Integrated Cartographic and Algorithmic Approach for Road Network Database Generalisation

Kazemi, S.,[1] Forghani, A.,[2] Cho, G.[1] and McQueen, K.[1]
[1]Faculty of Education, Science, Technology and Mathematics, University of Canberra (UC), Bruce ACT 2601 Australia, E-mail: sharon.kazemi@canberra.edu.au
[2]School of Civil and Environmental Engineering, The University of New South Wales (UNSW) Kensington Sydney NSW 2033 Australia

Abstract
Custodians of Geographic Information Systems (GIS) databases currently provide timely and high quality spatial products to users, maintaining multiple databases at different scales for different uses. These custodians are also committed to maintaining and updating the currency of cartographic data sets. Maintaining multiple databases is resource-intensive, time consuming and cumbersome. Cartographic generalisation has primarily been used to derive smaller scale map products from these databases. The practice is based on a cartographer's skill and incurs a high cost. Thus, new approaches for automating the generalisation of spatial data to produce highly varied sets of versatile, multipurpose map products need to be developed. This paper presents a generalisation methodology to derive multiscale spatial data through an evaluation of standardised mapping software that was used as a testbed based on the principles of generalisation. It focuses on integration and utilisation of generalisation operators in order to generalise a road network database, in order to produce small scale maps at 1:500,000 and 1:1,000,000 from 1:250,000 national topographic data that led to the development of a framework for derivative mapping concepts.

1. Introduction

National mapping agencies (NMAs), spatial data providers and map producers have been maintaining multiple databases at different scales for different uses. They are also committed to maintaining a set of cartographic data with synchronised updates. Maintaining multiple databases is resource-intensive, time-consuming and cumbersome (Arnold and Wright, 2005). This is a major challenge for NMAs and other map producers (Sarjakoski et al., 2002). A seamless master database should offer capabilities to derive various types of maps at different scales, e.g. at scales ranging from 1:250,000 to 1:10,000,000. Derivative mapping has been identified as an active research and development area for NMAs, industry and academia to accomplish the requirements for evaluation and validation of generalisation (Muller et al., 1995 and Visvalingam and Herbert, 1999). Therefore this research has focused on application of generalisation tools to the development of a detailed generalisation framework for deriving multiscale spatial data. More particularly, the research has focused on utilising the generalisation operators, combined with the skills of a trained cartographer, to generalise a road network database from 1:250,000 national topographic data and in order to produce smaller scale maps at 1:500,000 and 1:1,000,000.

1.1 Relevant Work on Thresholding for Line Simplification

Thresholding techniques have been applied in both image processing and map generalisation. When using the ArcGIS™ ArcToolbox Generalise tool, an appropriate weed threshold level must be determined. For the appropriate tolerance selection Shea and McMaster (1989) argue: '*...the input parameter (tolerance) selection most probably results in more variation in the final result than either the generalisation operator or algorithm selection.*' To decide how much information to discard while retaining useful information is a difficult task in line simplification. In fact, it is a very difficult and tedious task to set an appropriate threshold (Forghani, 2000). The basic question remains: what is the best threshold? Generally, the threshold value is determined by a given application. Mackaness et al., (1998) stated '*...the*

ISSN 1686-6576 / © Geoinformatics International

loss in total line length proceeds at a rate that is approximately linear to the increase in tolerance. Low tolerances produce high rates of change in vertex numbers and low changes in total line length and using higher tolerances increasingly fewer vertices are removed for a rapidly increasing loss in total line length.' The threshold value depends mostly on map content, map quality, map geometry, and the amount of noise present in the database. For operational production cartographic environments, generalisation capabilities of current GIS software tools such as Generalise™ are likely to use automated threshold selection in order to allow the complete cartographic production workflow to be carried out in a uniform environment (Hardy et al., 2004). Line simplification algorithms are mainly based on geometric principles (Cromley, 1992) with the reduction rate affected by a predefined tolerance (Nakos, 1999). The current way of selecting tolerance is by trial and error. Different levels of tolerance, ranging from 10m to 70m, were tested in the course of road network generalisation. Tolerances of 25m and 35m were acceptable for the 1:500,000 and 1:1,000,000 data, respectively. Similar tolerances have been used in other studies (Visvalingam and Williamson, 1995). In practice, generalisation involves thinning out coordinates (vertices) to simplify line strings. In this way, redundant points along lines can be efficiently removed and the changes of line shape are minimised. Approximately 25% of vertices in this study were eliminated, while the length of the lines did not change significantly. From a geometry perspective, a vertex (plural vertices) is a special kind of point that describes the corners or intersections of geometric shapes. Vertices are commonly used in computer graphics and cartographic mapping to define the corners of surfaces (typically triangles) in 3D models, where each such point is given as a vector. The *Pointremove* operator (based on the DP algorithm implemented in ArcGIS™) proved to achieve satisfactory line thinning/ results (Lee, 2004). It removes vertices quite effectively, but produces angularity along lines that is not very pleasing from an aesthetic point of view (Lee and Hardy, 2006). In addition, in some instances, cartographic manual editing is required. The *Bendsimplify* operator is designed to preserve cartographic quality. It removes small bends along the lines and maintains the smoothness (Lee, 1992). Bends are defined as sections of curves between two inflection points. Applications of these algorithms are well documented in the literature (e.g. Jenks, 1989, Mroz et al., 1996 and Nakos, 1999). It would be interesting to compare the performance of the two in future research. For example, at 1:500,000 the *Bendsimplify* operation resulted in reduction of 760 vertices (after generalisation) out of 4,750 vertices (before generalisation) over a subset of road dataset. Wang and Muller (1998) introduced the idea of an iterative bend simplification approach using Generalise™ *Bendsimplify* operator that detects bends in lines to drive the line generalisation process; it was proposed as a form of structure-based generalisation. The operation reduces the number of vertices significantly. This paper gives a brief discussion of a conceptual generalisation framework that forms the methodology for road network generalisation with special emphasis on simplification. This is followed by the analysis of results and evaluation of derived maps based on the *Radical Law* approach (Pfer and Pillewizer, 1966 and Muller, 1995). Finally, it concludes with discussion and indicated research directions.

2. Methodology

2.1 Software and Generalisation Tool

The ArcToolbox Generalise™ tool offers a set of line simplification algorithms such as *Pointremove* that uses the *Douglas-Peucker* (DP) algorithm, which keeps the so-called critical points that depict the essential shape of a line and removes all other points. The algorithm connects the end nodes of an arc with a *trend line*. The distance of each vertex to the trend line is measured perpendicularly. Vertices closer to the line than the tolerance are eliminated. The line (arc) is divided by the vertex farthest from the trend line, which makes two new trend lines. The remaining vertices are measured against these lines, and the process continues until all vertices within the tolerance are eliminated (Mroz et al., 1996). The DP algorithm attempts to preserve directional trends in a line using a specified tolerance factor (Taylor, 2005) by removing vertices from the selected lines and simplifying the line bends. The DP algorithm is one of the most popular and accurate generalisation algorithms available (Jenks, 1989) and has been widely used for many cartographic applications such as coastline generalisation (Dutton, 1999 and Nakos, 1999), river generalisation (Rusak and Castner, 1990 and Moreno et al., 2002), and road generalisation (Kreveld and Peschier, 1998 and Skopeliti and Tsoulos, 2001). The area includes a range of road types described as part of the 1:250,000 national topographic data specifications. Consideration is given to show appropriate level of detail on the maps rather than producing the maps to the scale.

