# CARTOGRAPHIC LINE SIMPLIFICATION: A FORMAL ROLE WITHIN DIGITAL CARTOGRAPHIC PRODUCTION

Thesis submitted for the degree of Doctor of Philosophy
at the University of Leicester

*by*

**Ali M. A. Al-Ghamdi**
(B.A.; Diploma; M.A. (*honour*))

**Department of Geography**
**University of Leicester**

**March 1998**

## ABSTRACT

This study examines the role of cartographic line simplification in traditional map production, and explores how that role can be transformed to digital map production. Whilst previous studies theoretically recognised that simplification is only a sub-process within the general context of generalisation, they, in practical terms, have inappropriately utilised digital simplification algorithms. For example, they confuse the role of simplification with that of generalisation. Consequently, there has been little, if any, progress in the field of formalising the process of cartographic line simplification, so as to be able to perform a truly digital cartographic simplification consistent with the requirements of cartographic generalisation. Recently, there have been calls to study cartographic processes before contriving new algorithms. This study is, therefore, a response of such calls, and proposes a novel scheme by which the process of line simplification is re-examined in both the traditional and digital realms. The proposed scheme consists of three consecutive logical stages. The first stage is concerned with examination of the definition of the traditional line simplification. The second stage is concerned with evaluation of a typical widely-used digital simplification algorithm, in this case, the Douglas-Poiker algorithm, according to its underlying design specifications. The third stage involves searching for cartographic quality in the output of the algorithm, assisted by post processing by a Cubic Spline smoothing routine.

Overall, a formulation of the cartographic role for the two simplification algorithms in digital cartographic generalisation is presented. The formulation can serve as a practical solution for an objective use of the two algorithms within digital mapping systems during digital cartographic productions. The study also shows that the process of simplification is a complex process, which is like any context-dependent generalisation process. Further effort will be required, however, to achieve a sound exhaustive understanding of the concept and practice of line simplification and hence its formulation. Furthermore, the optimal goal of this work is to provide an operational model for cartographic line simplification, and present a feasible methodology with which researchers can examine other generalisation processes.



*Key words: cartographic line simplification, cartographic line generalisation, digital cartographic line simplification, digital cartographic line generalisation.*

## ACKNOWLEDGEMENT

I would like first to thank my supervisor Dr. Peter Fisher for his guidance, and encouragement during my entire study, which will last long after the finer details of this thesis. I also thank Dr. Alan Strachan for his discussion and comments on the first draft of this work. I also thank the following people in the Department of Geography at Leicester University for helping me either directly or indirectly: Prof. Andrew Millington; David Orme; Bill Hicken; and the staff in the secretary office. I am also grateful to Dr. Arlond Mathew, and Ms Suzan Hanrahan in the Computer Centre at Leicester University. I am indebted to my government of Saudi Arabia for granting me a scholarship for pursuing my graduate studies. I am grateful to the educational attaché and his team at the Saudi Cultural Bureau in London for their constant support.

I am extremely grateful to my family in Riyadh and Leicester for their unabated support, patience, and encouragement. I am also touched by the encouragement and praying of relatives in my community in Saudi Arabia.
I will not forget the support of my friends in Saudi Arabia, Libya, Oman, and Switzerland.

*To my brother:*

***Said Ma'adhah Al-Ghamdi***

LIST OF CONTENTS

LIST OF FIGURES

CHAPTER TWO



CHAPTER THREE



CHAPTER FOUR

CHAPTER FIVE

CHAPTER SIX

LIST OF TABLES

CHAPTER FOUR



CHAPTER FIVE

CHAPTER SIX

# CHAPTER ONE: INTRODUCTION

## 1.1 Background:

Cartographic line simplification is essential within the process of traditional cartographic generalisation. It has attracted considerable attention in the cartographic community in the movement towards automation. Map generalisation is a complex process during cartographic production, and its execution tends to be left to the subjective decisions of the experienced cartographer. It is made up of a set of ill-defined tasks which has made it difficult for cartographers and non-cartographers to develop sound objective methods in the realm of digital mapping.

Simplification algorithms have been proposed since the late 1960s. Yet, users of digital mapping systems questioned the products of those algorithms in terms of their cartographic quality. Cartographic research in generalisation in the past three decades has focused on algorithm development and assessment, error analysis, formal description of map feature geometry, and the development of logical models. Buttenfield and McMaster (1991) provide a chronological summary of the trends of research in the field of automated generalisation. They identified three phases of research. The first, which dated from approximately 1960 to 1975, focused upon algorithm development with particular emphasis on algorithms for line simplification. In the second , the late 1970's and the 1980's, assessment of algorithm efficiency became an increasing concern. In this period, most research viewed geographical phenomena in isolation, ignoring the need to integrate generalisation procedures. In the third phase, the 1990's concern about map generalisation continues to develop with primary concern for formalising cartographic knowledge through development of comprehensive models, and application of expert systems and other knowledge based techniques.

Introducing formal or objective methods for cartographic generalisation is one thing, but providing a reliable cartographic quality is another. And in order to provide successful cartographic results in this context, a precise and comprehensive understanding of the traditional cartographic knowledge itself is a prerequisite. Shea (1991) stresses that in order "to intelligently integrate that knowledge in a digital environment, algorithm designs, implementation strategies, and control techniques for generalisation operations must consider the geographical implications of generalisation decisions" (p. 3).

As Shea (1991) indicated, the automated generalisation of point and area features has been addressed by several authors, but has resulted in few significant achievements (Monmonier 1983; Chrisman 1983; Lichtner 1978; Topfer and Pillwizer 1966). Many researchers have proposed theoretical models for digital generalisation (McMaster and Shea 1988; Brassel and Weibel 1988). Such models provide a foundation which furthers an understanding of the complexity of the generalisation process. Some have argued that formalising the subjective elements in generalisation is difficult and tends to sacrifice unique and creative aspects of map making (Beard, 1991a), but many agree that the formulation of rules and their implementation is one of the most difficult challenges in the cartographic research agenda now and in the future.

Many algorithms developed to date in this field are generally concerned with vector-mode line generalisation. Joao (1994) identifies two main reasons for this. Firstly, as Muller (1991b) indicates, approximately eighty percent of all features on a typical medium-scale topographic map consist of lines. Secondly, as Weibel (1986) suggests, automated generalisations of line features are problems of lower complexity compared to problems that involve, for example, area features. The present thesis, argues that development and progress in digital mapping have lacked effective contributions from cartographers.

In this brief outlook of the formalisation of cartographic knowledge, it is worth mentioning a significant achievement, which is an interactive generalisation program, called MGE Map Generalizer which is incorporated within the Intergraph Geographic Information System (GIS) software. Other GIS programs usually support minimal interactive and non-interactive generalisation tasks such as data reduction, smoothing, selection, amalgamation, and displacement. However, the success, if any, is primarily dependent upon the experience of the operator.

## 1.2 The problem:

There are various impediments to formalising knowledge in the realm of cartographic generalisation. One of the most effective discussions is given by Muller *et al.*, (1995b). They summarise critical hypotheses (p. 13), and they referred to the lack of understanding of the generalisation process itself, and what can be achieved in a digital context. They highlight the problem of evaluation by researchers who confuse the objectives and characteristics of model-oriented versus cartographic generalisation. A classical example is that many researchers evaluate data reduction algorithms for which either design specifications have not been stated by the authors of those algorithms, or where the algorithms are intended for generalisation. Muller *et al.*, (1995b) highlight a very significant point, that "... there is too little interaction between the computer experts and expert cartographers. Those who are working on the automation of generalisation do not know how to generalise, and those who would know how to generalise are not being asked (at least the right questions)" (p. 13). They also point out other problems, but those can be interpreted as a consequence of this root problem .

Recently, there has been a step towards gathering information and gaining an insight on the generalisation processes. A method of 'reverse engineering' is currently being examined at NCGIA Buffalo, and the University of Zurich. This approach involves taking existing map series at different scales and systematically analysing the depiction of features in order to gather information about generalisation (Muller *et al.*, 1995b).

The present thesis argues that in order to provide sound production and evaluation guidelines to users of digital algorithms for cartographic generalisation, a thorough understanding of the traditional cartographic processes should be the first step. This not only yields rules that can be formally executed, but also provides a framework within which the existing theoretical and practical methods can be appropriately re-evaluated. The merits of this approach lie in attempting to raise issues highlighting the two essential characteristics of cartographic generalisation which are *comprehensiveness* and *variability*. It is shown that previous studies have inappropriately focused on tests of limited types of data which have resulted in unwarranted conclusions. For example, a cartographic simplification that might seem appropriate for one particular representation may be unsuitable for another. It may, therefore, be inappropriate to apply a general rule. Cartographic generalisation involves several processes; e.g., selection, simplification, displacement, exaggeration and typification (Robinson *et al.*, 1995). Simplification is, therefore, a sub-process that can be applied to point, line and area features. Cartographic line simplification is a process involves preservation of line feature characters whilst removing redundant details, determined by the cartographer (Robinson *et al.*, 1995). The process applies to the details as well as the number of the lines. As implied, simplification incurs perceptible changes in the processed lines. Since line simplification is the subject of this thesis, the premise here is that line simplification is a sub-process within cartographic generalisation; complex, and context-dependent process. It is thus important to realise that digital formulation of this process should actually embody within its mechanism these cartographic characteristics. Cartographers should provide a major contribution to the process of understanding cartographic generalisation so that an acceptable operational formalisation of knowledge can result.

## 1.3 Objectives of the study:

The study aims to examine cartographic line simplification within the experimental framework outlined above. That is, studying line simplification within traditional and digital realms, and assessing the extent to which digital simplification techniques could have adhered to cartographic principles. A variety of specific objectives can be itemised:

1- To look in a greater depth at cartographic line simplification in a scale-dependent context, which provides insight into how this process is conceived and executed in both traditional and digital mapping (vector mode) contexts;
2- To re-evaluate the widely used point data reduction algorithm, known as the Douglas-Poiker algorithm, in a scale-dependent context and according to the design specifications of the algorithm as stated by its authors;
3- To identify cartographic quality in digital algorithms during a scale-dependent cartographic simplification, by combining different digital methods.
4- To examine the relationship between graphic reduction and spatial resolution, and to assess the effects of that relationship on cartographic simplification and how this relationship can be quantified.
5- To explore the role of cartographic simplification of line features in the context of cartographic production of paper maps.

Unlike previous studies, the ultimate goal of this study is to present a detailed account of cartographic line simplification. Specifically, it aims to provide guidelines about how the findings can be formalised, using the Douglas-Poiker algorithm, and the Spline smoothing routine, so as to be able directly to implement them in a true digital cartographic production.

## 1.4 Thesis outline:

Chapters 2 and 4 examine line simplification within traditional cartography, whilst Chapters 3, 5, and 6 examine line simplification in the digital realm.

Chapter 2 begins with a discussion of definitions and then moves onto the characteristics of traditional cartographic generalisation.

Chapter 3 explores the implementation of simplification in a digital environment. Algorithms that have been designed and used for line simplification, and the requirements for conducting that simplification are reviewed.

Chapter 4 tries to identify the role of simplification as a process integrated within cartographic generalisation. For this purpose, four types of map data are examined, at various scales, ranging from 1:25,000 to 1:1,000,000. They were chosen to provide a proper analytical context within which the relationship between cartographic generalisation and simplification can be closely examined.

Chapter 5 re-evaluates the Douglas-Poiker algorithm. An analysis of the algorithm is presented on the basis of its design specifications as outlined by its authors. A direct user input parameter is proposed for performing objective scale-dependent simplification. The relationship between graphic reduction and the resulting simplification is, also, examined and a method is proposed for quantifying this relationship.

Chapter 6 explores cartographic quality in a two stage process where the Douglas-Poiker algorithm is followed by Cubic Spline smoothing. A model is proposed for performing a scale-dependent cartographic line simplification. This model is based on the perceived role of simplification in an automated generalisation system where there are other processes involved. The chapter questions some theoretical and practical issues of simplification within this context.

Chapter 7 lists conclusions highlighting the most important finding of the thesis and refers to the need for future research.

# CHAPTER TWO: THE DISTINCTION BETWEEN CARTOGRAPHIC GENERALISATION AND SIMPLIFICATION

This chapter distinguishes between cartographic generalisation and simplification. The chapter commences with a definition of generalisation in traditional cartography in terms of execution, characteristics, and process types. Factors and quality controls of generalisation are presented. The role of line simplification as a sub-process of generalisation is explored. The chapter concludes by highlighting the significance of distinguishing between line simplification as a single integral process and line generalisation as a comprehensive cartographic practice.

## 2.1 Definition of Cartographic Generalisation

In its basic function and objective, cartographic generalisation is defined in the Multilingual Dictionary of Cartographic Terms as "the selection and simplified representation of details appropriate to the scale and/or purpose of the map." (ICA 1973, p.137). The first published work that addressed the problem of cartographic generalisation was in the early twentieth century by the German cartographer Max Eckert. During this period Eckert introduced the concept of a *scientific* cartography (McMaster and Shea, 1992). In his paper 1907, Eckert indicated that cartographic generalisation bridged between the *artistic* and *scientific* side of the field. This paper was translated and published in the Bulletin of American Geographical Society (1908) (McMaster and Shea, 1992). It was not until the 1940's that other significant contributions to the generalisation process were introduced by researchers such as Wright (1942).

Cartographic generalisation is an essential process in map production. All maps are only abstracted representations of phenomena on the ground, and the more generalised a map the more distant from reality it becomes (Joao, 1994). Geographical details are generalised to reduce complexity and so improve legibility at smaller scales, and the smaller the scale the greater is the generalisation. Figure 2.1 illustrates the effect of scale reduction with and without generalisation on a portion of a large-scale map (McMaster and Shea 1992). The large-scale map at the top illustrates part of a complex map. At the bottom left it is reproduced with no generalisation but at a reduced scale. A significant increase in the visual complexity is shown as a result of the increase crowding of features. The map on the bottom right shows a generalised representation in order to limit this crowding of features within the available map space.

There are various definitions for cartographic generalisation which indicate the wide range of views and approaches to the subject. Beard (1988) presents an elaborate discussion for different definitions and suggests four basic approaches to the definition of generalisation:

- descriptions of how the process is done, or the procedures it consists of;
- definitions of the standard limits of the process;
- descriptions of its characteristics; and
- definitions in terms of purpose.

Although there is no unanimity among cartographers as to the definition of cartographic generalisation, there is a general consensus on what it does. "Part of the reasons for so many definitions and interpretations of generalisation could be due to the lack of any 'rules' for the process (Thapa, 1988 p.187). The following sections will illustrate the general characteristics of this process.

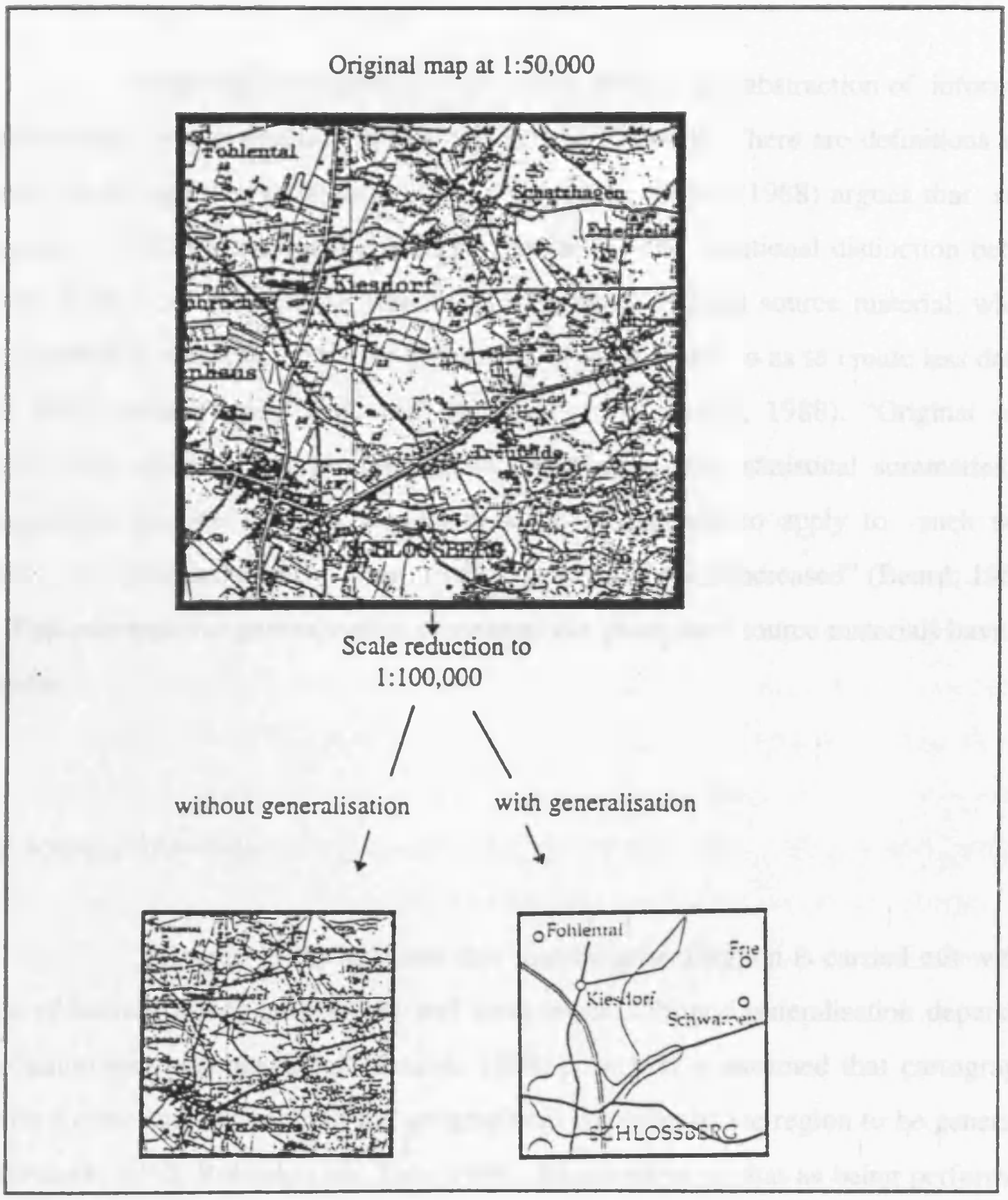



Figure 2.1: Generalised and ungeneralised representations at smaller scales
(after McMaster and Shea 1992).

## 2.2 Execution of Generalisation

### *2.2.1 Compilation:*

Compilation is regarded as the first reduction and abstraction of information in preparation for reproduction in map form (Beard, 1988). There are definitions which attempt to distinguish generalisation from compilation. Beard (1988) argues that such a distinction is difficult and for the most part arbitrary. The traditional distinction between the two is that compilation involves the processing of original source material, whereas generalisation is concerned with the processing of large scales so as to create less detailed maps from detailed ones (Hole and Campbell, 1985; Beard, 1988). "Original source material may exist in several forms including tabular data, statistical summaries, and photographic satellite imagery. If generalisation is assumed to apply to such source material, the complexity of the generalisation process is much increased" (Beard, 1988, p. 54). This assumes that generalisation processes take place once source materials have been compiled.

### *2.2.2 Scope of knowledge:*

Beard (1988) indicates that manual generalisation is carried out within a scope of knowledge at both global and local levels. "Proper generalisation depends on information and understanding" (Keates, 1989, p. 41). It is assumed that cartographers acquire a clear understanding of the geographical character of the region to be generalised (Pannekoek, 1962; Robinson and Sale, 1969). Beard refers to that as being performed at the global level, although this level of knowledge is not always attained. The scope of knowledge for cartographers at this level is limited, since in large map production agencies, cartographers perform generalisation for maps of areas of which they may have no previous knowledge. As Keates (1989) emphasises, "it is usually true that it is in the generalisation of less familiar environments that the process is most prone to error, essentially as a consequence of ignorance" (p. 41). The local level is the situation where pen is set to paper, where attention focuses on the characteristics of individual objects and their relationships to others. "At the most basic level the cartographer's attention is focused on a cone of vision a few centimetres around the tip of a moving pen. Small adjustments and conflicts are resolved within this limited scope. It must be remembered that even with this narrower focus the cartographer still retain an overall comprehension of the larger regional character "(Beard, 1988, p.65).

*2.2.3 Subjectivity:*

Cartographic generalisation is known to be a subjective process. It is "an ambiguous, creative process which lacks definitive rules, guidelines or systematisation" (Brophy, 1973, p.300). Pannekoek (1962) points out that "the form a cartographer eventually chooses depends on his personal insight, his feeling for the terrain forms, his decision is accordingly a subjective one"(p. 56). Although, some have attempted to produce objective methods, such as the required number of features at reduced scales (e.g., Topfer and Pillewizer, 1966), they can only be described as theoretical attempts (see section 2.4.3). In fact, such efforts do not explain how cartographers generalise features. Implemented as a strictly objective process, generalisation creates another set of problems, especially in the domain of automation (Beard, 1988). Lundquist (1959) cautioned against rigid standards which would not allow expression of locally important or unique features. He provided an example of settlements to illustrate this point. "If the objective criteria used to select settlements for depiction on a map was population size, small but locally important centres in rural or otherwise less populated regions would be left out" (p. 55). Although others would argue that rules can be modified to accommodate some situations, there are cases in which rules do not yield optimal results. Many researchers believe that generalisation could not be ascribed exclusively to absolute subjective or objective approaches, since both aspects are necessary to the generalisation process (Brophy, 1973; Steward, 1974; Nickerson and Freeman, 1986; Beard, 1988;

McMaster and Shea, 1992). As Brophy (1973) suggests, objective or quantitative methods can easily lend themselves to automation while subjective ones (such as control) may not.

## 2.3 Factors in Generalisation

### *2.3.1 Scale:*

The scale of the finished map has a major effect on the amount of generalisation that is required. Consequences of generalisation are most evident at smaller scales. There is a range of generalisations which would suit any single scale; i.e., will be neither too detailed nor too general for that scale (Robinson *et al., 1995).* Scale determines the type of the generalisation process required. Robinson *et al.,* (1995) indicate, that at large scales, generalisation processes such as classification and symbolisation are the most required, while at smaller scales, processes such as simplification, exaggeration, and classification become the most important. Keates (1989) points out that there are large-scale maps (such as those at 1:10,000 scale) which are derived from even larger scale sources, and there are basic derived maps at 1:50,000, 1:100,000, and, even, 1:250,000 scale. He argues, "it is hardly reasonable to suppose that because a map at 1:25,000 scale has been produced from an 'accurate' survey it does not involve generalisation" (p. 38). Each map scale requires rules for feature transformations. The effect of scale on generalisation can not be easily separated from other related factors such as graphical limits and map purpose.

### *2.3.2 Graphical limits:*

As the scale is reduced, so is the map space available for feature representation. Map symbols representing features can not be reduced in proportion, as this would lead to illegibility. "Available map space is reduced by the square of the difference in linear scales. For example, a region at 1:25,000 will only occupy one fourth as much map space when mapped at 1:50,000" (Robinson *et al.,* 1995, p. 454) (see section 5.4.1). Keates (1989) points out that legibility depends on symbol size, form , and colour, which in turn affect contrast. He indicates, that although there may be some rules regarding symbol representation at different scales, they can not be consistently applied, since they vary with the presumed importance and actual size of the feature. He provides an example where "at a scale of 1:25,000 a main road is enlarged more than a canal, because the road requires a complex symbol to represent its characteristics and classification, and is also judged to be more important. But a footpath is enlarged by a greater ratio, because the minimum size of legible symbol required is large compared to the width of a footpath" (p. 39). Robinson *et al.,* (1995) classified graphic limits to two groups: 1) technical limits set by the cartographer's tools, and 2) perceptual limits of the human eye. They indicate, that the ability to form a symbol from the basic graphic elements (point, line, and area) depends on three types of limitations. Physical limits are imposed on the graphic elements by the equipment, materials, and cartographer's experience. Physiological and psychological limits are imposed by the map user's perceptions and reactions to the primary visual variables. "For example, a line twice as wide as another will usually look that way, but a circle with twice the area of another will look significantly less than twice as large" (Robinson *et al.*, 1995, p.459). Figure 2.2 is presented by Keates (1989) and shows building outlines on a large scale plan with a scale reduction of twenty five times. In the figure, legibility becomes increasingly difficult with the decrease in scale, as the map space becomes increasingly smaller.

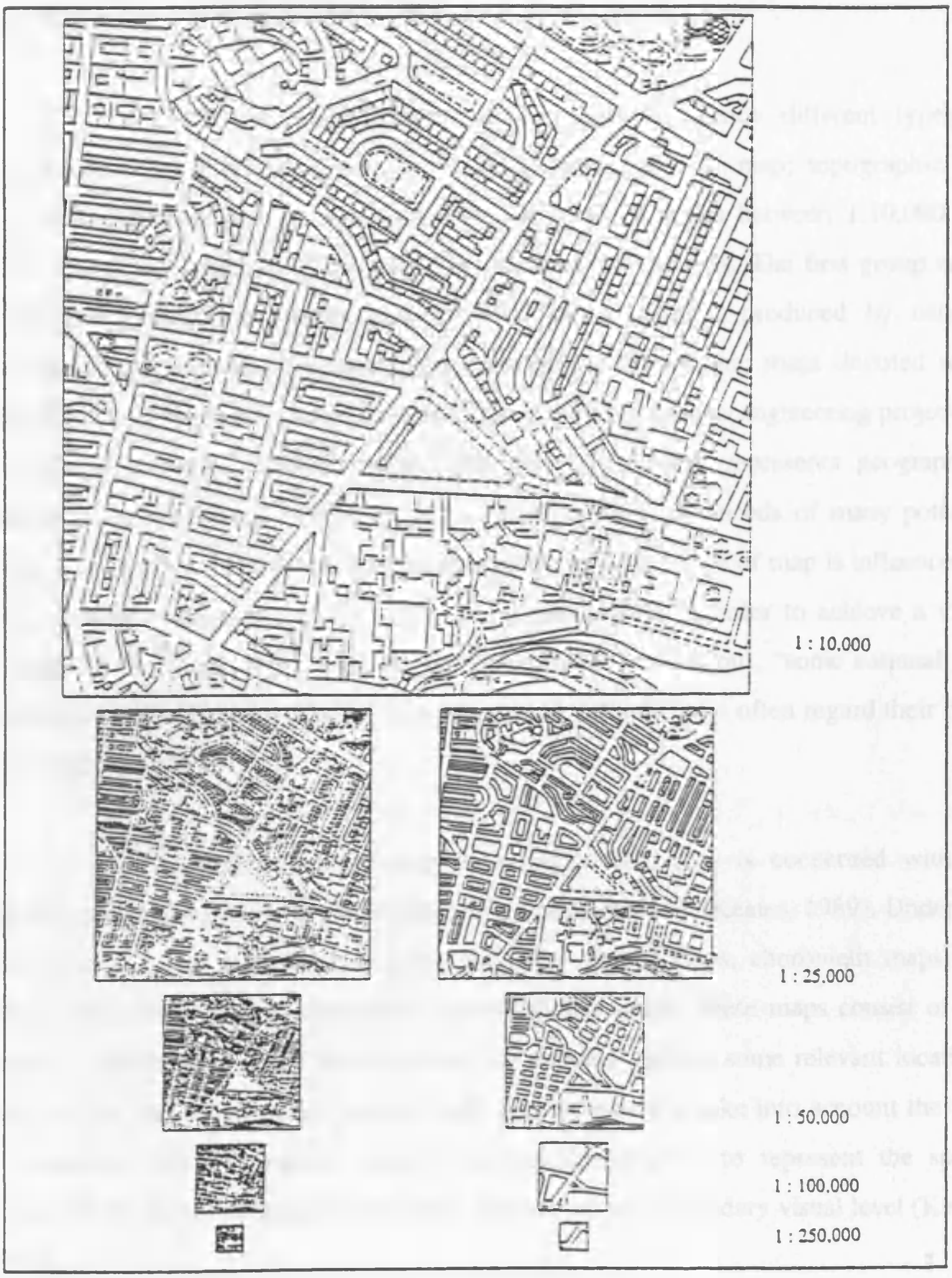



Figure 2.2: Effects of graphical limits on generalisation (after Keates, 1989).

*2.3.3 Purpose:*

There are various types of map purpose, hence different types of generalisation are required. Generally, there are two types of map; topographic and thematic. Topographic maps are normally in a range of scales between 1:10,000 and 1:250,000 (Keates, 1989). These maps fall into two main groups. The first group is the major group which represents standard map series, usually, produced by national surveying organisations. The other group consists of topographic maps devoted to a particular area, which may represent some special purpose, such as engineering projects or recreational activities (Keates, 1989). This type of mapping represents geographical features at greater detail since they are designed to serve the needs of many potential users. The degree of detail and level of generalisation of this type of map is influenced by many factors, yet it is essential to treat all features equally in order to achieve a visual balance for the whole representation. As Keates (1989) points out, "some national map series are noticeably more detailed than others, and yet map users often regard their maps as normal" (p. 148).

The other type of map is thematic. This type is concerned with the representation of a particular phenomenon or some aspect of it (Keates, 1989). Under this general term, maps such as road maps, oil maps, tourist maps, choropleth maps, and others can be described as thematic or special-subject maps. These maps consist of two types of information: 1) the reference base information such as some relevant locational information, and 2) the subject matter itself. It is necessary to take into account the need for contrast with the graphic elements during generalisation to represent the special subject of the maps, keeping the reference information at a secondary visual level (Keates, 1989).

## 2.4 Processes of generalisation

There are numerous definitions for describing the variety of generalisation procedures which, in part, as Stewart (1974) suggests, is a reflection of the richness of the English language rather than an expression of fundamentally different tasks. Several researchers have attempted to summarise the processes usually involved in generalisation. As Beard (1988) points out, several different terms can describe the same process, the opposite case also exists, since the same term can have different meanings in the minds of different cartographers. In this thesis, eight major processes of generalisation are identified: selection, classification, simplification, exaggeration, combination, displacement, symbolisation and induction. These processes are discussed in greater detail, below.

### *2.4.1 Selection:*

Selection of features is an intellectual preliminary process of describing which classes of features will be necessary to serve the map's purpose (Robinson *et al.*, 1995). In a small scale road map, for example, only major roads will be selected while other minor roads will be disregarded. Proper selection depends on the cartographer's ability to understand the information to be represented on the map according to the purpose of the map.

### *2.4.2 Classification:*

Classification is concerned with grouping, ordering, and scaling features by their attribute types and attribute values (Robinson *et al.*, 1995). This process is essential since it is difficult in practical terms to symbolise every individual value (Shea and McMaster, 1989). As Robinson *et al.*, (1995) suggest, it is an intellectual process, and there are two ways for performing classification on maps: 1) allocating similar qualitative attributes, such as land use or vegetation, into categories (e.g., cropland, forest), or values of quantitative attributes into defined groups (e.g., 1,2,3,4,5,6,7,8,10 to 1-5, 6-10), and 2) modifying the attribute value at a selected location to create a 'typical' feature for portrayal on the map. Robinson *et al.*, (1995) refer to one manipulation in classification as clustering which is often necessary when numerous discrete features characterise a distribution but, at the reduced map scale, it would be impossible to portray every individual feature.

