Chapter 9
Evaluation in Generalisation

Jantien Stoter, Xiang Zhang, Hanna Stigmar and Lars Harrie

Abstract This chapter presents the context, the issues and the research associated with the evaluation of map generalisation output as well as of map readability. Two main approaches of evaluation are described, i.e. visual and quantitative evaluation. Visual evaluation is subjective, qualitative, and time-consuming, while it is argued that quantitative evaluation is only appropriate for assessing specific aspects. Since automated evaluation is becoming very important in the field of automated generalisation, this chapter further explores the topic of automated evaluation. The previous frameworks for automated generalisation are reviewed and the three main components of automated evaluation are explained. Related to automated evaluation of generalisation output are formulas to automatically evaluate map readability. These are also discussed. This chapter ends with three case studies. The first Case study identifies and evaluates generalised building patterns. It demonstrates the three-step approach of data enrichment, data matching and constraint evaluation. The second Case study deals with formulas to automatically evaluate map readability and the third Case study carries out a comprehensive evaluation demonstrating the main aspects described in this chapter. Both visual and quantitative evaluation are applied of which the last one includes the three main components of automated evaluation. The chapter closes with conclusions and highlights research issues in evaluation.

J. Stoter (✉)
Delft University of Technology and Kadaster, Delft, The Netherlands
e-mail: j.e.stoter@tudelft.nl

X. Zhang
Wuhan University, Wuhan, China
e-mail: xiang.zhang@whu.edu.cn

H. Stigmar · L. Harrie
GIS Centre, Lund University, Lund, Sweden
e-mail: hanna.stigmar@gis.lu.se

L. Harrie
e-mail: lars.harrie@nateko.lu.se

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_9,

9.1 Introduction

Evaluation in map generalisation is the process of examining and checking whether the desired characteristics of the resulting data are satisfactory for a given task. Quality evaluation in map generalisation differs from measuring the absolute loss in data quality, which is not *per se* a problem in map generalisation. It is unavoidable that map generalisation will change spatial data in terms of data quality (Müller 1991). However, these changes may be very helpful given a specific map purpose and scale (a fitness for use concept). Apart from measuring the quality of generalisation output, evaluation also provides users with necessary information to understand the ability and limitations of a system or data in use.

It is acknowledged that evaluation is one indispensable component of a holistic generalisation process in both manual and (semi-)automated generalisation. This is because assessment of generalisation results can help to tune parameters and to select the best algorithms or sequence of algorithms for the generalisation process (Mackaness 1991; Weibel 1991, 1995; Ruas and Plazanet 1996; João 1998; Weibel and Dutton 1998; Galanda 2003).

Evaluation of a generalised map includes two parts. The first part considers the representation of features and the second part concerns the readability of the map. Therefore, there is an overlap between the generalisation evaluation methods and the method to define map readability analytically, e.g. by using map readability formulas (Rosenholtz et al. 2005; Stigmar 2010).

This chapter presents the context, the issues and the research that has been carried out in the domain of evaluation in map generalisation and map readability and extends the work of Zhang (2012). First the purposes of evaluation in generalisation are explained (Sect. 9.2). Then the two main approaches of evaluation are described in Sect. 9.3. The quantitative approach gets further attention in Sects. 9.4 and 9.5 describing frameworks for automated evaluation. Related to the automated evaluation are map readability formulas, which are discussed in Sect. 9.6. Sections 9.7–9.9 present three case studies on the evaluation of generalised maps. The chapter closes with conclusions and further research.

9.2 The Purposes of Evaluation

Evaluation of generalised maps is beneficial for different user groups. The evaluation helps cartographers in detecting errors and it thus facilitates the interaction between human and computer. For system designers, evaluation returns useful feedback, which improves the development of automated generalisation. For data users, evaluation results provide knowledge on the content of the data in use.

Earlier research on quality evaluation in map generalisation (Weibel 1995), suggested that evaluation can be made at three different stages in the generalisation process: *a priori*, *posteriori* and *ad-hoc* evaluation. Likewise, Mackaness and

Ruas (2007) summarised three scenarios where the evaluation may be applied: evaluation before, during and after generalisation. These three scenarios are motivated by different purposes:

Evaluation for tuning: acts as a support for setting parameters and tuning generalisation systems prior to processing.

Evaluation for controlling: acts as an intermediate step to the generalisation process. It can be used for (1) identifying conflicts and errors and (2) improving the process by triggering algorithm(s) or modifying the sequence of operations.

Evaluation for assessing: evaluation of generalisation output after the process is complete.

This chapter specifically focuses on the evaluation carried out during and after the generalisation process. These two evaluations are detailed in the subsections below.

9.2.1 Evaluation for Controlling the Generalisation Process

Earlier theoretical analysis on evaluation for controlling a holistic generalisation process was made by McMaster and Shea (1992), who suggested using 'cartometric evaluation' to answer the question of when to generalise. They further suggested to use a series of 'spatial and holistic measures' (e.g. measurement of density, distribution, shape) to detect undesired situations (conflicts) and to select operations to resolve those identified conflicts. Later, Ruas and Plazanet (1996) and Weibel and Dutton (1998) proposed similar generalisation models which treat automated evaluation as an integral part of the generalisation process, based on the idea of cartographic constraints (i.e. map specifications, see also Sect. 9.4). In their proposed models, the evaluation is used for structure recognition, conflict detection and quality assessment (AGENT 1998). Brazile (2000) designed an architecture for automated generalisation incorporating quality assessment for controlling the whole process. At the moment, an implementation for those proposed conceptual models is missing, due to the complexity of the process and technological limitations.

With the development of constraint-based approaches, more and more advanced techniques are being adapted from other domains in order to control generalisation. Multi-Agent System (MAS) for example is a technique to implement constraint-based generalisation, in which evaluation can be performed before and after each step of the generalisation in order to reach the optimal solution among a set of constraints. The research of Galanda and Weibel (2002) and Galanda (2003) focused on the generalisation of categorical (land use) data using MAS techniques. The evaluation mainly focused on the calculation of some basic metric and topological constraints, such as maintaining minimum distance and avoiding self-intersection. Based on measured values and their discrepancy to the goal values, the violation (or satisfaction) of a constraint can be calculated and a list of generalisation plans can be automatically proposed. On the other hand,

optimisation techniques were also widely used to implement constraint-based generalisation, (for example Harrie 1999, 2001; Harrie and Sarjakoski 2002; Sester 2005).

In summary, quality evaluation is an indispensable component to a successful automated generalisation system. Evaluation as part of a generalisation system has the advantage of being able to direct the system to its final goal (optimal solution). But it has two disadvantages. First, evaluation criteria (i.e. constraints) may well be the same as the constraints used to steer the generalisation process. In such a case, generalised data usually satisfy the evaluation criteria. Consequently the evaluation as such may give over-optimistic results (Harrie and Weibel 2007), because the data may violate other important constraints that are not defined. Another disadvantage is that the comparison between different generalisation outputs (i.e. evaluation for grading) performed on different systems is not possible for systems with self-evaluation capabilities (Ruas 2001).

9.2.2 Evaluation for Assessing Generalisation Output

Evaluation for controlling the generalisation process is not a good option for assessing the overall quality or for making comparison between different solutions. To achieve the goal of assessing the final output of generalisation, an independent evaluation is necessary. According to Bard and Ruas (2004), evaluation for assessing can be further divided into three types:

Evaluation for editing: focused on detecting cartographic errors and inconsistencies. It is performed at the final stage of the generalisation process and is used to aid subsequent manual or interactive editing, where necessary.

Evaluation for grading: attempts to find an aggregated value that reflects the overall quality of the generalised data and compares different generalisation solutions in order to determine the optimal solution or to identify poor generalisation solutions given specific generalisation tasks.

Descriptive evaluation: provides general information of the modifications performed on the generalised data. Such information can be used to improve the quality description (e.g. metadata) of the final products (e.g. what has been removed, or emphasised? How much has the data changed?)

Little research relates directly to the evaluation of generalised data at the end of the process. The first contribution to this topic was that of Ehrliholzer (1995). The important ideas emerging from this study were: (1) the use of three levels of abstraction (i.e. element, group and map level) for assessment, (2) the need to standardise the scale of various criteria (e.g. a value from, say, 1 to 5) for comparison purposes, and (3) the combination of criteria to provide a summary of evaluation results. The study remains theoretical and only a few generalisation algorithms were evaluated. Other experiments such as the ones developed as part of the AGENT project (Ruas 2001; AGENT 1999) helped the community to develop various evaluation techniques.

Up to the present time, the most comprehensive study on the automated evaluation of generalised data was undertaken by Bard (2004a, b). One of the objectives of his study was to develop a synoptic evaluation framework in which the relative quality of generalisation outputs can be graded (using evaluation grading categories described above). The author identified a set of reference functions for automated evaluation and the interaction between object types, object characters, reference functions and threshold values of these functions according to a scale transition. The author evaluated map requirements on buildings to show how these functions work in the evaluation process. A more recent project exploring evaluation of generalisation output was the EuroSDR project (Stoter et al. 2009b; Stoter 2010), —described in Sect. 9.9.

Also the work of Touya (2012) is relevant. He examined the use of social welfare theory to aggregate single satisfaction values to global legibility assessment at the map level.

9.3 Visual and Quantitative Evaluation on Map Generalisation

Two approaches are commonly used as a basis for evaluation: visual evaluation and quantitative evaluation.

9.3.1 Visual Evaluation

Traditionally, cartographers undertake a visual evaluation of the quality of generalisation output. The judgement on the quality depends largely on knowledge and experiences of the specific cartographer. The visual evaluation approach persists in (semi-) automated generalisation. That is, the evaluation of maps from automated solutions still resorts to visual comparison with 'optimal' solutions (often a paper map), in order to validate the automatic solutions or algorithms and to identify cartographic conflicts.

Usually, visual evaluation is carried out by asking cartographic experts to grade the quality of solutions based on certain criteria (Weibel 1995; Mackaness and Ruas 2007). The expert judgement can be made on a global criterion such as 'maintenance of overall character of the original map', or the criteria can include a very detailed checklist. Grading can be used (such as an index of 1–5 representing the quality of generalisation from 'poor' to 'excellent'). The overall quality can then be determined based on the qualitative grading of different cartographers. This process is also called expert evaluation. The difficulty with expert evaluation is that the criteria may be too general and different experts may have different

opinions as to what makes a 'good generalisation'. Thus, a standard is needed for such an evaluation. One approach is to produce a comprehensive and common form of checklist and a questionnaire developed in close collaboration with different experts (Weibel 1995).

The visual evaluation is the simplest form of assessment to implement, but the most time consuming (Bard 2004a). Furthermore, visual assessment is subjective and comments obtained from experts might be very general or based on rather vague criteria (João 1998; Weibel and Dutton 1999; Brazile 2000; Ruas 2001; Kazemi et al. 2005; Mackaness and Ruas 2007).

9.3.2 Quantitative Evaluation

In addition to the visual assessment approach, many researchers have been looking for quantitative methods for evaluating the quality of generalisation. The idea is to identify one or more quality measurements to determine the fulfilment of the product with respect to map requirements. Some early studies concentrated on the quantification of content change, while others focused on the measuring of specific primitive geometries before and after generalisation.

The *radical law* is one of the earliest attempts to formalise the change of map content over scale transitions (Töpfer and Pillewizer 1966). The radical law formulates the number of map objects (symbols) that should be selected according to different scale transitions. As a result, the Law can be used as a criterion for controlling the number of items to be selected. However, the Law cannot address the question of where and which objects should be selected, and consequently it is not capable of controlling semantic and structural aspects of the generalisation process.

Important work was done to study extensions to the Radical Law, for example to focus computations on line length (Buttenfield et al. 2010, 2011). Brewer et al. (2013) 'inverted' the Radical Law to determine optimal target scales for a given number of feature labels. Both are examples of 'Generalisation for Controlling', discussed in Sect. 9.2. These studies have demonstrated how the Radical Law can be extended in order to work within automated environments.