2.2 Study Area and Roads Classification

The study area covers approximately 23,630 km² of Australia's capital city, Canberra Australia (Figure 1a). The area coverage of the chosen sites are in the longitude and latitude ranges 148° 42′ 7″ to 149° 25′ 56″ and 35° 55′ 35″ to 35° 7′ 22″ respectively, with an elevation of 550-700m above sea level. The study site was chosen because it provided a mix of different roads with a reasonable amount of feature density. The selection of the area was based on testing the generalisation concepts over an urban environment, since the density of roads was a determinant factor, and for ease of access to the datasets.

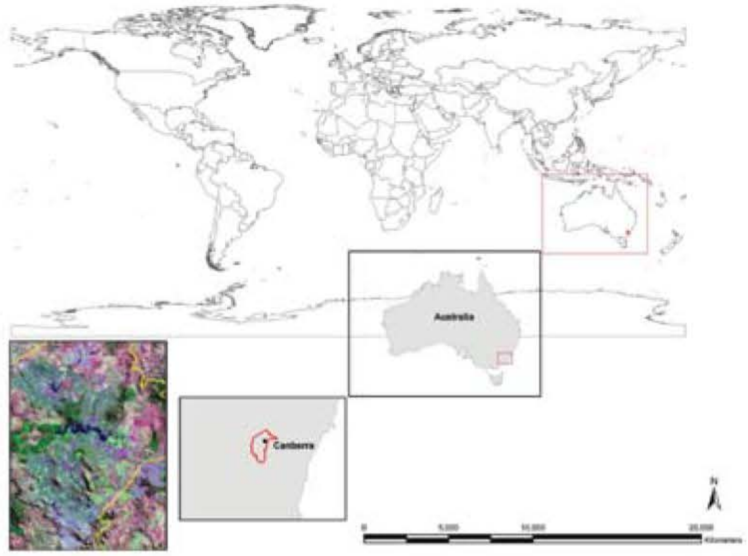

Figure 1a: Location of study area, Canberra, Australian Capital Territory (ACT) Australia; Study site is shown in Landsat ETM+ image (composite bands 6,4,2), courtesy of Geoscience Australia © 2001. The Landsat image is not to the same scale

Figure 1b Examples of roads within the study area (a) Corner of Hindmarsh Drive and Jerrabomberra Avenue, Symonston (Dual carriageway), (b) Commonwealth Avenue, Barton (Principal road), (c) Intersection of Melrose Drive and Yarra Glen Drive, Woden (Secondary road) and (d) Intersection of Leichhardt Street and Dawes Street, Kingston (Minor road)

The National Mapping and Information Group of Geoscience Australia (GA) produce 1:250,000 seamless topographic data and associated cartographic databases across Australia that is used in this research. Roads in the 1:250,000 national topographic data have varying widths and types (GA, 1999). In the classification, six road classes can be distinguished in the study:

- Dual Carriageway (DC) – divided highway, freeway (e.g. Federal Highway), tollway, and other principle roads with separated carriageways in opposite directions.
- Principal road (PR) – highways, main routes and principle connecting roads as described by the Australian Automobile Association (AAA) (both sealed and unsealed).
- Secondary Road (SR) – connecting roads (both sealed and unsealed) that provide a connection between main routes and principle connecting roads or between regional centres. All principal and secondary roads are shown including those in built-up areas.
- Minor Road (MR) – all other roads (both sealed and unsealed) that form part of the public roads system between principal roads and secondary roads.
- Vehicle Track (VT) – public or private roadways of minimum or no construction which are not necessarily maintained.
- Foot Tracks (FT) – these tracks are for pedestrian traffic only. Tracks with a length of less than1.25 kilometres have not been captured.

Principle roads (e.g. Northbourne Avenue) are located in the north and south of the study area. Because roads in the study area have different widths, the road vector coverage was overlaid on the fused ETM+ imagery (15m colour), and every road segment was labelled with a specific distance for its width based on visual assessment. The roads were classified into eight categories with buffer distances ranging from 1 to 6 metres. Principal roads (e.g. urban main roads) can be seen extending in two directions, starting from the north and extending to the southeast and to the southwest. The width of these is between 5 to 10 metres. Minor roads are concentrated mainly in the central part of the study area. The width of these streets is 4 to 8 metres. Minor roads are located mostly in the recreational areas. These roads include rural roads, access roads, and unpaved tracks.

The width of these roads varies from 2.5 to 5 metres. Figure 1b illustrates the different roads in the study area. The process of developing GIS data for this study is the most important step in developing a guideline for derivative mapping using software applications to generalise 1:250,000 national topographic roads to produce 1:500,000 and 1:1,000,000 roads spatial data and maps. This process is depicted in Figures 2 and 3.

2.3 Conceptual Framework

In order to gain an appreciation of the conceptual generalisation framework, it is worthwhile to look at the main steps of the research methodology in Figure 2. This is linked to the conceptual framework that is presented in Figure 3. To date, many frameworks/workflows have been developed for the generalisation of cartographic data (see Kazemi et al., 2004b). It seems that many of these frameworks are generic and do not offer a total solution for the operational environment. The motive of this research is to develop a workflow specifically for derivation of multiple scale maps from a master road network database. A large portion of this framework may be considered as generic too. However in this paper all the processes are considered specifically for road generalisation.

There are a number of properties that could be considered for generalisation of roads to maintain the legibility, the visual identity of each road segment and the pattern. Out of the six generalisation properties (reducing complexity, maintaining spatial accuracy, maintaining attribute accuracy, maintaining aesthetic quality, maintaining a logical hierarchy, and consistently applying generalisation rules) highlighted by Shea and McMaster (1989), the following four are taken into consideration in this research: *Congestion/Legibility, Coalescence, Imperceptibility* and *Length/Distance*. Some of the Shea and McMaster (1989) and Robinson et al., (1984) properties (e.g. length, connectivity, elimination) have been also used by Kreveld and Peschier (1998). According to Lee (2003), and collected feedback through a Cartographic Generalisation Survey (Kazemi et al., 2007a), generalisation of a network (for example, roads, rivers, or other linear features) requires at least six key operations/processes: *Classification, Selection, Elimination, Simplification, Typification, and Symbolisation* (Figures 4 – 6). They are therefore used in the generalisation framework of this research discussed in the next section.