*2.4.3 Simplification:*

In its simplest term, cartographic simplification is the practice of reducing the amount of information while maintaining the essential geographical character of the mapped phenomena. The larger the scale reduction, the greater the effect of simplification, and the greater the complexity of features, the greater the effect of simplification. Although simplification can be confused with selection, it is assumed that simplification is applied to a feature that has been already selected. For example, if a particular type of feature was selected to be shown on the map, a simplification process has to be applied where necessary in order to accommodate that feature within the available mapping space at the required scale. Within the simplification process, a process of selection-omission of characteristic features is required. For example, once line features have been selected for representation they have to undergo, where necessary, simplification so that their essential characteristics are retained whilst other unimportant ones are removed. This type of selection is applied within features, whereas the selection as a separate process is applied between features; i.e., choosing the type or class of features required for representation at a particular scale, which is a process performed during compilation. Robinson *et al.*, (1995) indicate that during the simplification process, omission or retention of the feature "depends on the relative importance of [that] feature in the visual hierarchy, the relation of that class of feature to the map's purpose, and the graphic consequences of retaining the feature" (p. 454). There have been attempts towards objective selection of features that have been retained at reduced scales (Topfer and Pilliwizer, 1966; Srnka, 1970; Cuenin, 1972). Topfer (Topfer and Pilliwizer, 1966) developed what he called the Radical Law, by which the number of features or items on a newly compiled map can be calculated. "Although the law's primary value is theoretical rather than practical, it is useful in several important ways. We can apply it to: (1) point feature sets (for example, towns on a road map), (2) linear feature sets, such as roads or streams, and (3) areal feature sets that consist of numerous small similar items within a region, such as lakes or islands" (Robinson *et al.,* 1995, p.454). In its basic form, the formula is expressed as follows:

$$N_f = N_a \sqrt{(M_a/M_f)}$$

where

$N_a$ = Number of objects on source map with a scale function $M_a$.

$N_f$ = Number of objects on derived map with a scale function of $M_f$.

In order to apply this expression successfully at small scales, where symbolisation takes up a much larger proportion of space, Topfer added two constants, $C_e$ and $C_f$, respectively which he called the constant of symbolic exaggeration ($C_e$), and the constant of symbolic form ($C_f$).

Thus:

$$N_f = N_a \, C_e \, C_f \sqrt{(M_a/M_f)}$$

Unfortunately, however, this method gives no indication as to which features should be selected, and which should be removed; that is left to the cartographer.

*2.4.4 Exaggeration:*

It is only possible to show features such as roads, buildings, and small streams at large scales (such as 1:2500) without greatly enlarging them (Robinson *et al.*, 1995). At small scales it is necessary to enlarge them in order to identify them and to be able to perceive their distinctive geographical character. This process is usually called emphasising. Robinson *et al.*, (1995) provide an example where at a scale of 1:25,000, a street which is 20 meters wide has to be included. If the street is shown true to scale, it would be symbolised by a line 0.8 mm wide. If reduced photographically to 1:100,000, the street symbol would be only 0.2 mm wide, whereas if reduced to 1:500,000, the lines probably would be imperceptible. At either reduced scale it has to be exaggerated since it has already been decided that it should be included.

*2.4.5 Combination:*

The process of combination refers to the practice of grouping or combining features of the same type. This process "takes place on the basis that [feature] extension over intervening spaces is theoretically possible. Therefore, this type of combination can not be applied to islands, which convert water to land" (Keates, 1989, p. 45). Obviously, as scale is reduced, the process of combination is progressively increased. The Figure 2.2 presented by Keates (1989), clearly shows how the building outlines are grouped into blocks, while major roads are exaggerated and minor ones omitted, until eventually all that is retained is a major road surrounded by continuous buildings.

*2.4.6 Displacement:*

Displacement is needed when the relative positions of features at small scale become imperceptible. The process of exaggeration leads in many situations to displacement. A classical example is an offset road junction where the relative position of minor roads have to be exaggerated and displaced. Displacement is required in the process of generalisation, and is cartographically justifiable, so long as it is controlled.

*2.4.7 Symbolisation:*

The process of symbolisation is concerned with transforming generalised features to graphic marks on the map (Robinson *et al.*, 1995). Some (e.g., Keates, 1989) argue that the whole process of generalisation is essentially a symbolisation process; hence it is not necessary to identify symbolisation as a sub-process within generalisation. Robinson *et al.*, (1995) indicate that graphic marks are for symbolising a series of facts, concepts, or the character of geographical distribution. They argue, that symbolisation is an intellectual process since cartographers may change a feature's dimensionality or measurement scale of a feature's attribute value (as in choropleth mapping).

*2.4.8 Induction:*

Robinson *et al.*, (1969) were the first to introduce the term induction, which is described as the process of making inferences from interrelationships among features on the map. They state that "cartographers have little control over induction When we apply induction or inductive generalisation, we extent the map's information content beyond its features" (p. 451). They provide an example where average January temperatures for a series of weather stations can be used to construct a set of isotherms, hence the isotherms allow inferences about probable January temperatures in areas <u>between</u> weather station locations, while the original data were only temperatures recorded at the weather stations. They also indicate that induction may extend beyond what was consciously added by the cartographer to the extent that the map user may further amplify the map's information. Proper induction would be necessarily encouraged by proper generalisation (Robinson *et al.*, 1995).

### *2.4.9 Examples:*

The diagrams in Figure 2.3 (after Shea and McMaster, 1989) show various examples which are graphic summaries of most of the generalisation processes discussed above. These graphic illustrations are self-explanatory, hence they are accompanied with no further discussion. The diagrams also reveal how generalisation processes can be perceived by different cartographers and researchers.

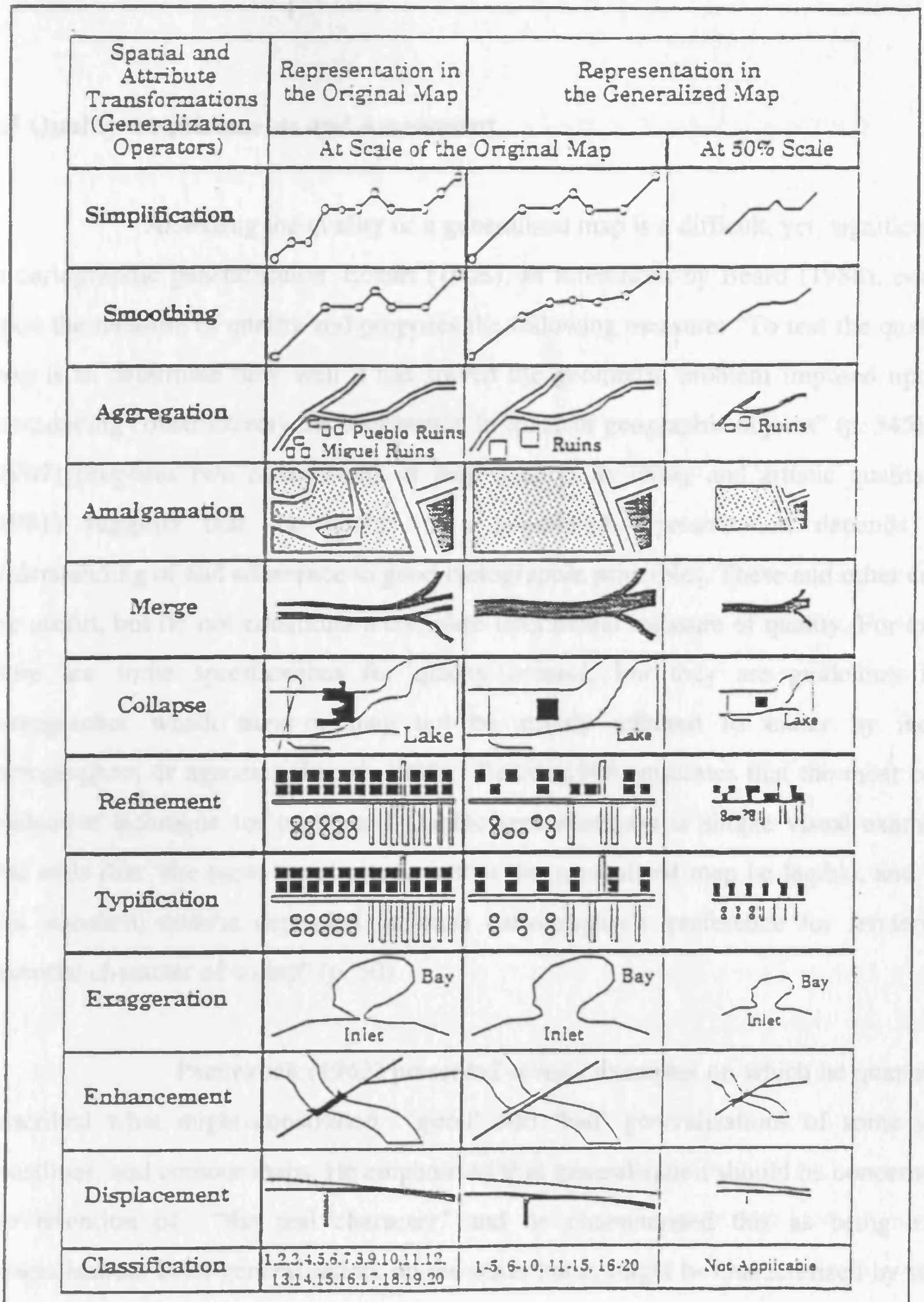



Figure 2.3: Cartographic generalisation processes (after Shea and McMaster, 1989).

## 2.5 Quality Requirements and Assessment

Assessing the quality of a generalised map is a difficult, yet, significant issue in cartographic generalisation. Eckert (1908), as referenced by Beard (1988), comments upon the measure of quality and proposes the following measure: "To test the quality of a map is to determine how well it has solved the geometric problem imposed upon it of reproducing constructively the distribution in space of geographic objects" (p. 345). Board (1967) proposes two components of map quality: its utility and artistic quality. Jenks (1981) suggests that the quality of a simplified representation depends on an understanding of and adherence to good cartographic principles. These and other concepts are useful, but do not constitute a complete operational measure of quality. For example, there are some specifications for quality control, but they are guidelines for the cartographer which may or may not be strictly adhered to either by individual cartographers or agencies (Beard, 1988). Beard (1988) indicates that the most common evaluative technique for traditional graphic representation is simple visual examination. She adds that "the most basic criteria is that the generalised map be legible, and beyond this standard, criteria depended on each cartographer's preference for rendering the essential character of a map" (p. 30).

Pannekoek (1962) presented several examples on which he quantitatively described what might constituted 'good' and 'bad' generalisations of some selected coastlines, and contour maps. He emphasised that generalisation should be concerned with the retention of "the real character" and he characterised this as being a sound generalisation. Poor generalisation, on the other hand, might be characterised by stressing non-essentials, depicting similar forms differently, or making different forms appear similar. Watson (1970) also compared a generalised map against an independent source of higher accuracy. His work might be regarded as being the first move towards an operational measure although it was still a visual check without quantifiable estimates of quality. He used a trigonometric survey map at 1:50,000 scale to test the accuracy and completeness of the 1:250,000 scale map series and generalised 1:500,000 scale photographically enlarged or reduced on film positives to a common scale. The maps were then superimposed and visually compared.

### *2.5.1 Quality and errors:*

In its broadest definition, quality is a measure of fitness for use. In this context, a generalised map should be evaluated on the basis of how well it addresses specific questions, or serves an intended purpose (Beard, 1988). Beard (1988) also emphasises that evaluation criterion must be flexible, since the requirements of individual applications of generalisation are very varied. The National Committee for Digital Cartographic Data Standards (NCDCDS) in the United States provides five quality components for the evaluation. These components are: positional accuracy, attribute accuracy, completeness, logical consistency and lineage. Each of these components, except lineage, can be tested at different levels of rigor. The four categories of tests recognised by the standard include: deduction, internal evidence, comparison to source and comparison to independent sources of higher accuracy (Beard, 1988). These components are nominally measured and the result is reported as right or wrong, and not all tests apply to each component as the accuracy of some components are not easily assessed (Beard, 1988). Pertinent to, or implied in, the subject of quality, is error. There are few references to error in traditional cartographic generalisation as it might be related to a belief among cartographers that generalisation is a deliberate change in the map features, hence any loss of accuracy is regarded as a mapping requirement, not error. Error in generalisation, or mapping in general, has received attention by cartographers and researchers recently as a result of the increase in the use of digital mapping techniques. Many have commented upon the definition and significance of error in generalisation (e.g., Jenks, 1981; Goodchild, 1980b, 1991; Chrisman, 1989; McMaster, 1986, 1987a, 1987b; Muller, 1987b; Beard, 1988; Joao, 1994). While many still argue about the significance of error in the context of generalisation, and whether it should be called error or controlled error, and how it can be quantified, others recognise that whatever the name there are still subtle and even serious side effects of generalisation on subsequent use and analysis of the resulting maps especially in GIS environments (Beard, 1988; Joao, 1994). Previous quality standards suggest that, so long as maps are generalised for production purposes (scale-dependent), rigorous assessment of absolute accuracy, as opposed to relative accuracy, of all features at every position on the map sheet is neither necessary nor possible, since visual perception is the key goal in map production. While digital mapping methods should minimise error and allow for its quantification, traditional mapping methods are commonly known to be inconsistent and have a fairly high potential for error. "There are no error free generalisation, but one generalised map may have fewer errors than others, depending on the quality of the source material and the processes used to generalise the map. Each generalisation procedure will, in fact, contribute different levels and components of error to the result" (Beard, 1988, p.86).

## 2.6 Line Simplification

### *2.6.1 Definition and characteristics:*

The importance of line simplification stems from the fact that almost eighty per cent of map objects are line features (Muller *et al.*, 1995b); hence it is important to achieve, at least, an acceptable level of understanding of the objectives of line simplifications during cartographic generalisation. This thesis argues that the process of 'dissection' of generalisation operations is assumed to be the first proper approach in order to understand their interrelationships. This approach, would yield fruitful results, as it will not only help provide standardised methods, but also to validate and improve these operations.

As noted earlier, simplification is only a process among other processes in cartographic generalisation. From the previous discussion it is implied that during line simplification particular types of line features (e.g., roads) can be selected and may well be subject to a further selective omission process (within the feature type; e.g., particular class of roads). The retained features would undergo simplification process of their details. The interrelationships between simplification and other processes at every map scale and purpose is almost unknown. It is surprising to witness a lack of interest among researchers, and particularly the enthusiasts for automation. The previous discussion of generalisation shows that simplification of point, line, and area features is perceived as being subjectively executed. The discussion, also, shows that simplification is neither a substitute for any generalisation process nor is generalisation. Furthermore, simplification can not be executed in isolation from other processes and the general mapping context. Thus it is a context-dependent process. Due to its close interaction with other processes, simplification is often erroneously regarded as synonymous with generalisation. Recently, Keates (1996) emphasises this important point. He indicates that

> "It is unfortunate that in current cartographic literature the term generalisation is often used to mean only simplification. It is unfortunate, because in many ways line simplification is the easiest aspect to understand, and also the easiest aspect for cartographer to deal with. In practice, generalisation involves five related procedures - selective omission, simplification, combination, exaggeration, and displacement - and in many cases they are all applied simultaneously. Although they involve familiar cartographic tasks, they also need to be understood by the map user" (p. 100).

While simplifying line features, the cartographer performs many tasks simultaneously; the important features of the line are selected, simplified, or exaggerated, and perhaps smoothed and displaced from other features (McMaster, 1987b). The subjective nature of manual line simplification has been commented upon by several researchers (e.g., Pannekoek, 1962; Robinson *et al.*, 1995; Zoraster *et al.*, 1984; McMaster, 1986, 1987a, 1987b, 1989; Thapa, 1988; Carstensen, 1988; Keates, 1989). Line simplification may therefore be described as an intellectual and manipulative activity (McMaster, (various dates), and because of its subjectivity, in terms of execution, manual simplification has criticised for its lack of theoretical foundation (Thapa, 1987), and its unpredictability (Pannekoek, 1962).

When simplifying a line feature, the cartographer is primarily concerned with the recognition and preservation of the line character. Different cartographers perceive line characters differently. Carstensen (1988) describes character as being "elusive, personal, and subjective, and a function of personal view". However, an understanding of the geographical nature of the region being mapped is essential to perform effective simplification (Pannekoek, 1962). It has been suggested that the subjectivity associated with manual simplification could be "removed through careful geographic training" (Raisz, 1962; Jenks, 1979). In fact, it can only be reduced, since the nature of manual processes can only be improved, but can never match their mechanical counterparts. Smith (1980) interviewed a few draughtsmen at the Ordnance Survey, and found that they were unable to define the rules by which they produce the 1:250,000 scale map series. They claimed that each line had to be considered on its own merits and in relation to the rest of the map (Smith, 1980). An alternative solution is to generalise linear features using automated methods (McMaster, 1987a).

The inconsistency of generalisation between cartographers is the result of many factors. Such factors may "include different human skills in drafting and checking map information, non-uniformity of geographic knowledge among cartographers, conditions of working environment, urgency of production, and the physical and mental being of the cartographer" (Steward, 1974, p.40). Because of this intrinsic inconsistency, even skilled cartographers would find it difficult to replicate precisely their results (McMaster, 1987a). An example, is provided by McMaster (1987a) illustrating the inconsistency of the manual line simplification process (Figure 2.4). The figure shows different versions of a section of Lake Ontario shoreline taken from three road maps of nearly the same scale. While the general form of the line is almost the same, the minor details are quite different. The figure confirms that inconsistency of manual line simplification tends to occur more frequently on minor or small details of the line feature.

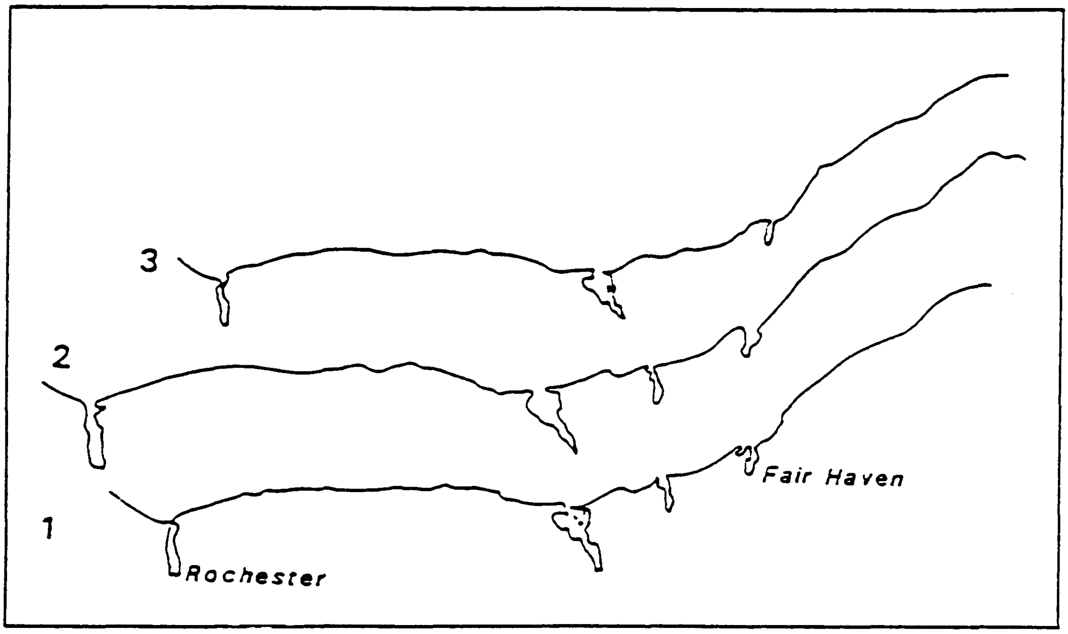



Figure 2.4: The inconsistency in manual line simplification (after McMaster, 1987).

*2.6.2 The perceptual model of line simplification:*

The previous discussion in this chapter refers to the terms of *minor* and *major* details of geographic lines. Although they are difficult to define practically, they are common descriptions of feature shape among cartographers. This general description of shape has to be carefully and appropriately maintained during generalisation of line features for different scales. Keates (1989) provides an excellent graphical example of what, in general, constitutes the character of geographic line (Figure 2.5). He indicates that an irregular line will suffer a progressive diminution in length as its minor irregularities are removed, while the important characteristics are retained, and what has to be avoided is the replacement of irregular lines by smooth curves. In his example, Keates (1989) shows that an irregular line can be described as having three major elements: its general direction, major forms, and minor forms (Figure 2.5a). "The major forms are marked by an angular shape. During generalisation, the process of simplification will lead eventually to the removal of the minor forms entirely, but the major forms should be retained as far as possible. Eventually, at a very small scale, only the major direction will be retained" (Keates, 1989, p. 44). Keates illustrates another example where generalisation progressively reduces irregular lines to regular ones (Figure 2.5b). He suggests, here, that irregular lines become regular or only the general direction will be retained. It is important to note that in this example Keates refers to the whole process of generalisation implying that simplification is a sub-process which is concerned with removing small details while maintaining large ones. Keates (1989) further demonstrates how cartographers simplify lines. "A cartographer can scan a section of a line, and can change from concentrating on the minor details to major forms at will, and can perceive a whole section of line simultaneously. Therefore, the operation of simplification can respond both to the detail of the line section and the overall line direction" (p. 45).

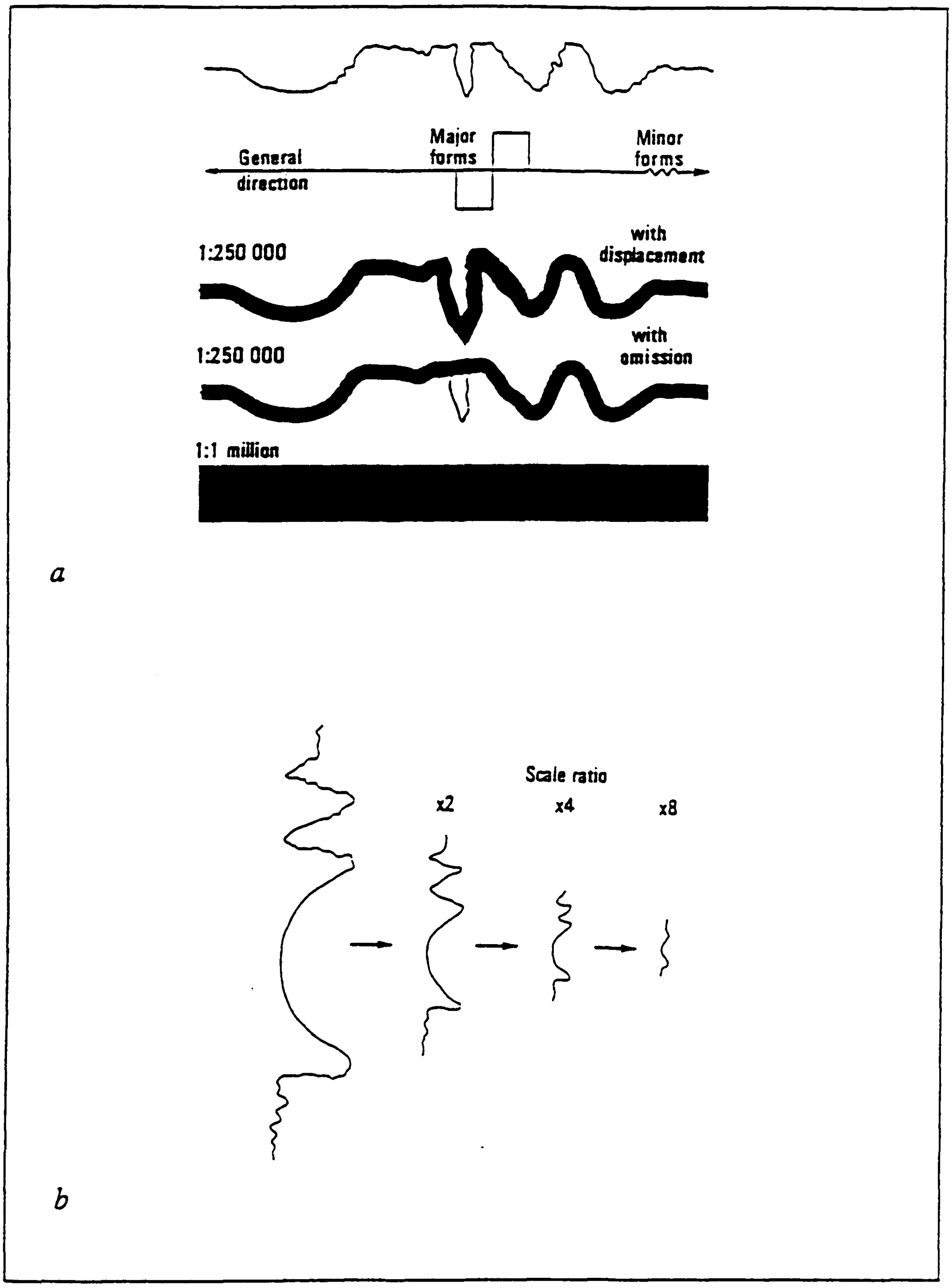



Figure 2.5: Character of geographic line and generalisation (a), and effect of geographic line reduction (b) (after Keates, 1989).

## 2.7 Summary

This chapter presented a theoretical overview of the traditional cartographic generalisation. The chapter covered a definition of generalisation in traditional cartography in terms of execution, characteristics, and process types. Factors and quality controls of generalisation were presented. Line simplification within generalisation was specifically discussed. The discussions indicated that line simplification in the traditional realm of cartography is a single integral process within line generalisation as a comprehensive cartographic practice, although it is difficult to draw a precise distinction between the two processes in practical terms. The following chapter is designed as a theoretical discussion of line features processing and simplification in the digital realm.

# CHAPTER THREE: DIGITAL LINE SIMPLIFICATION

This chapter reviews the implementation of simplification in a digital environment. Algorithms that have been designed and used for line simplification, and the requirements for conducting that simplification are reviewed. The chapter concludes with a discussion of the factors which often affect the evaluation process of digital simplification algorithms.

## 3.1 Spatial data capture and storage in digital cartography

Before reviewing the characteristics of digital line simplification, it is necessary to refer to the processes by which data are captured, edited, and stored. These processes are prerequisites to the process of line simplification in the digital context.

### *3.1.1 Digitising:*

The most commonly used method of capturing linear data is through manual digitising performed on digitising tables or tablets. The process results in a line being stored as a series of representative points. In manual digitising, line features are traced over the source paper material by a device called a cursor. The number and significance of the digitised points are determined by both the scale of the source material, the method of digitising, and the complexity of line features being digitised. Most of the manual digitisers can be operated in two modes: *point* and *stream* mode. In point mode, co-ordinates are only recorded when the operator gives a special signal, such as clicking a certain button on the cursor. In stream mode, co-ordinates are automatically recorded at given time or distance intervals. The spatial resolution of the resulting digital record is, therefore, determined by these intervals (Robinson *et al.* 1995), and by the speed at which the operator traces and digitises the source material.

Pertinent to digitising is classification or coding. In this process, the operator assigns feature codes to the captured line data. For example, line features can be assigned codes, indicating that they are roads, river, etc.

### *3.1.2 Editing, Cleaning and Weeding:*

Raw vector data, as they come from digitising processes, usually require considerable processing to bring them up to a cartographically acceptable standard. During digitising, errors may take several forms even with careful planning and execution (Robinson *et al.*, 1995). The geometry of digitised lines has to be checked for errors like crossed or missed arcs that have been digitised twice or more, etc. The process of editing can be done interactively. Jenks (1981) identifies two forms of data 'cleaning'; the cleaning through interactive display and correction, and the cleaning through the application of software. He suggests that physiological errors (involuntary muscular spasms) often result in spikes, switchbacks and knots (Figure 3.1) being recorded in the digital line, whereas line following psychomotor errors often results in significant displacement from the source line. Jenks (1979) claims that the prime objective of 'weeding and cleaning' should be to produce "an accurate, error free, minimum data file which when plotted should exhibit a nearly exact copy of the original manuscript line" (p. 211). Another advantage to be gained from this process, is the reduction of superfluous data for storage and processing. Jenks (1981) identifies this type of data reduction as a 'generalisation' in the imperceptible realm, where the map reader can not, yet, detect changes in the plotted line. Weeding is most needed on lines that had been captured through the process of *stream* mode digitising, especially using the process of time incremental digitising. There are routines that can be used to remove redundant digitising points, and for removing digitising errors, depending on the magnitude of the errors and the tolerance values used (e.g., Lang, 1969; Douglas and Poiker, 1973). Many GISs such as Arc/Info support facilities for detecting and cleaning more demanding errors such as switchbacks and spikes, and also weeding out superfluous data.

Jenks (1981) comments that, “error-free digital files rarely if ever exist and the concept of accuracy must be based on an understanding of the quality of the equipment and personal used to capture the data, and the cost involved in further editing” (p. 7). The accuracy of a digital representation is, therefore, dependent on the source quality, equipment quality, and the operator’s experience (Whyatt, 1991).

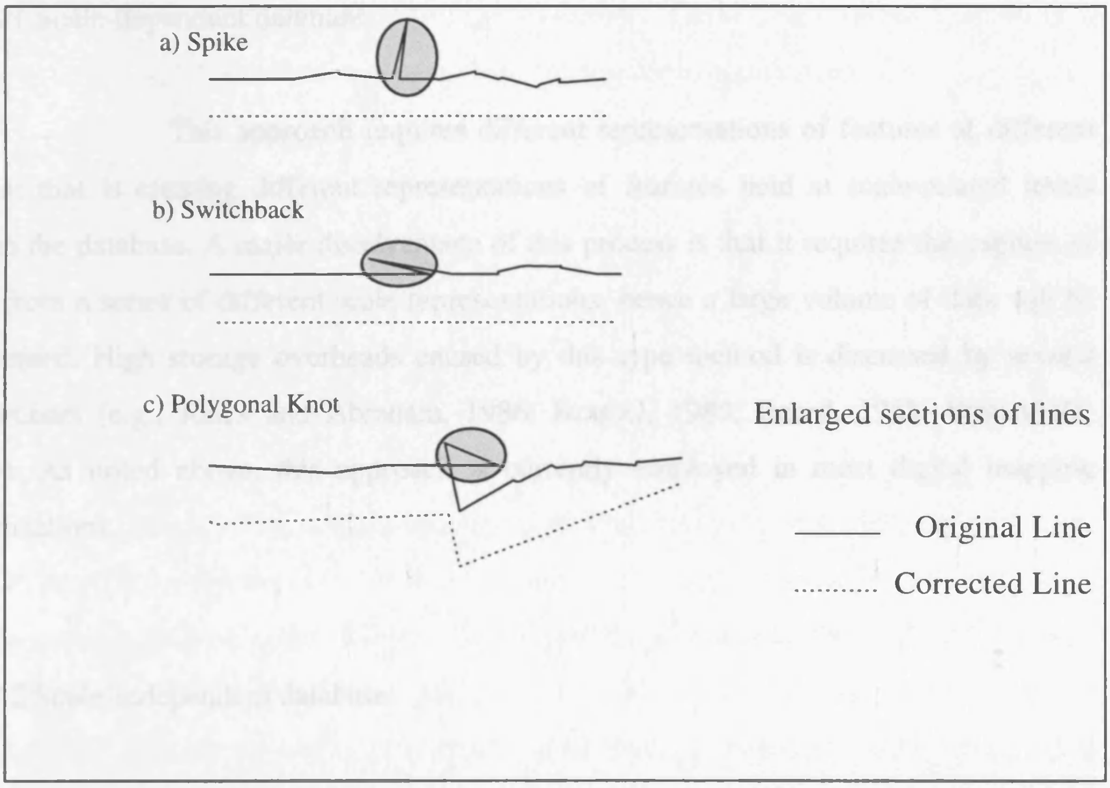



Figure 3.1: Typical digitising errors (adapted from McMaster, 1987).

*3.1.3 Data storage:*

After capture, linear data are normally stored in a topologically structured database for further cartographic processes such as analytical and communicational mapping tasks. The issue of captured data storage has significance to digital mapping users, for various reasons. For example, the process of data capture is known to be both expensive and time consuming. That has largely contributed to searching for alternative methods such as scale-independent databases. Because this particular approach proved to be difficult in terms of the implementation of appropriate data structure, current digital mapping organisations have to adhere to the process of scale-dependent databases (Whyatt, 1991). Both types of data storage are discussed further below.

