedings of the 2nd international symposium on spatial data handling, Seattle, Washington. International Geographical Union Commission on Geographical Data Sensing and Processing, Williamsville, NY, pp 537–556

Punt EM, Watkins D (2010) User-directed generalisation of roads and buildings for multi-scale cartography. In: 13th ICA workshop on generalisation and multiple representation, Zurich, 12–13 Sept 2010

Raposo P, Brewer CA, Stanislawski LV (2013) Label and attribute-based topographic point thinning. In: 16th ICA workshop on generalisation and multiple representation, Dresden, Germany, 23–24 Aug 2013

Ratajski L (1967) Phenomenes des points de generalisation. Int Yearb Cartography 7:143–151

Regnauld N, McMaster RB (2006) Creating a hydrographic network from its cartographic representation: a case study using Ordnance Survey Mastermap data. Int J Geogr Inf Sci 20(6):611–631

Regnauld N, McMaster RB (2007) A synoptic view of generalisation operators. In: Mackaness WA, Ruas A, Sarjakoski LT (eds) Generalisation of geographic information: cartographic modelling and applications, Elsevier Ltd, Amsterdam pp 37–66

Robinson AH, Sale RD (1969) Elements of cartography, 3rd edn. Wiley, New York

Roth RE, Brewer CA, Stryker MS (2011) A typology of operators for maintaining legible map designs at multiple scales. Cartographic Perspect 68:29–64

Ruas A (1995) Multiple paradigms for automated map generalisation: geometry, topology, hierarchical partitioning and local triangulation. In: Proceedings AutoCarto, Charlotte, USA, 12, pp 69–78

Ruas A (1999) Modele de generalisation de donnees geographiques a base de constraints et d'autonomie, These de doctorat. L'universite de Marne La Vallee, Paris

Ruas A, Duchêne C (2007) A prototype generalisation system based on the multi-agent system paradigm. In: Mackaness WA, Ruas A, Sarjakoski LT (eds) Generalisation of geographic information: cartographic modelling and applications, Elsevier Ltd, pp 269–284

Sarjakoski LT, (2007) Conceptual Models of Generalisation and Multiple Representation, In Mackaness WA, Ruas A, Sarjakoski LT (eds), Generalisation of Geographic Information: Cartographic Modelling and Applications, Elsevier Ltd, pp 11–35

Savino S, Rumor M, Zanon M, Lissandron I (2010) Data enrichment for road generalization through analysis of morphology in the CARGEN project. In: 13th ICA workshop on generalisation and multiple representation, Zurich, Switzerland, 12–13 Sept 2010

Savino S, Rumor M, Canton F, Langiu G, Reineri M (2011a) Model generalisation of the hydrography network in the CARGEN project. In: Ruas A (eds) Advances in cartography and GIScience (1):439–457, Lecture Notes in Geoinformation and Cartography, Springer, Berlin

Savino S, Rumor M, Zanon M (2011b) Pattern recognition and typification of ditches. In: Ruas A (eds), Advances in cartography and GIscience, Vol 1. Springer, pp 425–437 ISBN 9783642191428

Schylberg L (1993) Computational methods for generalisation of cartographic data in a raster environment. Doctoral Thesis, Department of Geodesy and Photogrammetry, Royal Institue of Technology Stockholm, Sweden, Photogrammetric Reports (60) TRITA-FMI Report, (1993:7)

Simley J, Doumbouya A (2012) National hydrography dataset—Linear Referencing. USGS Fact Sheet: 2012–3068

Stanislawski LV (2009) Feature pruning by upstream drainage area to support automated generalisation of the united states national hydrography dataset. Comput Environ Urban Syst 33:325–333

Stanislawski LV, Buttenfield BP (2011) A raster alternative for partitioning line densities to support automated cartographic generalisation. In: 25th International cartography conference, July 3–8, Paris

Stanislawski LV, Briat M, Punt E, Howard M, Brewer CA, Buttenfield BP (2012a) Density-stratified thinning of road networks to support automated generalisation for the na-tional map. In: 15th ICA workshop on generalisation and multiple representation, Istanbul, Turkey, 13–14 Sept 2012

Stanislawski LV, Doumbouya AT, Miller-Corbett CD, Buttenfield BP, Arundel-Murin ST (2012b) Scaling stream densities for hydrologic generalisations. Seventh international conference on GIScience, Columbus, Ohio, 18–21 Sept 2012

Stanislawski LV, Savino S (2011) Pruning of hydrographic networks: a comparison of two approaches. In: 14th ICA/ISPRS workshop on generalisation and multiple representation, Paris, France, 30 June–1 July 2011

Steiniger S, Burghardt D, Weibel R (2006) Recognition of island structures for map generalisation. In: International symposium on advances in geographic information systems, ACM-GIS 2006 pp 67–74

Steiniger S, Lange T, Burghardt D, Weibel R (2008) An approachfor the classification of urban building structures based on discriminant analysis techniques. Trans GIS 12(1):31–59

Steiniger S, Wiebel R (2007) Relations among map objects in cartographic generalisation. Cartography Geog Inf Sci 34(3):175–197