There was once a trend in utilising information theory (entropy) to model the changing levels of map content that arise from the generalisation process (e.g., Neumann 1994; Bjørke 1996, 1997, 2003; Li and Huang 2002). On the other hand, using information theory to evaluate the map content could be problematic. Mackaness and Ruas (2007) for instance argued that it is extremely difficult to formalise the change of map content caused by generalisation in terms of information entropy, since information content varies depending on the map purpose, application of map symbols and map reader's interpretation skills.

Earlier attempts at the evaluation of primitive objects focused on linear geometries (such as roads, rivers, contours, railways). McMaster (1983, 1987)

proposed measures to quantify the positional change in the vertices of lines before and after generalisation, based on the concept of 'vector displacement'. The main focus of this assessment is on geometric precision. A later work by Buttenfield (1991) expanded the description of linear features to five measures (length, width band, segmentation, and error variance), which makes up the 'structural signature' of the lines. The aim of Buttenfield (1991) was to formulate the changes of the 'structural signature' over scale transitions. These two contributions have inspired other attempts for describing the characteristics of linear features, including geometric and structural aspects (Plazanet et al. 1998; Wang and Müller 1998; Ai et al. 2000; Van der Poorten and Jones 2002). In addition, some studies have focused on evaluating the change of different types of linear objects in terms of line length and 'vector displacement' e.g. Barber et al. (1995), João (1998). Finally, several studies suggested a methodology for evaluating the quality of generalised lines based on the description of the structure and shape of the lines (Skopelity and Tsoulos 2000; Skopelity and Lysandros 2001).

Work evaluating polygonal objects is a more recent topic of research. Generally two feature types are treated separately: buildings and land use. These two types are differentiated by their different properties in shape and topology. The building features are anthropogenic constructs with regular shape and are usually spatially disjoint; whereas land use features have irregular shape and are adjacent to each other. Their evaluation therefore requires different approaches. Regnauld (1998) focused on the generalisation of buildings. He argued that there is a need to evaluate whether the generalisation results correspond to the following three aspects: (1) characteristics of buildings (e.g. size, ratio between the number of buildings at the target and initial scale), (2) spatial distribution of buildings (e.g. alignments and patterns), and (3) density of buildings. Other measures for evaluation of polygonal generalisation have been proposed by Peter (2001). Dettori and Puppo (1996) proposed the extension of quality evaluation of polygonal objects using topological qualities such as connectivity, and inclusion.

However, evaluation of primitive geometries is not enough, since multiple objects are involved when assessing the quality of a generalised map (Weibel and Dutton 1998). Relationships between objects (either within a thematic class or between different classes) should be formalised and taken into account by the evaluation process.

Visual evaluation is subjective, qualitative, and time-consuming, while quantitative evaluation has tended only to examine certain characteristics and is therefore only appropriate for assessing specific requirements. It is difficult to adapt and integrate these measures into a holistic automated generalisation process. Comprehensive solutions for quantitative, automated evaluation are needed and this remains an area of further research (Burghardt et al. 2008).

This chapter seeks to provide an overview of research in this area, and what remains in need of further attention.

9.4 Frameworks of Automated Evaluation

As suggested by Weibel (1995), the process of automated evaluation involves two tasks or components:

- Formalisation of map specifications: map specifications must be formalised to describe what is a 'good', 'acceptable', or 'poor' generalisation for a given application and mapping scale.
- Evaluation process: applying methodologies to determine the degree to which the specifications are respected for a given generalisation output. This process requires both measure for automated evaluation as well as data matching techniques to enable the automated evaluation.

These two components were also explored in the work of Bard and Ruas (2004). Mackaness and Ruas (2007) further pointed out that a challenge of automated evaluation is to decompose the evaluation process into a set of values that can be calculated automatically and that remain meaningful to users. It is important to know which characteristics are pertinent to a specific map task and scale, and how to precisely define them (formalisation). It is also important to develop various techniques to qualify and quantify the actual changes of characteristics of data objects, which are later compared with the formalised map requirements (evaluation process). In the AGENT (1998) project, the evaluation process was further divided into three steps: (1) measurement, which is used for the computation of raw parameters (e.g. area of buildings), (2) characterisation, which aims to interpret those measured parameters (e.g. using Min, Mean, and Max values of computed areas to classify buildings), and (3) evaluation, which determines the characterisation by comparing the measured value and the goal value (e.g. if a building is too big or too small with respect to the specifications).

Bard and Ruas (2004) proposed a comprehensive framework of automated evaluation of generalised data which extended the work of AGENT (1998). The components of their framework and the interactions among the components are shown in Fig. 9.1.

In their evaluation model, **Specification** (map requirements) contains lists of characteristics, reference functions, and parameters. The establishment of map specifications is referred to as the formalisation process. In Bard and Ruas's work, the specifications mainly contained formal requirements of buildings (e.g. size, shape, and orientation). Their focus is on the framework development and the evaluation process. In their model, the evaluation process is modelled as two classes, namely characterisation and evaluation. **Characterisation** holds the measures (algorithms) to calculate the properties of geographic objects (both initial and target objects), such as size, shape and orientation of buildings. **Evaluation** first uses the Compare() method to calculate the distance between measured values of a characteristic and the target value defined in the specifications. It then re-invokes Evaluate() to interpret the distance into a qualitative evaluation result according to the defined tolerance. For example, if the distance between the

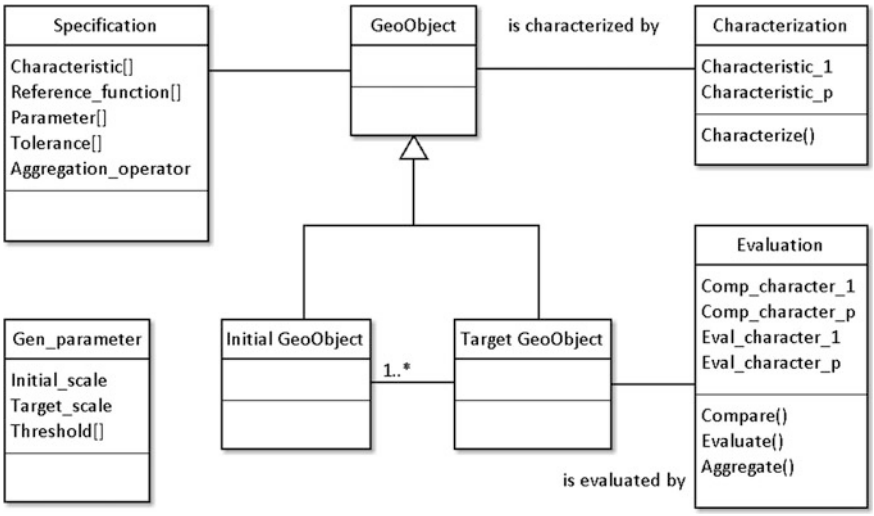

Fig. 9.1 A model to evaluate generalisation output by Bard and Ruas (2004)

measured size and target size of a building is 15 map mm^2, and the tolerance is 20 map mm^2, then the generalisation of this particular building is considered to be 'good' in terms of its size property.

It is noticeable that the evaluation process in this model involves measuring and comparing properties of objects at both the initial and target scales—that is to say, corresponding relations between target and initial objects are explicitly modelled. Burghardt et al. (2007) and Stoter et al. (2009a) also suggested that the evaluation of generalisation output should be based on the specification defined as a set of constraints, and partly on the initial dataset. As a consequence, data matching for automated evaluation of generalisation output is needed.

An improved framework for automated evaluation of generalised building patterns has been proposed in Zhang (2012). This framework is detailed further in the Case study in Sect. 9.7.

9.5 Components of Automated Evaluation

Starting from these frameworks, the following sections further explain the three main components of automated evaluation, namely (a) definition and formalisation of map requirements, (b) measures for automated evaluation and (c) data matching techniques. Together, these sections outline the major difficulties of the research on automated evaluation.

9.5.1 Defining and Formalising Map Requirements

The process of defining and formalising user requirements in map specifications has been described in Chap. 2. In an automated context, the challenge of this process is in specifying the map requirements in a computer understandable way (Burghardt et al. 2007; Stoter et al. 2009a). The challenge is in knowing which constraints are appropriate in order to assess the quality of generalised solution and how the constrained properties change at scale transitions. The former aims to select a set of constraints, whereas the latter aims to specify the goal values for the constraints (e.g. the minimum area of buildings after generalisation). Specifying goal values for constraints is however not always straightforward, since it depends (1) on the specifications of generalisation, (2) on the initial state of the property (Burghardt et al. 2007). Results from the EuroSDR project (see Sects. 2.5 and 9.9) show that constraints defined on two objects and on groups of objects are more difficult to interpret and evaluate by generalisation systems than those defined on individual objects. The formalisation of such constraints is something that requires further attention.

A closer look at the literature reveals that, low-level constraints such as those on one object (e.g. minimum dimensions) and between two objects (e.g. topology) are relatively easy to formalise. However, the constraint on general shape, which is also a low-level constraint, is harder to formalise. Shape is concerned with our ability to recognise the original form, silhouette or outline of objects and is different from positional accuracy (Müller et al. 1995). Shape needs to be preserved during generalisation as long as the details are still legible. In the constraint 'target shape should be similar to initial shape', shape is not a mathematically defined concept and the operation 'similar' is not well-defined either. Defining and precisely measuring shape is necessary for automated evaluation. The 'similarity' of two shapes can be formalised by specifying the scale-dependent modification to the measured values. A potential way of acquiring the scale-dependent knowledge is by studying map series at different scales (i.e. reverse engineering) as indicated in Harrie (2001). The results of the acquired change of a property can be stored in an 'evolution function' as proposed by Bard (2004b). This is a function that expresses how properties change over scale.

In addition to the shape constraint, the constraints dealing with higher-level concepts such as pattern, distribution, and network are also difficult to formalise. Most of these constraints define restrictions to the generalisation of groups of objects. These higher-level map requirements are important for defining the quality of generalisation output over large changes in scale, since the concepts describe important characteristics of the map at smaller scales and hence should be visually more discernable at these scales (Müller 1991; Ruas and Plazanet 1995; Mackaness and Edwards 2002; Mackaness and Ruas 2007).

At the conceptual level, Steiniger (2007) sought to formalise patterns using 'horizontal relationships', which describe the interactions between map objects

within the same map scale. His reason for formalising patterns was to be able to recognise patterns, which is the first step in automated evaluation. The recognition of patterns is viewed as a data enrichment process by some (Zhang 2012). A complete view of the formalisation of higher-level constraints for automated evaluation should also involve modelling the change of these concepts at scale transitions. This means that how the concepts modify themselves at scale transitions needs to be formalised. However, the knowledge of these constraints is rarely expressed in a clear and computer-readable way. Therefore, knowledge acquisition techniques are still needed to extract the knowledge for automated evaluation of generalisation output from various sources.

9.5.2 Measures for Automated Evaluation

As described in Sect. 9.4, automated evaluation is performed via three major tasks: defining and formalising map requirements; developing measures which automatically quantify the properties that need to be evaluated; and, evaluating by comparing between corresponding objects or with respect to a set of map specifications.

Measures play an important role in automated evaluation: they are used to detect the constraint violations and describe the severity of a violation. For example, a target building with an area slightly larger the size threshold is more acceptable than if it is smaller than the threshold. Using sensitivity ranges (tolerances), a measure can evaluate the degree to which a metric is violated.

This section further explains the concept of measures and the way they are used in automated evaluation.

A measure is defined in the AGENT (1999) project as a procedure to compute geometric properties or relationships, which is the basis for evaluating characteristics of map objects as well as for assessing the need for generalisation (as in the case of conflict detection) and the success of generalisation (that is, evaluation of generalisation output). Measures consist of four main aspects. Consider a distance measure for example, which has an operational aspect (a distance metric), a mathematical aspect (such as *Hausdorff* distance), an algorithm (implementation of *Hausdorff* distance) and a measured value. A measure can be used to quantify a property of an object with respect to a constraint. A measure may consist of a formula or involve a complex algorithm including computation of representations and auxiliary data structures. Representations here describe the geometry of objects derived by a transformation from 2D Euclidean space to another reference-system (e.g. Fourier transform) or by a transformation between dimensions (e.g., a collapse operation). Auxiliary data structures capture the relationships among a group of objects within higher-level concepts such as alignments and networks. Examples of these data structures are Delaunay triangulation and Minimum Spanning Tree (MST) (AGENT 1999).