3.1.3.1 Scale-dependent database:

This approach requires different representations of features at different scales; that is creating different representations of features held at scale-related levels within the database. A major disadvantage of this process is that it requires the capture of data from a series of different scale representations, hence a large volume of data will be duplicated. High storage overheads caused by this type method is discussed by several researchers (e.g., Jones and Abraham, 1986; Brassel, 1985; Beard, 1988; Buttenfield, 1989). As noted above, this approach is currently employed in most digital mapping organisations.

3.1.3.2 Scale-independent database:

As indicated above, this method is regarded as a solution to the problem of data storage. Using this approach, data have to be captured and stored once, at a single high resolution, and various reduced scale representations can be derived through automated generalisation. However, it is an ideal solution, and it will remain a research goal. Its flexibility stems from the fact that "the user is not restricted to working at specific pre-determined scales. Furthermore, since data is stored only once, storage overheads may be minimised if appropriate data structures are used"(Whyatt, 1991, p. 28). Jones and Abraham (1986, 1987), and van Oosterom (1989, 1991, 1993) produced a scale-independent data structure for the storage and subsequent retrieval of digital line data. "This structure, termed a line generalisation tree, or simply a line tree, is generated by using the Douglas-Poiker algorithm to classify the internal points of a line into a specific number of levels of scale-related significance. The highest level of the tree contains the most significant, smallest scale representation, while lower levels provide progressively greater intermediate detail. The tree is traversed to whichever level is adequate for the scale requirement, and thus only those points required are accessed, while data duplication between levels is avoided" (Jones and Abraham, 1987, p.36). However, this approach does not address the wider context of generalisation, although it can be regarded as a reasonable attempt at multiple representations from a single detailed database.

## 3.2 Digital Simplification Algorithms

### *3.2.1 Characteristics:*

Digital line simplification algorithms are primarily designed for data reduction prior to storage, i.e. for weeding purposes. Only secondarily, have they been applied to cartographic generalisation (Beard, 1988). The aim of these algorithms is to eliminate those points considered insignificant while retaining those deemed to be significant according to some pre-defined mathematical criteria. Unlike the manual simplification process where lines are simplified within the cartographic context, digital methods are objective and conducted in a serial digital mode. Since the process of simplification within the process of generalisation lacks clearly defined rules, development of objective algorithms has proved highly problematical (Whyatt, 1991). "The multitude of algorithms that have been developed undoubtedly reflects the general lack of agreement on the most suitable criteria for simplification" (Whyatt, 1991, p. 34). Although, Whyatt stated the above argument, Whyatt himself confuses simplification with generalisation. His work did not cartographically distinguish between what constitutes simplification and what constitutes generalisation, a misconception, under which he evaluated the Douglas-Poiker algorithm. Several algorithms exist, which vary in complexity, in terms of geometry and computation. Generally, they are two types: filtering and smoothing routines. McMaster (1987a) classifies the first type of algorithms into five categories as follows:

1. Independent point algorithms do not account for the mathematical relationships with the neighbouring co-ordinate pairs, i.e. operate independent of topology, e.g. $n^{th}$ point routine (Tobler, 1964, 1966).

2. Local processing routines utilise the characteristics of the immediate neighbouring points in determining the processes of selection and rejection; for example, angular change between points, distance between points (McMaster, 1987a).

3. Constrained extended local processing routines search beyond immediate co-ordinate neighbours and evaluate sections of the line. The extent of search depends on criteria such as distance, angular or number of points; for example, Lang algorithm (Lang, 1969), Opheim algorithm (Opheim, 1982), Johannsen algorithm (Johannsen , 1973), Deveau algorithm (Deveau, 1985), and Roberge algorithm (Roberge, 1985).

4. Unconstrained extended local processing routines search beyond co-ordinate neighbours and evaluate sections of the line, but the extent of search, here, is constrained by geomorphological complexity of the line, not by the algorithmic criteria; for example, Reuman-Witkam algorithm (Reuman and Witkam, 1974).

5. Global routines unlike the previous algorithms use a holistic approach in which the entire line, or specific line segment are considered in processing. The algorithms of this type select critical points iteratively on the basis of user-defined tolerance values; for example, Douglas-Poiker algorithm (Douglas and Poiker, 1973), and Ramer algorithm (Ramer, 1972).

All the algorithm types described above, are designed to weed from the line redundant or unnecessary points in a sequential mode (McMaster, 1989). The process of weeding here refers to the application of this type of algorithm for eliminating the superfluous data captured in the digitisation process. The aim is to represent the digitised lines with the minimum number of points. This process should incur no perceptual change on the lines at the representation scale of the source lines. Furthermore, the same weeding can be used to produce scale-dependent databases from single databases that already have been undergone a weeding processes after they have been digitised. This is an important process in digital mapping, since removing or weeding out redundant data is necessary either at the digitising scale or any subsequent representation scales (see the advantages of simplification algorithms in section 3.2.2). On the other hand, cartographic line simplification implies a perceptual change in the lines, due to increased reduction of line details (points). Chapter 5 will focus on data reduction for weeding purposes, while Chapter 6 will focus principally in the process of cartographic line simplification. Unlike these filtering algorithms, smoothing routines geometrically shift points in an attempt to plane away small perturbations while preserving only the most significant trends of the line (McMaster, 1989; McMaster and Shea 1992). These techniques produce aesthetically pleasing line shapes. McMaster (1989) classifies smoothing algorithms into three categories. They are as follows:

1- Basic averaging techniques:
e.g., - *three-point weighted-moving averaging*
- *five-point weighted-moving averaging*
- ........
- *distance-weighted averaging*
- *slide averaging*

2- Epsilon filtering:
e.g., - *Brophy algorithm*
3- Mathematical approximation:
e.g., - *local processing: e.g., Cubic Splines*
- *Extended local processing: e.g., B-Splines*
- *Global processing: e.g., Bezier Splines*

The pattern of displacement resulting from any of these algorithms is different. Whilst the filtering algorithms create displacements through the removal of points (details) from line features, either in the perceptible or imperceptible realms, displacements from the smoothing algorithms are caused by the process of shifting some of the points along the original line to new positions. On the other hand, the displacement from the filtering algorithms (based on the principle of selective point exclusion) occurs where the actual exclusion takes place in the line. However, calculation of the resulting areal displacements from both types of algorithm is achieved through calculating the polygonal displacements between the simplified and original lines (McMaster, 1986). Unlike both types of displacement, there is another type of displacement which is associated with intentional displacement. In this process the resulting displacement is produced from changing the position of all points along the original line into new positions. This type of displacement usually occurs during generalisation in which the position of a line has to be shifted to a new one, normally for legibility purposes.

The following Figures are to illustrate generally typical digital simplification algorithms. Figures 3.2, 3.3, and 3.6 represent the filtering routines which do not involve geometric shifting of co-ordinates. Figures 3.4 and 3.5 represent typical smoothing routines which produce aesthetically pleasing shapes by calculating new co-ordinates through which the resulting line passes, and so is smoother. The Douglas-Poiker algorithm is discussed in greater detail in Chapter 5, while the Cubic-Spline smoothing process is discussed in Chapter 6. Unlike manual simplification, digital simplification algorithms work at point level and seek to preserve the characteristic features of a line through the selection or processing of what is commonly termed Critical Points, which are further discussed in section 3.2.3. Whilst manual simplification tends to reduce line details and performs the smoothing process simultaneously, digital algorithms typically operate in a sequential manner on lines in isolation.

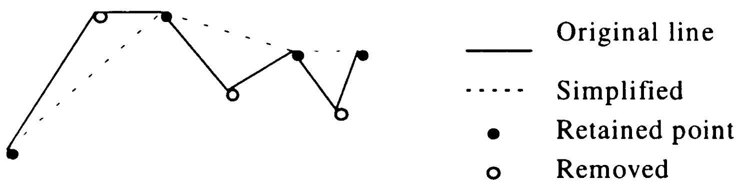



Figure 3.2: N-$^{th}$ point simplification algorithm.

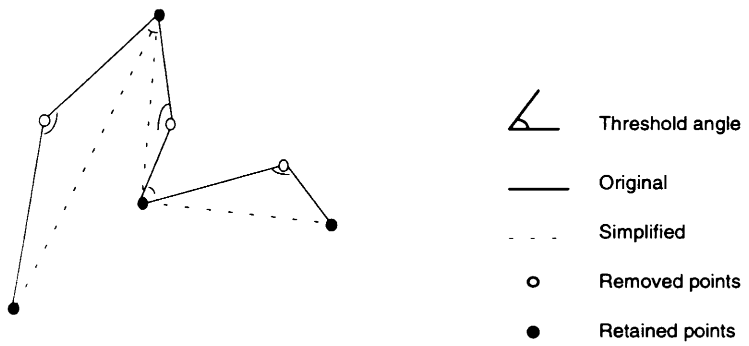



Figure 3.3: Angle threshold algorithm. Vertices with angles greater than tolerance are removed.

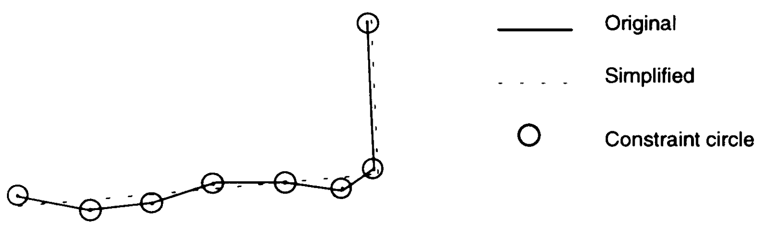



Figure 3.4: Williams point relaxation algorithm. (after Zoraster *et al.*, 1984).

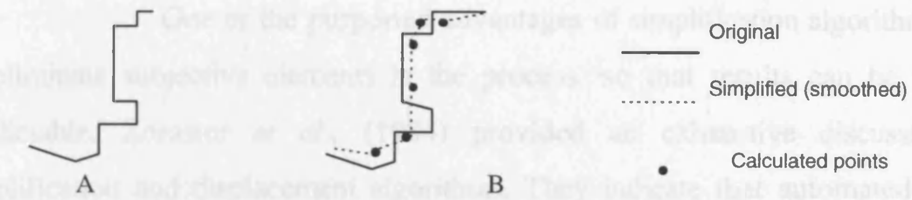



Figure 3.5: Line smoothing processes by averaging routines. In B the original line (A) is smoothed by averaging the co-ordinate values of every three pairs. (after McMaster, 1989).

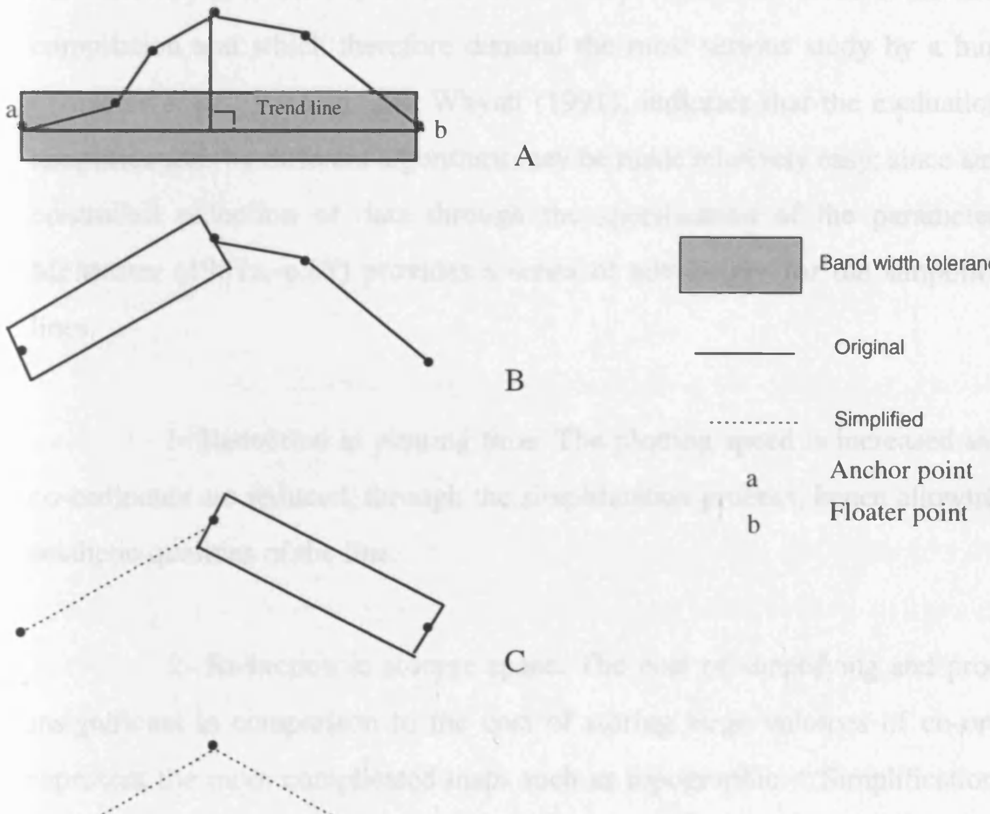



Figure 3.6: The Douglas-Poiker simplification algorithm. In A, a line (trend line) is constructed by connecting the first (anchor) and last (floater) point. Perpendicular distances are calculated from the intermediate points to the trend line. A point falling greater than the tolerance is retained, and becomes the new floating point. The process repeats with floating point moving toward the anchor.

*3.2.2 Advantages and limitations of digital simplification algorithms:*

One of the purported advantages of simplification algorithms is to reduce or eliminate subjective elements in the process, so that results can be consistent and predictable. Zoraster *et al.,* (1984) provided an exhaustive discussion of digital simplification and displacement algorithms. They indicate that automated generalisation techniques could be used to impose overall accuracy. They indicate that automation would affect the process of generalisation in two ways. "First, it would provide a level of standardisation which should make error detection easier. Second, it might be able to automatically indicate those areas of the map which had created the most difficulty in compilation and which therefore demand the most serious study by a human inspector" (Zoraster *et al.,* 1984, p. 20). Whyatt (1991), indicates that the evaluation of equivalent simplifications by different algorithms may be made relatively easy; since automation yields controlled reduction of data through the specification of the parameters to routines. McMaster (1987a, p.87) provides a series of advantages for the simplification of digital lines.

1- Reduction in plotting time. The plotting speed is increased as the number of co-ordinates are reduced, through the simplification process; hence allowing for improved aesthetic qualities of the line.

2- Reduction in storage space. The cost of simplifying and processing lines is insignificant in comparison to the cost of storing large volumes of co-ordinate pairs to represent the more complicated maps such as topographic. "'Simplification may reduce a data set by up to 70 % without changing the perceptual characteristics of the line'" (Jenks, 1979, p. 214).

3- Faster vector to raster conversion. A simplified co-ordinate set will result in a faster conversion from vector to raster mode.

4- Faster vector processing. Vector processing such as translation, rotation, scaling, cartometric analysis and symbolisation methods require less time if the data set is simplified (McMaster, 1987a).

Roberge (1985) claims that simplification process can be used to filter data for creating "data display". The level of simplification determines the size of data display files; high levels of simplification produce small display files, whereas low levels of simplification produce large display files. He indicates that faster display refresh rates can be achieved on CRT-based devices, since a reduced display file occupies less space; hence less time is needed to display it.

As early as 1969, Lang (1969) indicates that simplification algorithms are needed in order to prevent lines from merging together and producing dark smudges if they were plotted at reduced scales in an unsimplified form.

An effective visual communication at particular scales requires that redundant details be removed. The relationship between feature details and scale has proved difficult to quantify. In this context, digital line data are represented by co-ordinate pairs and in order to reduce them to a particular scale, the Radical Law (discussed in Chapter 2) has been utilised by many workers. The validity of this approach to the problem of scale-dependent simplification is further discussed in Chapter 5.

A major limitation of simplification algorithms is their underlying computation concept in which lines are treated as isolated objects; with no consideration of the geographic and cartographic contexts of the features being processed. Some workers have expressed concern that systems which strive for objectivity would necessarily encounter problems. This is true in the context of cartographic generalisation; since many of its processes are matters of judgement on a case by case basis. Rhind (1973) emphasises that automated generalisation, while based on manual processes, should not be a direct translation of them. This thesis argues that the main dissatisfaction and perceived limitation of the digital simplification algorithms is largely affected by the variability in their implementation and interpretation. For example, the confusion between simplification and generalisation, has erroneously led to evaluation of these algorithms on confused bases. As noted in Chapter 2, simplification does not imply generalisation. This thesis maintains that this problem extends beyond its philosophy or semantics, since evaluation of digital algorithms, in general, should be based upon an appropriate and pertinent cartographic theoretical foundation.

Preserving the character of a feature is the prime concern of digital simplification algorithms, and defining that quantitatively has proved difficult. The importance of critical point detection for the preservation of line character has been addressed in the disciplines of psychology, image processing, pattern recognition, computer graphics, and cartography (Thapa, 1988). The characteristics of critical points in the context of line simplification is reviewed in detail in the following section.

### 3.3 Critical Points

Critical points or characteristic points of a fine feature are those points that have the highest informative value in terms of shape description and recognition. Workers in psychology, computer graphics, pattern recognition and cartography have long been concerned with how people recognise shapes and what shape features facilitate this recognition. In psychology, Attneave (1954) indicates that information is concentrated at points when direction changes, that is at angles and sharp curves. These points or angles are the distinctive features the brain employs to store and recognise shapes. Attneave (1954) illustrated this through an example by abstracting the outline of a sleeping cat from points of maximum curvature (Figure 3.7).

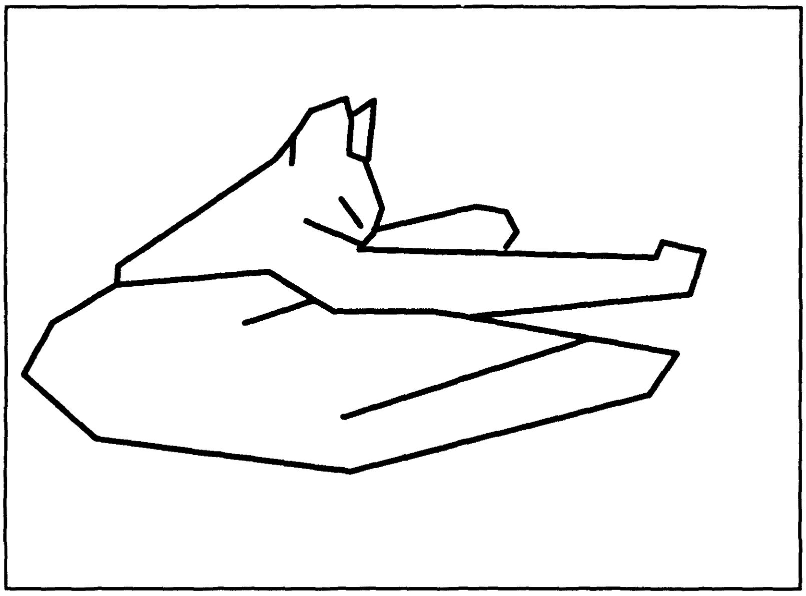


Figure 3.7: Attneave 's example (a sleeping cat) of Critical Points recognition (after Attneave, 1954).

In the field of pattern recognition, Freeman (1978) provides more specific examples of these characteristic points, including maxima, minima, points of inflection, discontinuities in curvature endpoints, intersections and points of tangency. In cartography, several workers have commented on this subject, (e.g., Douglas and Poiker, 1973; Arnheim, 1974; Jenks, 1979, 1981; Marino, 1979; White, 1985). Arnheim (1974) notes that generalisation occurs spontaneously in all perception, and all that remains in memory are simplified images.

Jenks (1979, 1981) identifies two types of characteristic points: 1) geometrically critical points, which are traditionally used to provide caricatures of lines, and 2) those points that bear some economical or cultural significance; for example, intersections of political boundaries. These two types are further discussed below in the context of digital simplification.

### *3.3.1 Geometrically Critical Points:*

Several algorithms have been proposed to objectively select critical points and to locate points of maxima on the digital line. Ramer (1972), and Douglas and Poiker (1973) were the first to comment upon the importance of selecting the geometrically critically points of line features during a digital simplification process. The Ramer algorithm (1972) and Douglas-Poiker algorithm (1973) are classic examples based on this principle.

Marino (1979) conducted an empirical study for determining whether cartographers and non-cartographers selected similar critical points when simplifying lines. In her study, the respondents were asked to mark points on various geomorphological types of lines at three levels of simplification. Marino's significant contribution in this context, is to identify (geometrically) critical points in the process of line simplification. It is an objective process, though manually performed, for finding out what in particular makes *the* geographic lines more informative to the map readers.

Pertinent to the importance of critical points, is the issue of the hierarchy of this importance; since it is assumed that critical points have varied significance at different levels of scales or representation. Boyle (1970) as referenced by Thapa (1988), suggests that the points which are regarded as more important (i.e., critical points) during the process of line generalisation should be retained. He further indicates that these points should be hierarchical and should be assigned weights (1-5), so that cartographers can easily select appropriate critical points for reduced scale representations. Marino (1979), indicates in her study (mentioned above) that a hierarchy of significance existed between the points selected by the respondents. On this assumption, Jones and Abraham (1987), proposed a method of data storage, so as critical points are identified by the Douglas-Poiker algorithm and subsequently stored in scale-related levels. At the highest level of representation, the smallest subset of critical points are stored, while less important points are stored at the lower levels, and are used in conjunction with the critical points higher up the data structure for producing more detailed representation (i.e., for larger scale representations). Muller (1987b) questions the validity of the hierarchical approach. He argues, that in certain cartographic situations points selected for small scale representations can not always be a subset of those used in larger scale representations. Thapa (1987) also points out that critical points detection is useful only for data compression and line generalisation when the change in scale between the original and generalised maps is modest. He further presents an original finding in which he indicates that at large scale reductions, the concern is to capture the basic shape of the line and not necessarily its critical points. His observation supports the discussions which are presented in Chapters 2 and 4; namely, that the cartographer's approach to preserving shape is not focused on absolute positions of particular points, but rather around these positions. Paradoxically, Thapa (1987, 1988) devised a simplification algorithm on the above assumption for generalisation purposes, claiming that the superiority of the algorithm lies in its ability to perform data reduction and line generalisation in one step. As noted in the previous chapter, line generalisation can not be solved by a single process such as simplification, since the maintenance of feature shape at high level representations is a context-dependent process.

Douglas and Poiker (1973) comment upon the possibilities of feature-based line simplification, as this would "...come closest to duplicating the task as performed by an experienced cartographer" (p. 114). Monmonier (1986) suggests that features can be assigned tags to their individual points, and that these points can be aggregated into specific geomorphologic features, which could, then, be simplified by an appropriate algorithm. He claims that this approach would maintain the geographic aspect of lines at reduced scales.

To summarise, many digital line simplification algorithms are based on the assumption that geographic lines consist of critical points and can have varied levels of importance levels. In order to preserve those points objectively, various approaches are employed, and the one that has been most widely used is the one which considers the entirety of the line according to a user-defined tolerance.

### *3.3.2 Culturally Critical Points:*

Jenks (1981) was the first to introduce the term of culturally critical points. He explains that certain points have to be preserved during simplification; since the points often bear geographical significance (e.g., economical, political, and cultural). During a manual simplification, the cartographer consciously considers retaining such points, since he is aware of the importance of these points to the information of the map. In the digital mode, these points may or may not be preserved, depending on their position and coding (labelling) during the process of digitising. Whyatt (1991) indicates that culturally critical points have received little research attention in digital cartography. He suggests that the preservation of such points can be performed through the use of database techniques. In fact, database routines can easily detect points such as intersections, but according to Whyatt the problems of a complex nature may only be resolved through the development of intelligent routines. Whyatt describes these routines as being able to consider the contextual and geometrical contexts of the lines being simplified. Although, it is true that, cultural points are a problem facing cartographic line simplification as well as generalisation, the solution by Whyatt is an ideal solution and remains a research goal that can be applied to both processes.

### 3.4 Evaluation of Digital Simplification Algorithms

As Jenks (1979, p.220) emphasises, "my concern is not for the future of computer cartography but for the quality of the maps computer cartographers produce. Quality comes, not from machinery, but from cartographically logical and soundly based algorithms". Jenks (1979) was one of the first to question the products of digital simplification algorithms. In a later date, Zoraster *et al.*, (1984) provided an extensive study of digital simplification, displacement, and smoothing algorithms. They point out that the relative merit of individual algorithms had not been studied in great detail. They indicate that it is difficult to measure the quality of digital simplifications. They, also, comment upon the requirement for producing visually pleasing results by algorithms. The complex question of what constitutes a cartographically sound algorithm, has resulted in many algorithms being devised, the proponents of each of which claim success. But, almost, all the authors of these algorithms have paid no consideration as to the distinction between the process requirements of data reduction, simplification, and generalisation. McMaster (1986) proposes that three aspects of algorithm evaluation should be addressed. They include:

1- What is the measured planimetric accuracy of the simplification?
2- What is the perceptual accuracy of the simplification?
3- What is the cost efficiency of the simplification?

These three evaluation requirements are regarded as being an account of both objective (mathematical) and subjective (perceptual) measurements. Whyatt (1991) points out that the factors that inhibit the simplification process, tend to influence the thoughts of researchers who have developed measurements for the purpose of evaluation. He explains that the nature of the manual manipulation upon line information during generalisation and how it is important to perform that in strict accordance with accuracy guidelines as specified by the mapping agency. Preservation of accuracy necessarily implies the preservation of shape, and the key methodology to preservation of feature shape is the maintenance of its critical points. Therefore, algorithms such as the Douglas-Poiker algorithm which tend to maintain such points are most likely to cause least displacement (McMaster, 1987a).

### *3.4.1 Quantitative measures (Mathematical):*

McMaster (1986, 1987a, 1987b) is a pioneer in the quantitative aspects of research. He developed thirty measures for evaluating the efficiency of the simplification process. These include both single attribute measures (e.g., length) and displacement or comparative measurements. He claims that single attribute measures can be used to determine the mathematical characteristics of a line before and after simplification, whilst displacement measures can be used to quantitatively compare a simplified line with its original. McMaster (1987b) used six of the original thirty measures for evaluating the performance of nine simplification algorithms. These measures are described below. He noted that there was not a significant difference between most of the algorithms tested when using two single attribute measures: percent change in angularity and percent change in curvilinearity. However, four algorithms were identified as "superior" on the basis of the four displacement measures used: total areal of displacement, total vector of displacement, areal index, and vector index. According to McMaster (1986), these six measures out of the thirty measures are statistically found to be the most significant, while others are shown to be redundant. These measures are:

1. **Percentage change in the number of co-ordinates** is not particularly useful by itself, but can be used in conjunction with other measurements. "A combination of co-ordinates per inch with another measure enables a useful standardisation in making comparisons across lines" (p. 115).

2. **Percentage change in standard deviation of the number of co-ordinates per unit length** measures the consistency of co-ordinates along a line indicating whether a simplification algorithm has uniform density, after removal of co-ordinates, in relation to the original.

3. **Percentage change in angularity (PCANGLE)** evaluates reduction in the irregularities of a line after simplification. A better simplification algorithm should retain more of the larger angular changes. This measure is mathematically expressed as the sum of consecutive vectors on the simplified line divided by this sum on the base line:

$$\mathbf{PCANGLE = ABC} \sum_{i=1}^{m-1} \mathbf{angs(i)} \;/\; \mathbf{ABC} \sum_{j=1}^{n-1} \mathbf{ango(j)} \times \mathbf{100.0}$$

where $m$ is the number of co-ordinates on the simplified line, $n$ is the number of co-ordinates in on the original line, **angs** is the angle of change between two consecutive vectors on the simplified line, and **ango** is the angle of change between two consecutive vectors on the original line.

4. **Total vector displacement per inch (TLVD)** evaluates the perpendicular distance between disregarded co-ordinates on the original line and the newly calculated simplified segment. It is expressed as follows:

$$\textbf{TLVD} = \left(1n \sum_{i=1}^{nv} \textbf{vls(i)}\right) / \sum_{j=1}^{n-1} \textbf{slo(j)}$$

where $nv$ is the number of vector displacements between a line and its simplified version, **n** is the number of co-ordinate pairs on the original line, **vls** is the length of an individual vector segment, and **slo** is an individual segment length on the original line.

5. **Total areal displacement per unit length (TAPD)** is computed as the sum of all displacement polygons standardised by the length of the original line. It is expressed as:

$$\textbf{TAPD} = \left(1n \sum_{i=1}^{np} \textbf{as(i)}\right) / \sum_{j=1}^{n-1} \textbf{slo(j)}$$

where $np$ is the number of polygonal displacement between a line and its simplified version, **n** is the number of co-ordinates on the original line, **as** is area of an individual polygon, and **slo** is an individual segment length on the original line.

6. **Percent change in number of curvilinear segments (PCCS)** is assumed to provide details on the nature of the reduction the number of curvilinear segments during simplification. It can be expressed as follows:

$$\textbf{PCCS} = \sum \textbf{a} \; / \; \sum \textbf{b} \times 100.0$$

where **a** is the number of curvilinear segments on the simplified line, and **b** is the number of curvilinear segments on the original line. Unlike the angularity measure (**PCANGLE**), curvilinearity evaluates the generalised trend of the digital line.

McMaster (1986) states that "hypothetically, better simplification algorithms, as a general rule, will eliminate a significant number of co-ordinates on straight line segments and, simultaneously, will retain a greater number around curves, resulting in a large difference in average density" (p. 109). In his research, McMaster suggests that tolerance band based algorithms (i.e., global algorithms that consider the entirety of the lines being processed) produce the most accurate representations of lines in simplified form. He indicates that the Douglas-Poiker algorithm is both mathematically and perceptually superior" (p. 109). McMaster's research has been criticised for two reasons. Firstly, he based all of his research on the study of lines in isolation. Secondly, he repeatedly used the same relatively simplistic set of test data (Whyatt, 1991). Furthermore, Whyatt (1991) indicates that "mathematical measures may not be used to measure the quality of highly simplified output. Visual evaluative techniques must be used to measure the performance of algorithms in this context" (p. 218).

Muller (1987a) utilised the theory of Fractals as an alternative mathematical evaluation of simplification algorithms. Others such as Dutton (1981), Buttenfield (1985), and Goodchild (1980a), have commented upon this approach on the basis that a cartographic line exhibits statistical self-similarity when measured at various map scales. Hence, the fractal dimension of self-similar lines should be preserved after simplification. Some researchers, such as Carstensen (1988), argue that geographic lines do not exhibit self-similarity in the natural world; hence fractal dimensionality is not preserved during the process of manual simplification. In his study, Muller (1987a) measured the fractal dimensionality of two lines that had been simplified by a number of different algorithms, and found that the original lines and their simplified counterparts had different fractal dimensions. Muller, therefore, concluded that the measure could not be used to measure the quality of simplified representation.

### *3.4.2 Perceptual Evaluation:*

Very few researchers have exclusively addressed the issue of perceptual evaluation of digital simplification algorithms. Marino (1979) indicates that a number of critical points must be maintained during simplification in order to preserve the general character of the original line. White (1985) conducted a study where respondents were asked to identify a fixed number of critical points on three lines so that "perceptual base lines" were created from the most commonly selected points. These base lines were, then, compared to their equivalent digital simplified versions produced by three algorithms: $n^{th}$-point elimination algorithm, a perpendicular algorithm, and the Douglas-Poiker algorithm. Four measures were used in this comparison. They are as follows:

1- Common point comparison.
2- Areal offset.
3- Mean number of times respondents judged particular points to be critical.
4- Visual comparison.