Stoter J, Baella B, Blok C, Burghardt D, Duchêne C, Pla M, Regnauld N, Touya G (2010) State-of-the-art of automated generalisation in commercial software. European spatial data research report, p 270 (2010)

Stofer S, Scheidegger C, Clerc P, Dietrich M, Frei M, Groner U, Jakob P, Keller C, Roth I, Vust M, Zimmermann E (2012) Nationales Daten- und Informationszentrum der Schweizer Flechten—SwissLichens. Datenbankauszug vom 25. September 2012 (SST_20120925). Eidgenössische Forschungsanstalt WSL, Birmensdorf

Strahler AN (1952) Hypsometric (area-altitude) analysis of erosional topology. Geol Soc Am Bull 63(11):1117–1142

Sugarbaker LJ, Carswell WJ (2011) The national map: U.S. geological survey fact sheet 2011–3042. pp 4 http://pubs.usgs.gov/fs/2011/3042. Accessed 29 Jan 2013

Thomson RC, Brooks R (2002) Exploiting perceptual grouping for map analysis, understanding and generalization: the case of road and river networks. Graphics recognition algorithms and applications, Lecture notes in computer science 2390: 148–157

Thomson RC, Brooks R (2007) Generalisation of geographical networks. In: Mackaness WA, Ruas A, Sarjakoski LT (eds) Generalisation of geographic information: cartographic modelling and applications, Elsevier Ltd, Boston pp 255–267

Thomson RC, Richardson DE (1995) A graph theory approach to road network generalisation. In: Proceedings 17th international cartographic conference, Barcelona, Spain, 3–9 Sept 1995

Thompson RC, Richardson DE (1999) The 'Good Continuity' principle of perceptual organisation applied to the generalisation of road networks. In: Proceeding of the 19th international cartographic conference, Ottawa pp 1215–1225

Töpfer F, Pillewizer W (1966) The principles of selection. Cartographic J 3(1):10–16

Touya G (2007) River network selection based on structure and pattern recognition. In: Proceedings 23rd International cartographic conference, Moscow, Russia

Touya G (2010) A road network selection process based on data enrichment and structure detection. Trans GIS 14(5):595–614

Touya G, Duchêne C, Ruas A (2010) Collaborative generalization: formalization of generalization knowledge to orchestrate different cartographic generalization processes. GIScience 2010, Zurich, Switzerland, 14–17, Sept 2010

USGS (1985) Standards for 1:100,000-scale quadrangle maps. National mapping program technical instructions. U.S. Geological survey, Department of the Interior

USGS (2006) The best practices data model—The national map, March 1 2006, http://services.nationalmap.gov/bestpractices/model/acrodocs/Poster_BPTrans_03_01_2006.pdf. Accessed 5 Feb 2013

van Oosterom P (2005) Variable-scale topological data structures suitable for progressive data transfer: the GAP-face tree and GAP-edge forest. Cartography Geogr Inf Sci 32(4):331–346

van Oosterom P, Meijers M (2011) Towards a true vario-scale structure supporting smooth-zoom. 14th ICA/ISPRS workshop on generalisation and multiple representation, Paris, 30 June–1 July

Verdin KL, Verdin JP (1999) A topological system for delineation and codification of the earth's river basins. J Hydrol 218:1–12

Wallace RM, Tarboton DG,Watson DW, Schreuders KAT, Tesfa TK (2010) Parallel algorithms for processing hydrologic properties for digital terrain. GIScience, Zurich, 14–17, Sept 2010

Weibel R (1991) Amplified intelligence and rule-based systems. In: Buttenfield BP, McMaster RB (eds) Map generalisation: making rules for knowledge generation. Longman, London

Weibel R (1992) Models and experiments for adaptive computer assisted terrain generalization. Cartography Geogr Inf Syst 19(3):133–153

Weibel R, Burghardt D (2008) Generalisation on-the-fly. In: Shekhar S, Xiong H (eds) Encyclopedia of GIS. Springer, New York, pp 339–344

Weiss R, Weibel R (2013) Road network selection for small-scale maps using an improved centrality approach. In: 16th ICA workshop on generalisation and multiple representation, Dresden, Germany, 23–24 Aug 2013

Werder S, Kieler B, Sester M (2010) Semi-automatic interpretation of buildings and settlement areas in user-generated spatial data. In: Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, pp 330–339. doi: 10.1145/1869790.1869836

Wu H, Kimball JS, Li H, Huang M, Leung LR, Adler RF (2012) A new global river network database for macroscale hydrologic modeling. Water resources research, vol. 48, W09701, pp 5 doi:10.1029/2012WR012313

Yager DB, Hofstra AH, Granitto M (2012) Analyzing legacy U.S. geological survey geochemical databases using GIS—Applications for a national mineral resource assessment: U.S. Geological survey techniques and methods 11–C5, pp 28

Zhang X, Ai T, Stoter J (2010) Characterization and detection of building patterns in cartographic data: two algorithms. The international archives of the photogrammetry, remote sensing and spatial information sciences 38(II):261–266

Zhou Q, Li Z (2012) A comparative study of various strategies to concatenate road segments into strokes for map generalization. Int J Geogr Inf Sci 26(4):691–715