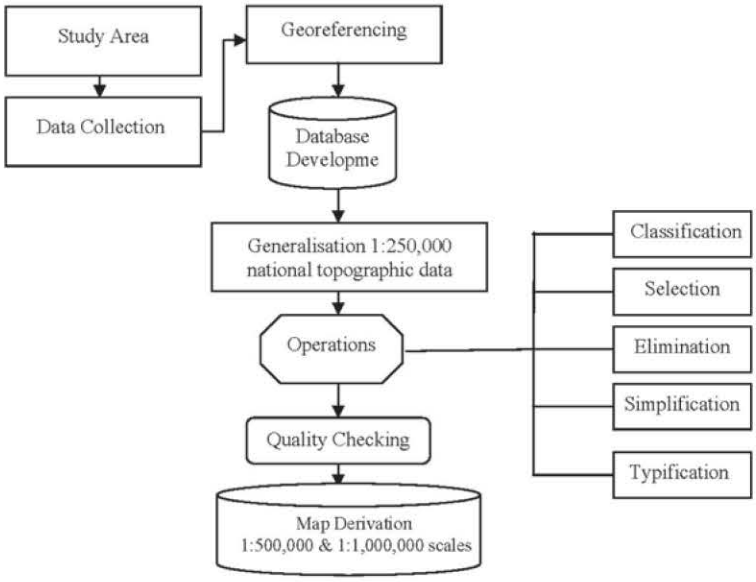

Figure 2: Schematic representation of the research methodology

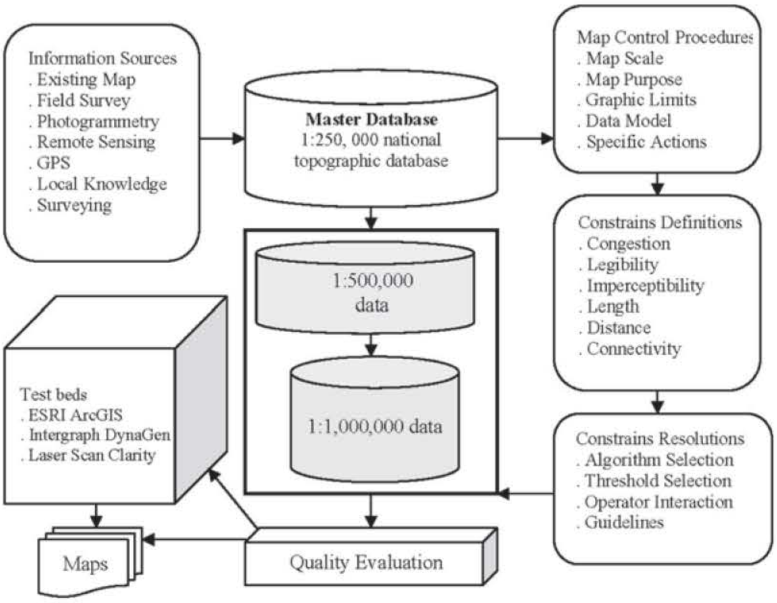

Figure 3: Framework for road generalisation

The generalisation operators of ArcToolbox Generalise™ were tested for the generalisation of roads by employing the conceptual generalisation framework for derivative mapping, but the results showed that the algorithms of *Douglas-Peucker* built into ArcToolbox Generalise™ (*Bendsimplify and Pointremove)* do not support dynamic generalisation.

However, the way to work around this problem is to use Intergraph DynaGen™ generalisation system which enables the derivation of a multiscale database from a master database. The author has demonstrated that testing and incorporating of the DynaGen™ generalisation system enhanced the conceptual generalisation framework (Kazemi and Lim, 2005b).

2.4 Road Generalisation

The following operations are required to generalise a road network (Lee, 2003) (the outputs of generalisation workflow applying these operations are shown in Figures 4 – 6):

- *Classification:* a good classification of roads makes the selection easier and more accurate. This step identifies objects that are placed in groups according to similar properties. It also reduces the complexity and will improve the organisation of a map. For example, roads were categorised into 5 classes with widths ranging from 2m to 12m.
- *Selection:* only certain road classes are selected for inclusion at the target scale. For instance, all vehicle tracks are dropped from the 1:250,000 road database when deriving 1:500,000 scale roads data.
- *Elimination:* mapped features that are not relevant to the map's purpose (Kazemi et al., 2004a), e.g. a road branch shorter than a certain length or small road segments that can cause conflict in the final map and are not significant for representation on the map - can be eliminated. GA (2000) applied this operation to delete small buildings, short roads, and small villages in a generalisation of 1:250,000 scale data to generate the 1:1,000,000 *Global Map*. The Global Map Australia 1:1,000,000 (2001) is a digital spatial dataset covering the Australian continent and island territories at 1:1,000,000 scale which consists of eight spatial layers in vector form (administrative boundaries, drainage, transportation and population centres) as well as incorporating raster images (elevation, vegetation, land cover and land use). This is created from GA's 1:250,000 national topographic data. The United States Geological Survey (USGS) provided the raster images.
- *Simplification:* selected roads can be simplified to reduce the details. This research used the *Pointremove* and *Bendsimplify* operators to remove vertices from the selected lines and to simplify extraneous bends of roads.
- *Typification:* this function is not directly available in ArcToolbox Generalise™, but a manual cartographic editing approach was used. The basic aim is to reduce the density of features and simplify the distribution and the pattern of the network. The result should be a connected and less congested network that represents a similar pattern at the smaller scale.
- *Symbolisation:* the systematic process of creating graphic marks, which represents the objects and features to be mapped. For example, roads are symbolised according to their class.

Generalisation operators are applied in order to derive small scale maps in this investigation. To commence generalisation, necessary objects in the map should be chosen. The next step is to apply simplification, which may include smoothing of curves and straightening paths. The objective of simplification is to increase the legibility of the map by removing unnecessary details of the road network. In this simplification process the road line features are simplified to remove extra vertices in the road segments. ArcToolbox Generalise™ (*Pointremove*) is used to simplify the lines for this process, and is based on the DP algorithm that removes vertices from the road linear features with a user-defined tolerance. To simplify the line bends, the bend operator is used to omit extra bends. Tolerances for each scale of the roads were based upon comparison of the generalisation results with cartographic products of the same scale.

2.5 Road Generalisation Assessment Methods

In this study an enriched database (1:250,000 national topographic data) was used as the main reference dataset. In total 1,202 road segments were present in the 1:250,000 national topographic data database over the study area. Changes in the representation of road features at 1:250,000, 1:500,000 and 1:1,000,000 scales due to generalisation were assessed using the *Radical Law* (e.g. Pfer and Pillewiser, 1966 and Muller, 1995) and an interactive accuracy evaluation method (Figures 7 – 12). The road type was also considered a major index of importance. During the transformation of road representation from the 1:250,000 to 1:1,000,000 scales some of the less important road segments were deleted. At each scale the number of road segments classified by road types was counted and the results were graphically illustrated (Section 4).

3. Results and Discussion

All vehicle tracks disappeared at the 1:1,000,000 scale. In addition there was a 61% reduction of minor roads at 1:500,000 and a further 45% reduction of minor roads at 1:1,000,000 (Figure 7).