White concluded that the Douglas-Poiker algorithm produced the best "generalisation"; 45 % the original set points were selected by both the study participants and the algorithm. However, none of the algorithms could match points denoted as critical by respondents. White (1985) suggests that algorithms should be generated with more attention to map user's perceptions of importance than to the algorithm's more mechanical efficiency. This conclusion is in contravention of the basis upon which she conducted her study; since she compared simplification algorithms with manually generalised lines. As can be noted from her statement above, she inappropriately referred to the results of the algorithms in her study, as generalisation. It is important to note, that all of those who have quantitatively commented upon the quality of the digital simplification algorithms, have, also, included perceptual evaluations of those algorithms.

### *3.4.3 Economic Considerations:*

There has been little research interest in the issue of algorithm efficiency in terms of computing or CPU time. Morrison (1975) briefly compared the Brophy smoothing algorithm with the Douglas-Poiker algorithm in terms of reading and reducing the Sardinia coastline at two different scales. His analysis was inconclusive; although the Brophy algorithm appeared slightly faster than the Douglas-Poiker routine. Roberge

(1985) developed an algorithm -a modification of the Reumann-Witkam algorithm- and compared it with the Williams routine (Williams, 1978). His algorithm proved more efficient at multiple data reductions. McMaster (1986) points out that evaluation of simplification algorithms based on CPU efficiency is an important area for future research. Zoraster *et al.* (1984) state that an algorithm that performed "perfect" line simplification would be unacceptable if it required an excessive amount of computer resources. However, such a statement can not be without reservation; since so long as cartographic quality is concerned, the additional cost of utilising such algorithms would be insignificant, given the ever increasing quality of both the computer software and hardware.

### 3.5 Summary

It is widely acknowledged that there is a need for well-defined objective bases for producing cartographically acceptable digital simplifications. While the objective measures address the quantitative aspects, such as how much detail to retain or remove, the subjective measures address the qualitative aspects, such as perceptual or aesthetic properties. Therefore, introducing well-defined rules can not be without underlying subjective elements. This thesis proposes that these subjective elements should be addressed first during the process of evaluation of digital simplification, since it is a presentational problem. Such a proposal is supported by the fact that much of the available objective rules are only derived from what once were subjective judgements. Thus, evaluating digital simplifications should, first, adhere to subjective criteria, and only then, to quantitative measures. By a quick review of the literature on this topic, it can be noted that all the previous evaluative studies have had their limitations (Whyatt, 1991). According to those evaluations, algorithms appear to have varied or/and contradicting merits and limitations. Depending on the evaluation context, the variability in the processes of implementation and interpretation of those algorithms by many researchers makes it quite difficult, in many situations, to build much confidence on each others conclusions, hindering the way towards achieving sound results in the field of formalisation of this method. This thesis maintains that each evaluation should seek to systematically answer the following questions:

1) *What is the role of the cartographic process itself within generalisation, in this case simplification?*
2) *What is the aim of the underlying design specifications of the digital routine?*
3) *Is there a cartographic quality in the resulting digital product?*

Several impediments towards an effective evaluation of the digital simplification algorithms can be summarised as follows:

1- Misconception and confusion of cartographic generalisation terms, as between simplification and generalisation by most of those who have actually conducted evaluation studies of line simplification.

2- All previous evaluations have applied the algorithms to rather simplistic and isolated test data; that is the wider cartographic context has not been considered.

3- Variability in the interpretation of the purpose of the algorithms (Visvalingam *et al.*, 1991), and as a consequence, the results tend to be inconclusive or misleading.

4- Variability in the objectives of cartographic generalisation. Different map purpose and scale requires appropriate generalisations; hence the complexity inherent within generalisation processes makes it difficult to be grasped and broken into rules that can be easily translated into digital formats.

Unless the role of each generalisation process in the traditional realm is clearly understood, the advance towards a cartographically acceptable digital quality will never be realised. This study aims to first provide a cartographic context within which simplification can be explored, understood, and where possible defined, so that the available digital simplification procedures can be appropriately tested. The following chapter is, therefore, designed as an empirical study aimed at examining the process of line simplification within generalisation more closely. Guided by the three questions proposed above for an effective evaluation, the widely used Douglas-Poiker algorithm is chosen to be tested in Chapter 5 on the context of its design specifications, while this algorithm combined with another smoothing routine are cartographically tested in Chapter 6.

# CHAPTER FOUR: AN EMPIRICAL ANALYSIS OF CARTOGRAPHIC LINE SIMPLIFICATION

Further to the theoretical background in Chapter 2, this chapter attempts to present a practical example of an *in-depth* study of the role of line simplification within generalisation. This study aims to explore some factors affecting line simplification during manual generalisation. Many factors are known to influence generalisation as a whole (e.g., Keates, 1989), but the three key factors studied here are: feature character, map purpose, and mapping techniques. Due to the nature of cartographic generalisation, it is important at the outset to point out the difficulty in isolating either simplification from other generalisation processes or isolating the relative influence of the factors affecting line simplification. However, this should not negate an attempt to unravel this complexity. The relationship between simplification and other generalisation processes is the focus of the test. It is aimed specifically at exploring the process by which cartographers conceive and execute line simplification during traditional cartographic generalisation sessions; hence it does not attempt to present an exhaustive account of the generalisation processes.

## 4.1 Data description and test objectives

To investigate feature character as a factor in simplification requires a digital database with multiple line types derived from multiple source scales. Furthermore, the investigation requires one or more feature types for examination in two or more study areas. For this purpose, naturally occurring lines as opposed to culturally occurring line features are chosen from upland and lowland terrains. The upland headwaters of a drainage basin will have many first order streams making high complexity in the drainage system (e.g., Leopold, *et al.*, 1964; Gordan, *et al.*, 1992). An upland stream usually tends to flow round obstacles formed by outcropping resistant rocks, creating relatively complex fluctuations or abrupt indentations. In contrast, a river in a lowland area wanders in a series of sweeping meanders over a broad almost level valley. A river in such an area has been described as being in its mature stage, as compared to the first order streams which represent the youthful stage (Monkhouse, 1975). Therefore, lowland and upland areas were examined to yield study areas 1 and 2 which show contrasting line types, the one made up of broad sweeping forms and the other of more frequent changes of direction, and generally greater complexity. The influence of map purpose on line simplification is to be investigated through a discussion of a map of specific map purpose (route mapping). The effect of mapping techniques by different mapping organisations is addressed through comparison of similar types of lines in maps produced by different mapping agencies. The fourth and final study area was mapped by the Swiss mapping agency, and is compared with others.

Table 4.1 and Figures 4.1 and 4.2 show descriptions and locations of the study areas. The first area is a lowland area where the height is around sea or marshland levels, and over which the selected features (naturally occurring lines) are characterised by rather low frequency sinuosity. The area depicts a portion of the estuary of the River Humber, in the east of England, covering an area of 20 by 20 km. Three scales of representation are considered: 1:25,000, 1:50,000, and 1:250,000. The prime objective of the test is to examine the effect of line feature type, as opposed to any other factor, on simplification during a typical generalisation for topographic mapping.

The second area is chosen to represent relatively high complex line forms (drainage network). In this case, the area is located in the Lake District, in the north west of England, where elevation ranges from 200 to 3000 feet, covering 10 by 10 km. Three scales of representation are considered for this area: 1:25,000, 1:50,000, and 1:250,000. The test should explore how such complex features are typically simplified compared to the simplification required for less complex ones such as those in the first study area. It should be pointed out that complex terrain does not imply that all the geographic features within it are necessarily complex. However, drainage features were found to be more intricate and complex than in study area 1.

The source map series of the first and second study areas is the Ordnance Survey, whereas the source of the road map series of the third study area is the Geographer's A-Z Map Co. Ltd., which are based upon the Ordnance Survey maps. All maps of the three study areas date between 1974 and 1991. The third area is located in the West Midlands, England, covering an area of 30 by 50 km. The area is characterised by a relatively low level terrain, around 400 feet. This study area is chosen to examine principally the effect of map purpose on simplification; in this case, the map purpose is route mapping. There are four scales of representation for this area: 1:158,400, 1:200,000, 1:730,000, and 1:1,000,000.

The fourth study area is located in the north of Switzerland, showing a small portion of the River Aare with an area of only 7.5 by 6 km (Figure 4.2). By comparison with other study areas, this test examines the effect of mapping methods within different mapping organisations on simplification. Four scales of representation are considered for this area: 1:25,000, 1:50,000, 1:300,000, and 1:500,000. These maps are from the topographic map series of Switzerland, produced by the Office Federal de Topographie, Bern, Switzerland (1964-1994).

| Study Areas | Location | Type of Features and Coverages | Scale Original | Scale Generalised | Source |
|---|---|---|---|---|---|
| **Study Area 1** | East of England | One coverage: different line and area features. | 1:25,000 | 1:50,000<br>1:250,000 | Ordnance Survey, Britain<br><br>(1974-1991) |
| **Study Area 2** | North West of England | Drainage Coverage | 1:25,000 | 1:50,000<br>1:250,000 | |
| | | Transport Coverage | 1:25,000 | 1:50,000<br>1:250,000 | |
| **Study Area 3** | West Midlands, England | Transport Coverage (roads) | 1:158,400 | 1:200,000<br>1:730,000<br>1:1,000,000 | Geographer's A-Z Map Co. Ltd., production based on Ordnance Survey maps |
| **Study Area 4** | North of Switzerland | Drainage Coverage | 1:25,000 | 1:50,000<br>1:300,000<br>1:500,000 | Office Federal de Topographie, Bern, Switzerland<br>(1964-1994) |
| | | Transport Coverage | 1:25,000 | 1:50,000<br>1:300,000<br>1:500,000 | |

**Table 4.1: General description of the line features obtained for this study from four study areas.**

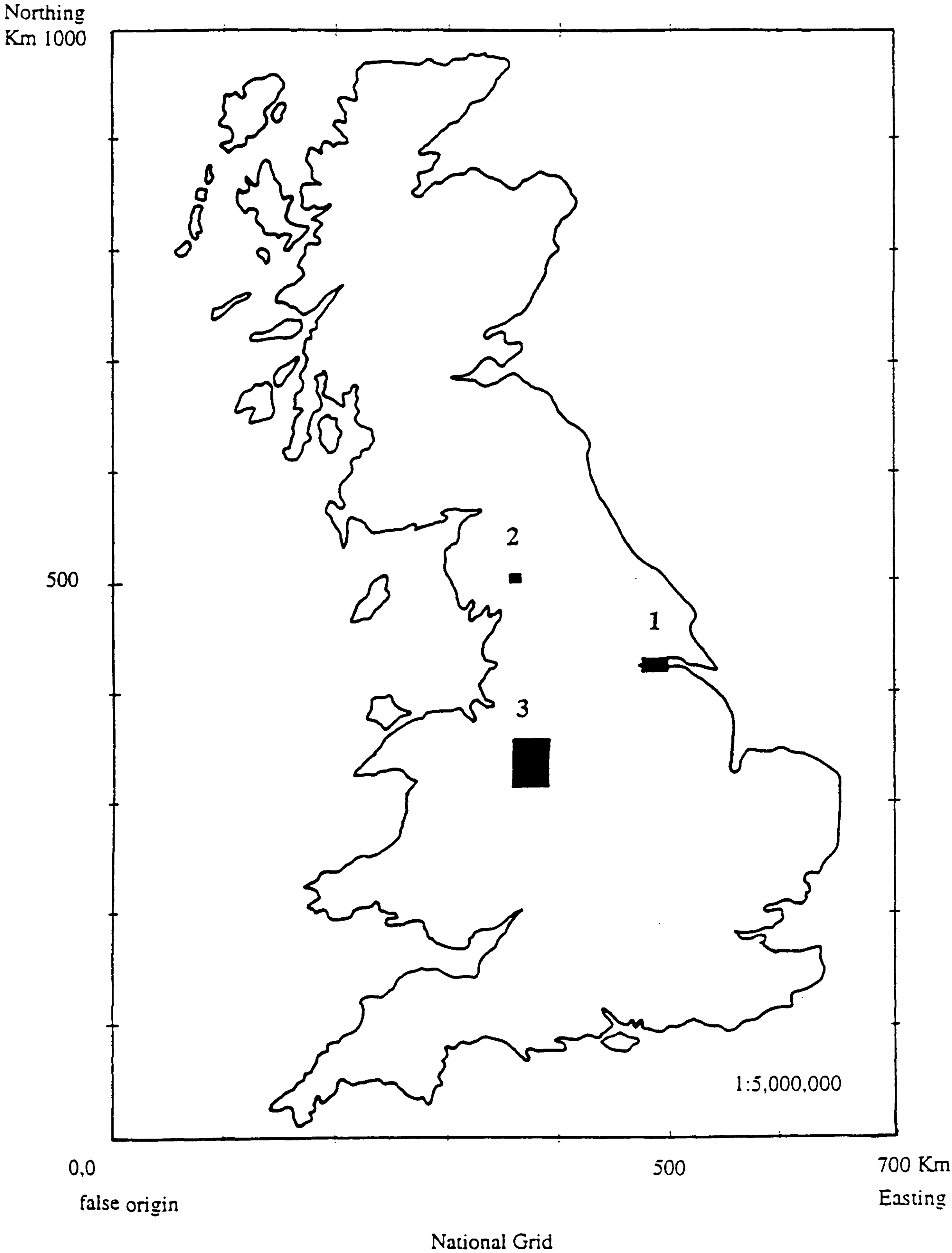



Figure 4.1: Locations of study areas, 1, 2, and 3 (source: Ordnance Survey, Southampton, Great Britain).

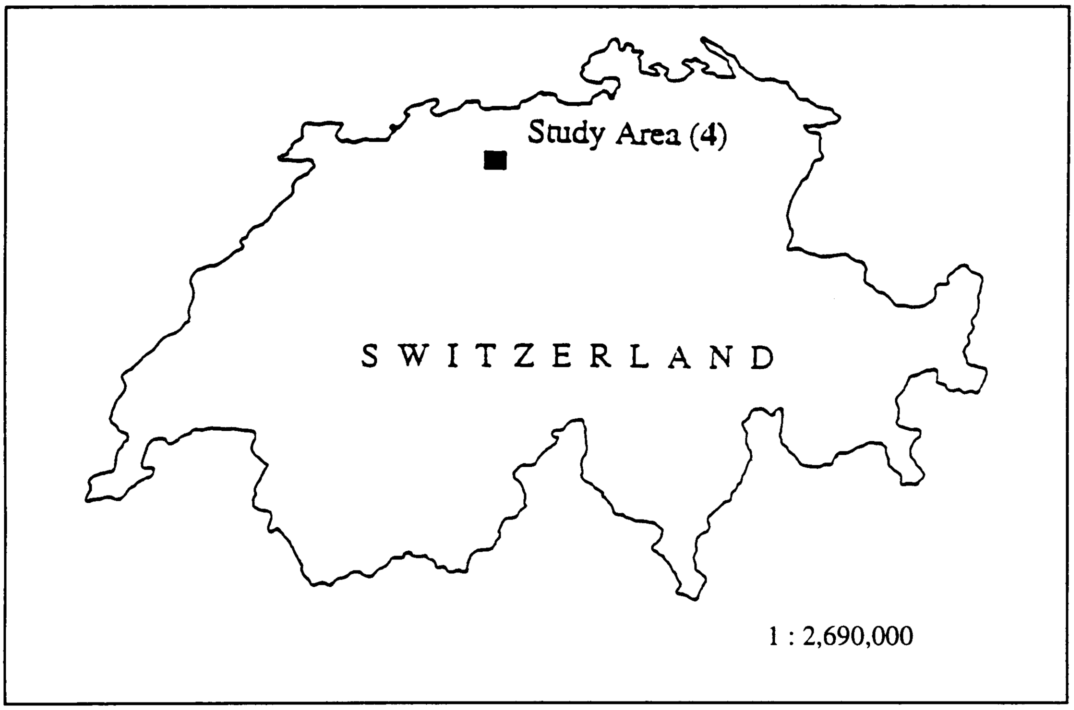



Figure 4.2: Location of the fourth study area, North of Switzerland (source: Office Federal de Topographie, Wabern, Bern, Switzerland).

## 4.2 Methodology

In order to facilitate comparison of different study areas and scales, all the selected features have had to be converted into digital forms through manual digitising. The selected line features were digitised manually at all the scales mentioned above. In some cases, that required two layers to be created, for example one for the drainage networks and another for the transport networks. The Arc/Info GIS mapping software was utilised for both the digitising and analysis processes. For all the four experiments, the largest scale maps of 1:25,000 (for study areas 1, 2, and 4) and 1:158,400 (for study area 3) were referred to as the original and chosen to be the reference against which other smaller scale maps (generalised) were compared. Although, the generalised maps might not be derived directly from these large scale maps, this methodology should suffice for the purpose of this study. This is for two main reasons: 1) these large scale maps are either slight modifications of larger scale maps (e.g., between 1:10,000 and 1:25,000 scales) on which the reduced scales were then based, or the derived scales were directly based on them; and 2) this particular analysis is not aimed at presenting a detailed account of every positional divergence between line features at each scale, but it is primarily concerned with the general perceptual change in shapes of line features at different, decreasing scales.

As seen in Chapter 2, simplification of line features involves the process of selecting the characteristic features of lines according to the mapping context which may involve both the number and details of lines. The present analysis focuses on the processing of line details of the selected line features, whereas the process of removal of the lines on the generalised maps is classified as "*selective omission*" (Keates, 1989). It is the purpose of this study to examine how details within line features are processed in relation to the mapping context, especially the focus of the next chapters will be processing line details digitally, whereas the process of selecting and deleting the entire line can be seen as another process in either the traditional or digital realms.

*4.2.1 Analysis and types of measurements:*

Whilst Arc/Info, used for the analysis, provides powerful tools for mapping purposes, it provides an effective means for conducting analyses that require considerable calculations in terms of line properties such as length and area, and an effective visual communication, in terms of comparison analysis. In such systems, maps at different scales can be overlaid for comparison, so that feature displacement can be analysed both qualitatively and quantitatively. Furthermore, this medium of analysis provides a powerful visual comparison technique by which line features can be overlaid and displayed in different colours. In this particular analysis when these coloured lines show complete or some conformity they would appear in black or a dark colour, otherwise the colours assigned to the features would be displayed. The comparisons took place at the scale of the generalised maps.

Whilst it is difficult in practice to isolate generalisation processes, processes such as smoothing, exaggeration, and displacement can usually be associated with the process of manual line simplification. That is, during manual simplification of line details, the cartographer smoothes out line details (hence, causing displacement) and where necessary performs a limited exaggeration in places where the line curves become smaller than the width of the drawing pen. On the other hand, exaggeration and displacement can be totally independent of line simplification, and they are processed in relation to the mapping context. Simplification can be described as either being systematic or consistent, in terms of execution, indicating a visually balanced simplification, or it can be inconsistent or variable, and only through displacement between the simplified and original line features one can judge the simplification type. Displacement is therefore given particular attention and such a distinction will be exemplified where possible in the analysis. For practical and analytical purposes, generalisation processes other than simplification (associated with the type of displacement and exaggeration discussed above) are discussed and listed under the term generalisation, while simplification describes the processing of line details, and displacement accounts generally for the displacement (either independent or associated with simplification) assessing the type of simplification. It is again important to bear in mind that such a distinction between generalisation processes is difficult in the practical sense.

While it is important to study visually the process of line simplification at each generalised scale and explore its interrelationship with other generalisation processes, it was necessary to account for some quantitative measures such as the total line length. Measuring the total line length (in metres) of the digitised line features was rapidly and accurately facilitated by Arc/Info. Because the study primarily focuses on qualitative measures, quantitative measures other than total length were ignored. First, quantitative methods lack the ability for addressing the question of the interaction between complex processes that are guided by subjective rules such as generalisation processes. Second, quantitative measures such as areal displacement can be used to generally account for the displacement which occurs between generalised cartographic features irrespective of the sources behind their displacement, they are best suited where one particular type of feature such as lines are to be compared and thoroughly studied within a certain context of analysis, for example, studying different digital line simplifications. Third, unlike other measures, the total line length measure is best suited to quantitatively observing the effects of line simplification, since simplified line features necessarily undergo reduction in their lengths. It is acknowledged that line features have varying lengths at different scales. Penck (1894) showed that the coastline of Istria varied in length from 223 km on the 1:75,000 map to 105 km at the 1:1 million map. He explained that this decrease was due to the removal of variations or indentations within of the coastline. Other coastlines have been measured; for example, Hakonson (1978) measured the shorelines of Scandinavian lakes. Mandelbrot (1967, 1982) has proposed the fractal dimension (D) to measure this property. For a line, D can be between 1, representing a completely smooth line, and 2 for a completely complex line that covers an area. Geographical line features fall in between these two values. The significance of this concept is to give a measure of line complexity independent of scale. This implies that lines at different scales have the same D value and are self-similar. On the other hand, line features that are not self-similar are termed as scale-dependent. Some researchers argue that Fractals may be the "most important single parameter of irregular cartographic features" (Goodchild and Mark, 1989, p. 267). Others argue against this idea (e.g., Krantz, 1989; and Batty, Fotheringham and Longley, 1989). Goodchild (1980a) and Muller (1986) showed that D for geographical features is variable across scales, contrary to what has originally been expected. Muller (1986) found in his study that the fractal dimension decreased with scale; i.e., at smaller scales D was smaller as the line features were smoothed out. Given disagreement over the importance of fractals when measuring lines, total line length has been chosen as a measure for the line features in this study, and it is expected that line features at smaller scales will have simplified (smoother) shapes. Regarding the presentation of the analyses in this chapter, for each analysis, the results and are presented first followed by the related figures.

### *4.2.2 Digitising error:*

It is essential not to ignore digitising errors during examining the generalisation effects especially when the perceptible deviation between the generalised and original features is the focus of the study. When estimating the reliability of vector digitised line data, it is commonly assumed that the true position of the line (source) lies within the error band of the digitised line (Blackmore, 1984). This band is also known as the Perkal epsilon Band (Perkal, 1966). As Goodchild (1988) indicates, researchers proposed uniform, normal, and bimodal distribution of error across this band. Whyatt (1991) suggests that, this concept provides some basis for estimating the position of the true line at locations between the digitised points. He indicates that, at present no sound basis for modelling this error exists, and workers tend to assume that a bivariate normal distribution of error exists when estimating the position of a true point. In the context of line simplification, "positional accuracy is less important than the relative position of points describing the shape of features along the line" (Whyatt, 1991, p.96).

Line features were digitised for this study at various scales, giving differing complexity, and symbolisation. Their symbolisation ranged in width from approximately 0.10 to 0.35 mm. According to the *error band* concept and based on similar studies (e.g., Joao, 1994), the lowest digitising accuracy value for the 0.10 mm line would be equal to half this width, since at worst a digitised point would be placed on the edge of the line rather than on the centre. Therefore, it is reasonable to assume that all point data for this line may only be captured to an accuracy of +/- 0.05 mm (map unit). In order to provide an average accuracy for the whole digitised features, the mid point of the range of all the line widths is chosen to provide an average digitising accuracy. Therefore, the representative average value between the 0.10 and 0.35 is 0.225 mm, and hence, the average accuracy of the digitising performed in this study is +/- 0.1125 mm. Since the comparison was conducted at the reduced scales (generalised maps), the perceptual displacement that was greater than this error was, therefore, of particular significance, and has a large likelihood of being caused by the generalisation process. Furthermore, it was important to give an account of what might be the digitising error given similar environmental conditions. For this purpose, an empirical test was performed to indicate generally the variability of the manual digitising of single line features exercised in this study. That required a repetition of a digitising process three times, at different time intervals. That was considered to be helpful in giving an indication of the extent of the uncertainty of the manual digitising performed. In this case, line features of the first study area were digitised three times (Figures 4.3a and 4.3b). The figures show that the displacement between the three maps at the scale of 1:50,000, is variable from one place to another. At some locations, the displacement is hardly perceptible or imperceptible (owing to the dominance of black as opposed to either red, green, or cyan in the figures), whereas at others it appears perceptible. At one particular place the displacement becomes exceptionally large, this can be seen in the displacement in Figure 4.3b between the lines representing a railway, in the top left of the map. In general, it can be concluded that the resultant displacement is within the range of the digitising accuracy value calculated above. It is expected that such a displacement will obviously be further imperceptible at smaller scale representations. It is crucial to remember that digitising error variability associated with tracing single lines (e.g., coastlines), like the exercise here, is expected to be less than the error from digitising double lines (e.g., lines representing roads) like the line features in the third study area, where the centre lines are approximated. In this work, the digitising errors associated with the road coverages are larger than the estimated general error calculated above (+/- 0.1125).

There are some caveats about calculation of digitising errors in general terms. It is essential to bear in mind that during the analysis of generalisation effects the deviations between features would be compared to a general estimation of digitising errors. This implies that within the deviation all possible digitising errors (calculated, known and unknown) are included. Two facts should, therefore, be born in mind; 1) the digitising error might be smaller or greater than this defined average, and 2) if there was no perceptual displacement at the reduced scale, it is not an indication that there was no digitising error. It is important to indicate that digitising errors might have led to another complication. This is usually in the form of rounding errors that result from the process of mistransformation of the digitised maps from the local digitising table units (e.g., mm) to a geographical co-ordinate system. The error is characterised as being systematic, however, if such an error occurred that does not automatically indicate a digitising error, since it may well be due to an actual deviation resulting from the generalisation. The analysis should therefore indicate where possible the likelihood of the occurrence of the digitising errors based on comparison with the actual source digitising materials. Given this nature of digitising errors (at least the method of calculation of average figures), one cannot actually ignore reference to the source digitising materials just because a figure was included in the analysis. Thus, the significance of the calculation of the digitising performed is to indicate that where the perceptible deviation is larger than the calculated error of +/- 0.1125, it would be reasonable to deduce the error from the deviation in order to provide fair statement about the generalisation that took place, although reference to the original digitising materials is necessary. But where the displacement is roughly this error, especially the one associated with road maps, the deviation between features is as likely to be due to generalisation as to digitising errors and only reference to the source digitising materials should determine how much digitising errors could actually (though generally) contribute to the deviation between features.

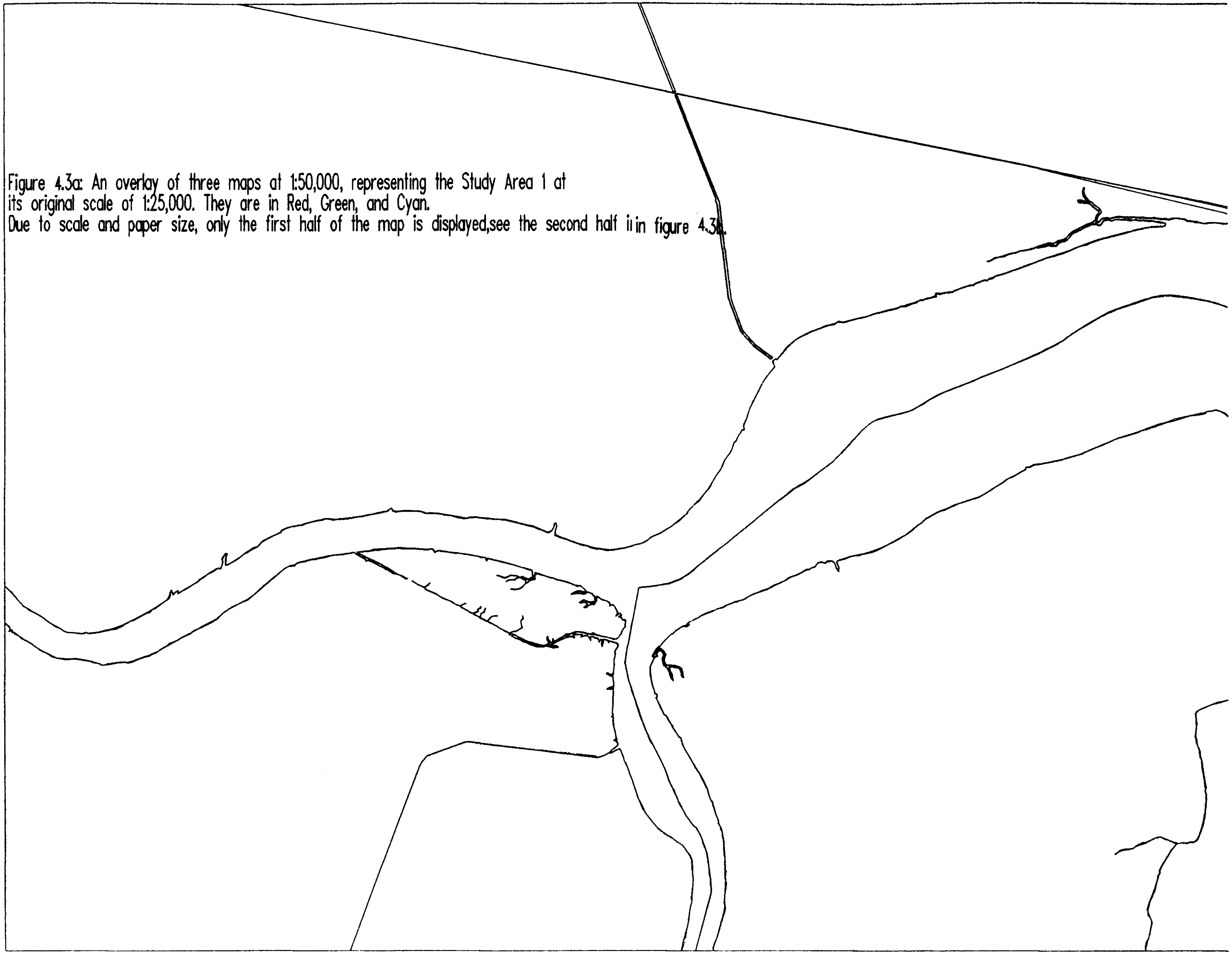


Figure 4.3a: An overlay of three maps at 1:50,000, representing the Study Area 1 at its original scale of 1:25,000. They are in Red, Green, and Cyan.
Due to scale and paper size, only the first half of the map is displayed,see the second half ii in figure 4.3b.