Mathematical measures can be exact or approximate, according to the implementation algorithm. For example one can use an approximate *Hausdorff* algorithm but also precise ones such as Euclidean distance to calculate the distance between two lines. Computing a precise *Hausdorff* distance is time-consuming. A measure may have either a precise or approximate value, e.g. numeric [e.g. distance (a, b) = 3.2], ranked distances, (as in the earliest versions of the Douglas-Peucker simplification algorithm), or boolean [e.g. disjoint (a, b) = false]. Mathematically precise measurements may not be adequate to describe higher-level concepts fully such as patterns and structures using a single value. In many cases they have to be characterised geometrically using auxiliary data structures such as MST (Regnauld 1996).

McMaster and Shea (1992) described several measures in digital cartography for determining when to apply generalisation operations: for example density, distribution, length and sinuosity, shape, distance, Gestalt, and abstraction. To support quality evaluation, Weibel (1995) proposed a classification of measures into global, geometric, topological, and software-related categories. At present, the most comprehensive classification of measures for quality evaluation is proposed by Mackaness and Ruas (2007). It has two aspects:

(1) A measure can be internal or external:

   (a) Internal: measure of objects at the same scale (within a dataset)
   (b) External: measure of objects between two scales (before and after generalisation).

(2) A measure can be Micro, Meso or Macro:

   (a) Micro: measure on individual objects or parts of an object
   (b) Meso: measures on groups of objects
   (c) Macro: measures on all objects of a feature class.

As defined in AGENT (1999), there are two classes of external measures. An implicit external measure is computed by comparing two internal measurements calculated at the initial and target scale (e.g. change rate of size by dividing the two areas). Explicit external measures compare two objects at different states (e.g. *Hausdorff* distance for describing the shape change of lines at two scales). The measures at micro- and meso-level are in line with the map requirements of one object, between two objects or on groups of objects. For instance, measures at the micro-level relate to position, orientation, size, and shape constraints on one object. Measures at the meso-level deals with topology, spatial arrangement/distribution, and density constraints on groups of objects.

Shape describes the geometric form of individual spatial objects (Wentz 1997) and should be maintained during the generalisation process. General shape is difficult to measure. There are many internal shape measures (e.g. sinuosity, complexity, elongation, compactness, concavity, shape index) AGENT (1999). The measures describe certain aspects of shape using scalar values.

While all of the individual metrics can be computed, they do not readily summarise to a general shape measure that is readily understand. The work of Skopelity and Tsoulos (2000) and Skopelity and Lysandros (2001) advanced the formalisation of shape metrics in an important way. They reduced these structure metrics to only three primary metrics using Principle Component Analysis. Similarity measures of shapes have also been studied. For example, Veltkamp and Hagedoorn (1999) summarised several techniques for shape matching (e.g. Fréchet distance, Hausdorff distance, turning function, signature function). Ai et al. (2008) proposed a shape descriptor for polygonal objects based on Fourier transform. These measures can be used to compare the similarity of shapes at different map scales.

For each of these measures we might ask: Which is the best-suited measure for the description of general shape in map generalisation? How do we combine different measures together to describe general shapes? If we can answer these questions then shape measures can be incorporated into the evaluation process (Burghardt et al. 2008).

High-level concepts or phenomena, such as road and water networks, building alignments and patterns are especially important at smaller scales. A definition of these concepts is hard to come by: What are the characteristics of these high-level concepts? Can these characteristics be measured? Do these measures really reflect the nature of the concepts in the experts' mind? How do these high-level phenomena change at scale transitions? These questions are pertinent to automated evaluation. Measuring higher-level concepts might involve complex computations of auxiliary data structures, such as Delaunay triangulation, Minimum Spanning Tree, and Voronoi diagram.

Various authors have examined how these high level concepts can be measured. Regnauld (1998) used MST in conjunction with road networks to measure the characteristics of groups of buildings. Christophe and Ruas (2002) used a straight-line scanning technique and statistical analysis to detect building alignments and measure their distribution. A density measurement based on skeletonisation of gap space was developed to evaluate the spatial distribution of the density of features (Zhang et al. 2008). For network features, 'continuity' is regarded as a measure for evaluating the quality of generalised network features (Edwardes and Mackaness 1999), as is "connectivity" (Li and Zhou 2012). Other methods for describing network features utilise graph theory (Mackaness 1995; Jiang and Claramunt 2004; Heinzle et al. 2005, 2006, 2007).

Despite these efforts, it is not yet known which one is suited best for modelling a certain characteristic with respect to a constraint. Relatively little work has been done in formalising the modification of these high-level concepts at scale transitions. Initial work on formalising and evaluating high-level concepts has been done by Zhang (2012) who studied the automated evaluation of building patterns in generalised maps based on a classification of building alignments (see Sect. 9.7).

9.5.3 Data Matching Between Initial and Target Data

One difficulty of automated evaluation is that the reference data (which is more detailed) and the data being evaluated are at different scales (Bard and Ruas 2004; Mackaness and Ruas 2007). Consequently, corresponding relations identifying objects that represent the same real-world entities (an object or a group of objects) at different scales are hard to establish (Bobzien et al. 2008). In order to measure the properties of spatial entities at different scales and to compare the measured values of both entities, the relations between entities at the two scales should be explicitly known.

In the generalisation domain, this kind of relation is referred to as a vertical relation, and obtaining this relation-information is considered as a data enrichment task (Neun 2007). The vertical relation has three different classes: 1-to-1 relation, *n*-to-1 relation and *n*-to-*m* relation (Ai and van Oosterom 2001). The 1-to-1 relation is the simplest relations that can be identified during generalisation. The automated evaluation against preservation constraints on one object (e.g. absolute position, orientation, shape) depends on the 1-to-1 relation of corresponding objects. The 1-to-1 relation usually results from generalisation operations such as simplification, exaggeration and displacement (Neun and Steiniger 2005). The n-to-1 and n-to-m relations result from applying generalisation operations such as aggregation and typification.

Mao et al. (2012) proposed the usage of Attributed Relational Graphs (ARG) as a similarity measure of n-to-1 and n-to-m relationships. In their work they used the method for comparing original and generalised city models consisting of 3D buildings. The approach used was that the original buildings were described by one ARG (graph) and the generalised by another ARG where the nodes contained visual features of the buildings (height, ground plan area, etc.) and the edges relationship between the buildings (e.g. distance between buildings). To model the similarity between the original and generalised buildings then become an ARG matching problem. A common method to perform ARG matching is the Nested Earth Mover's Distance (NEMD) method (Kim et al. 2004). This type of approach is common in pattern recognition, but more studies are required to evaluate its usefulness in cartographic generalisation.

The vertical relations of topographic data at different scales are not an explicit and standardised part of most data models. The links have to be created by means of automatic matching techniques. The primary objectives of data matching between different levels of detail are for establishing MRDB and for spatial database integration (Bernier and Bédard 2007). Various studies on matching network data such as roads can be found in the literature (Devogele et al. 1996; Zhang and Meng 2007; Mustière and Devogele 2008). It is clear that automated evaluation against the constraints of 1-to-1 will benefit from a data matching process relationships (e.g. preservation of absolute position, orientation). But it is not yet clear the extent to which data matching is required for automated evaluation and how this works for n-to-1 and n-to-m relations.

Some work has evaluated generalisation results by automated comparison to a previous accepted cartographic solution, or benchmark dataset. Stanislawski (2009) presents a Coefficient of Line Correspondence (CLC) in a gridded solution to map the distribution of how well a generalised set of lines matches benchmark lines for the same area. The method does not relate between the two sets, but rather marks features in both sets as either matching or not matching. A confidence interval is generated for the study area CLC value (Stanislawski et al. 2010). The CLC has been used to evaluate different generalisation solutions for hydrography (Stanislawski and Buttenfield 2011; Buttenfield et al. 2011). The CLC is analogous to a "similarity" measure (Tversky 1977; Willett 1998), which has been used to compare graph structures and was recently used by Li and Zhou (2012) to compare a generalised road network against a benchmark. The CLC has been used in conjunction with raster-based line density differencing to quantitatively and visually evaluate how well original and generalised hydrographic lines conform to terrain-derived drainage lines (Stanislawski et al. 2012, 2013). Raster density differencing is a relatively fast automated method to map the distribution of local differences in the context of two sets of vector data at various resolutions.

9.6 Map Readability Formulas

The evaluation of a generalised map requires several balanced perspectives. Preservation constraints are needed to protect the presence and arrangement of objects, and this must be countered by readability constraints to ensure that the map is readable. Readability constraints are linked to map readability formulas that are analytical expressions that describe how legible a map is. They can be used to evaluate any type of map. Map readability formulas are useful in several mapping applications, for example when combining geographic data from several sources to create real-time maps. In the context of map generalisation, map readability formulas are vital for defining map readability constraints.

Readability formulas are currently not as well developed for maps as they are for written text. Readability formulas were introduced in the 1920s and although criticised (Chomsky 1957) for not reflecting the full semantics of deep meaning, they are still used. Several hundred readability formulas have been published and used for predicting the difficulty of reading texts. In order to create readability formulas for maps, the cartographic variables need to be converted to measures in a similar manner as the linguistic variables have been. The following section presents a number of proposed measures for map readability. In Sect. 9.8, a Case study concerning the suitability of these measures is presented.

9.6.1 Measures for Map Readability

There are several measures for map readability. They can be categorised into the following types: the amount of information, the spatial distribution, object complexity and graphical resolution. Some of these measures can be used for other purposes such as for automated evaluation (see Sect. 9.4) and as constraints to trigger generalisation.

The amount of information measured refers to the number and size of map objects. Some measures that have been proposed for map readability are: the number of objects (e.g. Phillips and Noyes 1982; Wolfe 1994; Oliva et al. 2004), the number of objects of a particular type (Töpfer and Pillewizer 1966), the number of vertices (Woodruff et al. 1998; MacEachren 1982; Stigmar and Harrie 2011), the number of nodes, links and areas (MacEachren 1982), and the amount of occupied space (Frank and Timpf 1994; Oliva et al. 2004).

Measures of *spatial distribution* are determined by the density and distribution of map objects. This type of measure can reveal if objects on the map are too densely positioned, and if the map gives a homogeneous or heterogeneous impression. Some of the measures proposed include: the distribution of objects (MacEachren 1982), their symmetry and organisation (Oliva et al. 2004), entropy measures for objects and points (Bjørke 1996; Li and Huang 2002; Harrie and Stigmar 2009), homogeneity and the number of neighbours (Li and Huang 2002), and the density of objects (Hangouët 1998).

Measures of *object complexity* are determined by the shape and size of map objects. Examples of such measures are: sinuosity (João 1998), total angularity (McMaster 1987), and line connectivity (Mackaness and Mackechnie 1999; Fairbairn 2006; Li and Zhou 2012).

Measures of *graphical resolution* are determined by the size and colour of symbols. Examples include contrast of the visualised objects (Eley 1987; Oliva et al. 2004), contrast in brightness and hue (Rosenholtz et al. 2005; Stigmar et al. 2013) and line thickness (Spiess 1995).

A single measure cannot describe the readability of a map, however. We need information about how many objects there are, how they are distributed, and the graphical properties of their symbols. That is, we need to develop a synthesis of measures to establish readability formulas.

The measures above are concerned with the graphical properties of the map objects. Critics argue that these types of quantitative measures do not sufficiently reflect the true readability as they do not capture the whole meaning of the objects, as interpreted by the map reader. However, one can also argue that the meaning of the map objects can be very subjective, and that they therefore are the carriers of the meaning, the graphical symbols, that should be the focus of any measure. When the quantitative measures and readability formulas that best reflect the readability have been found, the next challenge is to quantify semantic aspects into measures and add to the readability formulas.