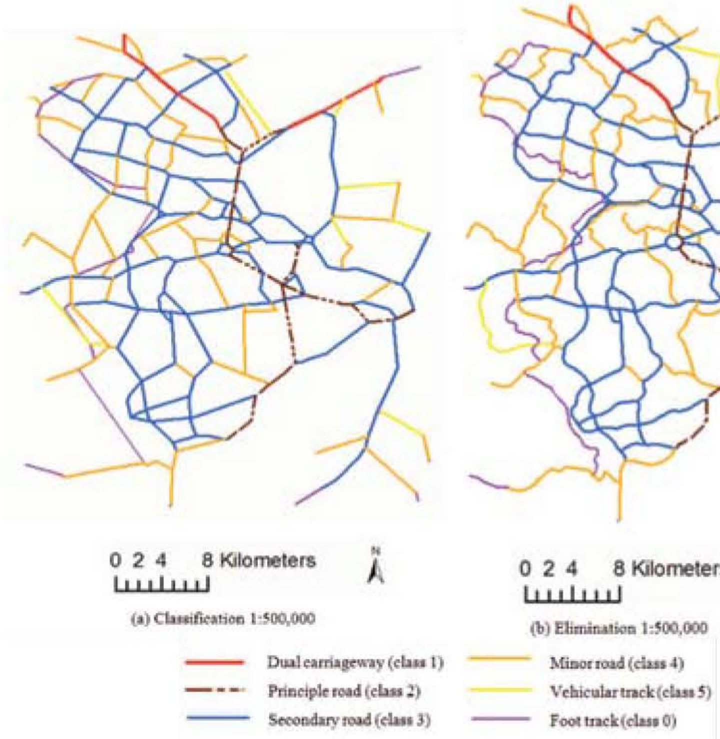

Figure 4: Road generalisation outputs (stage 1); deriving 1:500,000 maps from the source data at 1:250,000; (a) Classification and (b) Elimination

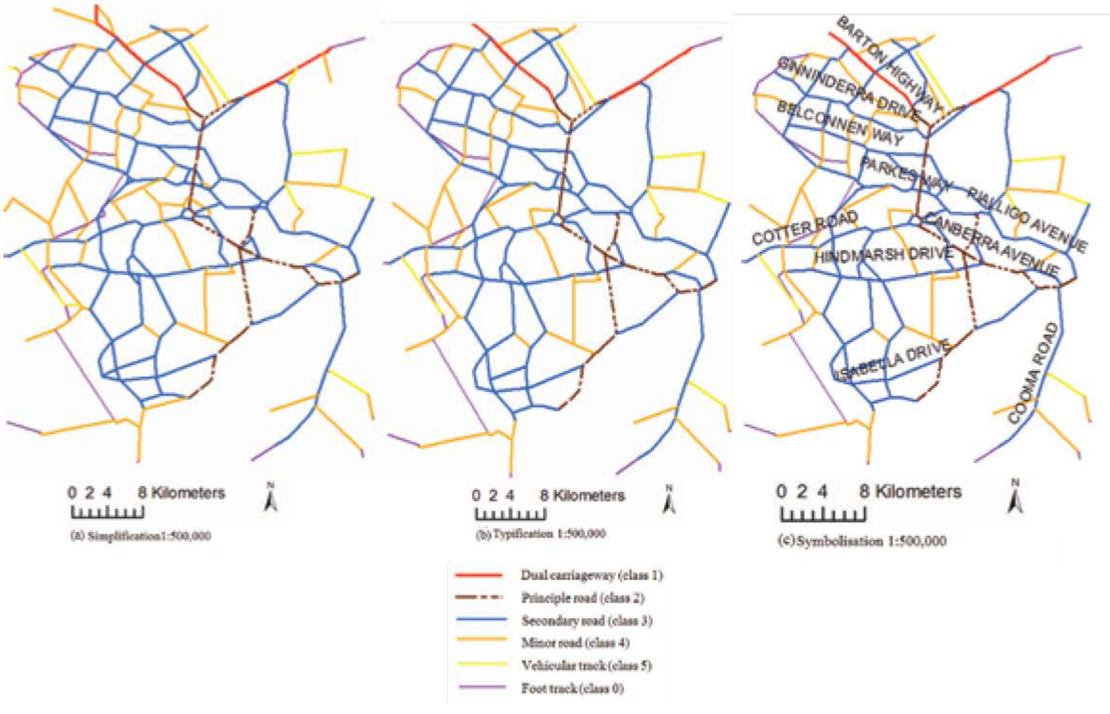

Figure 5: Road generalisation outputs (stage 2); deriving 1:500,000 maps from the source data at 1:250,000; (a) Simplification, (b) Typification and (c) Symbolisation

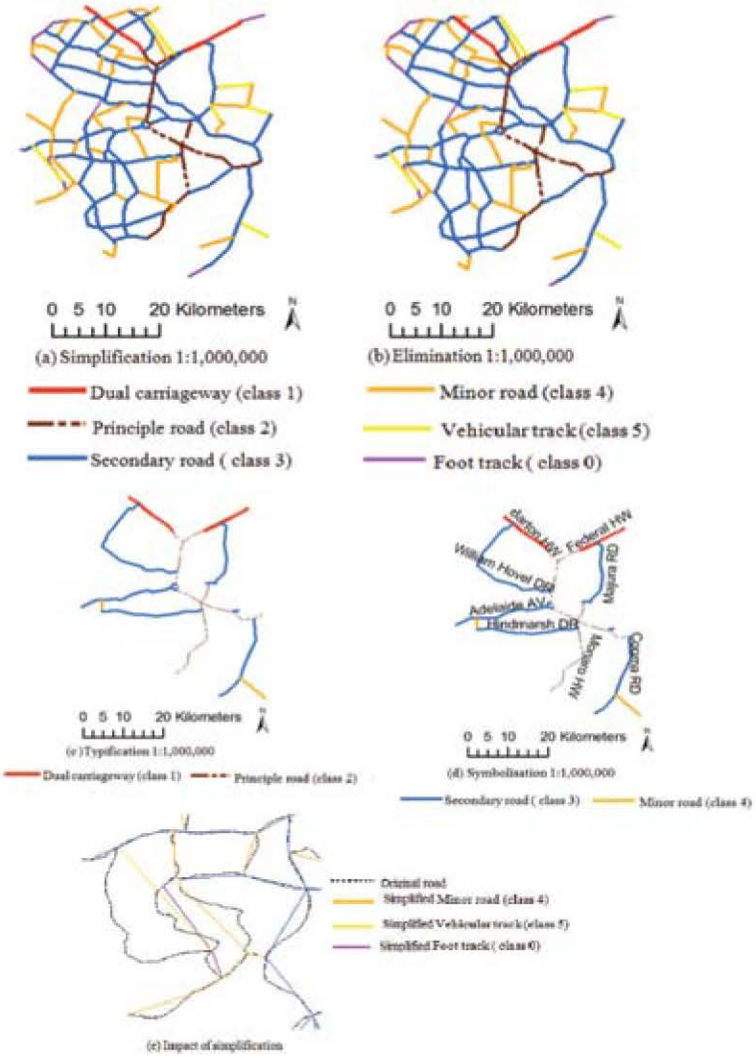

Figure 6: Road generalisation outputs; deriving 1:1,000,000 maps from newly derived 1:500,000 scale data; (a) Simplification, (b) Elimination, (c) Typification, (d) Symbolisation and (e) a zoom area from middle-left of (a) showing the impact of simplification on a portion of the road network