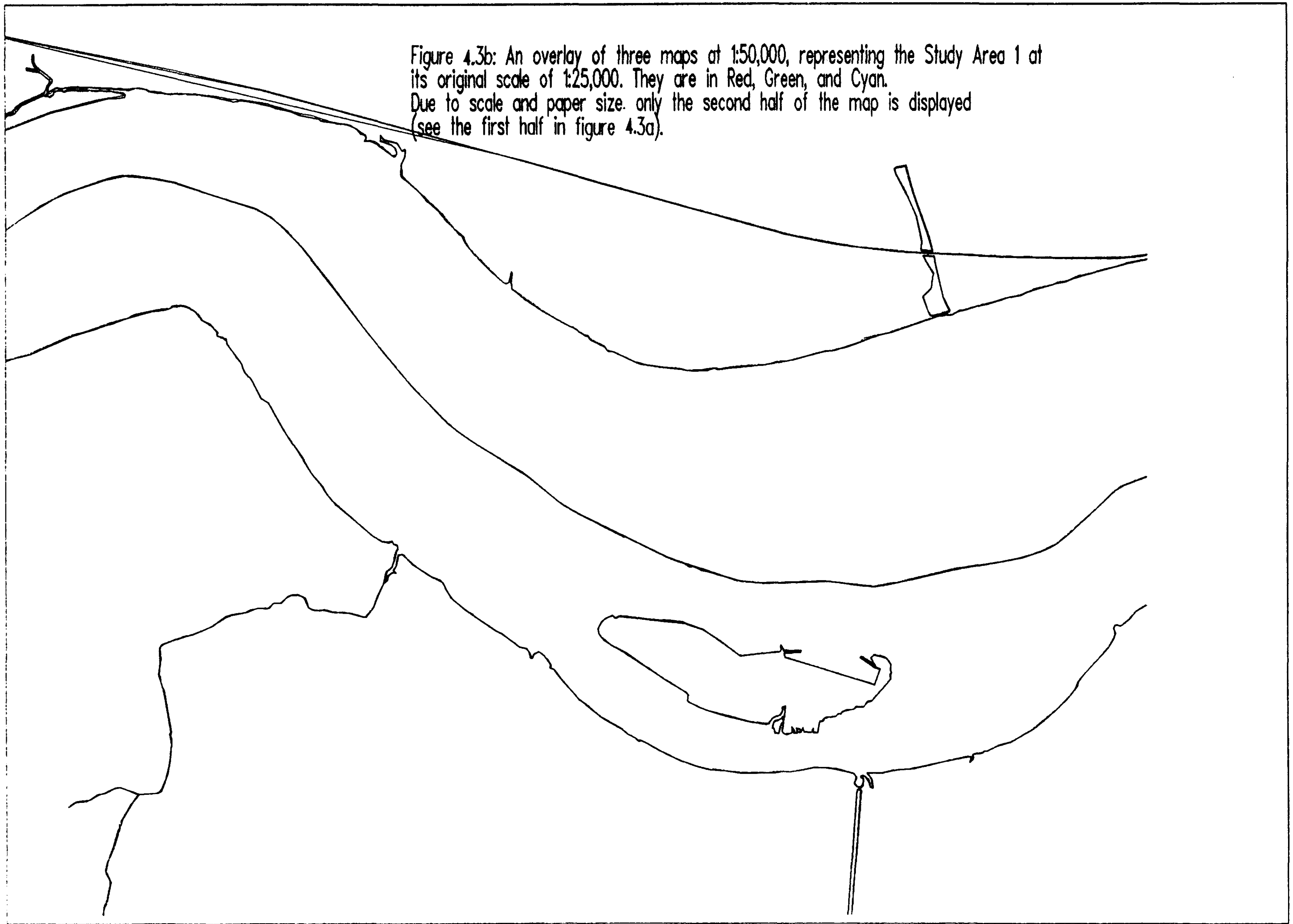


Figure 4.3b: An overlay of three maps at 1:50,000, representing the Study Area 1 at its original scale of 1:25,000. They are in Red, Green, and Cyan.
Due to scale and paper size. only the second half of the map is displayed (see the first half in figure 4.3a).

## 4.3 Results

*4.3.1 Study Area 1:*

The selected features are shown in Figure 4.4, consisting of the estuary of the River Humber, some selected canals, streams, a railway, an island, and two other area features.

4.3.1.1 Simplification and other generalisation processes:

The 1:50,000 generalised map (in Green) in Figures 4.5a and 4.5b shows that line features were only slightly smoothed if compared to the 1:25,000 original map (in Red). The visual comparison reveals that there is a close similarity between feature details at the two scales. Table 4.2 shows, that the total line length of features was reduced by 6.53 %. Generally, reduction in the total length is a reflection of both the reduction in minor details and number of features. In this case, the reduction was due to a small reduction in the minor details, as can be seen from Figures 4.5a and 4.5b.

| Map scale | Total line length (m) | % Reduction |
|---|---|---|
| 1:25,000 (original) | 142,359.95 | - |
| 1:50,000 | 133,067.87 | 6.53 |
| 1:250,000 | 98,418.53 | 30.87 |

**Table 4.2: Total line length of features in original and generalised maps (Study Area 1).**

The 1:250,000 generalised map (in Green) shows that the process of simplification was further increased (Figure 4.6). While simplification seemed to be almost the only generalisation process that took place upon features at the scale of 1:50,000, at the scale of 1:250,000, other generalisation processes were also performed. Figure 4.6 shows such processes which include: omission (e.g., O), exaggeration (e.g., E), and displacement (e.g., D). As can be seen from Table 4.2, the reduction in the total line length was a 30.87 %. This reduction was largely caused by the reduction in the number of features, and partly by the reduction in the small details

#### 4.3.1.2 Displacement:

The feature displacement shown between the generalised maps and original was the result of various factors. However, digitising error and process error (inconsistent generalisation) were the prime source for such displacement. Visual comparison of the map features in Figures 4.5a, 4.5b, and 4.6 clearly show that in most parts the displacement exceeds the defined digitising error (+/- 0.1125 mm). Regarding the execution of simplification, the comparison reveals inconsistency which could have been better controlled. For example, the simplification of the Island (Figure 4.5b) could have had its character better preserved, since the map scale and purpose require a more faithful representation. The term of inconsistency as opposed to objectivity will be used where necessary through the whole analyses describing the process of generalisation. It refers to the variability that characterises the manual work of cartographers, as mentioned previously. The ideal simplification might be conducted in a way that the simplified line when overlaid on its original would lie in the centre of that of the original. On the other hand, deliberate and controlled displacement of features during generalisation is a quite common, yet, understandable practice. As noted above, simplification is shown to be the main process performed at the scale of 1:50,000, and because it was minimally applied upon the line details, the relatively large displacement which resulted is beyond the effect of the simplification; hence, it is not justifiable. That is, according to the level of simplification performed at this scale, the resulting displacement is, therefore, expected to be minimal, compared to the displacement shown.

One way of comparing displacement between generalised features and their originals across scales is to reproduce all features to a common scale. However, the degree of alignment between the generalised and original lines at reproduction scales and the degree of simplification performed may indicate whether the displacement is expected to be larger or smaller across the scales. Figure 4.6 shows that the coincidence between the generalised and original lines between the 1:25,000 and 1:250,000 maps is generally no better than that between the 1:25,000 and 1:50,000 maps. However, the displacement at the 1:250,000 is expected to be larger than that at 1:50,000, since the generalised line features at the scale of 1:250,000 were largely simplified and displaced. Factors such as map purpose, scale, and graphic limits required features to be displaced and exaggerated (e.g., D, and E). Over simplification at this scale led to an excessive displacement in one particular place (SD).

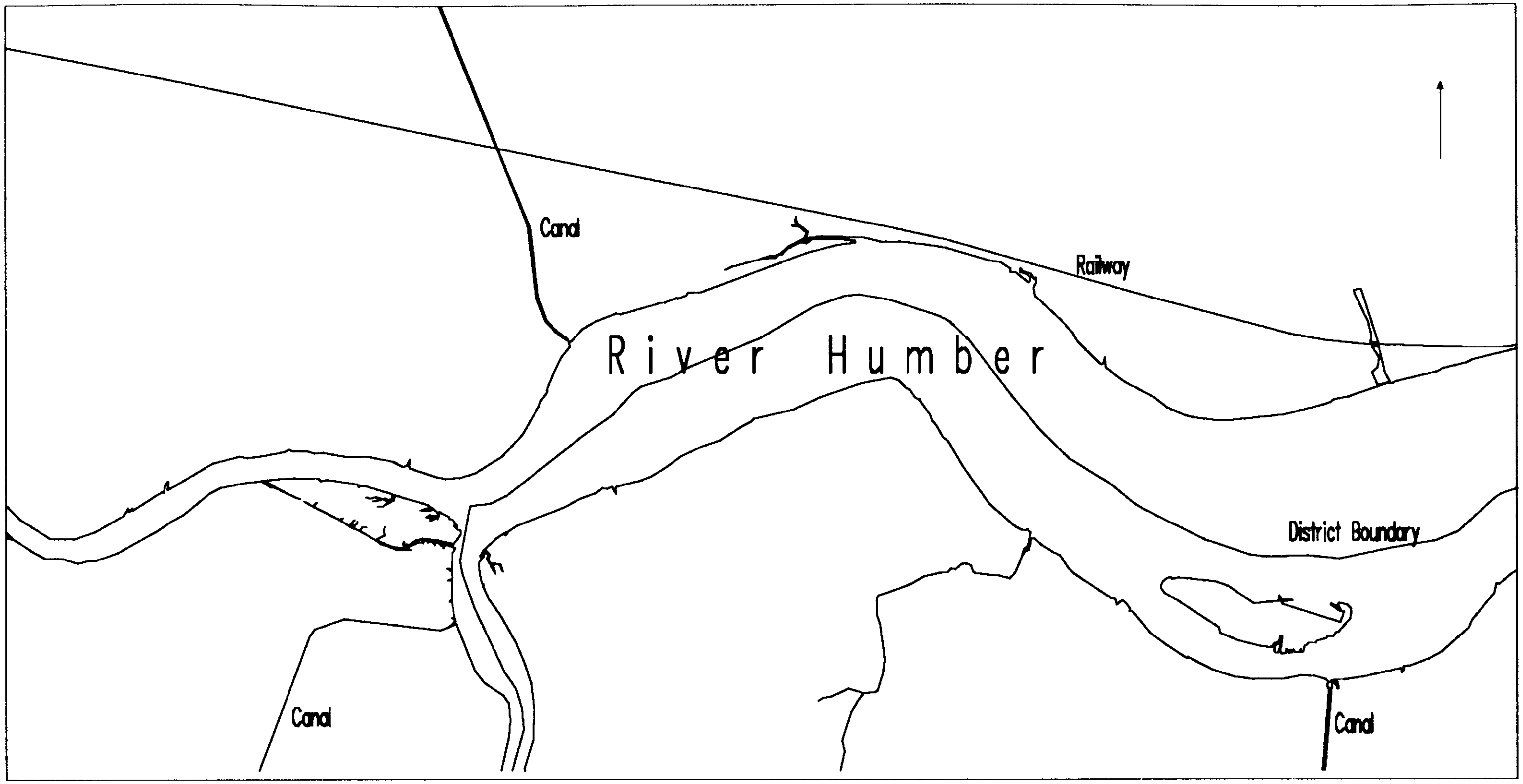



Figure 4.4: A digital representation of the study area (1). The 1:25,000 original map at 1:100,000 depicting a large portion of the estuary of the River Humber, east of England, Great Britain (OS production).

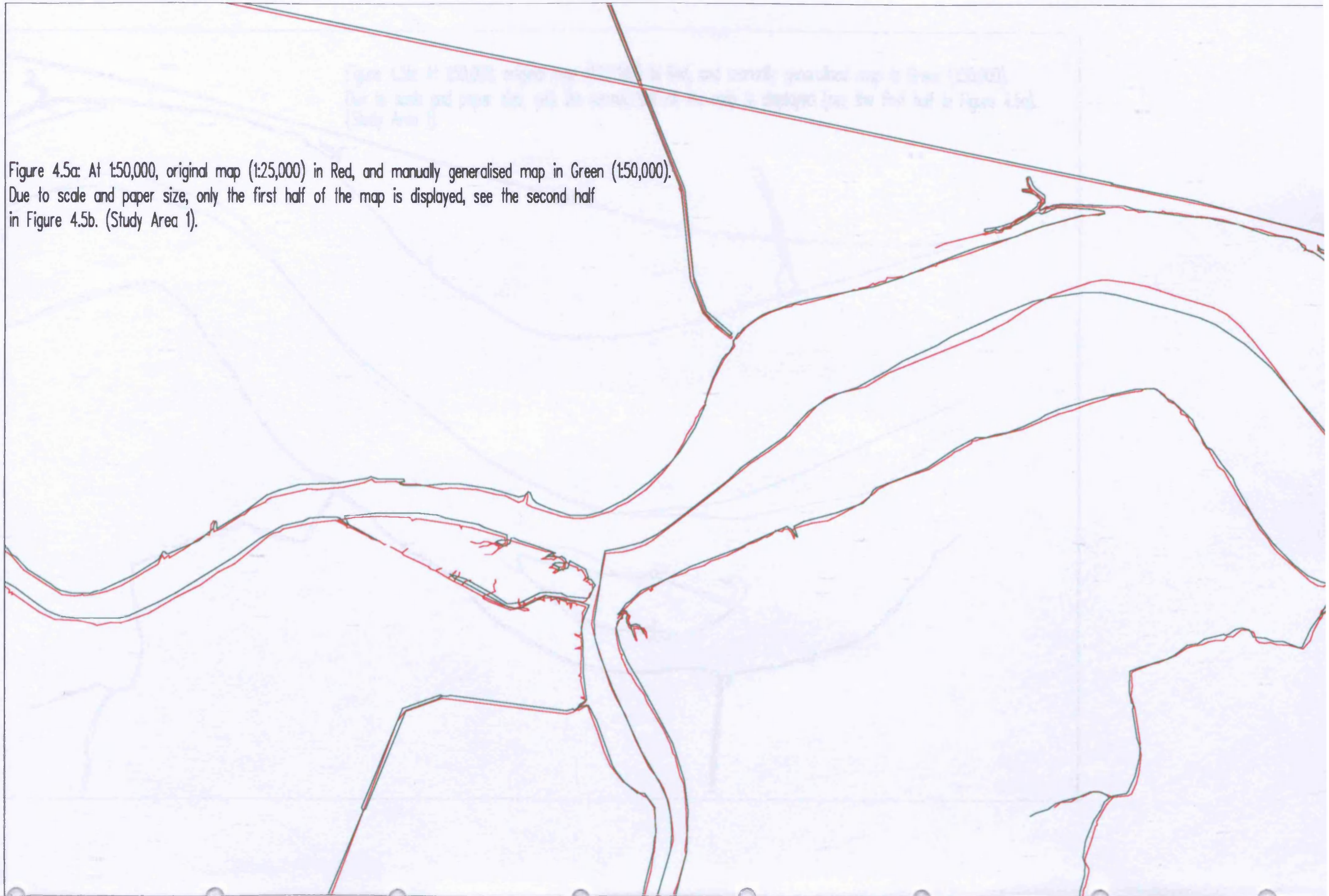


Figure 4.5a: At 1:50,000, original map (1:25,000) in Red, and manually generalised map in Green (1:50,000). Due to scale and paper size, only the first half of the map is displayed, see the second half in Figure 4.5b. (Study Area 1).

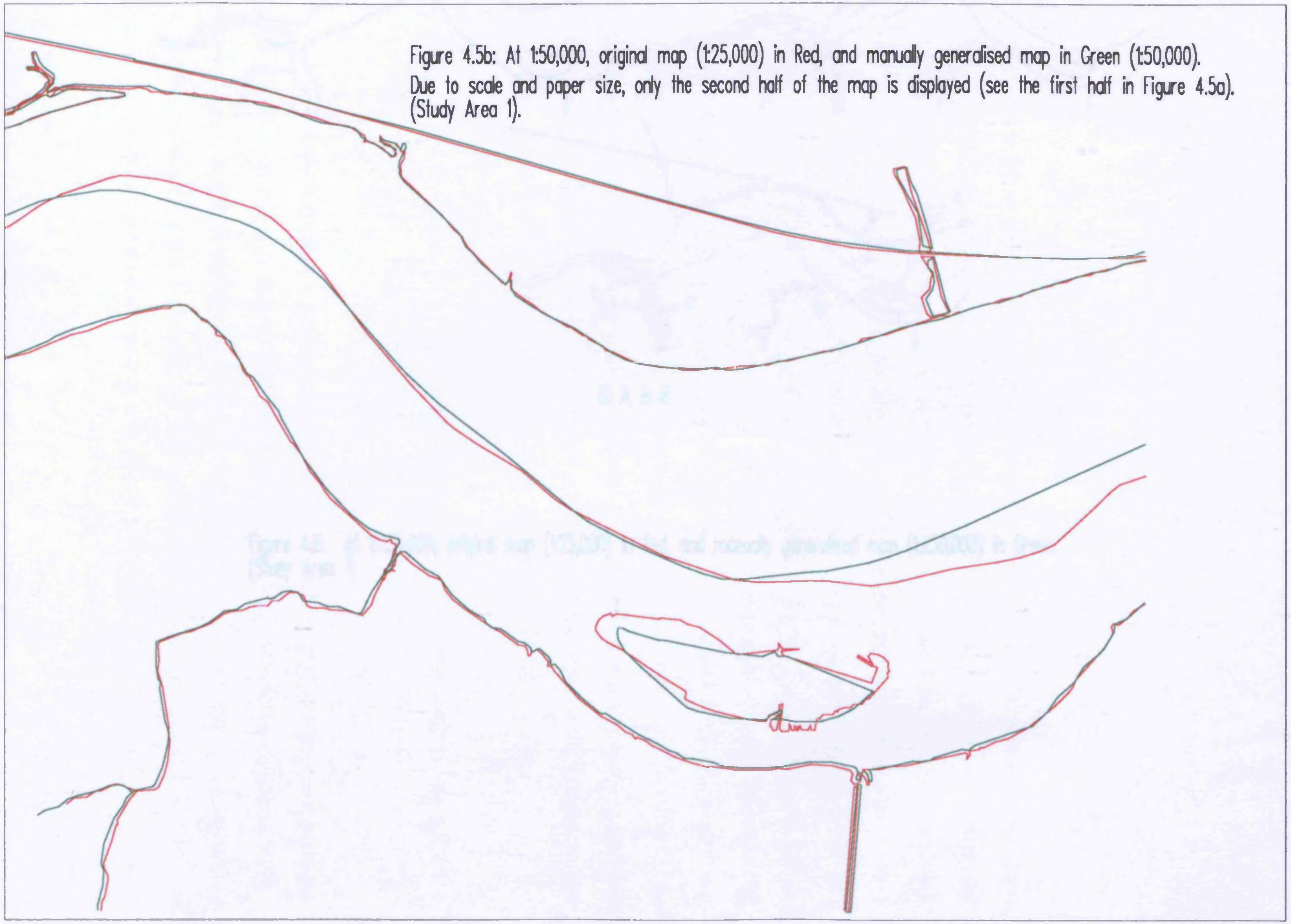


Figure 4.5b: At 1:50,000, original map (1:25,000) in Red, and manually generalised map in Green (1:50,000). Due to scale and paper size, only the second half of the map is displayed (see the first half in Figure 4.5a). (Study Area 1).

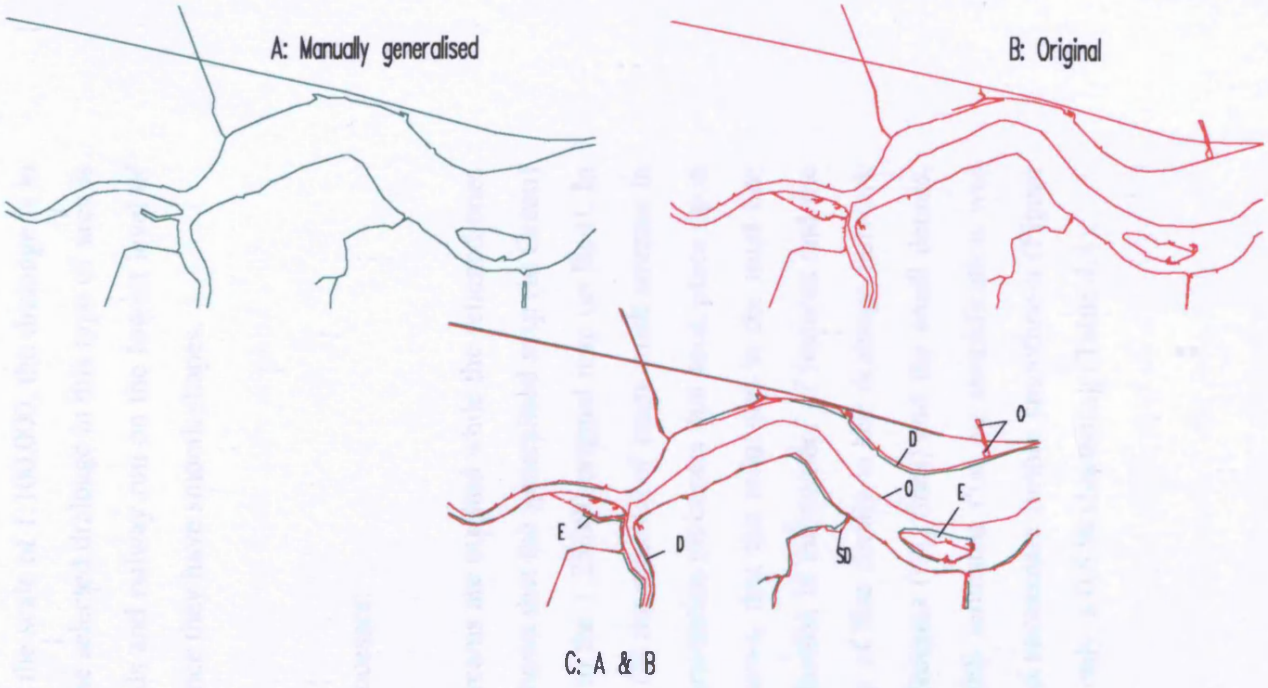



Figure 4.6: At 1:250,000, original map (1:25,000) in Red, and manually generalised map (1:250,000) in Green. (Study Area 1)

*4.3.2 Study area 2*

In this study area, three types of line features were selected: drainage, roads, and a railway. Figure 4.7 shows these features at the scale of 1:100,000, the drainage is in Green and the roads and railway in Red. While the selected drainage in this type of area is characterised by high sinuosity, the selected roads and railway run on the lowest level of the terrain where elevation is around 200 feet, hence they have smooth shapes.

4.3.2.1 Simplification and other generalisation processes:

In the 1:50,000 scale map, many streams are omitted while the retained ones can be seen to have been simplified. Figure 4.8 shows that the generalised map (in Green) appears in smoother line shapes when compared to the 1:25,000 original map (in Red). In the figure, simplification (e.g., SD) and omission (as the removal of many small streams in the generalised map) seemed to be the main generalisation processes that took place upon the drainage network at this scale. Table 4.3 shows, that the reduction in the total line length was a 45.98 %. This reflects both the reduction in the number of features and the reduction in small details. Whilst the complexity of line details in the drainage network necessitated a reduction in both the number of features (i.e., lines) and the small details, the roads and the railway at this scale (1:50,000) indicated that no simplification was performed, since their characters appeared not to necessitate further smoothness (Figure 4.10). For this reason, their total length suffered, only, a 0.6 % shortening (Table 4.4).

| Map scale | Total line length (m) | % Reduction |
|---|---|---|
| 1:25,000 (original) | 234,964.89 | - |
| 1:50,000 | 126,929.55 | 45.98 |
| 1:250,000 | 59,096.81 | 74.85 |

**Table 4.3: Total line length of drainage network (Study Area 2) in original and generalised maps.**

| Map scale | Total line length (m) | % Reduction |
|---|---|---|
| 1:25,000 (original) | 34,196.65 | - |
| 1:50,000 | 33,994.26 | 0.6 |
| 1:250,000 | 29,880.83 | 12.63 |

**Table 4.4: Total line length of transport network (Study Area 2) in original and generalised maps.**

At the 1:250,000 map, some generalisation processes such as displacement, exaggeration, omission, and simplification were performed upon the drainage network. As Figure 4.9(c) shows, all small streams which appear in the original (in Red) were removed while other features were displaced (e.g., D). There is a consistent simplification of feature details, in terms of a uniform execution, which appeared to be performed upon all retained features, as can be seen more clearly in Figure 4.9 (b). As a result of the reduction in both the number of features and feature detail, the total line length of the network at this scale was reduced by a 74.85 % (Table 4.3). The roads and railway at this scale underwent some generalisation processes; namely, omission (e.g., O), exaggeration and displacement (e.g., ED in Figure 4.11). As the Table 4.4 shows, their total line length was reduced by 12.63 %. This reduction was, therefore, caused by the omission of features, but not by the simplification (in terms of removal of line details), since the feature characters of the this coverage are essentially smooth lines.

4.3.2.2 Displacement:

On the 1:50,000 generalised map, rivers and streams underwent displacement from their original position. This can be explained in relative terms, since there were many known and unknown factors which contributed in varying degrees to this displacement. Figure 4.8 shows two types of displacement: consistent; i.e., systematic, as in (Da), and inconsistent, i.e., variable, as in (Db). The inconsistency here relates to the process of simplification in the sense that the displacement resulting from the simplification is unpredictable and hence variable from one place to another due to the process of manual simplification. Variable or inconsistent displacement is more related to the process simplification more than to an independent displacement determined by the mapping context such as symbolisation requirements, where the pattern of the displacement is usually uniform or systematic. It is visually evident that both displacements largely exceed the digitising error (+/- 0.1125 mm). Considering the nature of manual generalisation in traditional cartography, it is difficult to perform an absolute alignment between generalised features and their originals, however, such a deviation has to be controlled; i.e., be appropriately sensitive to the cartographic requirements such as map purpose, scale, and symbolisation. The consistent or systematic displacement, here, indicates a systematic error, since visual inspection of the whole generalisation context for all the features on the source paper material revealed that there was no cartographic requirement for such displacement. It is not clear whether this error was due to the process of mistransformation during digitising or due to the generalisation process. Figure 4.12 shows the importance of deliberate displacement, yet controlled, for graphic clarity during symbolisation (e.g., D); i.e., allowing for a map space for representing the generalised features legibly at the reduced scale. Simplification of the rivers at this scale largely contributed to the inconsistent or variable type of displacement, (see Figure 4.8).

The roads and railway at this scale, underwent deliberate displacement for symbolisation purposes (e.g., ED, in Figure 4.10).

Comparison of the drainage network at the two scales of 1:50,000 and 1:250,000 shows that the alignment between the generalised lines and their originals at the scale of 1:250,000 is more than that between the line features at the scale of 1:50,000. However, the displacement between the line features at the scale of 1:250,000 is larger than that at the scale of 1:50,000, as would be expected. There was some necessary displacement, as in (D) in Figure 4.9 (C), for symbolisation purposes. This was largely due to simplification. Apart from the displacement needed for symbolisation purposes (e.g., D in Figure 4.9 (C)), simplification was the prime source for the displacement between the rivers at the original and reduced scale of 1:250,000. Figure 4.9 (C) therefore shows how the retained rivers were simplified and hence displaced. Regarding the roads and railway, their displacement was largely due to the process of exaggeration, as in (ED) in Figure 4.11 (C). This figure shows how it was necessary to exaggerate and displace features in order to represent them legibly at this scale. Figure 4.13 emphasises how different features were generalised while their character and relative position were maintained.

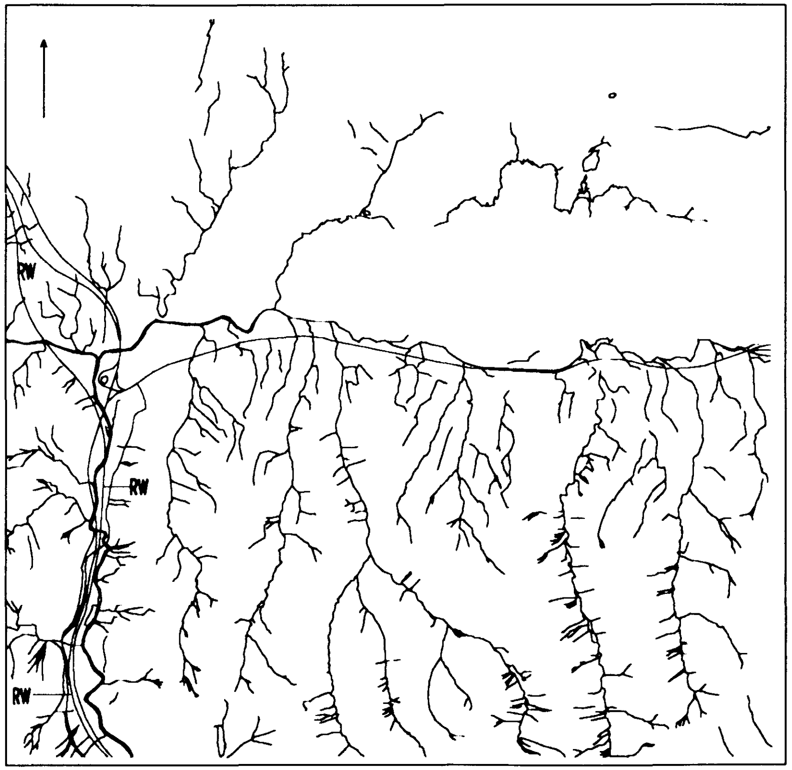



Figure 4.7: A digital representation of the study area 2, located in the north west of England, (OS production). Original map features (1:25,000) at 1:100,000, drainage is Green and roads and railway (RW) in Red.

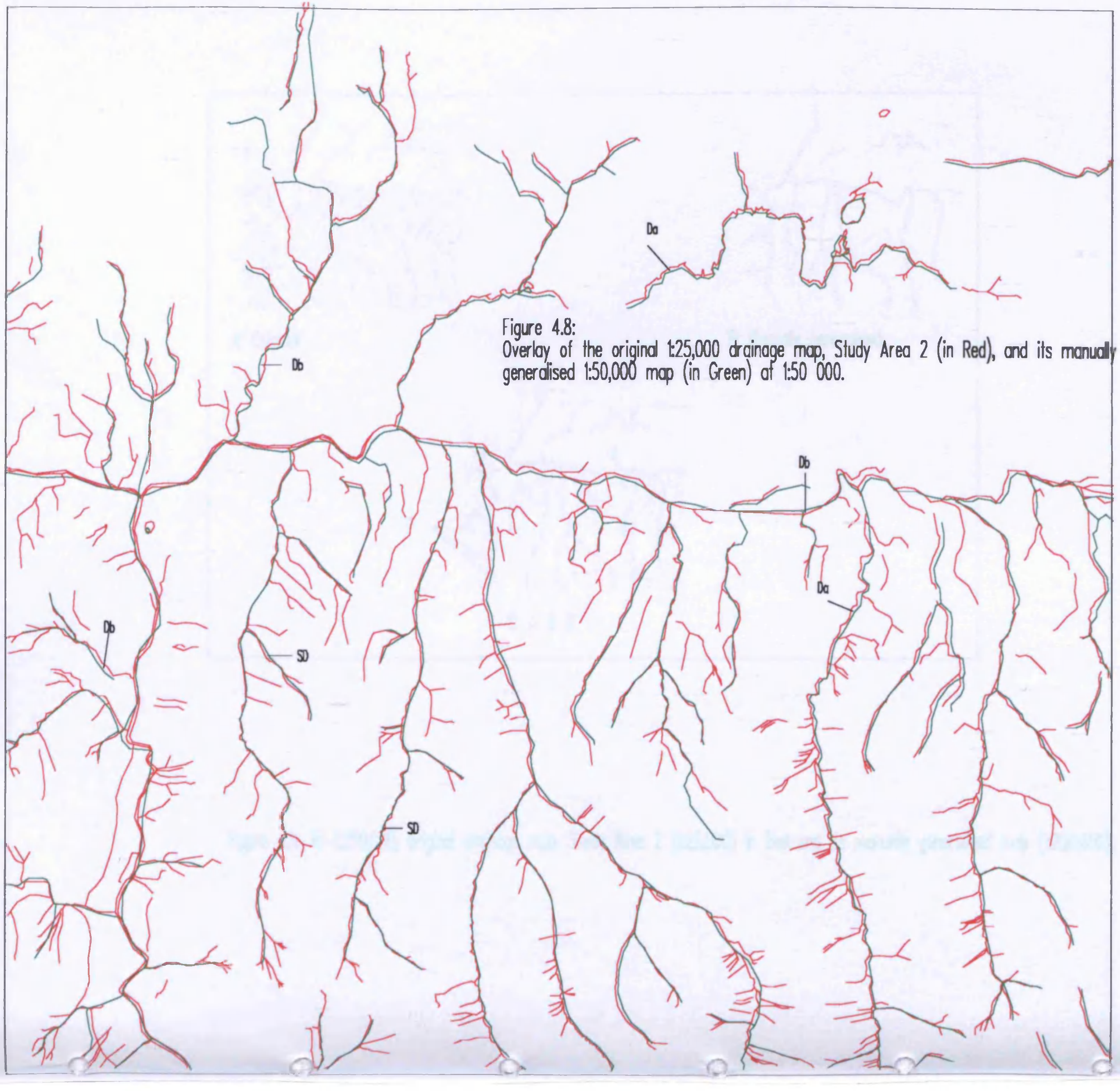



Figure 4.8:
Overlay of the original 1:25,000 drainage map, Study Area 2 (in Red), and its manually generalised 1:50,000 map (in Green) at 1:50 000.