9.6.2 User Studies for Evaluation of Measures for Map Readability

User studies have assessed the usefulness of these measures. Ideally map readability formulas should be based on them. The general approach of these studies has been to create a number of map samples. A set of measures for each map sample then estimates the degree of map readability. Then user study evaluates the readability of the map samples. Finally, an evaluation compares the analytical map readability with the user studies. Based on this evaluation, conclusions can be drawn if any of the measures, or syntheses of the measures, are successful in predicting the readability of the map.

There have been a number of user studies performed on map readability measures. Phillips and Noyes (1982) tested the complexity of the topographic basis of geological maps. The aim was to enable recommendations for ways of improving 1:50,000 scale geological maps. In this test, five versions of the same map with different topographic bases were compared for map reading performance. They found that the amount of information on a map reflects the performance. The reduction of points produced the largest increase in map reading performance. The results also supported the idea that similar symbols tend to each other, both concerning symbol type (points vs. points, lines vs. lines, etc.) and colour.

Rosenholtz et al. (2005, 2007) present a feature congestion index, which is a synthesis of measures for colour, contrast, and orientation. Their measure of feature congestion is based on the local variability of certain key features. Different features are combined at each point, which results in a clutter map, and combined across an area to obtain a clutter measure (in the form of a raster map). Rosenholtz et al. (2005) tested their feature congestion index by asking 20 users to rank the clutter on 25 different paper maps. The maps were compiled at different scales and used different symbologies. A statistical evaluation showed an agreement in ranking and perceived clutter. Rosenholtz et al. (2007) found that identification of objects and the time taken for search were both depended on the clutter measure.

Stigmar (2006) studied subjective preferences for maps with a specific interest in the amount of information (here represented by the number of vertices in the objects). Twelve test participants were asked to rank a number of test maps according to preference for use as navigational aids in mobile map services. The result showed that while the amount of information was important, participants preferred the presence of "all" map objects over a smaller amount of information. As long as "all" objects were present, the maps were considered better even if they were generalised to a larger degree. In Stigmar and Harrie (2011), 17 measures related to the amount of information, spatial distribution and object complexity were evaluated. The aim was to see whether they could be used to describe map legibility. Twelve test participants were asked to discuss the legibility of a number of test maps and to rank some maps according to their perceived legibility. The results showed that some measures of the amount of information and spatial distribution corresponded well to the participants' opinions. The measures of

object complexity did not show the same correspondence. In Stigmar et al. (2013) the studies were extended to include syntheses of measures (see Sect. 9.8).

Most map readability formulas are based upon the amount of information and spatial distribution of information rather than map design. It is an intriguing question as to how map design really affects map readability. For example, Raposo and Brewer (2011) made an user study of map design on top of orthoimages. Their study revealed that the information content affects map reading more than the design of the map. They found that map design alone did not significantly affect map reader's impressions, performance or certainty.

The use and design of map readability formulas is still a research issue. One question is to ask, how general the map readability formulas could be? The user studies reported above use small map samples, and it is not certain that formulas based on small sample sizes can be extended to larger maps. Another difficulty is that geographic data do vary substantially between data sets, which make it difficult to create general formulas. On the other hand, it would be very useful to have map readability formulas to trigger, control and evaluate the map generalisation process. The research on the issue has been quite limited so far; to really state the value of map readability formulas would require more user studies and probably the design of new formulas.

9.7 Case Study I: Automated Evaluation of Generalised Building Patterns

Xiang Zhang

The aim of this Case study (also reported in Zhang et al. 2013b) is to demonstrate how preservation constraints can be evaluated automatically for groups of objects (e.g. building alignments). As previously noted, preservation constraints (i.e. constraints that keep important characteristics of geographic phenomena) and constraints for groups of objects (i.e. constraints that deal with groups such as networks and alignments) are more difficult to formalise and evaluate. This is because corresponding relations between source and target objects and patterns, which are implicit in the datasets, have to be made explicit in order to support automated evaluation. In this Case study, the components of pattern recognition and data matching will be discussed.

9.7.1 Methodology

The evaluation methodology can be described as a three-step approach starting from data enrichment through to data matching and to constraint evaluation). For more details the readers are referred to Zhang (2012). Before presenting the

Table 9.1 Harmonised constraints on alignments (after Stoter et al. 2009a)

Constraint ID	Geometry type	Kind of group	Kind of objects of the initial data composing the group	Constrained property	Condition to be respected
CI	Point/ polygon	Alignments	[N/A]	Alignment	Alignment should be kept
C2	Point/ polygon	Alignments	[N/A]	Alignment orientation	Target orientation should be similar to initial orientation
C3	Polygon	Building alignment	Buildings aligned	Spatial distribution	Target distribution should be similar to initial distribution

three-step evaluation, map specifications related to the generalisation of building groups are analysed.

The specifications (i.e. constraints) used in this Case study are based on the generic constraints from the EuroSDR project on automated generalisation (Sect. 9.9). The EuroSDR constraints were harmonised based on expert surveys intended for capturing NMA specifications across different mapping organisations and universities in Europe. These constraints are regarded to be generic for many countries.

Table 9.1 lists three group-level constraints that are relevant to building features, though they are also applicable to other feature classes that contain groups of objects. These constraints describe existence (C1), orientation (C2) of alignments, and spatial distribution of buildings in alignments (C3).

It is useful to further distinguish two levels of constraints. Entity-level constraints (e.g. C1) are in effect when the 'constrained property' in Table 9.1 (e.g. alignment in C1) applies to the entities themselves. Another example of this type of constraint is: 'important buildings should be kept'. Property/relation-level constraints evaluate the property of, or relation between, entities rather than the entities *per se* (e.g. target building shape should be similar to initial shape). C2 and C3 are property-level constraints.

In general, automated evaluation of preservation constraints should be carried out by linking and comparing target datasets with source datasets (Burghardt and Schmid 2010). The above distinction implies two schemes in which the comparison can be performed. In one scheme, property-level constraints perform the comparison by checking the states of the entities that are kept in the target dataset and referring back to their initial states. In the second scheme, the evaluation of entity-level constraints starts from the source dataset and checks if initial entities are still represented in the target dataset. In this case, any deletion of entities (alignments in this Case study) will be penalised by entity-level constraints.

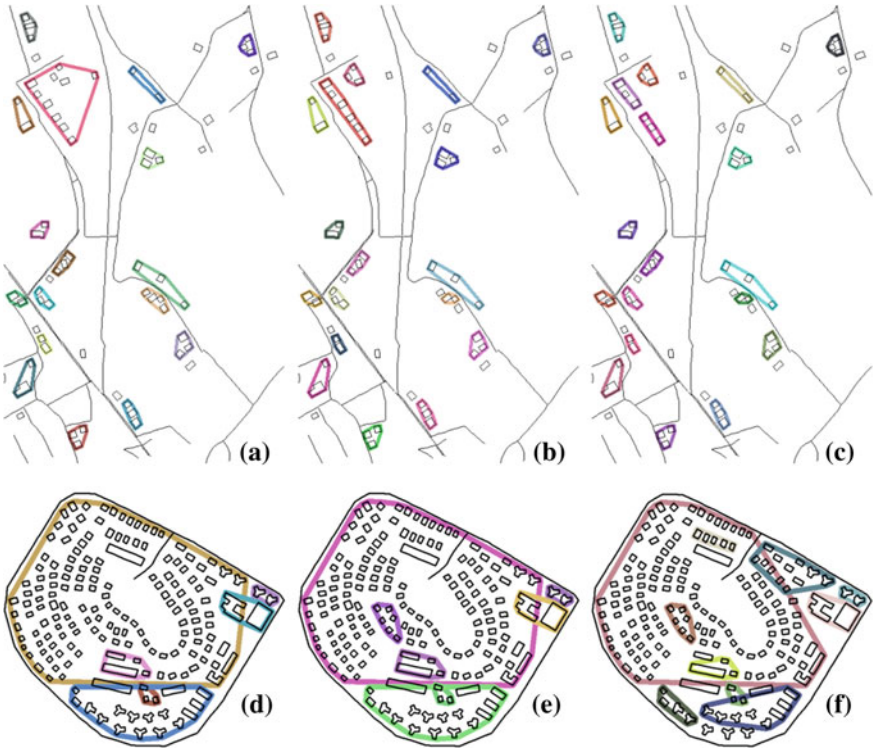

Fig. 9.2 Building clustering based on the method of Zahn (1971) with $n = 3$ (*left column*), 2 (*middle*) and 1 (*right*) (as the parameter n decreases more finer-grained groups are detected; a random colour scheme is used to delineate individual clusters; *Data*: *Top* Kadaster, NL, *Bottom* Shenzhen, China)

Building patterns discussed here are local arrangements of building groups, to which the human eye is usually attracted. These are important characteristics to retain on topographic maps. Typical forms of local building patterns include linear and nonlinear clusters. The former can be divided into collinear, curvilinear and align-along-road alignments; whereas the latter can be refined into grid-like and unstructured clusters (Zhang et al. 2013a; Anders 2006).

Building clusters are usually detected using proximity graphs, such as minimum spanning tree, relative neighbourhood graph and Delaunay triangulation (Regnauld 1996; Anders 2003). Regnauld (1996, 2001) automatically recognised building clusters for contextual building generalisation using a well-known graph-theoretic point clustering algorithm (Zahn 1971). However, recognising useful patterns is more than just identifying clusters, as is shown in Fig. 9.2. The clustering algorithm cannot explicitly recognise more refined building patterns such as alignments, and especially in densely populated areas the detected patterns are less meaningful (Fig. 9.2d–f). More details of the parameters used in the clustering can be found in Zahn (1971) and Zhang (2012).

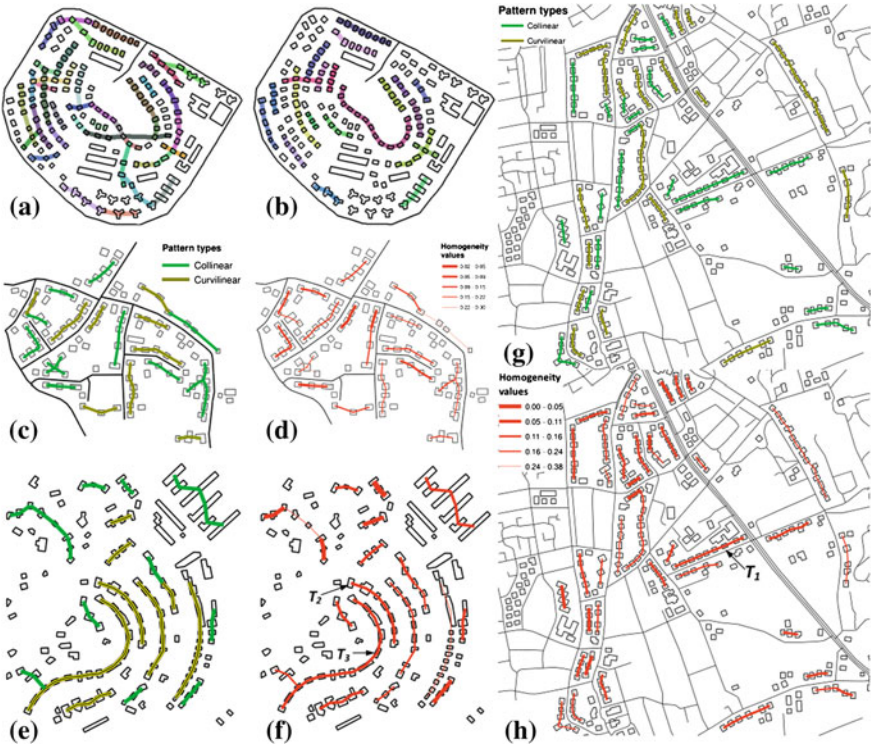

Fig. 9.3 Building alignments recognised from topographic datasets: results by a collinear algorithm (**a**) and a curvilinear algorithm (**b**); combined results of various algorithms which are visualised by pattern types (**c**), (**e**) and (**g**), and by pattern regularities (**d**), (**f**) and (**h**)

The above clustering method is based solely on a distance (proximity) relationship and ignores other important factors (e.g. size, orientation, shape, good continuity) that may impact upon the grouping process. Zhang et al. (2012) therefore proposed algorithms that recognise align-along-road and unstructured clusters. An algorithm that detects grid-like patterns was previously described by Anders (2006) for grid pattern typification. Later, Zhang et al. (2013a) proposed a recognition framework that integrates aspects of computational geometry, graph-theoretic concepts and theories of visual perception. In this framework several algorithms are used in parallel to detect collinear and curvilinear alignments, after which a mechanism was designed to combine conflicting results to improve the recognition quality. Figure 9.3 illustrates some of the recognition results.