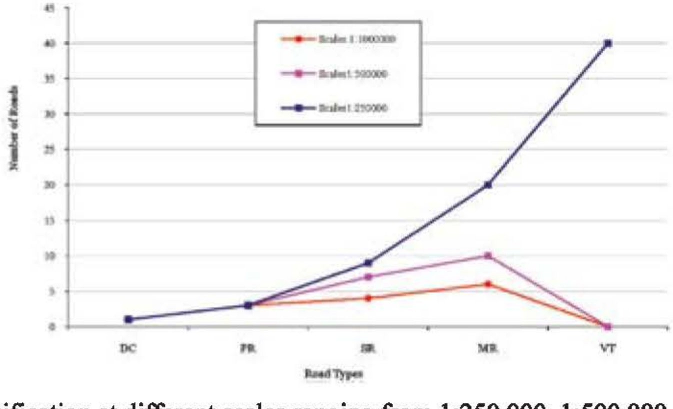

Figure 7 Road classification at different scales ranging from 1:250,000, 1:500,000 and 1:1,000,000 and number of road segments in different classes

For example at 1:500,000 scale, there are 10 Minor Road (MR), 8 Secondary Road (SR), 3 Principal Road (PR), and 1 Dual Carriageway, but no Vehicle Tracks (VT). The results of map derivation were considered using two methods: the *Radical Law* and an interactive accuracy evaluation method. The *Radical Law* determines the retained number of objects for a given scale change and the number of objects of the source map (Nakos, 1999). The original 1:250,000 national topographic data roads have 1,202 segments and the derived roads have 856 segments before the line simplification operation. The equation (1) computes the number of roads in the map after the line simplification operation:

$$n_T \approx n_S \sqrt{S_S / S_T}$$

Equation 1

where n_S is the number of objects in source dataset; S_S is source scale; S_T is scale after transformation; and n_T is number of objects in the dataset after transformation.

- Number of lines in 1:500,000 scale roads map => $n_T \approx 1202\sqrt{250/500} = 849.9$
- Number of lines in 1:1,000,000 scale roads map => $n_T \approx 856\sqrt{500/1000} = 605.3$

The numerical value 849.9 represents the predicted number of segments after generating the 1:500,000 scale road map, and the value 605.3 predicts the number of segments after generating the 1:1,000,000 scale map. These results have also been analysed visually since evaluation is a critical element of semi-automated road generalisation. Cartographic knowledge, such as the presentation of features on the map and the contextual information of features, has been combined with generalisation operators to derive multiple scale maps. This includes cartographic knowledge and information about roads and surrounding areas. In order to perform an intelligent and efficient generalisation framework, the processes of combination, omission and simplification should address a linkage between having a good knowledge of geography and a sense of proportion. Design and representation of map aspects are considered in relation to text portrayal. Map symbols must exhibit a number of key characteristics, including a) designative, reflecting the features portrayed by other map symbols; b) analytical, linking features on the map with their attributes and analysing relationships amongst features; c) positional, describing or confirming the location of features; and d) informative, giving a description of the nature of the source data (Fairbairn, 1993). Thematic properties such as length, width, and connectivity were also used. Although these measures are not directly available in the Generalise™ tool, the roads database specification, local and cognitive knowledge, and other ancillary information were employed to characterise these thematic properties. The information was used to determine which road should be deleted or merged in the generalisation process in order to retain a certain percentage of roads at a smaller scale. In this regard, Choi and Li (2001) believe that incorporation of thematic information can control geometric operations (e.g. simplification) and topological information (e.g. connectivity). Thematic information can be used to validate the geometric operation and to control the quality of map generalisation. For example, the use of Elimination involves removal of very short branches and unimportant roads, which is mainly based on cognitive knowledge. The generalisation processes such as Elimination and Typification results in a reduction of the complexity and the density of roads (Figures 8 – 10 examples of the effects of various tolerance levels). The results were evaluated through cartographic visual interpretation The Generalise™ tools (Bendsimplify and Pointremove) did not smooth some portions of roads satisfactorily. Previous studies (e.g. Visvalingam and Williamson, 1995) confirmed that the DP algorithm, by applying very low tolerances, resulted in removal of an appropriate number of vertices. It was found that generalisation of roads requires a very large volume of processing to transform data from 1:250,000 scale to 1:500,000 and 1:1,000,000 scale maps. The DP algorithm eliminates shorter branches and results in simplification of the road network, but also improves the preservation of the characteristics of the structure. The impact of generalisation tolerance can be seen in various portions of the roads (e.g. see circle around Antill Street) database; road network is overlaid on merged Landsat 7 ETM+ Panchromatic (15m) and Multispectral (30m) imagery (composite bands 1, 2, 3). Figure 9 and Figure 10 compare the road network before simplification (blue lines) and after simplification (those lines shown in colours other than blue) and show the results of derived maps at 1:500,000 and 1:1,000,000 scales. Then the road network was symbolised.