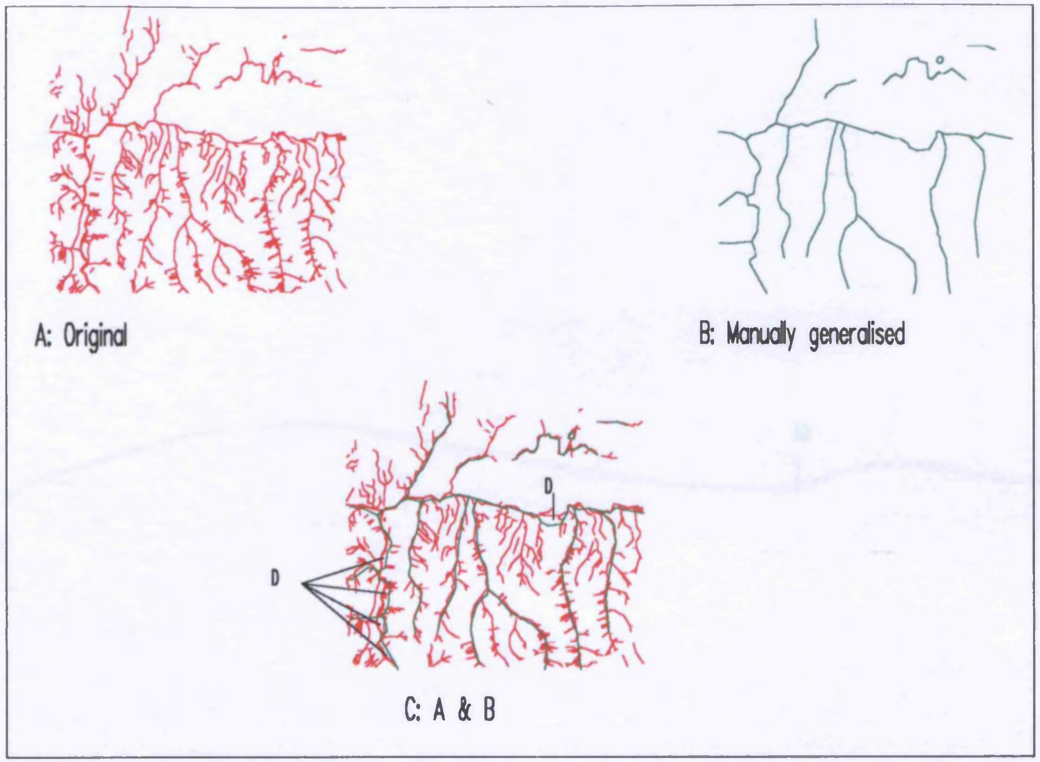



Figure 4.9: At 1:250,000, original drainage map, Study Area 2 (1:25,000) in Red and its manually generalised map (1:250,000).

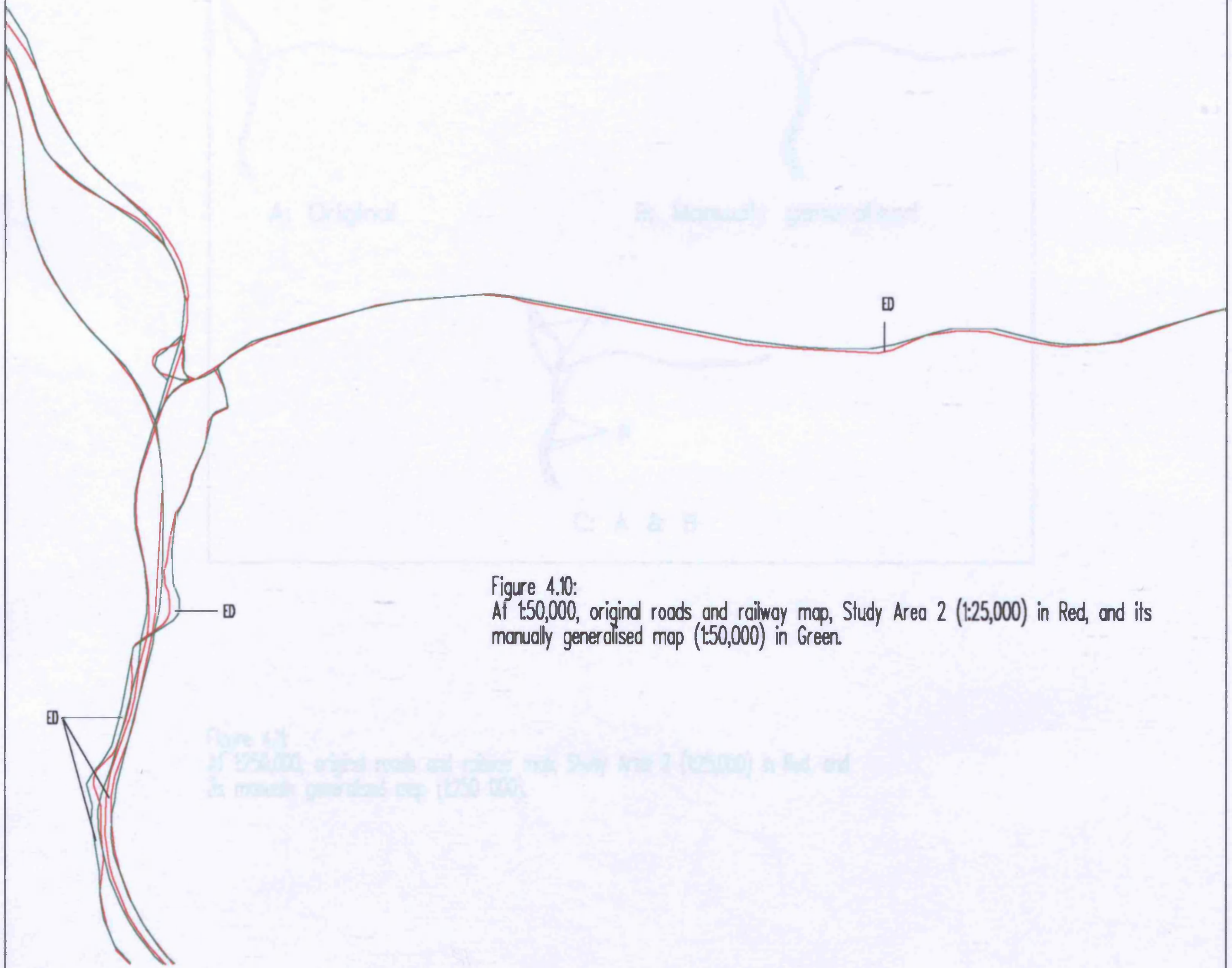



Figure 4.10:
At 1:50,000, original roads and railway map, Study Area 2 (1:25,000) in Red, and its manually generalised map (1:50,000) in Green.

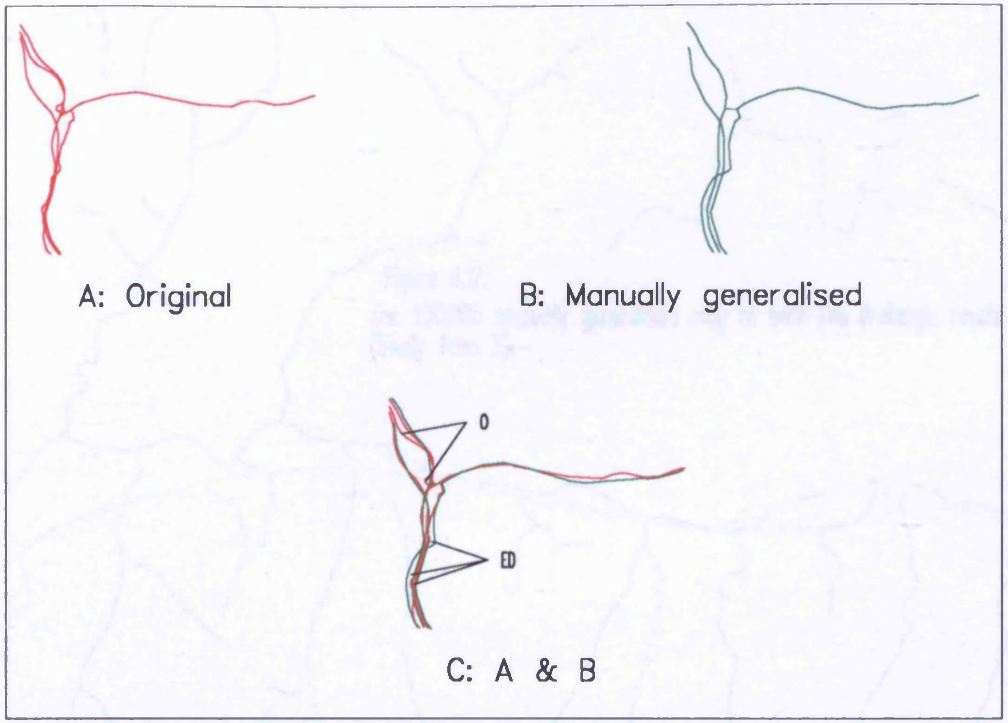



Figure 4.11:
At 1:250,000, original roads and railway map, Study Area 2 (1:25,000) in Red, and its manually generalised map (1:250 000).

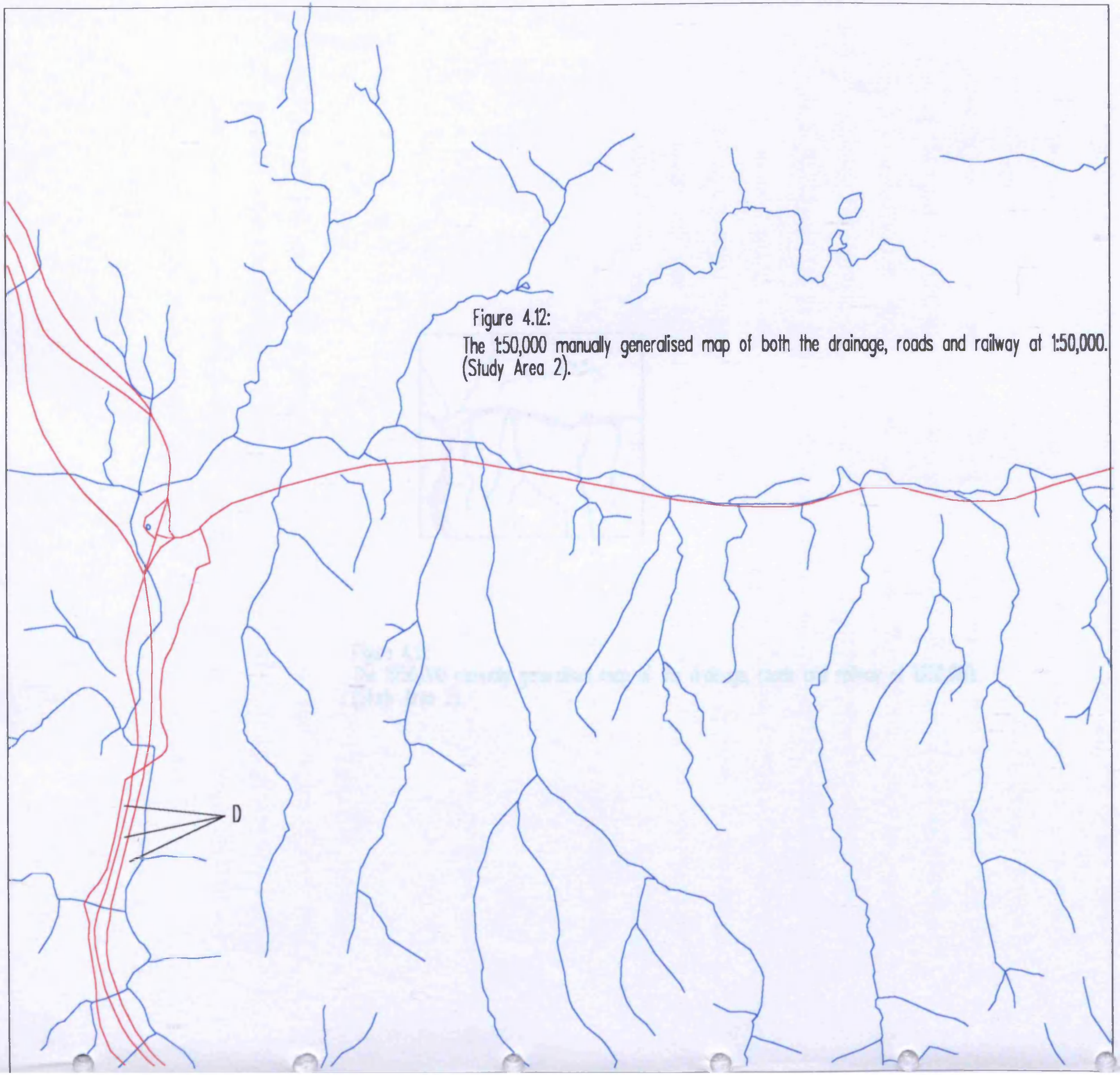



Figure 4.12:
The 1:50,000 manually generalised map of both the drainage, roads and railway at 1:50,000.
(Study Area 2).

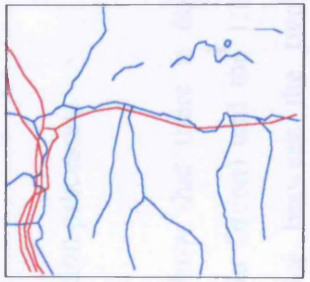


Figure 4.13:
The 1:250,000 manually generalised map of the drainage, roads and railway at 1:250,000. (Study Area 2).

*4.3.3 Study Area 3:*

The features selected are only the road network, representing three classes of roads: Motorway (M 6), Primary Routes, and A Roads (Figure 4.14). The study area does not include all the network; only some of the three types of roads that were selected. They were therefore the only ones that were traced and digitised at the scales of 1:158,400, 1:200,000, 1:730,000, and 1:1,000,000.

4.3.3.1 Simplification, and other generalisation processes:

Figures 4.15a and 4.15b show that there is only a small displacement between the 1:200,000 generalised map (in Green) and the 1:158,400 original map (in Red). This is due to the small difference between the two map scales; hence no generalisation appears to have been performed on the roads, and the observable displacement is assumed to be due to digitising error. The explanation for such a deviation is due to the process of tracing and digitising assumed centre lines within the double lines representing roads at both the two scales, although, the process of following these unmarked centrelines was conducted through a careful approximation, but any error that might be introduced in digitising can not be ruled out. The precise contribution of the digitising and generalisation processes to this displacement is difficult to quantify, but, through careful visual examination of the source paper maps and Figures 4.15a and 4.15b, it can be generally inferred that the digitising process could be the prime source behind the displacement. It is important to note that much of this type of displacement, exceeds the average digitising error defined earlier (-/+ .1125). Table 4.5 indicates that there was no reduction in the total length of features at the 1:200,000 map. Instead, the roads appeared to be longer at this scale compared to the 1:158,400 scale, being a + 0.08 % difference. This clearly suggests that there was no simplification, since the measure of total length is very sensitive to simplification.

In the 1:730,000 map, roads underwent simplification and other generalisation processes. In Figure 4.16, visual comparison between this map (in Green) and the original (in Red) reveals that there was a minimum simplification performed which resulted in displacement (e.g., SD). Other processes such as exaggeration, which led to displacement (e.g., ED), and omission (e.g., O) were performed for legibility purposes. On the other hand, there is a systematic displacement which is assumed to be the result of digitising (e.g., DD), caused by a mistransformation from the local digitising table co-ordinates to the proper geographical co-ordinate system. In general, not much simplification was performed at this small scale, indicating the importance of details in such a type of mapping. This is emphasised by the modest reduction of the total length, as a 3.8 % reduction (Table 4.5).

| Map scale | Total line length (m) | % Difference from original |
|---|---|---|
| 1:158,400 (original) | 367,191.45 | - |
| 1:200,000 | 367,466.18 | + 0.08 |
| 1:730,000 | 353,238.97 | - 3.80 |
| 1:1,000,00 | 349,588.21 | - 4.80 |

**Table 4.5: Total line length of features in original and generalised maps (Study Area 3).**

At the scale of 1:1,000,000, the roads underwent further simplification and other generalisation processes. Figure 4.17, band C, shows, that the generalised map (in Green) appeared to have undergone simplification, which resulted in displacement (e.g., SD). The magnitude of this displacement can be regarded as small if compared to a simplification required for a representation at the same scale but for another map purpose. This generalised map bears a close resemblance to the 1:730,000 map. There is a displacement which resulted from the process of digitising (e.g., DD), which appears in a random pattern. Although, it is difficult to quantify such an error, visual observation and the measurement of total line length, especially when there is no reduction in the number of features, can indicate whether simplification took place. The other generalisation processes were mainly: omission (e.g., O), and exaggeration and displacement (e.g., ED). The total line length at this scale, again, emphasises the minimal simplification performed, only a 4.80 % reduction (Table 4.5).

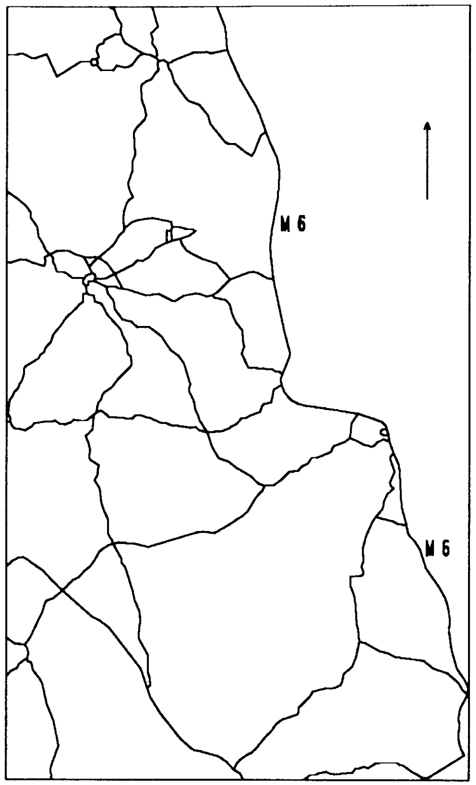



Figure 4.14: A digital representation of the study area 3. The 1:158,400 original map at 1:500,000 showing a road network north west of Birmingham, England, Great Britain (production of Geographer's A-Z Co Ltd, based on OS maps). The far right line represents part of the M 6 motorway while other lines on the left represent different roads of two classes.

Figure 4.15a: Overlay of the original 1:158,400 map (in Red), and its manually generalised 1:200,000 map (in Green) at 1:200,000. Due to scale and paper size only half of the map is displayed, Figure 3.15b shows the other half. (Study Area 3).

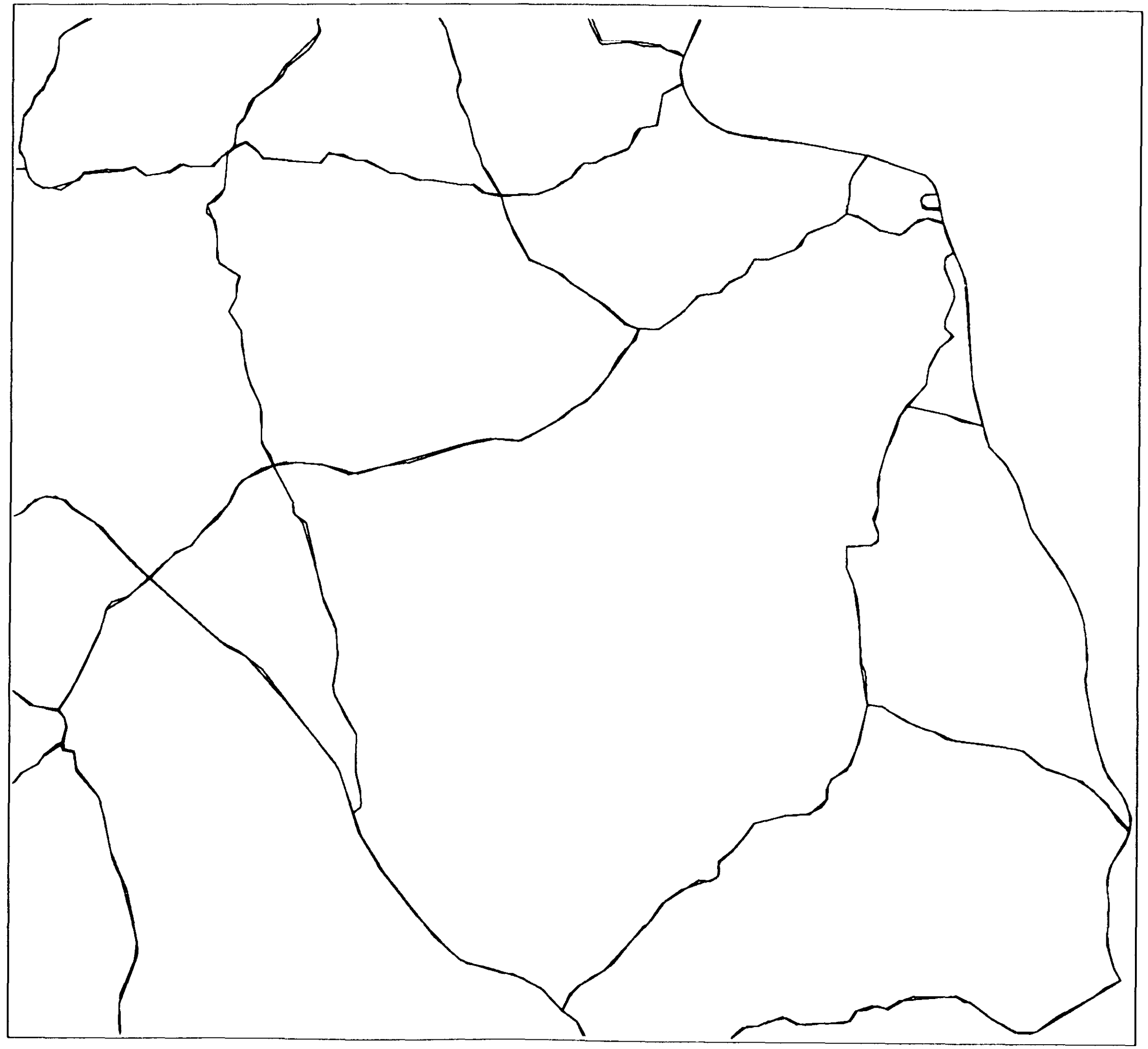


Figure 4. 15b: Overlay of the original 1:158,400 map (in Red), and its manually generalised 1:200,000 map
(in Green) at 1:200,000. Due to scale and paper size only half of the map is displayed, Figure 3.15a shows the other half.
(Study Area 3).

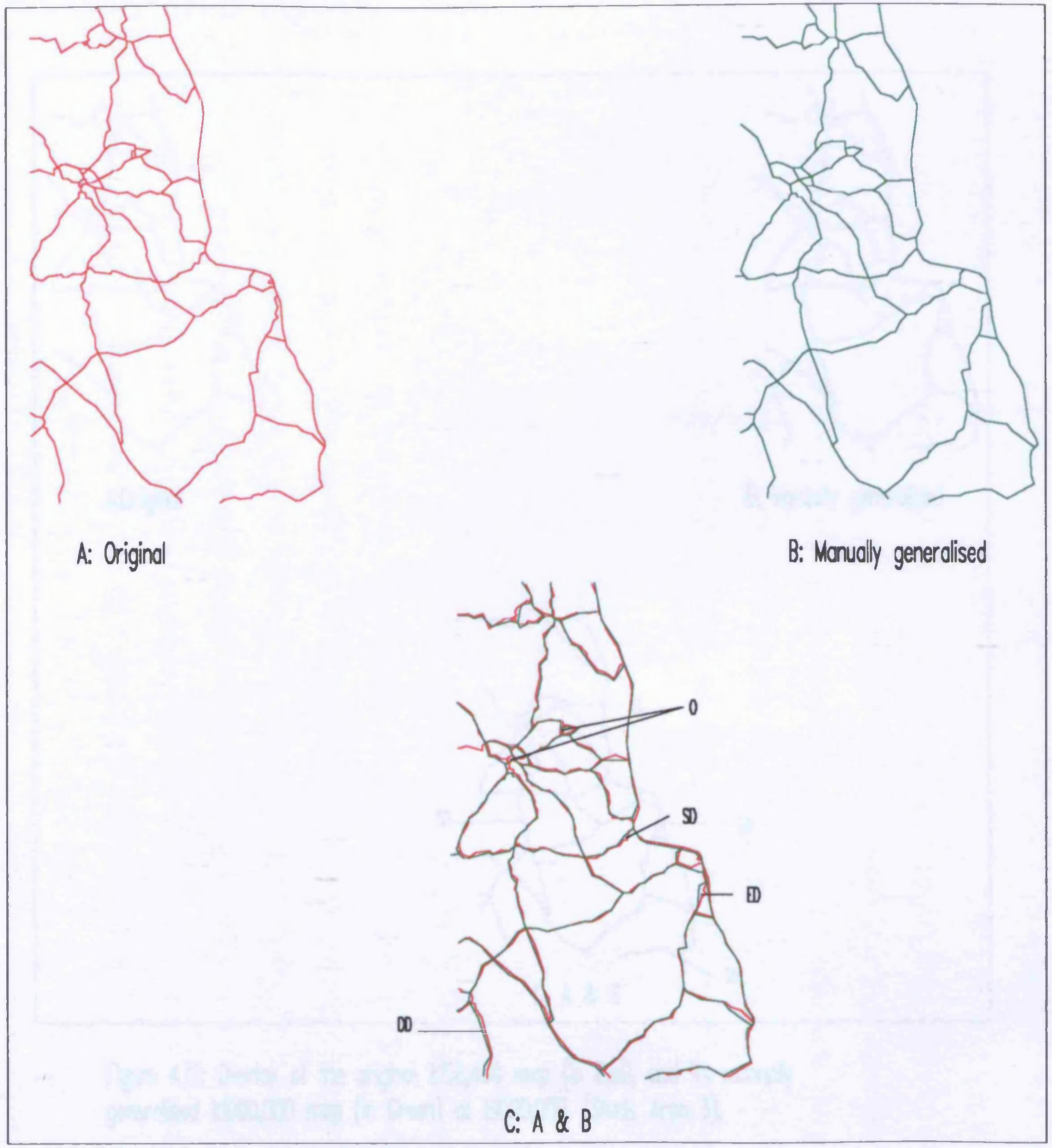



Figure 4.16: Overlay of the original 1:158,400 map (in Red), and its manually generalised 1:730,000 map (in Green) at 1:730,000. (Study Area 3).

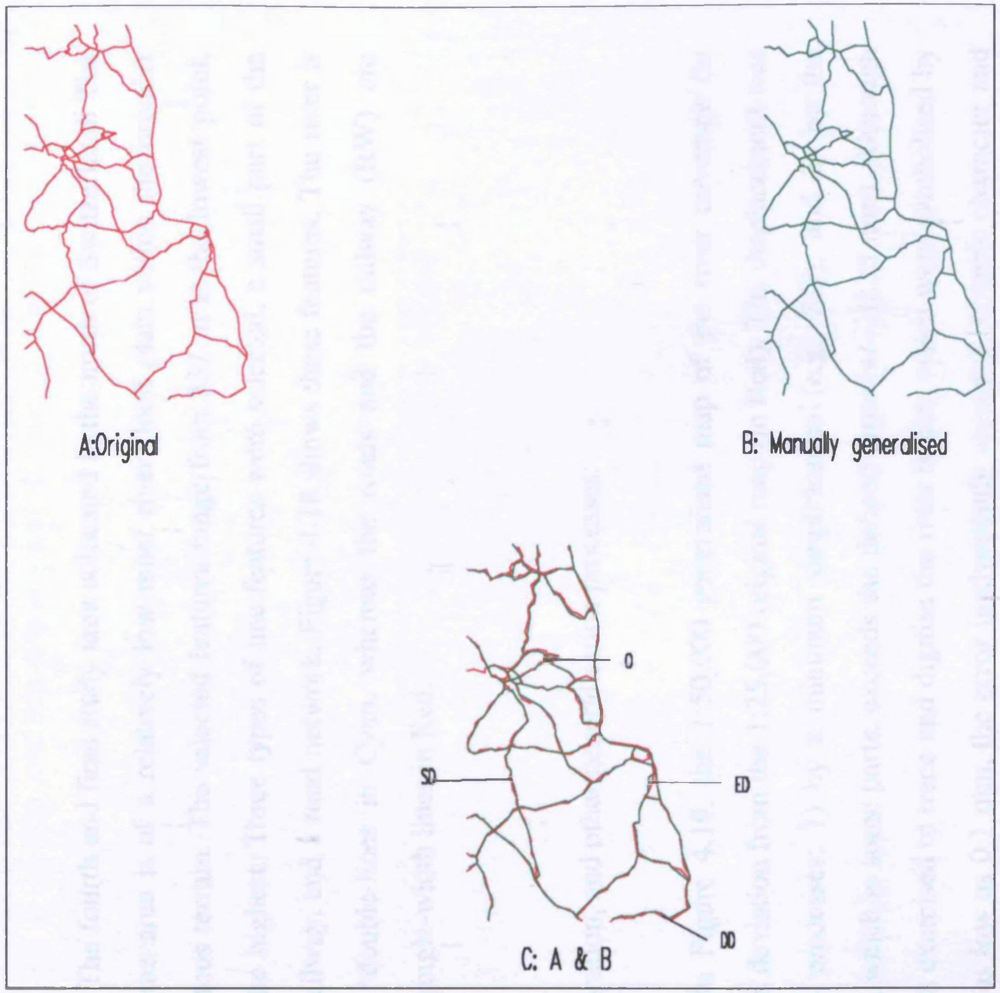



Figure 4.17: Overlay of the original 1:158,400 map (in Red), and its manually generalised 1:1000,000 map (in Green) at 1:1000,000. (Study Area 3).

*4.3.4 Study Area 4:*

The fourth and final study area is located in the north of Switzerland. The topography of the area is of a relatively low relief river flood plain within the broader Swiss mountainous terrain. The selected features range from 337 m at the lowest point, and 409 m at the highest. Three types of line features were selected, a small part of the River Aare, a railway, and a road network. Figure 4.18 shows these features. The river is represented by double-lines in Cyan, whereas the roads and the railway (RW) are represented by single-width lines in Red.

4.3.4.1 Simplification, and other generalisation processes:

In Figure 4.19, the 1:50,000 generalised map of the river coverage (in Green), shows a deviation from the 1:25,000 original map (in Red). The displacement was caused by two processes: 1) by a minimum simplification (e.g., SD), and 2) by the digitising error, which in some parts, exceeds the defined error (+/- .1125 mm). Although utmost care was exercised to trace and digitise the river banks, which were symbolised by a line width of as low as 0.1 mm, the error inadvertently occurred. Feature character and measurement of total length should make the task of observation much easier. In this case, the shape character of the river indicates no complexity in terms of minor details, and only a very small change appeared to be required. This is emphasised by the very small reduction in the total length, which was only a 0.15 % (Table 4.6). Close visual examination of both Figure 4.19 and the source material indicates that the resultant displacement was largely due to the digitising error (e.g., DD), and partly due to generalisation processes, such as deliberate displacement (e.g., D), and simplification (e.g., SD).