The recognition framework described in Zhang et al. (2013a) deals explicitly with alignments and also improves the general clustering method (Figs. 9.3a–b and 9.2d–f). For the evaluation in this Case study, this recognition framework was applied with the same parameter values to both source and target datasets for matching and evaluating building alignments. According to the constraints

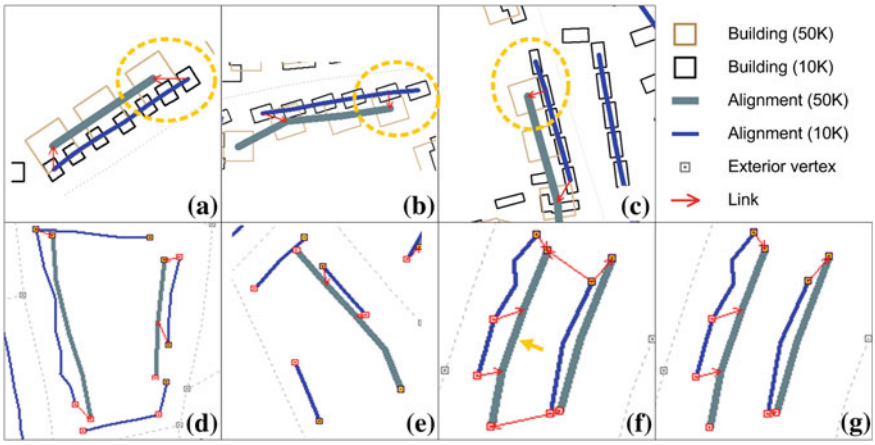

Fig. 9.4 Example results of alignment matching: (**a**)–(**c**) details of the sub-alignment matching; (**d**)–(**e**) elimination of irrelevant candidates; (**f**)–(**g**) removal of inconsistent many-to-many correspondences; reprinted by permission of the publisher (Taylor and Francis Ltd, www.tandfonline.com)

described in Table 9.1, orientation and homogeneity were measured for the detected alignments. Since building alignments are regular distributions of buildings, regularity (measured in terms of homogeneity) can be used to measure the spatial distribution of the buildings.

In order to automatically evaluate building patterns, corresponding patterns in the source and target scales must be identified. This is completed through an automated data matching process. The matching is based on the geometric representation of detected alignments (each alignment is represented by a line connecting the buildings). Because the many-to-many and partial correspondence between source and target alignments is a side-effect of generalisation, a matching process using sub-alignment elements (i.e. buildings in alignments) is designed, such that it is possible to identify which buildings in the alignments correspond to each other. This can be one building to one building, but also one building to two or more buildings, which is facilitated through the sub-alignments. The sub-alignments help to identify parts of the alignments that correspond exactly to each other. Additionally, a confidence indicator (CI) is introduced to document the reliability of the evaluation caused by the partial matching. If two alignments properly match, CI equals to 1 (Fig. 9.4a); otherwise CI decreases as the exact corresponding parts become smaller. For a detailed description of the matching algorithm, the partial correspondence and the confidence indicator see Zhang et al. (2013b). Figure 9.4 shows some of the matching results.

After data enrichment and matching processes, generalised topographic data can be automatically evaluated against the constraints (C1–C3) defined in Table 9.1. CV is used to denote the degree of constraint violation, where $CV \in [0, 1]$ with $CV = 1$ indicating a complete violation and $CV = 0$ meaning that the

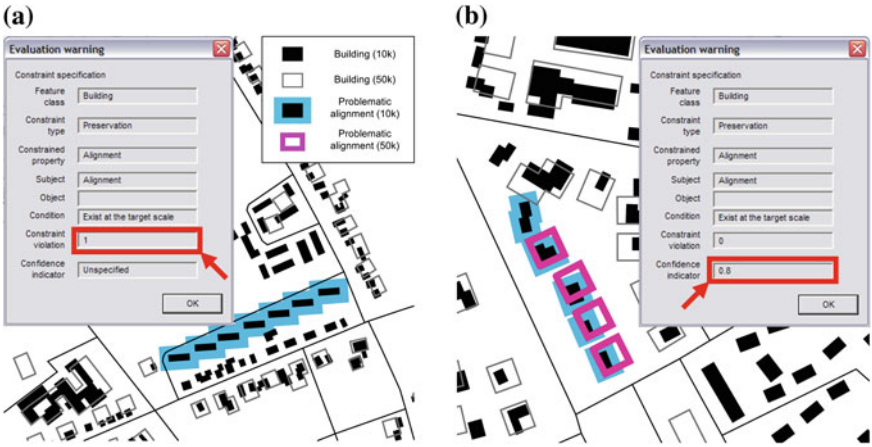

Fig. 9.5 Violation to constraint C1 (**a**) and an alert ($CI = 0.8$) of a less reliable evaluation (**b**); reprinted by permission of the publisher (Taylor and Francis Ltd, www.tandfonline.com)

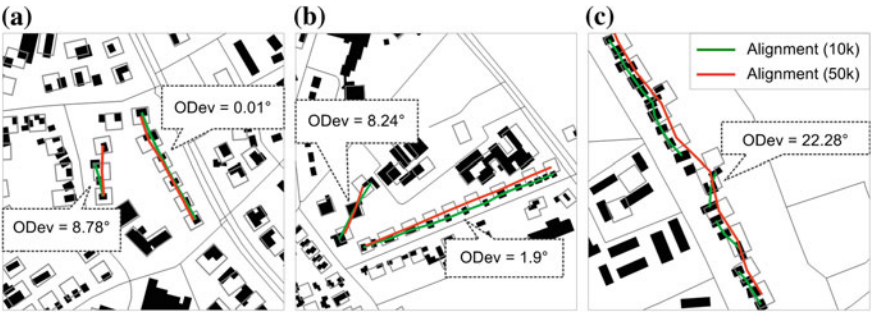

Fig. 9.6 Evaluation of C2 (ODev is defined as the deviations between alignment orientations); reprinted by permission of the publisher (Taylor and Francis Ltd, www.tandfonline.com)

constraint is not violated at all (or is satisfied). The three constraints were evaluated for individual alignments in the source and target datasets, and were aggregated to obtain an overall view of each constraint for a whole dataset. C1 is evaluated by calculating the ratio between matched target alignments and total source alignments. C2 is evaluated by calculating the deviation between orientations of corresponding alignments; the deviation is compared with a threshold to map it into a violation range [0, 1]. C3 is evaluated in a similar way by calculating the deviation between homogeneities of corresponding alignments. Individual assessments of a constraint are aggregated by weighted average. The weight is determined by each alignment: the one that is more homogeneous and contains more buildings is regarded to be more significant in the aggregation. The following figures visualise some of the evaluation results obtained in this Case study. More details can be found in Zhang et al. (2013b).

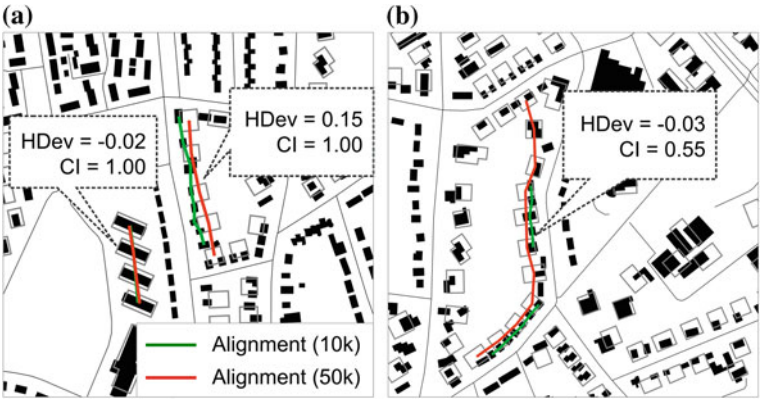

Fig. 9.7 Evaluation of C3 (HDev is defined as the deviation between homogeneities of matched alignments); reprinted by permission of the publisher (Taylor and Francis Ltd, www.tandfonline.com)

Figure 9.5a illustrates how an alignment in the source data is not kept in the target data, which constitutes a complete violation to C1 (i.e. alignment should be kept). While the source alignment is kept in the target data in Fig. 9.5b (CV = 0), an alert (CI = 0.8) indicates a less reliable evaluation caused by the partial correspondence. This can be used by cartographers to further check and improve their generalised maps (there may or may not be a problem).

Figures 9.6 and 9.7 visualise the evaluation results for C2 and C3. The confidence indicator in Fig. 9.7b shows that the result is less reliable, because a short collinear alignment is compared with a much longer and more sinuous curvilinear alignment at the target scale. Figure 9.8 visualises the evaluation of C3 for manually and automatically generalised data. It shows that more violations can be found in the automatically generalised data (alignments highlighted in green do not violate C3 while those in red and orange significantly violate C3). Further results are presented in Zhang et al. (2013b).

9.7.2 Discussion

Results in Fig. 9.8 confirm the assumptions made of the datasets. That is, the interactively generalised data has a higher quality than the automatically generated data when measured against the constraints in Table 9.1. As this Case study only concerns the evaluation of generalisation output, the process as to how the generalisation is performed is out of the scope.

This application case shows that after recognising building patterns, matching alignments between source and target datasets, the three preservation constraints on groups of objects can be evaluated automatically. The three-step evaluation

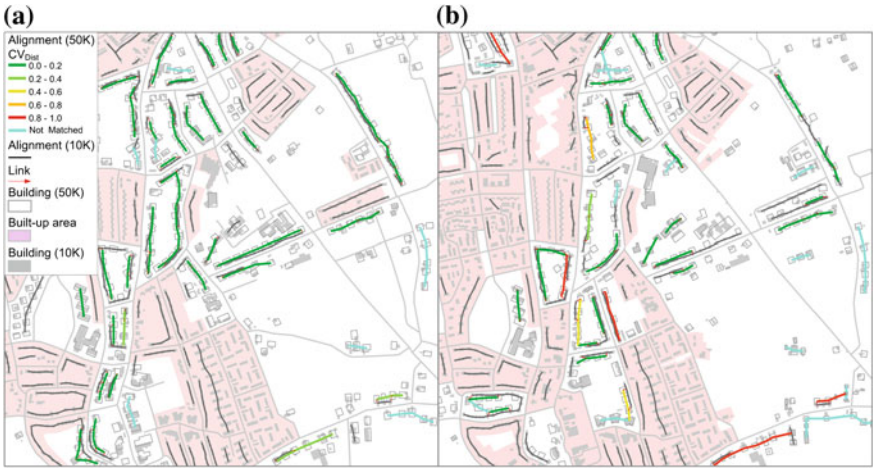

Fig. 9.8 Constraint violations are visualised for interactively (**a**) and automatically generalised data (**b**): target alignments are matched to and evaluated against C3; reprinted by permission of the publisher (Taylor and Francis Ltd, www.tandfonline.com)

approach should provide a general framework for automated evaluation of generalisation output, to compare different solutions and to summarise their data qualities. With this approach it is also possible to validate and evaluate existing and new quality measures by applying it to both interactive and automatically generalised data, which is otherwise not possible for evaluations before and during generalisation.

9.8 Case Study II: Map Readability Formulas

Hanna Stigmar and Lars Harrie

The aim of this study was to develop a map readability formula based on synthesis of measures. For details of the study see Stigmar et al. (2013). A total of 350 map samples were used in the study (Fig. 9.9). These were derived from a geographic database covering the vicinity of Helsingborg, Sweden, consisting of layers in the scale range 1:10,000–1:50,000. The map regions were chosen to present relatively homogeneous areas, and represented most typical types of areas in the map, and at the same time varying in map legibility. The maps were compiled using three levels of detail by selecting data layers with different resolutions. The maps were symbolised by two different symbologies, one traditional Swedish symbology (TS) and one "new" symbology (NS) developed by the Swedish National Land Survey.