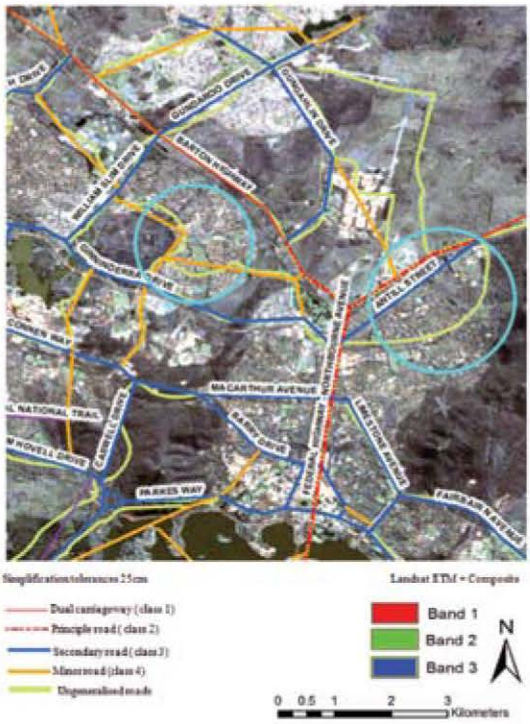

Figure 8: The impact of generalisation using 25m tolerance

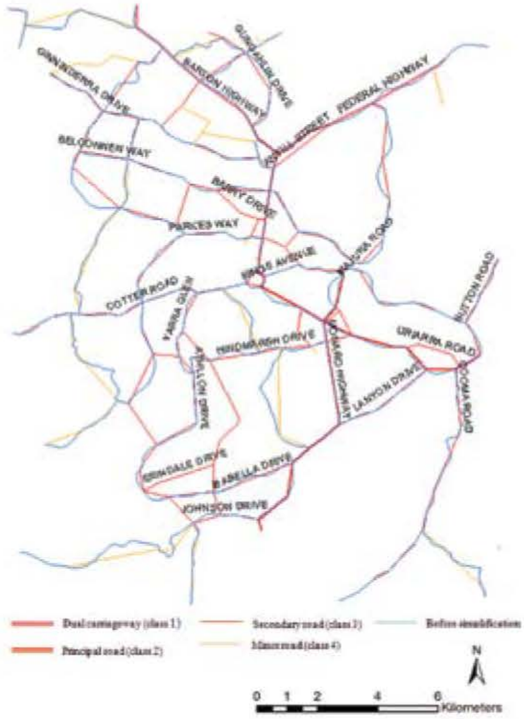

Figure 9: The impact of generalisation tolerance of 45m

The reduction of the 1:500,000 scale map is less readable than the 1:1,000,000 map. The objects are too small and too close to be readable. By comparing the two, some objects have been enlarged. At a higher level, some objects have been removed, but use of space has been preserved at the 1:500,000 scale. Qualitative or quantitative assessment of generalisation results is necessary. In this study the derived maps were compared with existing maps. This involved experienced cartographers carrying out a visual assessment/inspection of the results. From the cartographic intuition and perspective, through consultation with experienced GIS operators/cartographers, evaluation by independent assessment helped to define the best levels of tolerance ranging from 10m to 70m that produced satisfactory simplification outputs. The tolerance value (positive number greater than zero) determines the degree of line simplification. It reflects the distance that two points in a road segment can be apart and still be considered the same to produce generalised data. It requires setting the tolerance equal to or greater than the threshold of separation (the minimum allowable spacing between graphic elements). The threshold value relates to the map coordinate system.

The tolerance value is expressed as a number of metres. Various tolerance parameters (values) were tested in the course of road network generalisation. Based on several tests, it became possible to optimise the process of map derivation by semi-automated cartographic and database generalisation by means of simplification and thresholding to derive small scale maps from a master database. Two different values based on the introductory experiment demonstrated that threshold parameter values of 25m and 35m were the most useful tolerance levels to be applied for the 1:500,000 and 1:1,000,000 maps, respectively. Similar tolerance levels (20m using Douglas-Peucker and 40m with Reumann-Witkam algorithms) have been reported by Nakos (1999) to generalise a coastline database of Greece. The Reumann-Witkam simplification algorithm uses two parallel lines to describe an area of interest (AOI) after calculating the original slope of the AOI, the line is processed successively until one of the edges of the search corridor intersects the line. Figures 9 and 10 illustrate the effects of thresholding versus original data where the impact of the *Bendsimplify* operations with a tolerance of 25m and 35m to produce spatial data products at the 1:500,000 and 1:1,000,000 scale respectively.

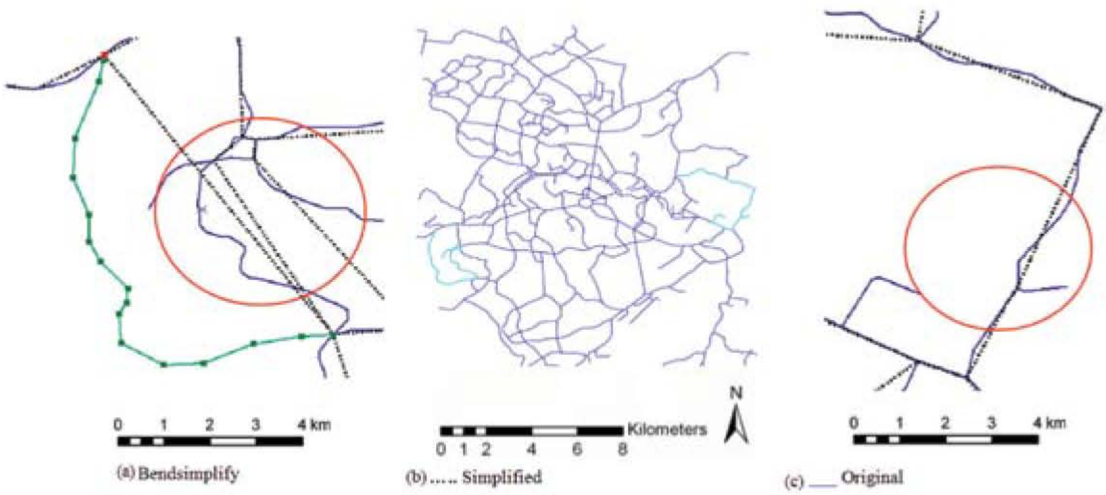

Figure 11: The impact of simplification using Bendsimplify as seen in a portion of the road database

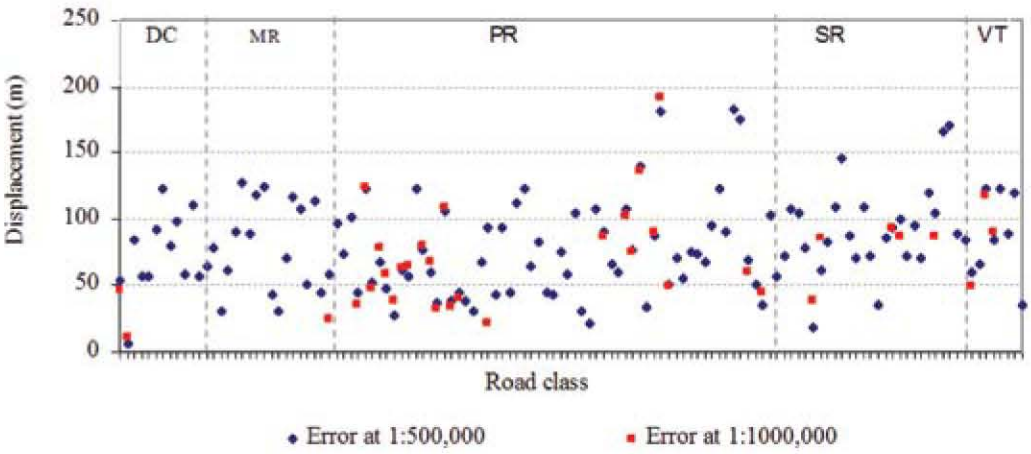

Figure 12: Assessment of the average shape changes caused by the simplification operation in relation to the selected control points, showing the displacement values of the coordinate differences on the derived roads map. Blue legend (dot) shows the displacement errors for derived 1:500,000 product from the original data (TOPO250K), whereas the red legend (dot) shows the displacement errors for derived 1:1000,000 product from the original data (1:500,000 scale)