At the same scale (1:50,000), the roads and the railway reveal almost the same pattern observed in the river coverage. Figure 4.20 indicates that no simplification took place. This is due to the fact that the character of the features did not require simplification at this scale. The resulting displacement in the roads was largely due to the digitising process, since, only, the centre lines of these features were captured. On the other hand, the railway suffers less displacement since it was possible to trace its character, which was symbolised by a single line. Therefore, the reduction in the total line length was only 0.05 per cent.

| Map scale | Total line length (m) | % Difference from original |
|---|---|---|
| 1:25,000 (original) | 38,139.29 | - |
| 1:50,000 | 38,083.09 | - 0.15 |
| 1:300,000 | 18,823.22 | -50.65 |
| 1:500,000 | 19,079.67 | - 49.98 |

**Table 4.6: Total line length of the river in original and generalised maps (Study Area 4).**

| Map scale | Total line length (m) | % Difference from original |
|---|---|---|
| 1:25,000 (original) | 32,379.26 | - |
| 1:50,000 | 32,363.98 | - 0.05 |
| 1:300,000 | 30,074.73 | - 7.12 |
| 1:500,000 | 28,568.95 | - 11.77 |

**Table 4.7: Total line length of the roads and the railway in original and generalised maps (Study Area 4).**

At 1:300,000, the river principally experienced a typification process. That is, the river at this scale (Figure 4.21) was symbolised by a single line (in Green) as compared to the original (in Red). The overlay of both representations (C), shows that there is some displacement (e.g., D) to accommodate the surrounding features such as roads (see Figure 4.23(A)). The overlay indicates that there was some simplification which resulted in a displacement (e.g., SD). It is difficult to detect the digitising error that might be responsible for the displacement at this scale. The total line length largely reflects the reduction in the number of features (50.65 %), from two lines at the original scale (1:25,000) to one line at this scale, contributing at least a 50.00 % reduction, whereas other processes might only have contributed to the remaining 0.65 % reduction (Table 4.6).

At the same scale, Figure 4.22 shows that the roads and the railway experienced generalisation processes, such as omission (e.g., O), typification (e.g., OT), and exaggeration and displacement (e.g., ED). The character of these features (as being smooth) might be the most likely reason for not undergoing further smoothing; hence no simplification was performed. There was a small reduction (7.12 %) in the total line length (Table 4.6).

At the scale of 1:500,000, the river experienced no further generalisation, compared to its representation at the 1:300,000, except the width of its symbolisation became larger, as shown in the original material. In Figure 4.21, the overlay of the original and the generalised representation of the river exhibits some displacement. This is due to symbolisation requirements, as can be clearly demonstrated in Figure 4.23(B). The total line length points to an unusual effect, if compared to that at the scale of 1:300,000. Since there was no simplification applied, it is expected that the reduction in length at this scale (49.98 %), would be at least the same if not more than that at the scale of 1:300,000 (50.65 %). This can be explained by the fact that if there was no simplification performed upon feature details during the generalisation process, the total line length would necessarily be invariant at the reduced scale(s). Thus, and given the consequences of the digitising error, it is therefore possible to assume that this effect is more likely to be due to the digitising error which occurred at either the scale of 1:300,000 or 1:500,000, or both.

In regard to the roads and railway at this scale (1:500,000), Figure 4.22(F) shows that there was simplification (e.g., SD), displacement (e.g., D), omission and typification (e.g., OT), and omission (e.g., O). The resulting displacement was due to symbolisation and legibility (see Figure 4.23(B)). Because, there was some reduction in minor details and smoothness of shape, i.e., simplification, and some reduction in the number of features, the reduction in the total line length was 11.77 %. Generally, there was a small deviation between all the generalised features at both scales (1:300,000 and 1:500,000), indicating the relatively high level of detail.

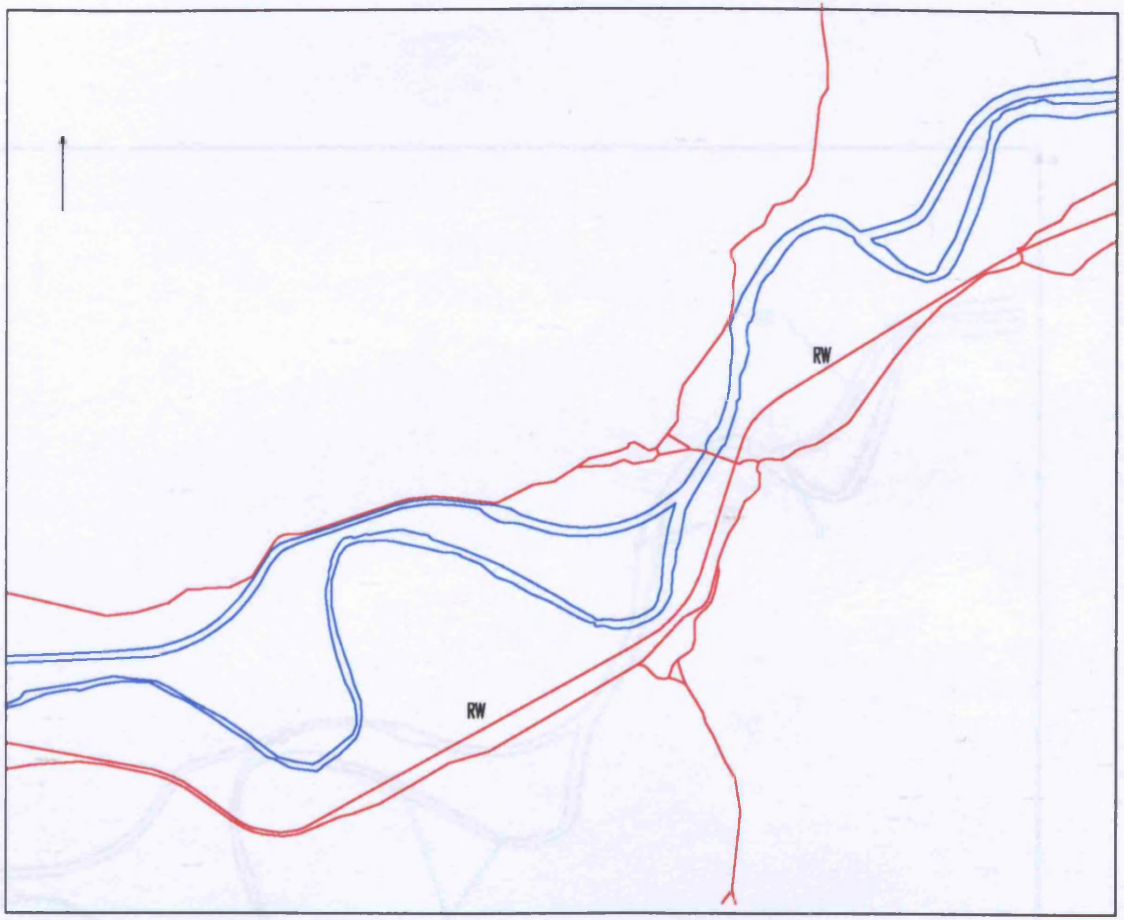



Figure 4.18: A digital representation of the study area 4. The 1:25,000 original map at 1:50,000 depicting a road network railway and part of the River Aare, north of Switzerland, (production by Office federal de topographie, Bern, Switzerland). The river is represented by a double-line symbol (in Cyan), and the railway is denoted by RW, while other lines represent the road network, (both the railway and roads are in Red).

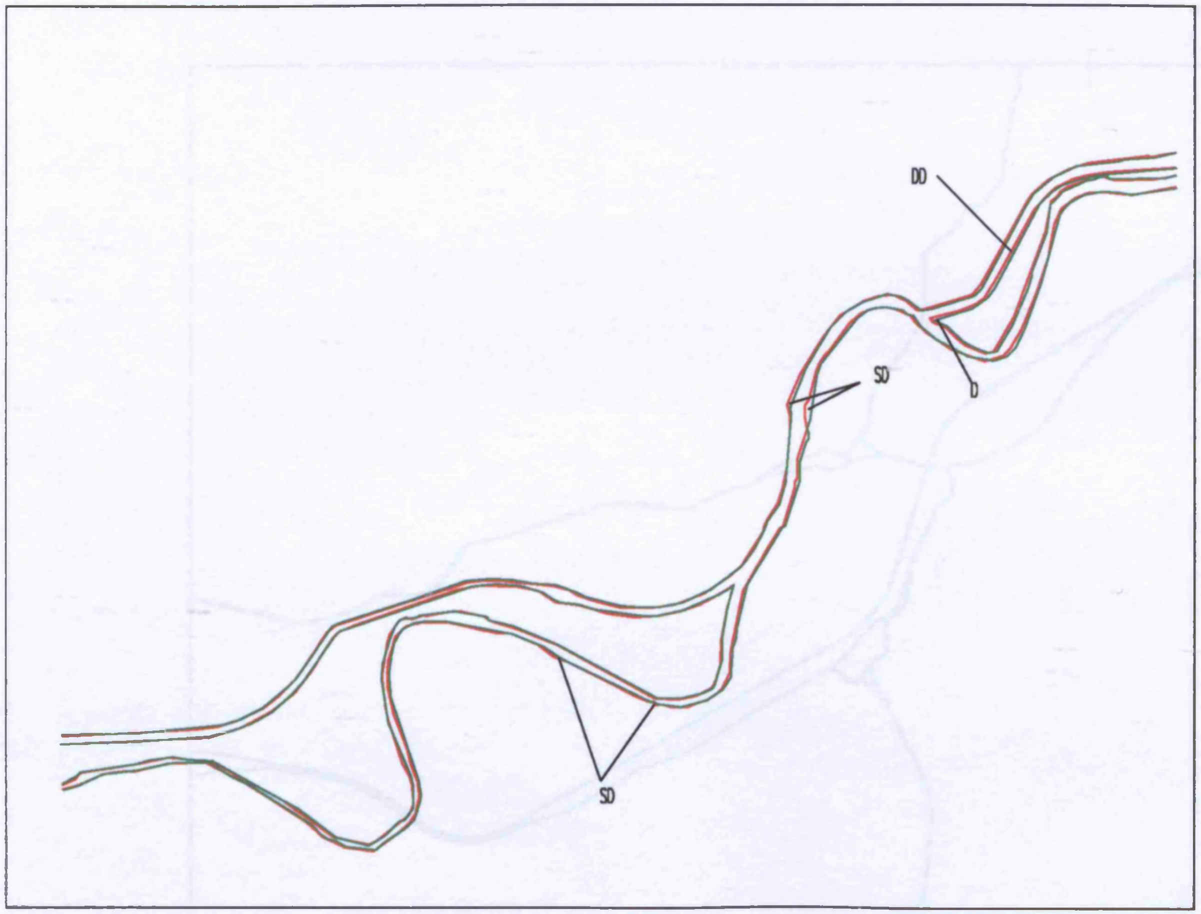



Figure 4.19: Overlay of the original 1:25,000 river map (in Red), and its manually generalised 1:50,000 map (in Green) at 1:50 000. (Study Area 4).

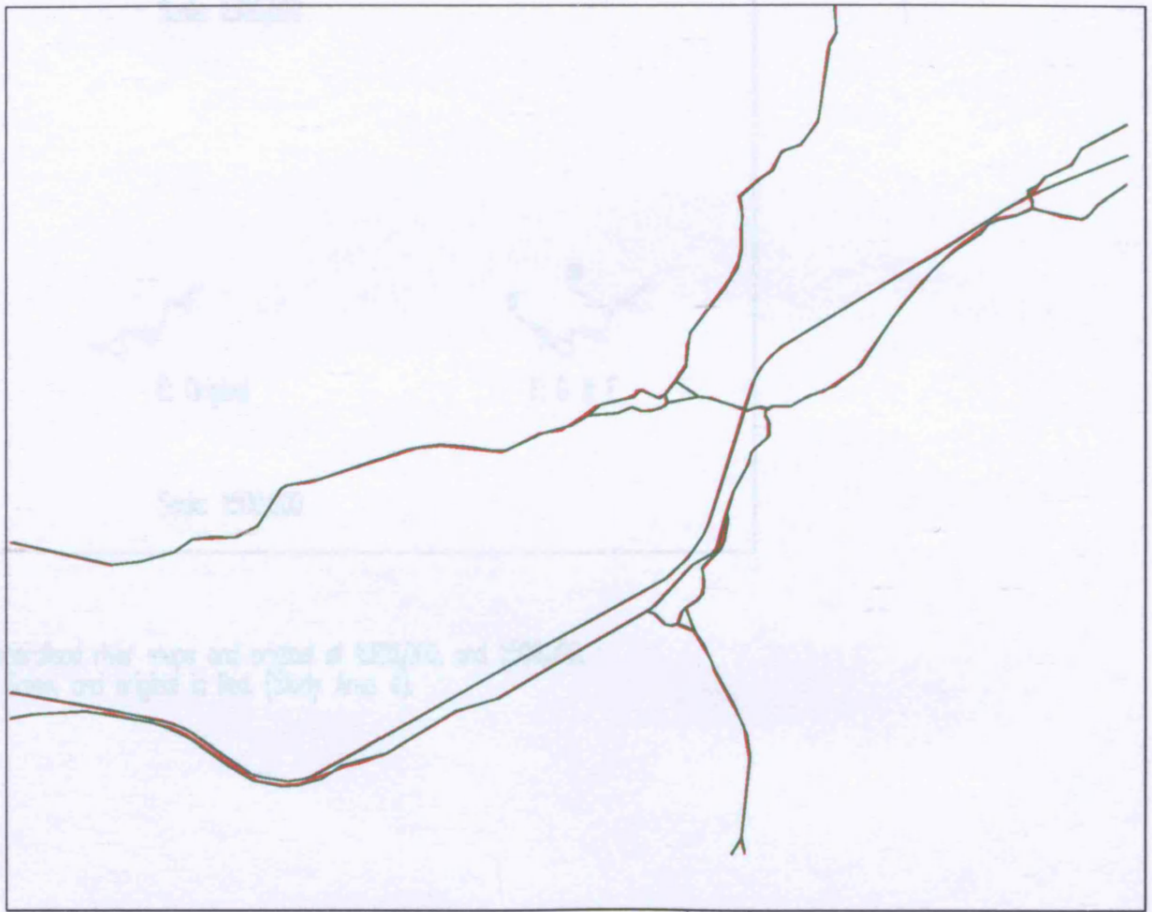


Figure 4.20: Overlay of the original 1:25,000 roads and railway map (in Red), and its manually generalised 1:50,000 map (in Green) at 1:50 000. (Study Area 4).

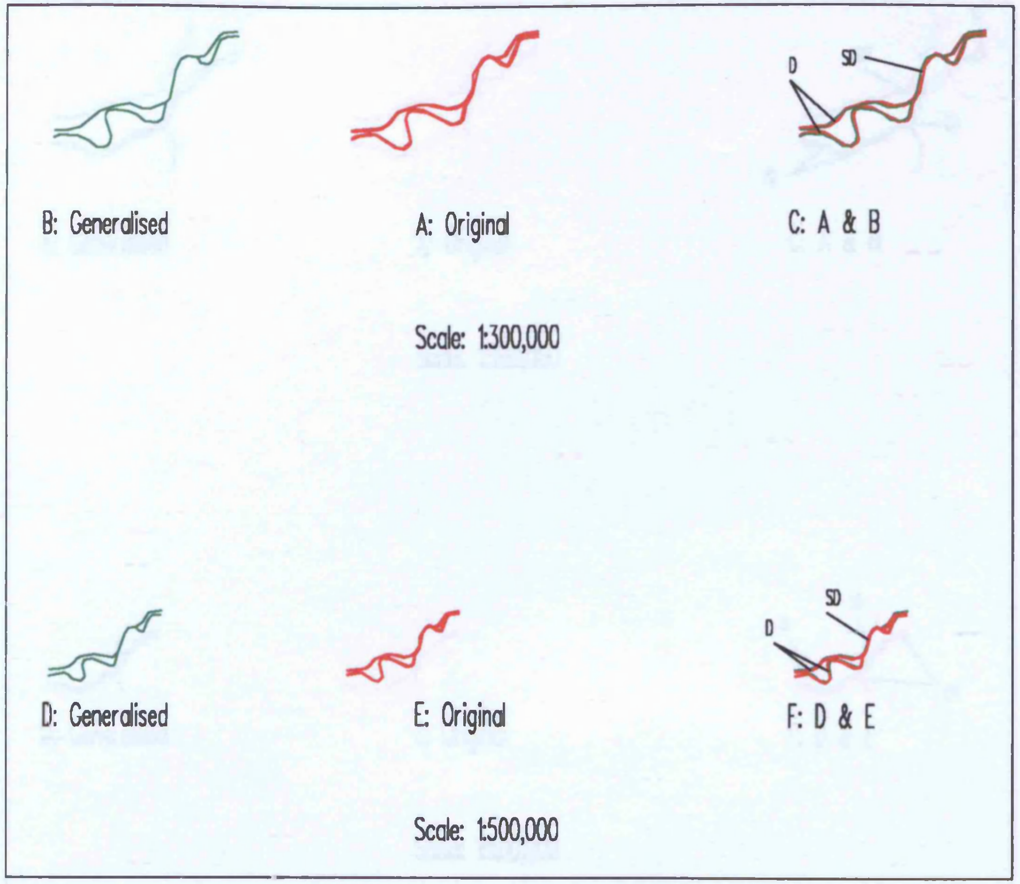



Figure 4.21: Manually generalised river maps and original at 1:300,000, and 1:500,000. Manually generalised in Green, and original in Red. (Study Area 4).

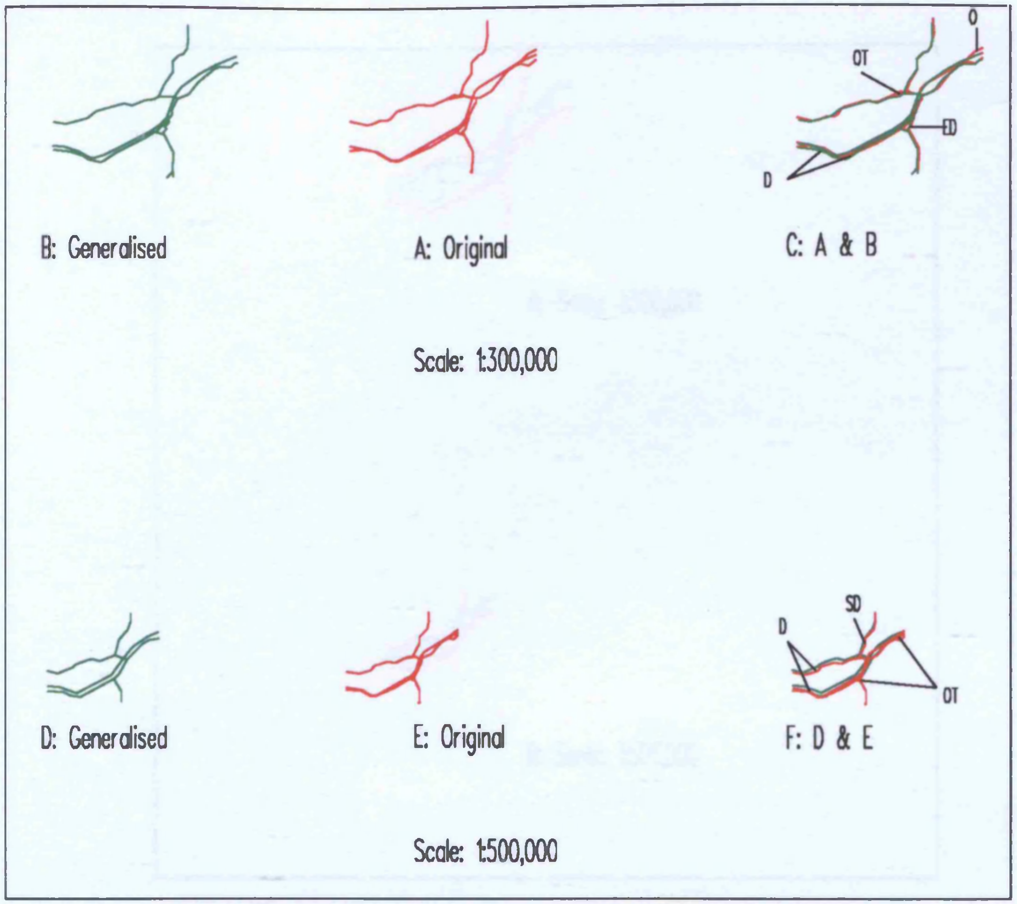



Figure 4.22: Manually generalised transport maps and original at 1:300,000, and 1:500,000. Manually generalised in Green, and original in Red. (Study Area 4).

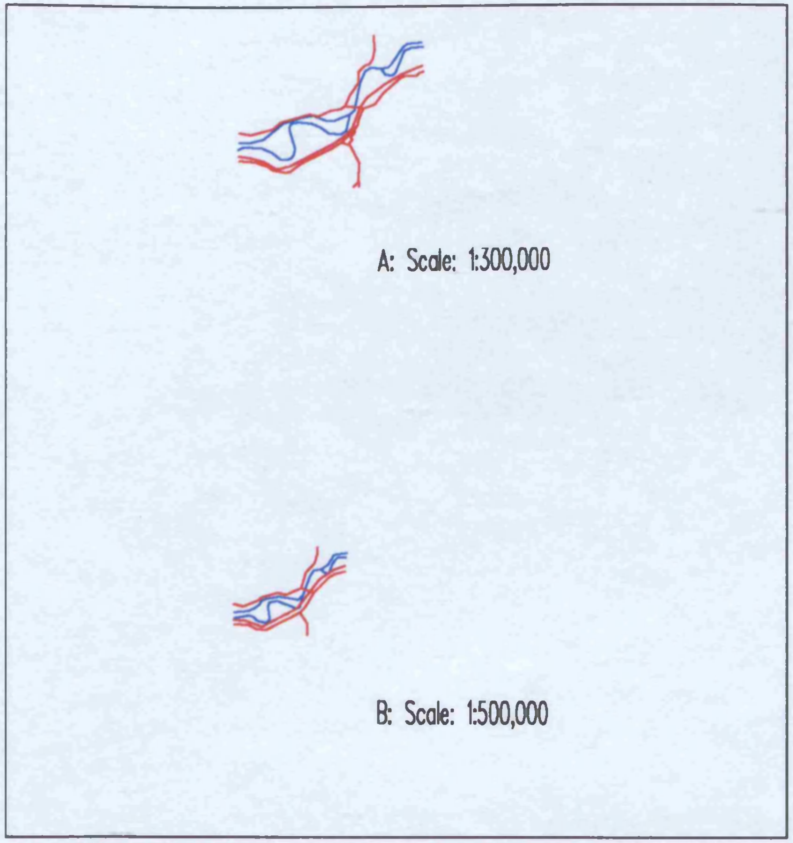



Figure 4.23: Overlays of manually generalised maps at the scales of 1:300,000 and 1:500,000. The river is in Cyan, and the roads and railway are in Red. (Study Area 4).

## 4.4 Discussion

4.4.1 *Effects of feature character on line simplification:*

There is no doubt that feature type or character is an important factor that influences line simplification. Discussion of the effect of this factor on line simplification in these analyses requires that reference to the general requirements that cartographers bear in mind during line simplification is presented first. It is commonly acknowledged that line details should be processed judiciously and according to the mapping context. However, there are general rules. If the line was relatively less complex in details then reproducing this line to medium scales for topographic mapping purposes requires that the line should retain some of its details (e.g., Bertin, 1983; Swiss Society of Cartography, 1987). This is important since elimination of all the details would render the line as a smoother curve that perceptually differs from its original character at the scales. If this line was to be reproduced to small scales, it is unlikely that retention of such limited details would be possible. This is largely due to the fact that such details would be difficult to maintain at small scales, unless the details bear geographical significance in this case they have to be accentuated by exaggeration (Keates, 1989). It is also expected that if the line in question was to be in a subordinate level of importance, for example, during thematic mapping, it is most likely that most or all of its details would be eliminated. In contrast, if the line was to be represented in a thematic map where most or all the details are important, then such details would be retained and accentuated (Keates, 1989). The other example would be what if the line was relatively complex? It is reasonable to expect that not all the details have informative value, and the requirements referred to above are applied but the reduction of details here would be more than if the line was less detailed. Complex lines undergo greater simplification compared to less complex ones. It is a general requirement but it is left to the cartographer to determine according to his own judgement the extent of the reduction. The question now is how this factor (feature type) affected the line simplification in the maps of the study areas presented in this chapter, given the above general requirements? In study area 1 and 2, the analyses indicated that line details were less simplified in study area 1 compared to study area 2. Given the same map purpose, the influence of feature character was largely seen to be the prime factor influencing line simplification in the maps of these two study areas compared to the maps of study areas 3 and 4. In study area 1, line features were relatively less detailed compared to the line features (drainage network) in study area 2; hence limited simplification was performed on lines in study area 1 compared to the relatively greater simplification in study area 2. However, the influence of feature type on line simplification is implied in the other maps of study areas 3 and 4, but it is most observed in the comparison of the maps of study areas 1 and 2, and for which the choice of the later study areas was made.

4.4.2 *Effects of map purpose on line simplification:*

Since it proved difficult to find the same line types at the same or even similar scales and for different purposes, discussion of the effect of this factor (map purpose) involved reference to general requirements in thematic mappings as well as comparison to hypothetical situations, given the fact that the map purpose of the maps of the third study area was thematic. The discussion provided for study area 3 showed less simplification of line details was performed especially at the smallest scales. Given the above general requirements regarding how relative importance of line details affects line simplification, in other words the effect of map purpose, line details of the line features in study area 3 were not largely removed especially at the last two scales. This is generally due to three reasons. They are in the following order: 1) the purpose of the maps (all are route maps), and 2) compared to the maps in the previous study areas there is relatively little difference between map scales, and 3) line character. Whilst it is difficult to attribute the factor of map purpose to the simplification at the scale of 1:200,000, the effect of this factor can be seen with less difficulty at the last two smaller scales, especially at the scale of 1:1,000,000. For, example, at the smallest scale (Figure 4.17) the generalised lines (in Green) are expected to undergo more simplification if they were to be represented in a general map at the same scale where roads are classified as important as other features.

Given the facts that the scale, here, is very small, and the generalised lines in the source paper map are symbolised by larger and double line widths in which the lines appear highly detailed, it is, thus, a valid assumption to attribute the limited simplification largely to the purpose of the map more than to the other factors. Furthermore, had the purpose of the map been not thematic, it is, in fact, expected to find a larger difference between the generalised lines at both small scales, compared to the small difference shown here. However, it is expected to observe more simplification of line details if the lines were more complex, yet it is important to realise that the above argument would be valid; i.e., the simplification was largely influenced by the map purpose. Had the lines been more detailed, it is expected to find that the effect of mapping purpose on line simplification would be much clearer at all scales, since the lines would still be seen as relatively detailed. Thus, it can be concluded that the first factor as opposed to the other factors was the most important behind this limited application of simplification to line features in the generalised maps of this study area, especially at the smallest scales. Furthermore, and as is the case with the generalisation processes in the previous tests, simplification was performed as an integrated process among other generalisation processes.

#### 4.4.3 *Effects of mapping techniques on line simplification:*

Comparison of different mapped features by different mapping agencies or even individuals can help indicate the relative influence of the factor of mapping techniques on line simplification. The effect of this factor would ideally be much observable if the same line features were simplified by different mapping organisations or individuals given the same map purpose and other factors. Again, and for the same reasons provided above, this is difficult to achieve. However, the choice of study area 4 representing a Swiss cartographic product as a different mapping agency from the Ordnance Survey the producer of the maps in the previous study areas, is seen as reasonable for an empirical comparison highlighting the difference between the simplification of line details in the two cartographic products.

Although the selected lines in this study area are relatively less detailed than some features in the previous study areas, a close visual examination of the generalisation would indicate that some minor details were not smoothed out. For example, the upper right and middle parts of the generalised river line at the scale of 1:50,000 (Figure 4.19) could have been represented by a smoother shape, yet it showed that the small minor details at these locations were retained. Furthermore, compare this effect with the generalised map at 1:250,000 of study area 1 (Figure 4.6), especially the line details in the line representing the island in the Humber Estuary. Given the terrain, distribution and size of settlements in the British maps (especially study areas 1 and 2) and the Swiss maps, and given the same purpose of topographic mapping in study areas 1, 2 and 4, the Swiss maps show that line simplification has generally been performed more judiciously compared to the British maps, especially the drainage network in study area 2 where a lot of details were removed at both scales of 1:50,000 and 1:250,000. This conclusion conforms to the observation of the relatively high level of feature details portrayed in the Swiss topographic map series. The reader of maps of this type of terrain expect as much details as possible, especially with reference to the importance of tourism industry in this country.

## 4.5 Summary

### *4.5.1 Line simplification within generalisation is complex:*

The empirical tests revealed difficulty in trying to identify where the simplification process started and where it stopped. In other words, the interaction between line simplification and other generalisation processes is unpredictable or not clearly defined. The distinction between the two processes is, therefore, less definite in practical terms. For example, during simplification, the cartographer would find it difficult to maintain the character of the feature without a minimum exaggeration or typification. This result can only be due to the nature of the generalisation process.

Generally, it was found that the greater the scale reduction the more generalisation processes are required, and vice versa. This implies that, the role of simplification is more likely to be more evident at small scales as opposed to large and medium scales. At small scales, the cartographer would focus on the general form(s) of the line being represented more than on its minor details.

### *4.5.2 Factors influencing line simplification are intricately linked:*

The analyses indicated that the interrelationship between the three factors looked at in this chapter and other factors are difficult to ignore and isolate. However, the analyses generally indicated that feature character is the most important factor influencing the process of simplification of line details, especially in study areas 1 and 2. However, the effect of this factor is also implied in study areas 3 and 4. Map purpose and mapping techniques were the focus of the analyses in study areas 3 and 4, respectively. The influence of these two factors can not be ignored in study areas 1 and 2. It proved necessary to present general conclusions as to what might have contributed most to the simplification process in each test. Thus, it can be concluded that the relative influence of factors such as these discussed here on line simplification may be ranked. For example, given a generalised map, the influence of feature character is to be ranked as the most important factor followed by map purpose, and mapping techniques. That is, the simplification was not influenced only by a single factor.

A number of difficulties were found in isolating the factors influencing line simplification. Three key reasons can be identified. First, isolation of every relative influence of each factor is difficult to achieve. Influence of factors of only map purpose and mapping technique may, for example, be isolated only if the actual same line features were cartographically processed for different map purpose and by different mapping techniques, and by multiple cartographers in each organisation. In fact, it is difficult to come across the same line features in multiple scale maps and for multiple map purposes and processed by different mapping agencies. Second, it is important to bear in mind that line simplification is only a single context-dependent process. As such, the cartographer manipulates line details according to the mapping context (scale, purpose, etc.), and above all this manipulation is greatly influenced by the cartographer's judgement. Third, due to the subjective nature, and hence elusiveness of manual simplification, it is unlikely that even a more rigorous analysis would yield complete answers as to what and where all factors affect simplification, since 'full' answers to such questions are actually in the mind of the cartographer who performs the simplification (see Chapter 2). Therefore, a thorough analysis of the effects of all possible factors that may influence line simplification is well beyond the scope of this thesis, due to its technical constraints imposed by the research work and topic.

The next part of the thesis (Chapters 5, and 6) will discuss empirically the process of digital line simplification and to what extent this process could produce cartographic quality according to the cartographic context discussed above.

# CHAPTER FIVE: EVALUATION OF THE DOUGLAS-POIKER ALGORITHM

This chapter aims to re-evaluate the Douglas-Poiker algorithm. An analysis of the algorithm is presented on the basis of its design specifications as outlined by its authors. A direct user input parameter is proposed for performing objective scale-dependent database simplification. Also, the relationship between graphic reduction and the resulting simplification is examined and a method is proposed for quantifying this relationship.