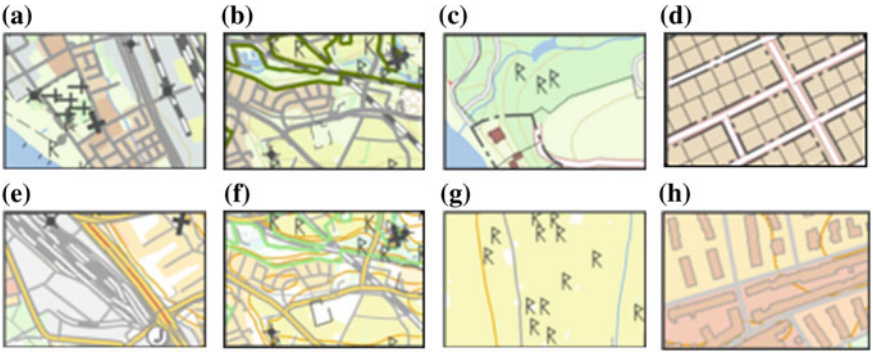

Fig. 9.9 Eight map samples used in the study. The maps are symbolised by two different symbologies, (NS and TS), at two different scales (1:50,000 and 1:10,000) and with three different levels of detail (LOD 1-3). Maps *b* and *f* only differ in symbologies

9.8.1 Methodology

The general methodology of this study was to compute several readability measures for each map sample. Then *analytical readability values* were computed using syntheses of the measures (using four different techniques). A user study was performed to obtain *perceived readability values*. Finally, in the evaluation step, the analytical and the perceived readability values were compared for each map sample.

For each map sample several readability measures were computed. Using the classification of measures given in Sect. 9.6 the following measures were used. Measures of the amount of information were: *number of objects*, *number of vertices, object line length, object area and number of object types*. Measures of spatial distribution were: *spatial distribution of objects, spatial distribution of vertices, degree of overlap, semantic homogeneity*, *number of neighbours*, *local density and proximity indicator*. Object complexity included five measures: *object size, line segment length, angularity*, *line connectivity* and *polygon shape*; while measures of graphical resolution were two: *brightness difference* and *hue difference*. For definition of the measures see Harrie and Stigmar (2009) and Stigmar et al. (2013).

The user test was designed as a web-distributed questionnaire in which participants were asked to asses the level of map legibility of the map samples as "very difficult to read", "difficult to read", "easy to read" or "very easy to read". The participants were also asked to rank some maps according to how difficult they were to read. A total of 214 participants were included in the study. Most of these were GIS professionals or GIS/geography students. Based on the user test each map sample was assigned a perceived readability value and classified as either readable or non-readable.

The evaluations were performed first for single measures and then for syntheses of measures. Four methods for syntheses were tested: manual interpretation of threshold values, multiple linear regression, support vector machines (Vapnik

1979; Tso and Mather 2009) and a supervised artificial neural network method (Biased ARTMAP, see Carpenter and Gaddam 2009).

9.8.2 Result

For the evaluation of single measures the results showed that measures of the amount of information provided the best correlation with perceived map legibility. The best correlation was given by *number of object types*, *number of vertices* and *object line length*. For the measures of spatial distribution, *proximity indicator* and *degree of overlap* showed the best correlation, while no measures of object complexity and graphical resolution alone showed a statistically significant correlation.

The first syntheses method was manual interpretation of threshold values. By manual inspection of scatter plots for various measures, we obtained a number of thresholds for non-readable maps. These are listed in Table 9.2.

A readable map was defined as a map in which all the values were below these thresholds. When all the maps were classified analytically a comparison was made with the outcome of the user test. The result revealed 85 % agreement between the classified map samples. Maps that were not correctly classified were visually investigated. Most of these maps were rather similar, and could be categorised into two groups: maps containing a high density of lines (see the example in Fig. 9.10, left), and maps with dense and overlapping point symbols in a small area (Fig. 9.10, centre). There was one exception to these categories; a map with many small, rather similar-looking area objects forming tessellations (Fig. 9.10, right).

For the three analytical syntheses methods the following measures were selected:

m_1 number of vertices
m_2 number of object types
m_3 degree of overlap of disjoint objects
m_4 brightness difference
m_5 object size.

The values of these measures were then used to classify the map samples into the classes readable or non-readable. The outcome of this classification was then compared with the user test (Table 9.3).

9.8.3 Discussion and Conclusion

The manual interpretation of threshold values gives good result. However, this method is rather limited as it is labour intensive. We would recommend using *manual interpretation of threshold values* for obtaining knowledge about the

Table 9.2 Thresholds for non-readable maps found in the manual interpretation of threshold values

Measure	Threshold
Number of object types	>17
Number of objects (for area tessellations)	>11 cm^{-2}
Object line length (for area tessellations)	>17 cm^{-1}
Object line length (for continuous fields)	>4 cm^{-1}
Object line length (for all objects)	>27 cm^{-1}
Number of vertices (for continuous fields)	>70 cm^{-2}
Number of vertices (for all objects)	>450 cm^{-2}
Proximity indicators	>80 pairs
Degree of overlap (for disjoint objects)	>3
Angularity (maximal)	>40 cm^{-1}

Table 9.3 Result of the evaluation syntheses of measures

Measures used	Multiple linear regression	Support vector machine	Biased ARTMAP
m_1	75	74	69
m_1, m_2	77	76	65
m_1, m_2, m_3	79	82	66
m_1, m_2, m_3, m_4	78	79	71
m_1, m_2, m_3, m_4, m_5	81	83	67

The columns show percentage of correctly classified map samples

Dense lines | Dense/overlapping symbols | Small, similar-looking area tesselations

Fig. 9.10 Three examples of maps that could not be identified correctly with the help of threshold values

relationship between values of the measures and the perceived legibility. For this purpose a visual method is invaluable. For the automated syntheses methods we recommend support vector machine. This method separates classes (e.g. readable and non-readable maps) using hyperplanes. This is a property that corresponds with observations from the visual interpretation of the outcome of the user study (if one or several readability measures are violated then the map is regarded as non-readable) (Stigmar et al. 2013).

From this Case study it can be concluded that several challenges need to be addressed to develop good map readability formulas that can support readability

constraints in map generalisation. First, measures need to be improved. Syntactic measures that can cope with properties as in Fig. 9.10 are needed. Well-defined semantic readability measures that can be implemented through automated approaches need to be identified and developed. Additional studies should address how to synthesise these measures to classify whether a map is readable. Finally, and probably most importantly, more user preference studies and usability studies are needed to evaluate the use of these map readability formulas and their usefulness in map generalisation context.

9.9 Case Study III: The EuroSDR Project

Jantien Stoter

The EuroSDR project was introduced in Sect. 2.5. This project on the state-of-the-art of automated generalisation is a cooperative project among Universities, NMAs, and institutes across Europe (Stoter et al. 2009a; Stoter 2010). The main objective of the EuroSDR project was to study the state-of-the art of available commercial generalisation software packages to generalise a complete map. The methodology of the project consisted of (1) setting up the test cases, (2) involving commercial generalisation software vendors into the project, (3) defining map requirements, (4) setting up test and (5) designing and executing the evaluation methodology. There were four source datasets (see Chap. 2: from ICC, IGN France, OSUK and Kadaster)) and four generalisation systems: ESRI Inc. (USA), University of Hannover (Germany), 1Spatial (United Kingdom) and Axes Systems (Switzerland). The tests were performed on versions that were commercially available in June 2007. Each generalisation test was carried out by up to four testers, which resulted in around 40 generalisation outputs. The following sections describe the evaluation methodology of the project and the main lessons learned from the evaluation.

9.9.1 Evaluation Methodology

Evaluation methodology in the EuroSDR project was related to the evaluation of generalisation output after the generalisation process. It contained three basic evaluation procedures (Fig. 9.11):

- Evaluation I: automated evaluation of generalised data depends on the specification of target dataset (defined as a set of constraints), the results and partly on the initial source dataset. This type of evaluation was undertaken to answer: how many constraints are solved or violated? To what extent is the generalisation output satisfactory with respect to the map specifications?

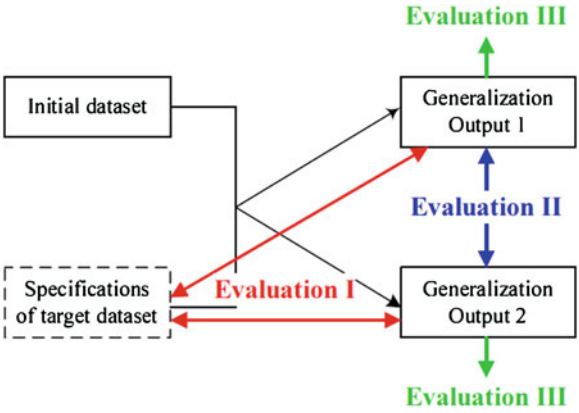

Fig. 9.11 Overview of the evaluation methodology developed in the EuroSDR project (Burghardt et al. 2007)

- Evaluation II: this type of evaluation compares the results of different outputs generalised by different software, in order to get insight into the generalisation process.
- Evaluation III: this type of evaluation requires experts to visually evaluate the final results (expert evaluation).

Within the EuroSDR project, automated evaluation was carried out only for constraints on minimum size and distance (Burghardt et al. 2008; Stoter et al. 2009a; Stoter 2010). Unfortunately, it was not possible to automatically evaluate all of constraints within the time frame of the project (Burghardt et al. 2008). However, Follow-up research was carried out to evaluate several preservation constraints on individual objects (Schmid 2008; Burghardt and Schmid 2010).

9.9.2 Lessons Learned from the EuroSDR Project

Usually, the preservation specifications in the EuroSDR project were more difficult to evaluate than the legibility specifications. This was partly due to weak formalisation of these constraints. Therefore, better understanding of preservation specifications is required in order to improve their formalisation as constraints as well as the measurement of constraint violation. This includes the concepts involved (i.e., how to mathematically describe "shape" on the basis of existing measures such as length-to-width-ratio, shape index, fractal dimension,) and of the changes allowed (how to mathematically describe acceptable modifications). Harrie (2001) obtained such information by studying existing maps at different scales. Another problem in evaluating preservation constraints is that the required correspondence with the initial data is not always easy to find, specifically when it does not concern a 1:1 relation.

The difficulty of evaluating preservation specifications was also encountered in the expert survey: it was often unclear whether a preservation constraint was

assessed as "good" because the system had carefully accounted for it, or because the system had simply ignored it and at the same time had not much altered the data during the generalisation process.

This was also true for legibility constraints. For example if the system removes all the elements less than a minimum size, instead of exaggerating them, the automatic evaluation of the minimum size constraint will give a "good" result, because the constraint is not violated, but the resulting map does not represent the situation very well.

Further research is needed on how far a violation of preservation constraints is deemed tolerable. Investigations on interactively generalised data showed that cartographers also tolerate violations of legibility constraints. For both legibility and preservation constraints, a formal description of tolerated violations is therefore required. Furthermore research on the weighting between constraints and constraint violations has to be carried out to guide the generalisation process and to get an overall evaluation result for the generalised map.