In Figure 11 (c), vertices in the original line are shown in green squares. Clearly, applying a higher threshold (threshold 35m) levels resulted in removal of extraneous vertices in the road segments that are shown here. It required manual editing to correct the over simplification, whereas in Figure 11 (a), the simplification (threshold 25m) led to satisfactory results from a cartographic perspective. Figure 11(a) shows an over simplification output. In the graph (Figure 11(a)), vertices in the original line are shown in green squares. Clearly, applying higher threshold (threshold 35m) levels resulted in removal of extraneous vertices in the road segments that are shown here. It required manual editing to correct the over simplification, whereas in Figure 11(c), the simplification (threshold 25m) led to satisfactory results from a cartographic perspective. Positional accuracy measures are used to evaluate the quality of road generalisation. Accuracy of simplified and generalised outputs is determined by comparing the positions of 50 well defined ground control points. The point locations are assessed on the generalised maps (1:500,000 and 1:1,000,000) and corresponding positions from the published 1:250,000 national topographic maps. The source map '1:250,000 digital topographic data' has a basic horizontal accuracy of approximately ±120 metres (GA, 2009). Cartographic generalisation of line and area features introduces errors into the derived map. Generalised output maps have been checked from multiple sources including Landsat ETM+ satellite imagery and published 1:250,000 digital topographic data. Positional accuracies for generalised 1:500,000 and 1:1,000,000 maps are approximately 160m and 190m respectively; thus measured errors are within mapping standards. In Figure 12 the results achieved for 1:500,000 and 1:1,000,000 scales are shown. Despite the fact that results from the Generalise™'s DP and *Bendsimplify* are satisfactory, they do not support dynamic generalisation. Therefore there is an opportunity to evaluate other generalisation systems, such as Intergraph's DynaGen™ and Laser-Scan's Clarity™ software, to derive multiscale spatial data products in future investigations. It is worth noting that the importance of expert system application in generalisation operations and cartographers' experience has often been ignored. Previous research of the lead author has highlighted the importance of integrating cartographers' knowledge with the generalisation system to facilitate the development of a powerful, flexible and robust expert system capable for semi-automated road network generalisation (Kazemi *et al.* 2005). A comprehensive evaluation of generalisation systems and their performance is essential to marry the cartographic knowledge of experts and bring this into a generalisation framework.

4. Summary and Conclusions

This paper presented a generalisation methodology to derive multiscale spatial data through an evaluation of ArcToolbox Generalise™ software. It focused on integration and utilisation of generalisation operators in order to generalise a road network database and produce small scale maps at 1:500,000 and 1:12,000,000 from 1:250,000 national topographic data. The derived maps are satisfactory when compared with the existing small scale road maps such as the Global Map at a scale of 1:1,000,000. It is suggested that a comprehensive evaluation of other generalisation systems and their performance is essential in order to integrate the cartographic knowledge from experts into a generalisation framework. Generalisation operators in the Generalise™ application were tested to generalise roads by employing the above methodology for derivative mapping. The method was empirically tested with a reference data set at 1:500,000 and 1:1,000,000 scales. According to visual interpretation and quantitative analysis, the results show that the derived maps are satisfactory when compared with the existing small scale road maps such as the *Global Map* at a scale of 1:1,000,000. The derived 1:1,000,000 map was then compared to the 1:1,000,000 *Global Map*. There were no existing 1:500,000 reference data to compare with the derived 1:500,000 map. As the methodology is only tested on roads it is worthwhile to extend it to other complex cartographic datasets such as drainage networks, power lines and sewerage networks. Additionally, various kinds of linear, areal and point cartographic entities (e.g. coastlines, rivers, vegetation boundaries, administration boundaries, land cover, localities and towers) should also be studied. To evaluate and validate existing generalisation tools the author's research focus was on the development of a detailed generalisation framework to derive multiscale spatial data. The work was concerned with the integration and utilisation of generalisation operators as well as cartographer's intuition/skills using the Generalise tool (and possibly DynaGen) software in order to achieve acceptable results. Previous studies (e.g. Mroz et al., 1996, Nakos, 1999, Lee, 2004 and Kazemi et al., 2009) have also achieved good results in automatic generalisation; however significant manual work was required since the *Pointremove* generalisation algorithm and the *Bendsimplify* algorithm do not support dynamic generalisation. Other generalisation systems such as Intergraph's DynaGen™ and Laser-Scan's Clarity™ support such generalisation and enable users to derive a multi scale database from a master database (Watson and Smith, 2004). Therefore, it is suggested that the proposed methodology be enhanced by testing and incorporating tools.

Acknowledgments

This study was sponsored by the Faculty of Engineering's PhD Scholarship and Supplementary Engineering Award at the University of New South Wales (UNSW). The lead author is grateful to several reviewers, in particular Professor Chris Rizos and Associate Professor Samsung Lim of Surveying and Geospatial Engineering of UNSW. They participated in various discussions on the topic of generalisation and provided constructive suggestions toward the lead author's doctoral research. The author also would like to sincerely thank the anonymous reviewers for their insightful comments and constructive suggestions of the original manuscript. GIS software applications were made available to this research as student licence and Geoscience Australia has provided various geospatial datasets and remote sensing imagery.

Special thanks are also extended to the University of California Berkeley for hosting the lead author as Visiting Scholar from 2006 to 2007. Professor Nitin Kumar Tripathi, Editor-in-Chief of International Journal of Geoinformatic (IJG) and Ms. Nitiporn Saardmoung, Administrative Secretary of IJG have kindly facilitated print of the paper into this issue of the Journal. The paper is extracted and modified from previous publications. Finally, our sincere thanks goes to Distinguished Professor Arthur Georges, Chief Scientist of Institute for Applied Ecology (IAE) at the University of Canberra for generous financial support toward this publication.

References

Arnold, L. M. and Wright, G.L., 2005, Analysing Product-Specific Behaviour to Support Process Dependent Updates in a Dynamic Spatial Updating Model. *Transactions in GIS,* Vol. 9, No. 3, 397-419.

Choi, Y. and Li, Z., 2001, Correlation of Generalisation Effects with Thematic Attributes of Cartographic Objects. Department of Land Surveying and Geo-Informatics, the Hong Kong Polytechnic University, Hong Kong, 1-26. Accessed online 1 June 2007: http://www.geo.unizh.ch/ICA/docs/Beijing-2001/papers/choi_li_v1.pdf

Cromley, R. C., 1992, Digital Cartography. Prentice Hall, Englewood Cliffs, New Jersey, USA.

Dutton, D., 1999, Scale, Sinuosity and Point Selection in Digital Line Generalisation. *Cartography and Geographic Information Science,* Vol. 26, No. 1,33-53.

Fairbairn, D. J., 1993, On the Nature of Cartographic Text. *Cartographic Journal.* Vol. 30, No. 2, 104-111.

Forghani, A., 2000, Semi-Automatic Detection and Enhancement of Linear Features to Update GIS Files. *Proceedings of the 6th ISPRS Congress, 16-23 July 2000, Amsterdam, The Netherlands,* CD-ROM procs, 10.

Geoscience Australia (GA), 1999, Topographic Data and Map Specifications (CD-ROM). Technical Specifications, Version 3.4, National Mapping Division, *Geoscience Australia.* Accessed online 5 December 2004: http://www.ga.gov.au/map-specs/250k100k/section3.

Geoscience Australia (GA), 2001. Helping the Global Community to Resolve Global Problems through Spatial Information. *Proceedings of the AURISA 2001 - the 29th Annual Conference of AURISA, 19-23 November 2001, Melbourne, Australia,* CD-ROM procs, 12.

Hardy, P., Briat, M. O., Eicher, C. and Kressmann, T., 2004, Database-Driven Cartography from a Digital Landscape Model, with Multiple Representations and Human Overrides. *Proceedings of the ICA Workshop on Generalisation and Multiple Representation, 20-21 August 2004*, Leicester, UK, CD-ROM procs, 12.

Jenks, G. F., 1989, Geographic Logic in Line Generalisation, *Cartographica,* Vol. 26, No. I, 27-42.

Kazemi, S., Lim, S., and Paik, H., 2009, Generalisation Expert System (GES): A Knowledge-Based Approach for Generalisation of Line and Polyline Spatial Datasets. *Proceedings of the Surveying & Spatial Sciences Institute Biennial International Conference, 28 September - 2 October 2009,* Adelaide, Australia, 717-729.

Kazemi, S. and Lim, S., 2007, Deriving Multiscale GEODATA from TOPO-250K Road Network Data. *Journal of Spatial Science,* Vol. 52, No. 1, 171-182.

Kazemi, S., Lim, S. and Ge, L., 2007, An International Survey Research: Cartographic Generalisation Practices in Mapping Agencies. *Proceedings of the SSC 2007 Spatial Intelligence, Innovation and Praxis: The National Biennial Conference of the Spatial Sciences Institute,* 4-8 May 2007, Hobart, Australia, 537-556.

Kazemi, S., Lim, S. and Ge, L., 2005, Integration of Cartographic Knowledge with Generalisation Algorithms. *Proceedings of the IEEE Geoscience & Remote Sensing Symposium*, 25-29 July 2005, Seoul, South Korea, 3502–3505.

Kazemi, S. and Lim, S., 2005, Generalisation of Road Network using Intergraph DynaGen System. *Proceedings of the SSC 2005 Spatial Intelligence, Innovation and Praxis: The National Biennial Conference of the Spatial Sciences Institute,* 12-16 September 2005, Melbourne, Australia, 1285-1296.

Kazemi, S., Lim, S. and Rizos, C., 2004a, A Review of Map and Spatial Database Generalisation for Developing a Generalisation Framework. *Proceedings of the XIX Congress of 2004 the International Society for Photogrammetry and Remote Sensing,* 12-23 July 2004, Istanbul, Turkey, 1221-1226.

Kazemi, S., Lim, S. and Ge, L., 2004b, Is Automated Generalisation There Yet? *Proceedings of the 12th Australasian Remote Sensing and Photogrammetry Association Conference, 18-22 October 2004,* Fremantle, Australia, CD-ROM procs, 12.

Kreveld, M. J. and Peschier, J., 1998, On the Automated Generalisation of Road Network Maps. *Proceedings of the 3rd International Conference on GeoComputation, 17 - 19 September 1998, University of Bristol, United Kingdom.* Accessed online 12 July 2006: *http://www.geocomputation.org/1998/21/gc_21.htm*

Lee, D., 2003, Generalisation within a Geoprocessing Framework. *Proceedings of the GEOPRO 2003 Workshop,* 4-5 November 2003, Mexico City, Mexico, 82-91.

Lee, D., 2004. Geographic and Cartographic Contexts in Generalisation. *Proceedings of the ICA Workshop on Generalisation and Multiple Representation,* 20-21 August 2004, Leicester, UK. *Accessed online 1 March 2006:* http://ica.ign.fr/Leicester/paper/Lee-v2-ICAWorkshop.pdf

Lee, D. and Hardy, P., 2006, Design and Experience of Generalization Tools. *Proceedings of AutoCarto 2006, June 25* 2006, Vancouver, Canada, CD-ROM procs, 11.

Mackaness, W. A., Edwardes, A. and Urwin, T., 1998, Self Evaluating Generalisation Algorithms for the Derivation of Multiscale Products from UKBORDERS Data of the Census. *Proceedings of the 3rd International Conference on GeoComputation, 17 - 19 September 1998, Bristol, UK, GeoComputation CD-ROM. ISBN 0-9533477-0-2.*

Moreno, M., Torres, M., Levachkine, S. and Fajard, I., 2002, The Automatic Generalisation of the Multiscale Geographic Information. *Proceedings of the 5th AGILE Conference on Geographic Information Science (AGILE 2002), 25-27 April 2002, Palma de Mallorca, Spain.* Accessed online March 2005: http://gis.esri.com/library/userconf/proc02/pap0173/p0173.htm

Mroz, M., Berger, K., Becherer, R. and Packman, J., 1996, The Defence Mapping Agency (DMA) World Vector Shoreline Plus (WVSPLUS) Revision and Conversion Project. Accessed online April 2005: http://gis.esri.com/library/userconf/proc96/TO250/PAP237/P237.htm

Muller, J. C., Weibel, R., Lagrange, J. P., and Salge, F., 1995, Generalisation, State of the Art and Issues. In Muller, J.c., Lagrange, J.P., Weibel, R. (editors) GIS and Generalisation, Methodology and Practice, Taylor and Francis, Bristol, Uk, 91-105.

Muller, J. C., 1995, GIS and Generalisation: Methodology and Practice. Editor(s): J-P Lagrange; R Weibel; J. Muller, University of Ruhr, GISDATA Series. 20 April 1995, Bochum, Germany, ISBN: 9780748403196.

Nakos, B., 1999, Comparison of Manual Versus Digital Line Generalisation. *Proceedings of the Workshop on Generalisation, 12-14 August 1999, Ottawa, Canada,* CD-ROM procs, 10.

Pfer, T. and Pillewizer, W., 1966, The Principles of Selection. *The Cartographic Journal,* Vol. 3, No. 1, 10-16.

Rusak, E. and Castner, H.W., 1990, Horton's ordering scheme and the generalisation of river networks. *Cartographic Journal,* Vol. 27, 104-110.

Sarjakoski, T., Lehto, L., Sester, M., Illert, A., Nissen, F., Rystedt, B. and Ruotsalainen, R. 2002, Geospatial Info-Mobility Services: A Challenge for National Mapping Agencies. *Proceedings of the International Symposium on Geospatial Theory, Processing and Applications,* 9-12 July 2002, Ottawa, Canada, CD-ROM procs, 5.

Taylor, D. R. F., 2005, Cybercartography: Theory and Practice, Taylor, Fraser, Elsevier, Amsterdam, The Netherlands.

Visvalingam, M. and Williamson, P. J., 1995, Simplification and Generalisation of Large Scale Data for Roads - A comparison of two filtering algorithms. *Cartography and Geographic Information Systems, Vol. 22, No. 4,264-275.*

Visvalingam, M. and Herbert, S., 1999, A Computer Science Perspective on the Bend Simplification Algorithm. *Cartography and Geographic Information Science,* Vol. 26, 253-270.

Wang, Z. and Muller, J.C., 1998, Line Generalisation Based on Analysis of Shape. *Cartography and Geographic Information Systems,* Vol. 25, 3-15.

Watson, P. and Smith, V., 2004, Interoperability of Agent-Based Generalisation with Open, Geospatial Clients. *Proceedings of the ICA Workshop on Generalisation and Multiple Representation,* 20-21 August 2004, Leicester, UK, CD-ROM procs, 5.