The project methodology used constraints both to direct the generalisation process as well as to determine to what extent the output maps meet the specifications. The evaluation, which integrates three methods, has shown that this approach has an important limitation: the results for individual constraints are not always a good indicator of the quality of the overall solution. This has various explanations. First, some constraints may have been violated deliberately to enable good results for other constraints, e.g., by allowing (slightly) more displacement to avoid overlap. Secondly, as was observed in the automated constraint-based evaluation of interactively generalised data, one should assess not only *if* a constraint was violated but also if the violation yields an unacceptable cartographic conflict. Third, very good results for one specific constraint (e.g., minimal distance between buildings) may coincide with bad results for another constraint (e.g., building density should be kept). Fourth, a non-satisfied constraint can be due to missing functionality in a system, but can just as well be due to imprecise constraint definition. And finally, as Harrie and Weibel (2007) observed, results of constraint-based evaluation heavily depend on the defined test cases: is the constraint set complete and evenly balanced, or does it contain many constraints for very specific situations? Although the expert evaluation did evaluate generalised outputs on individual constraints taking the specific context into account, future research seek to:

- improve generalisation models and constraints to enable taking the notion of flexibility of threshold values into account;
- express constraint satisfaction in values ranging from 0 to 1, instead of in Boolean values. Boolean values may be more appropriate to identify cartographic errors. They may, however, be less appropriate for assessing the evaluation output, because they do not provide information on the degree to which a threshold is ignored;
- validate the constraint approach by considering how to aggregate "constraint-by-constraint" assessments for global indicators of map quality, specifically by

better understanding their interdependencies and impact. This also raises questions on the domain of constraint satisfaction and violation values and on their weighting and prioritising to make different constraints comparable and to enable their aggregation to global indicators. These issues have previously been addressed in the domain of constraint-based optimisation (see Ruas 1998; Bard 2004b; Mackaness and Ruas 2007);

A future test should consider selecting a representative set of constraints to better evaluate generalisation functionalities in commercial systems.

The EuroSDR study concentrated on the question of whether commercially available solutions could meet the map specifications of NMAs defined as constraints. However, during our tests several other aspects were encountered that are relevant to assessing commercial generalisation systems. For example, the testers found that in some cases topological errors were introduced during the generalisation process, and that links between generalised and ungeneralised objects, required for automated evaluation, were not created in most of the outputs. These aspects should be addressed in future tests.

The tests also highlighted the difficulty of parameterising complex algorithms. In fact, several of our evaluations showed that some vendors' solutions are better than solutions generated by the testers from NMAs and universities which shows that mastery of the software is required to obtain the best possible solutions. Software systems could help the user in finding the best parameterisation, for instance by providing tools to support interactive parameterisation (e.g. providing default parameters), or by providing tools to select similar situations, which could be generalised with the same parameterisation or tools for situation dependent, automated parameterisation. Further research should highlight parameterisation possibilities as well as user friendliness.

In addition, a future test should address aspects not amenable to constraints. The constraint approach is based on the consequences of scale changes. According to Mackaness and Ruas (2007), this bottom–up approach might work better for small-scale changes. In contrast, a top–down approach that meets the consequences of (large-) scale reduction by choosing appropriate representations for phenomena might work better over larger scale changes where changes are much more fundamental. A future test can provide more insights into the appropriateness of both approaches for automated map generalisation. Indeed, it appeared that constraints on the final result are sometimes not sufficient to fully express without ambiguity what is expected. In some cases, specifying the expected transformation can help if this transformation is always the same and if it is well known (for example building sizes should not exceed a certain limit). However, fuzzy and incomplete constraints (for example the overall landscape structure should be respected) resulted in very different interpretations and solutions among the testers. This may require a different approach in defining the requirements for automated generalisation.

The limited sizes of the four test cases precluded addressing the problems of dealing with large amounts of data (computational complexity, potential memory

overflows that necessitate data partitioning, presence of numerous and various particular cases that make some algorithms fail). Future tests should define criteria as well as measuring tools to assess scalability of systems. Future tests should also quantify customisation possibilities. The most realistic way to address NMA-specific requirements may be to customise existing software. This requires facilities for writing extensions or for allowing integration with other systems.

9.10 Conclusions and Further Research

This chapter presented concepts and a review of manual and automated evaluation methodologies in the domains of map generalisation and map readability. Evaluation requires three components: (1) definition and formalisation of map requirements (2) measures for quantifying characteristics of map objects and patterns (also called data enrichment) and (3) data matching between corresponding objects.

From this chapter, we observe that the primary focus of automated evaluation research has been on evaluating constraints on one object as well as between two objects. This is evident in the evaluation model by Bard and Ruas (2004), where the model focuses on individual objects (Fig. 9.1). Less research attention has been paid to entities such as groups and feature classes and to characteristics of patterns and structures and on map readability issues for the whole map (e.g. expressing map readability formulas). This is mainly because constraints on groups, features, patterns, networks, and spatial distributions are hard to define (Burghardt et al. 2007; Stoter et al. 2009a). In addition, readability constraints have got more attention within automated evaluation research than preservation constraints. This is because preservation constraints are harder to evaluate. Readability constraints can be evaluated independently from the source data. In contrast, the evaluation of preservation constraints needs the generation and maintenance of complex (n-to-1 and n-to-m) relations between identical spatial entities.

Acknowledgments Many people contributed to this chapter. We would like to thank all of them, specifically the EuroSDR project team.

This chapter is written with the support of the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs (project code: 11300).

References

AGENT (1998) Constraint analysis. Technical Report A2, Agent Consortium. Available via http://agent.ign.fr/deliverable/DA2.html

AGENT (1999) Selection of basic measures. Technical Report C1, Agent Consortium. Available via http://agent.ign.fr/deliverable/DC1.html

Ai T, van Oosterom P (2001) Spatial relation abstract in map generalisation process. In: Aref W (ed) Proceedings of the 9th ACM international symposium on advances in geographic information systems (ACM-GIS 2001), Atlanta

Ai T, Guo R, Liu Y (2000) A binary tree representation of bend hierarchical structure based on gestalt principles. In: Proceedings of the 9th international symposium on spatial data handling, Beijing, pp 230–243

Ai T, Shuai Y, Li J (2008) The shape cognition and query supported by Fourier transform. In: Headway in Spatial Data Handling (SDH'08), Springer, Montpellier, pp 39–54

Anders KH (2003) A hierarchical graph-clustering approach to find groups of objects. In: The 5th workshop on progress in automated map generalisation, Paris, 2003. Available via http://www.geo.unizh.ch/ICA/docs/paris2003/papers03.html

Anders KH (2006) Grid typification. In: Progress in spatial data handling, pp 633–642

Barber C, Cromley R, Andrle R (1995) Evaluating alternative line simplification strategies for multiple representations of cartographic lines. Cartography Geogr Inf Sci 22(4):276–290

Bard S (2004a) Quality assessment of cartographic generalisation. Trans GIS 8:63–81

Bard S (2004b) Méthode d'évaluation de la qualité de données géographiques généralisées: Application aux données urbaines. PhD thesis, Université de Paris

Bard S, Ruas A (2004) Why and how evaluating generalised data? In: Developments in spatial data handling (SDH' 04), Springer, Leicester, pp 327–342

Bernier E, Bédard Y (2007) A data warehouse strategy for on-demand multiscale mapping. In: Mackaness WA, Ruas A, Sarjakoski LT (eds) Generalisation of geographic information: cartographic modelling and applications. Elsevier, The Netherlands, pp 177–198

Bjørke JT (1996) Framework for entropy-based map evaluation. Cartography Geogr Inf Sci 23(2):78–95

Bjørke, JT (1997) Map generalisation: information theoretic approach to feature elimination. In: Proceedings of the 18th international cartographic conference (ICC'97), vol 1. Swedish Cartographic Society, Stockholm, 1997

Bjørke JT (2003) Generalisation of road networks for mobile map services: an information theoretic approach. In: Proceedings international cartographic conference (ICA), Durban, 2003

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

Brazile F (2000) Semantic infrastructure and methods to support quality evaluation in cartographic generalisation. PhD thesis, Department of Geography, University of Zurich

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

Burghardt D, Schmid S (2010) Constraint-based evaluation of automated and manual generalised topographic maps. In: Gartner G, Ortag F (eds) Cartography in central and eastern Europe. Lecture Notes in Geoinformation and Cartography, Springer, pp 147–162

Burghardt D, Schmid S, Stoter J (2007) Investigations on cartographic constraint formalization. In: Proceedings of the 11th ICA workshop on generalisation and multiple representation, Moscow, Russia, 2007

Burghardt D, Schmid S, Duchêne C, Stoter J, Baella B, Regnauld N, Touya G (2008) Methodologies for the evaluation of generalised data derived with commercial available generalisation systems. In: Proceedings of the 12th ICA workshop on generalisation and multiple representation, jointly organised with EuroSDR commission on data specifications and the Dutch program RGI, Montpellier, France, 2008

Buttenfield BP (1991) A rule for describing line feature geometry. In: Buttenfield BP, McMaster RB (eds) Map generalisation: making rules for knowledge representation. Longman, London, pp 150–170

Buttenfield BP, Stanislawski LV, Brewer CA (2010) Multiscale representations of water: tailoring generalisation sequences to specific physiographic regimes. GIScience short paper proceedings. Available via http://aci.ign.fr/2010_Zurich/genemr2010_submission_11.pdf

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

Carpenter GA, Gaddam CS (2009) Biased ART: a neural architecture that shifts attention toward previously disregarded features following an incorrect prediction. Technical report CAS/CNS TR-2009-003

Chomsky N (1957) Syntactic structures. Mouton, Paris

Christophe S, Ruas A (2002) Detecting building alignments for generalisation purposes. In: Richardson DE, van Oosterom P (eds) Advances in spatial data handling (SDH 2002). Springer, Ottawa, pp 342–419

Dettori G, Puppo E (1996) How generalisation interacts with the topological and geometric structure of maps. In: Kraak MJ, Molenaar M (eds) Advances in GIS research II (SDH96). Taylor and Francis, Delft, pp 559–570

Devogele T, Trevisan J, Raynal L (1996) Building a multi-scale database with scale transition relationships. In: Kraak MJ, Molenaar M (eds) Advances in GIS Research II (SDH'96). Taylor and Francis, United Kingdom, pp 337–351

Edwardes A, Mackaness WA (1999) Modelling knowledge for automated generalisation of categorical maps: a constraint based approach. In: Atkinson P, Martin D (eds) GIS and geocomputation (innovations in GIS 7). Taylor and Francis, London, pp 161–173

Ehrliholzer R (1995) Quality assessment in generalisation: integrating quantitative and qualitative methods. In: Proceedings of the 17th international cartographic conference, Barcelona, 1995

Eley MG (1987) Color-layering and the performance of the topographic map user. Ergonomics 30:655–663

Fairbairn D (2006) Measuring map complexity. Cartographic J 43(3):224–238

Frank AU, Timpf S (1994) Multiple representations for cartographic objects in a multi-scale tree: an intelligent graphical zoom. Comput Graph 18(6):823–829

Galanda M (2003) Modelling constraints for polygon generalisation. In: Proceedings of the 7th ICA workshop on progress in automated map generalization, Paris, France, 2003

Galanda M, Weibel R (2002) An agent-based framework for polygonal subdivision generalisation. In: Richardson D, van Oosterom P (eds) Advances in spatial data handling (SDH 2002). Springer, Ottawa, pp 121–136

Hangouët JF (1998) Voronoi diagrams on segments: properties and tractability for generalisation purposes. Technical report for AGENT, Cogit, IGN

Harrie L (2001) An optimisation approach to cartographic generalisation. PhD thesis, Lund Institute of Technology

Harrie L, Sarjakoski T (2002) Simultaneous graphic generalisation of vector data sets. GeoInformatica 6(3):233–261

Harrie L, Weibel R (2007) Modelling the overall process of generalisation. In: Mackaness WA, Ruas A, Sarjakoski L (eds) Generalisation of geographic information: cartographic modelling and applications. Elsevier, Amsterdam, pp 67–87

Harrie L (1999) The constraint method for solving spatial conflicts in cartographic generalisation. Cartography Geogr Inf Sci 26(1):55–69

Harrie L, Stigmar H (2009) An evaluation of measures for quantifying map information. ISPRS J Photogrammetry Remote Sens 65(3):266–274. doi:10.1016/j.isprsjprs.2009.05.004

Heinzle F, Anders KH, Sester M (2005) Graph based approach for recognition of patterns and implicit information in road networks. In: Proceedings of 22nd international cartographic conference (ICC2005), La Coruna, 2005

Heinzle F, Anders KH, Sester M (2006) Pattern recognition in road networks on the example of circular road detection. In: Raubal M, Miller HJ, Frank AU, Goodchild MF (eds) Geographic, information science, vol 4197. Lecture notes in computer science, pp 153–167

Heinzle F, Anders KH, Sester M (2007) Automatic detection of pattern in road networks: methods and evaluation. In: Proceedings of joint workshop visualization and exploration of geospatial data (IAPRS), vol XXXVI—4/W45

João EM (1998) Causes and consequences of map generalisation. Taylor and Francis, London

Jiang B, Claramunt C (2004) A structural approach to the model generalisation of an urban streets network. GeoInformatica 8(2):157–171

Kazemi S, Lim S, Ge L (2005) Integration of cartographic knowledge with generalisation algorithms. In: Proceedings of IGARSS'05, vol 5, pp 3502–3505

Kim DH, Yun ID, Lee SU (2004) A new attributed relational graph matching algorithm using the nested structure of earth mover's distance. In: Proceedings of 17th international conference on pattern recognition (ICPR'04)

Li Z, Huang P (2002) Quantitative measures for spatial information of maps. Int J Geogr Inf Sci 16(7):699–709

Li Z, Zhou Q (2012) Integration of linear and areal hierarchies for continuous multi-scale representation of road networks. Int J Geogr Inf Sci 26(5):855–880

MacEachren AM (1982) Map complexity: comparison and measurement. Am Cartographer 9:31–46

Mackaness WA (1991) Integration and evaluation of map generalisation. In: Buttenfield BP, McMaster RB (eds) Map generalisation: making rules for knowledge representation. Longman, London, pp 217–226

Mackaness WA (1995) Analysis of urban road networks to support cartographic generalisation. Cartography Geogr Inf Sci 22:306–316

Mackaness WA, Mackechnie GA (1999) Automating the detection and simplification of junctions in road networks. GeoInformatica 3:185–200

Mackaness WA, Edwards G (2002) The importance of modelling pattern and structure in automated map generalisation. In: Joint workshop on multi-scale representations of spatial data, Ottawa, p 11

Mackaness WA, Ruas A (2007) Evaluation in the map generalisation process. In: Mackaness WA, Ruas A, Sarjakoski LT (eds) Generalisation of geographic information: cartographic modelling and applications., Series of international cartographic associationElsevier Science, Amsterdam, pp 89–111

Mao B, Harrie L, Ban Y (2012) Detection and typification of linear structures for dynamic visualization of 3D city models. Comput, Environ Urban Syst 36(3):233–244. ISSN 0198-9715. http://dx.doi.org/10.1016/j.compenvurbsys.2011.10.001

McMaster RB (1983) Mathematical measures for the evaluation of simplified lines on maps. PhD thesis, Universite du Kansas, Etats-Unis

McMaster RB (1987) Automated line generalisation. Cartographica 2(24):74–111

McMaster RB, Shea KS (1992) Generalisation in digital cartography. Resource publications in geography. Association of American Geographers

Muller JC (1991) Generalisation of spatial data bases. In: Maguire DJ, Goodchild MF, Rhind DW (eds) Geographical information systems: principles and applications. Longman, London, pp 457–475

Müller JC, Weibel R, Lagrange JP, Salgé F (1995) Generalisation: state of the art and issues. In: Müller JC, Lagrange JP, Weibel R (eds) GIS and generalisation: methodology and practice. Taylor and Francis, United Kingdom, pp 3–17

Mustiere S, Devogele T (2008) Matching networks with different levels of detail. Geoinformatica 12(4):435–453

Neumann J (1994) The topographical information content of a map/an attempt at a rehabilitation of information theory in cartography. Cartographica 31(1):26–34

Neun M (2007) Data enrichment for adaptive map generalisation using web services. PhD thesis, Department of Geography, University of Zurich

Neun M, Steiniger S (2005) Modelling cartographic relations for categorical maps. In: Proceedings of XXII international cartographic conference

Oliva A, Mack ML, Shrestha M, Peeper A (2004) Identifying the perceptual dimensions of visual complexity of scenes. In: Proceedings of the 26th annual meeting of the cognitive science society

Peter B (2001) Measures for the generalisation of polygonal maps with categorical data. In: 4th ICA workshop on progress in automated map generalisation, Beijing

Phillips RJ, Noyes L (1982) An investigation of visual clutter in the topographic base of a geological map. Cartographic J 19(2):122–132

Plazanet C, Bigolin NM, Ruas A (1998) Experiments with learning techniques for spatial model enrichment and line generalisation. GeoInformatica 2(4):315–333

Raposo PJ, Brewer CA (2011) Comparison of topographic map designs for overlay on orthoimage backgrounds. In: Proceedings of ICC, Paris, 2011

Regnauld N (1996) Recognition of building clusters for generalisation. In: Kraak MJ, Molenaar M (eds) Advances in GIS research II (SDH'96). Taylor and Francis, Delft, pp 4B.1–4B.14

Regnauld N (1998) Généralisation du bâti: Structure spatiale de type graphe et representation Cartographique. PhD thesis, Université de Provence, Aix-Marseille

Regnauld N (2001) Contextual building typification in automated map generalization. Algorithmica 30(2):312–333

Rosenholtz R, Li Y, Mansfield J, Jin Z (2005) Feature congestion: a measure of display clutter. In: Proceedings of SIGCHI conference on human factors in computing systems

Rosenholtz R, Li Y, Nakano L (2007) Measuring visual clutter. J Vis 7(2):1–22

Ruas A (2001) Automatic generalisation research: learning process from interactive generalisation. Technical report 39, OEEPE

Ruas A, Plazanet C (1995) Data and knowledge modelling for generalisation. In: Müller JC, Lagrange JP, Weibel R (eds) GIS and generalisation: methodology and practice. Taylor and Francis, United Kingdom, pp 73–90

Ruas A, Plazanet C (1996) Strategies for automated generalisation. In: Kraak JM, Molenaar M (eds) Advances in GIS research II (SDH'96). Taylor and Francis, United Kingdom, pp 319–336

Schmid S (2008) Automated constraint-based evaluation of cartographic generalization solutions. Master's thesis, Department of Geography, University of Zurich

Sester M (2005) Optimization approaches for generalisation and data abstraction. Int J Geogr Inf Sci 19(8):871–897

Skopelity A, Lysandros T (2001) A methodology for the assessment of generalization quality. In: Proceedings of the 5th ICA workshop on progress in automated map generalization. Beijing, China, 2001

Skopelity A, Tsoulos L (2000) Developing a model for quality assessment of linear features. In: Proceedings of the 4th international symposium on spatial accuracy assessment in national resources and environmental sciences, 2000

Spiess E (1995) The need for generalisation in a GIS environment. In: Müller JC, Lagrange J-P, Weibel R (eds) GIS and generalisation, Gisdata 1. Taylor and Francis, United Kingdom, pp 31–46

Stanislawski LV, Buttenfield BP, Brewer CA, Anderson-Tarver C (2013) Integration metrics for cartographic generalisation: assessment of 1:1,000,000-Scale hydrography and Terrain. In: 26th international cartographic conference, Dresden, 25–30 Aug 2013

Stanislawski LV, Buttenfield BP and Samaranayake VA (2010) Generalisation of hydrographic features and automated metric assessment through bootstrapping. In: Proceedings of the 13th ICA workshop on generalisation and multiple representation, Zürich, Switzerland, 2010

Stanislawski LV, Buttenfield BP (2011) Hydrographic generalisation tailored to dry mountainous regions. Cartography Geogr Inf Syst 38(2):117–125

Stanislawski LV, Doumbouya AT, Miller-Corbett CD, Buttenfield BP, Arundel ST (2012) Scaling stream densities for hydrologic generalisation. In: 7th international conference on geographic information science, Columbus

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

Steiniger S (2007) Enabling pattern-aware automated map generalisation. PhD thesis, Department of Geography, University of Zurich

Stigmar H, Harrie L (2011) Evaluation of analytical methods to describe map readability. Cartographic J 48(1):41–53

Stigmar H, Harrie L, Djordjevic M (2013) Analytical estimation of map legibility. Submitted

Stoter JE (2010) State-of-the-art of automated generalisation in commercial software. Number 58. Official publication—EuroSDR 2010. Available via http://www.eurosdr.net/projects/generalisation/eurosdr_gen_final_report_mar2010.pdf

Stoter JE, Burghardt D, Duchêne C, Baella B, Bakker N, Blok C, Pla M, Regnauld N, Touya G, Schmid S (2009a) Methodology for evaluating automated map generalisation in commercial software. Comput, Environ Urban Syst 33(5):311–324

Stoter JE, van Smaalen J, Bakker N, Hardy P (2009b) Specifying map requirements for automated generalisation of topographic data. Cartographic J 46(3):214–227

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

Touya G (2012) Social welfare to assess the global legibility of a generalized map. In: Proceedings of GIScience 2012, Springer, Berlin

Tso B, Mather PM (2009) Classification methods for remotely sensed data, 2nd edn. CRC Press, United States

Tversky A (1977) Features of similarity. Psychol Rev 84(4):327–352

Van der Poorten P, Jones C (2002) Characterisation and generalisation of cartographic lines using Delaunay triangulation. Int J Geogr Inf Sci 16(8):773–794

Vapnik (1979) Estimation of dependences based on empirical data. Nauka, Moscow (in Russian)

Veltkamp RC, Hagedoorn M (1999) State-of-the-art in shape matching. Technical report, Department of Information and Computing Sciences, Utrecht University

Wang Z, Müller JC (1998) Line generalisation based on analysis of shape characteristics. Cartography and geographic information science 25(1):3–15

Weibel R (1991) Amplified intelligence and rule-based systems. In: Buttenfield BP, McMaster RB (eds) Map generalisation: making rules for knowledge representation. Longman, London, pp 172–186

Weibel R (1995) Three essential building blocks for automated generalisation. In: GIS and generalisation: methodology and practice, vol 1 of Gisdata. Taylor and Francis, London, pp 56–69

Weibel R, Dutton G (1998) Constraint-based automated map generalisation. In: Poiker TK, Chrisman N (eds) Proceedings 8th international symposium on spatial data handling, IGU: Geographic Information Science Study Group, Vancouver, pp 214–224

Weibel R, Dutton G (1999) Generalising spatial data and dealing with multiple representations. In: Longley PA, Goodchild MF, Maquire DJ, Rhind DW (eds) Geographical information systems: principles, techniques, management and applications, geoinformation international, 2nd edn. Cambridge, pp 125–155

Wentz E (1997) Shape analysis in GIS. In: Proceedings of AUTO-CARTO 13

Willett P (1998) Chemical similarity searching. J Chem Inf Comput Sci 38(6):983–996

Wolfe JM (1994) Guided search 2.0: a revised model of visual search. Psychon Bull Rev 1(2):202–238

Woodruff A, Landay J, Stonebraker M (1998) Constant information density in zoomable interfaces. In: Proceedings of advanced visual interfaces'98

Zahn CT (1971) Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Trans Comput C-20(1):68–86

Zhang M, Meng L (2007) An iterative road-matching approach for the integration of postal data. Comput Environ Urban Syst 31:597–615

Zhang X (2012) Automated evaluation of generalised topographic maps. PhD dissertation, ITC, University of Twente. Available via http://www.itc.nl/library/papers_2012/phd/zhang.pdf

Zhang X, Tinghua A, Stoter J (2012) Characterization and detection of building patterns in cartographic data: Two algorithms. In: Advances in spatial data handling and GIS. Lecture notes in geoinformation and cartography, Springer, pp 93–107

Zhang X, Tinghua A, Stoter J, Kraak MJ, Molenaar M (2013a) Building pattern recognition in topographic data: examples on collinear and curvilinear alignments. GeoInformatica 17:1–33

Zhang X, Stoter J, Tinghua A, Kraak MJ, Molenaar M (2013b) Automated evaluation of building alignments in generalised maps. Int J Geogr Inf Sci. doi:10.1080/13658816.2012.758264

Zhang X, Tinghua A, Stoter J (2008) The evaluation of spatial distribution density in map generalisation. Int Arch Photogrammetry, Remote Sens Spat Inf Sci 37:181–188