## 5.1 Background and Reasons for Re-evaluation

The Douglas-Poiker algorithm (Douglas and Poiker, 1973) is probably the most widely used line simplification algorithm in digital mapping. This algorithm is supported by several computer packages, including GIS packages (e.g., Arc/Info), mapping packages (e.g., GIMMS, MAPICS), statistical packages (e.g., SAS/GRAPH), and on-line tutorials such as the Geographical Information System Tutor (GIST). It is often referenced in text books concerned with the subject of digital cartography (e.g., Clarke, 1990).

Although this algorithm is credited to Douglas (e.g., McMaster, 1986,1987a 1987b) or Douglas and Poiker (e.g., Douglas and Poiker, 1973) it should be noted that it was independently published by Ramer (1972). Ramer's description corresponds to the second method of implementation published by Douglas and Poiker (1973). Also, the same algorithm has been published by Duda and Hart (1973), and according to them, this algorithm was suggested by G. E. Forsen. Thapa (1988) explains that these publications were all made about the same time, but since publishing a book usually takes longer than publishing an article, it can be assumed that Forsen was the first to devise the technique. In this study the algorithm will be referred to as the Douglas-Poiker algorithm. That is for two reasons: 1) Douglas and Poiker did not refer to the other implementations of the method by other researchers, and it is assumed that Douglas and Poiker developed the idea independently, and 2) this algorithm is widely known and acknowledged in the digital mapping community, as the Douglas-Poiker algorithm.

The popularity of the algorithm is related to its ability in producing mathematically and perceptually good results, compared to other algorithms (McMaster, 1987a). Mathematically, it produces the lowest amount of displacement from the original line, whilst perceptually it tends to select critical points closest to those that might be selected by humans. Its ability in producing such results stems from its being a *global operator;* that is, it considers the entirety of the line being simplified. Also, the popularity of the algorithm is furthered by being controversial, in the sense that it has been differently interpreted and used. This, an unintended aspect, of the algorithm is the main impetus behind conducting this study.

## 5.2 Algorithm Characteristics

### *5.2.1 Original Algorithm description:*

Douglas and Poiker (1973) present two methods of point reduction based on the principle of selecting points with the highest offset from an anchor-floater line determined by a predefined tolerance. These two methods are described by the authors as follows:

"Method one begins by defining the first point on the line as an anchor and the last as a floating point. These two points define a straight line segment. The intervening points along the curved line are examined to find the one with the greatest perpendicular distance between it and the straight line defined by the anchor and the floater. If this distance is less than the maximum tolerance distance, then a straight segment is deemed suitable to represent the whole line. In the case when the condition is not met, the point lying furthest away becomes the new floating point. As the cycle is repeated the floating point advances toward the anchor. When the maximum distance requirement is met the anchor is moved to the floater and the last point on the line is reassigned as the new floating point. The repeat of this latter operation comprises the outer cycle of the process. The points which had been assigned as anchor points comprise the generalised line. Method two is exactly the same as the method one except that note is taken of all points which have been assigned as floaters on previous inner cycles. These are stacked in a vector. After the anchor is moved to the floating point, the new floating point is selected from the top of the stack, thereby, avoiding the necessity of re-examining all the points between the floater and the end of the line. This procedure usually results in the selection of a slightly greater number of points than the Method 1, but takes approximately 5 percent of the computing time and is thought to produce better caricatures" (Douglas and Poiker, 1973, p. 117).

Method 2 is clearly more effective, and, therefore, has been favoured by the majority of researchers in digital cartography (Whyatt, 1991). Figure 5.1 illustrates the concept underlying the algorithm, using Method 2. In Figure 4.1b, the first point is defined as the anchor (*a 0*) and the last point as the floater (*f 0*). They are connected by a line. The perpendicular distances to all intervening points are computed, and the point with the highest offset (*f 1*) is identified; since the distance of this point from the line exceeds the user-defined tolerance, it is selected as a new floater and placed in a stack. In the second iteration, the original anchor (*a 0*) is connected to the new floater (*f 1*), and offsets are calculated to intervening points. Since all points between the anchor (*a 0*) and the new anchor (*f 1*) fall within the tolerance, *f 1* is defined as the new anchor (*a 1*). At this stage, the algorithm takes *f 1* from the stack as the new anchor to be connected to *f 0* as the floater point. After calculating all offsets, a new floater (*f 2*) is selected; since it exceeds the tolerance, and stored in the stack, above *f 1* . The process continues until eventually the last point on the line is reached. Points previously assigned as anchors are connected by straight segments, producing a simplified line (Figure 5.1c).

According to Douglas and Poiker (1973) the prime purpose of the algorithm "is to reduce the number of points required to represent a line and to produce abstractions, or caricatures of the line in cases where these will suffice. In many cases, these could be considered to be perfectly adequate generalisation procedures" (p. 122).

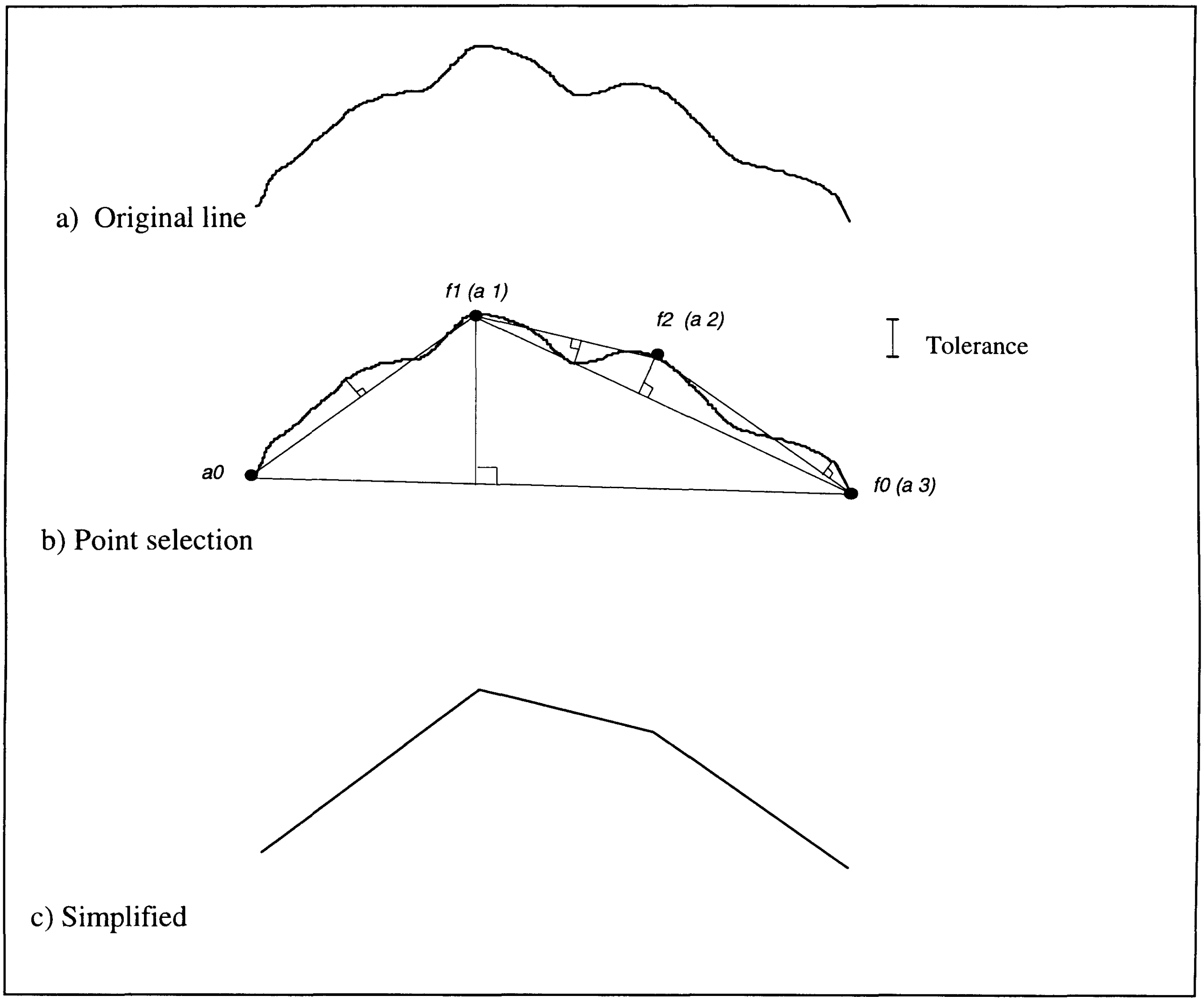



Figure 5.1: Description of the Douglas-Poiker algorithm.

*5.2.2 Alternative descriptions and implementations:*

Whyatt (1991) points out that the original description of Method 2 is slightly "ambiguous". Accordingly, some workers have offered their own descriptions of the method, to the extent that some appear to be erroneous. For example, Maguire (1986, p. 96) states that "if one of the perpendiculars exceeds the bandwidth (tolerance corridor) the point before the floater becomes the new floater and the new perpendiculars are measured". Different implementations of the algorithm have been proposed, which, when applied to standard forms of test data, produce different results (Whyatt, 1991). According to the algorithm (Method 2), points that have the highest offset values are regarded as both geometrically and perceptually the most important, while the least important points are those that have the lowest offset values. The implementation of the algorithm within the GIMMS mapping package (Waugh and McCalden, 1983) is an improvement in that it allows the user to assign points to nominal, scale related classes. Waugh and McCalden were the first to introduce this concept of point tagging in their mapping package in the *GENERAL command. The user is allowed to assign up to nine different tolerance values, which correspond to decreasing levels of simplification. Wade (Whyatt and Wade, 1988) provided a modification of the original algorithm to return the perpendicular offset values with the selected points. Wade had to modify this implementation to the offset storage strategy, so that the results could be consistent with those produced by the original.

Van Horn (1985) and Buttenfield (1986) have proposed improvements for the algorithm in order to correct the unbalanced simplification resulting from use of a single tolerance value with the original algorithm. Van Horn's method involved pre-processing of the points to reduce their precision by moving them to the nearest corner of a grid cell, the resolution of which is determined by the display device and the scale of the display. His approach succeeded in retaining some detail along what had previously been over-simplified sections of a coast line chosen as an example. Buttenfield suggests that the segmentation of a line into its constituent geometric types could facilitate the use of a set of suitable tolerance values to achieve more balanced simplifications at given scales.

## 5.3 Previous Evaluations:

Several researchers have commented upon the Douglas-Poiker algorithm. Depending on their data sets and evaluation criteria, some of those researchers, have expressed satisfaction while others have reservations. Their conclusions are reviewed under two main headings: merits and limitations.

### *5.3.1 Merits:*

Some workers have studied the relative merits of different algorithms in perceptual terms whilst others have offered mathematical measures for evaluating algorithms. White (1985) has compared lines that had been simplified by three simplification algorithms with manual simplifications of the same lines. The results indicated that the Douglas-Poiker algorithm produced perceptually similar lines to those that had been simplified manually. White used the number of critical points as the measure of the perceptual quality of the line. She found that 45 % of the original set of points were selected by both the study participants and the Douglas-Poiker algorithm. She concluded that the Douglas-Poiker algorithm was the most effective at selecting critical points. Zoraster *et al.,* (1984) have conducted an extensive review of digital algorithms for generalisation. They found that the "Douglas-Poiker algorithm proved superior in both approximating the lines obtained by manual generalisation and approximating the original curves" (p. 104). As indicated in the previous sections, the algorithm's ability in considering the entirety of the line being processed is the major asset of the algorithm, since it detects all critical points that mark significant changes in direction. It is, therefore, expected, that this would necessarily lead to favourable results, compared to any routines that do not work under the same principle. Perceptually, the algorithm produces good caricatures of lines; since it tends to include "maxima" and "minima" (Williams, 1978).

McMaster (1987a, 1987b) evaluated nine simplification algorithms, with six of the original thirty mathematical measures that he developed ( section 4.5.1). He concluded that the Douglas-Poiker algorithm performed well at all levels of simplification and it was both "mathematically and perceptually superior". Muller (1987a) compared seven algorithms using the fractal dimentionality of lines as a statistical measure of the line complexity. He found that the Douglas-Poiker algorithm produced the best preservation of the fractal dimentionality of the line which was simplified.

*5.3.2 Limitations:*

Several researchers express dissatisfaction about all or some aspects of the algorithm (e.g., Van Horn, 1985; Morrison, 1975; Dettori and Falcidieno, 1982; Monmonier, 1986; Thapa, 1988; Li and Openshaw, 1992; Muller, 1987a; Visvalingam *et al.,* 1991; Whyatt, 1991). Common concern is related to the problem of shape distortion resulting from gross levels of simplifications by the algorithm.

Monmonier (1986) and Thapa (1988) indicate that the Douglas-Poiker algorithm can only produce acceptable results when the reduction in scale is modest. Muller (1987a) noticed that the algorithm produced a very spiky simplified line. He, therefore, questioned McMaster's observations as to the ability of the algorithm in preserving the angularity of the original lines. He argues that, since the angularity measure is strongly influenced by the presence of spikes, the preservation of angularity can not be taken as an indication of the quality of simplification. Muller, also, questioned the measurement of total displacement proposed by McMaster; since it can not be used to determine whether an algorithm is capable of preserving the geometric shape of the original line, as two entirely different geometric shapes could result in the same amount of overall displacement.

Van Horn (1985) points out that the algorithm tends to retain spiky details while it over-generalises variations within the tolerance band, therefore, producing unbalanced simplification. Whyatt (1991) presents a more elaborate critical review of the algorithm. He points out that various implementations of the algorithm can produce different results. For example, different implementations using ad-hoc values dealing with some special geometric such as closed loops. Ramer (1972) proposes a solution for this case in which he suggests that any two distinct vertices could be selected arbitrarily for the initial anchor and floater. He suggests that the best choices would be two oppositely located extremal points (i.e., the furthest two points from each other in the curve), since he believes that the algorithm would eventually select these points anyway. In his implementation, he specifies the choice of the highest left-most point and the lowest right-most point for these extremal points. Visvalingam and Whyatt (1991) point out that the algorithm may produce significantly different results if the algorithm was written using single precision, and even if it was compiled with the double precision option. Huggins (1991) states that the arbitrary-precision arithmetic language 'bc' can be used to obtain precise results, when programs are written using single precision REALS while compiled using the double precision option. Forrest (1985) points out that floating point calculations are still very much machine dependent. However, the perceptual significance of this problem is less serious from a cartographic point of view. Visvalingam and Whyatt (1991), also, refer to the impact of digitising error on the algorithm's behaviour, as this would affect the selection of the most critical points. Again, although the effect of digitising error can not be ignored (especially during geo-cartographic analytical operations), such an impact is perceptually insignificant, provided the error was below the level of the visual limits. Although, Visvalingam and Whyatt (1991), themselves, indicate that the absolute position of the point is irrelevant at small scales, while it could matter at intermediate scales, cartographers do not consider such exact positions of points to be significant during either generalisation or simplification. Whyatt (1991) and Visvalingam and Whyatt (1991) emphasise the finding of Duda and Hart (1973) which states that the algorithm is strongly influenced by individual points and that a single 'wild' point can drastically change the final results. However, such an effect can only be perceptible when there is a small number of points representing a line feature, and/or the simplification was performed within a scale-independent context. Whyatt (1991) questions the concept of critical points, since the points selected by the algorithm are not always critical; for example, selecting points representing minor details at the expense of larger ones. Furthermore, Whyatt concludes that the algorithm is not even an optimal choice for performing weeding purposes. It can be inferred from this statement that Whyatt bases his judgement on the manual cartographic practice, where the cartographer selects the larger forms of a line to be simplified or generalised at the expense of minor or unimportant details. Thus, the algorithm is compared to a general, variable, and undefined manual approach, which tends to result in an over-critical and biased evaluation of the algorithm.

The observations are only valid within their evaluation contexts. As noted above, the quality is influenced by the limits of human perception and expectations. Visvalingam and Whyatt (1991) emphasise this very conclusion: "since this degradation of acceptability is not related to any variation in the inherent behaviour of the algorithm, it must be related to changing objectives and expectations" (p. 217). While the previous evaluations have concentrated on assessing the Douglas-Poiker algorithm on ill-defined and variable bases in the process of cartographic generalisation, the following sections are designed to re-evaluate the algorithm within a cartographic context (i.e., data and graphic reduction requirements) and according to its design specifications.

## 5.4 Parameterising the algorithm

The aim of this test is to evaluate the Douglas-Poiker algorithm according to the data reduction process for which the algorithm is principally intended. This evaluation is conducted in a scale-dependent context (i.e., at different levels of representation), so as the results are to be cartographically assessed. As mentioned above, the authors of the algorithm state that the algorithm is first designed to perform reduction of the number of points required to represent a line, irrespective of whether for cartographic simplification, generalisation, or polygonal approximation. The algorithm, also, is designed to produce caricatures of the line in cases where these will suffice, regardless of the mapping context, so long as the result can be accepted. The evaluation of the algorithm in this chapter is to assess the algorithm's ability in data reduction for weeding purposes, but not the ability to produce caricatures. In other words, the evaluation will be within a scale-dependent context in order to assess the algorithm's ability in weeding out redundant data at multiple reduced scales from single databases. Pertinent to this test, is the impact of graphic reduction on the resulting shapes. Given this influence, the test should explore the ability of the algorithm in reducing the number of points without causing shape distortion at target scales during multiple scale-dependent data reductions.

Before presenting the results and discussion, the significance of the graphic reduction factor is first highlighted. A formula to determine the tolerance used in the algorithm is, then, presented, followed by a description of the data sets used for this purpose.

### *5.4.1 Effects of graphic reduction on data:*

Graphic representations of features have to take into account the effect of areal scale determined by the rate of reduction or simply scale. As shown in Chapter 2, areal scale or the available map space, has a rate of reduction which is greater than that of linear scale. As can be seen from Figure 5.2, areal scale is reduced by the square of difference in linear scale. In this figure, an area of 100 square millimetres at the 1:25,000 scale, will only occupy one fourth as much map space when mapped at the 1:50,000 scale, that is 25 square millimetres.

| **Nominal Scale** | **Linear scale** (mm) | % Change | **Areal Scale** ($mm^2$) | % Change |
|---|---|---|---|---|
| 1:25,000 | 10 | | 100 | |
| 1:50,000 | 5 | 50.00 | 25 | 75.00 |
| 1:100,000 | 2.5 | 75.00 | 6.5 | 93.75 |

Figure 5.2: Types of map scales and rate of reduction.

So far as digital line simplification is concerned, point data representing digital lines have to account for the rate of areal reduction. Simplification routines are initially designed for data reduction of file sizes to a size in keeping with intended output, only later have they been used for scale-dependent simplifications. Workers with those routines have been searching for objective methods for performing scale-dependent simplification. The Radical Law's formula has, therefore, been widely utilised for determining the number of points with which to represent a line at a given scale. The application of this formula and its appropriateness are discussed in section 5.4.2.2. In this study, the effects of graphic reduction on the data reductions by the algorithm are numerically and perceptually assessed. Numerically, a novel approach is proposed to relate the effect of graphic reduction on the reduced data by the algorithm. It is based on the concept of areal scale, output resolution and the areal extent of the graphic representation of the data simplification resulted by the algorithm. The relationship between data, output resolution, and rate of reduction, is explored, and it is of great importance to data reduction, simplification and generalisation. This relationship is first addressed in detail.

5.4.1.1 Output resolution, data, and scale:

Often associated with scale is the term resolution. The resolution of a data set defines the smallest object or feature which is included or is discernible in the data (Goodchild, 1991). The relationship between scale or graphic reduction and resolution is significant; since for each map scale there is a limit at which an object can be shown. Therefore, objects at various scales will have different resolutions (i.e., minimum size) that can be shown true to scale on maps. For example, the smallest object is 1.25 metres long at the scale of 1:5000, while at the scale of 1:1,000,000 the minimum size available would be 250 metre. The process of data reduction can not, therefore, be performed without considering the resolution at particular scales. Also, the process should correspond to the resolution of the graphic representation medium. For example, data that are digitised on equipment with a resolution of 0.0001 inch and are to be plotted by equipment with a resolution of 0.001 inch, have 10 times more data that are redundant; since they can not be physically plotted (Robinson *et al.*, 1995).

There is a limited range of magnification and reduction that can be applied to a data set at which features still remain legible and aesthetically pleasing (Robinson *et al.*, 1995). "This range is about 2 x. For example, a 1:24,000 scale cartographic database could be displayed from a range of 1:10,000 to 1:50,000 and still be visually pleasing. Beyond that range, we must apply generalisation operations to the data being reduced in scale" (Robinson *et al.*, 1995, p. 252). Given these factors, databases are, therefore, being created to suit various resolutions or scale levels. The highest resolution or the largest scale is the geographical database, where the location of features is not influenced by cartographic requirements (Robinson *et al.*, 1995).

Jenks (1981) suggests a perceptual model of linear data simplification based on map purpose. This model contains three thresholds which subdivide the simplification continuum into four categories of representations, so that each category corresponds to potential map uses (Figure 5.3). Jenks claims that lines produced from minimally reduced point subsets can be used for the purpose of contour maps, whereas lines produced from moderately reduced point subsets can be used for the purpose of thematic maps. On the other hand, there is a level of point reduction beyond which lines are perceived to be different, and hence unacceptable. Jenks (1981) also provides physical limits of his model. He indicates that approximate limits can be expressed as follows:

| % Point Reduction | Cartographic Application |
|---|---|
| 100 - 40 | Topographic Maps |
| 40 - 20 | Thematic Maps |
| 20 - 5 | Cartograms |
| < 5 | No Use |

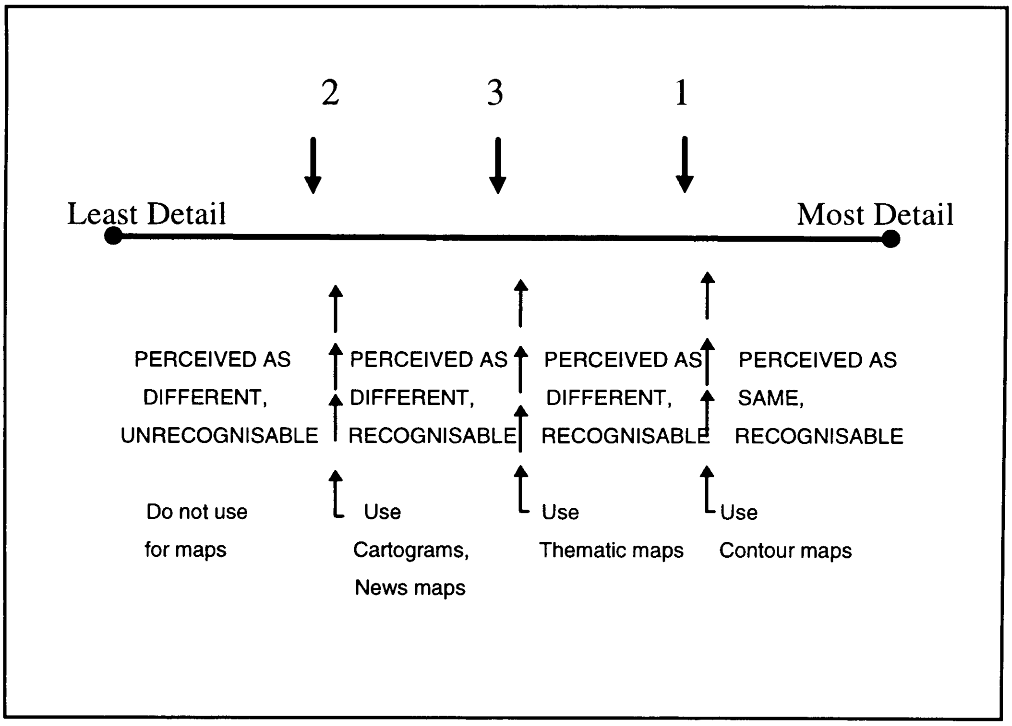



Figure 5.3: The Perceptual Model of Simplification proposed by Jenks (1981).

These limits are taken with caution; since they are theoretically general quantitative definitions of undefined cartographic processes by the author. This is because, cartographic line simplification does not necessarily imply data reduction; since the process deals with shape, but not directly with number of points. Also, data reduction for weeding purposes is independent from the map purpose. McMaster (1987a) proposed a modified version of the model proposed by Jenks (1981). This model involves three dimensional axes in order to include both a scale axis and output technology axis (Figure 5.4). The model emphasises the complex interrelationship between data, output resolution, and map purpose. In Figure 5.4, the model builds on that proposed by Jenks, which is the y-axis on this illustration. The second axis represents scale (x-axis), so the two edges of this axis will converge as scale is reduced, whereas the maximum scale reduction, indicated by a single point (or line in three-dimensions), will be a single point on the graphic output. The z-axis represents a technology continuum (McMaster, 1987a). McMaster indicates that although scale represents the single most important factor in selecting a tolerance parameter for the simplification algorithm, there is little or no work which deals with these problems. Furthermore, the selection of simplification tolerance is made more difficult by the fact that "each algorithm requires a unique set of parameters based on map purpose, scale, and technology restrictions" (McMaster, 1987a, p. 101).

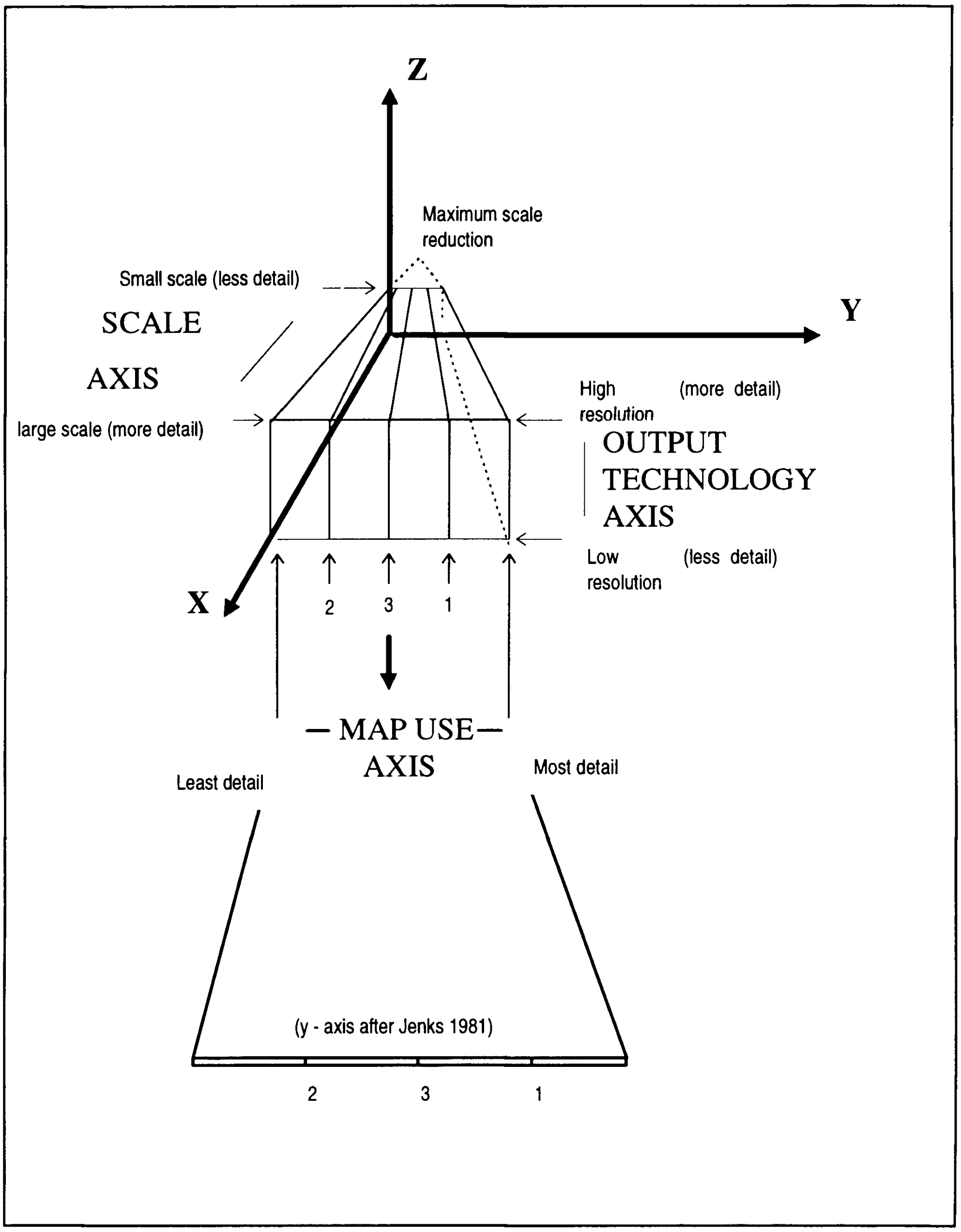



Figure 5.4: Model of simplification based on map use, scale, and output technology proposed by McMaster (after McMaster, 1987a).

It is important to note that areal scale imposes requirements upon line feature simplification and generalisation that differs form those upon data reduction. This is because, cartographic simplification and generalisation entails a perceptible change in feature shape and number of objects being processed, while data reduction is merely a process concerned with the removal of the redundant data in the imperceptible realm at the required graphic representations. The question, here, is what tolerance parameters might be regarded as suitable to weed out the superfluous data at a given scale of representation. So long as the aim of this chapter is to evaluate the Douglas-Poiker algorithm in the light of data simplification or reduction, which is the principal design purpose of the algorithm, the question is what tolerance parameters suit multiple data reductions for multiple scale representations. A proposed answer for this question is presented and discussed in the sections discussed below.

### *5.4.2 Proposed formulae:*

#### 5.4.2.1 Past proposed formulae:

The previous section indicates that digital data have to appropriately correspond to scale reduction. Many workers have used the Radical Law or modified versions of its basic expression as an objective method for reducing data in accordance with scale reduction. Some have built on this method and have introduced other modifications or other alternative ones in order to provide direct tolerance inputs to be used by digital algorithms (e.g., Abraham, 1988). For this purpose, Abraham (1988) developed a formula based on a method of sampling approximation which was based on a link he established between the tolerance used in the Douglas-Poiker algorithm and the number of points selected by the algorithm. This sampling allows the parameters in this relationship to be estimated for any given data set. It is important to indicate that Abraham did not provide his formula for data reduction but for 'generalisation'. He indicates that a pre-filtering tolerance equal to half the pen width is necessary. The formula is briefly expressed as:

$$T = G\,((Mf\,/\,Ma)\,\text{-}1)\,/\,b \qquad (1)$$

where $T$ is the tolerance, $G$ is the average distance between the digitised points at the source scale, $Ma$ is the denominator of the source scale, $Mf$ is the required scale, and $b$ is calculated for every data set by sampling, and the value of $1$ relates to a substitution of the variable (a) in an equation in of the form: y = 1 / (a+bx) which he found it to be a practical fit to the graphs describing the data sets he used. Abraham points out that the value of (a) is theoretically 1, since for zero tolerance all points would be selected.

There are limitations to formula 1, some of which are pointed out by Abraham, himself. These are summarised as follows:

1- The applicability of formula (1) is "restricted by a lower bound of y = 0.005, which corresponds to a scale change factor of 200" (Abraham, 1988, p. 77).

2- The formula tends to produce increasing numbers of 'cross-overs' for very large scale changes (Abraham, 1988).

3- The formula is not 'algorithm specific' (Abraham, 1988), but 'data specific'.

4- Abraham intended to use the Douglas-Poiker algorithm for generalisation purposes, but not for simplification.

5- His focus was only on isolated cartographic lines, and the wider implications of the application of the formula were unknown.

6- The reduced data sets resulted from the application of the formula in his work showed rather a filtering process more than a generalisation process.

Abraham (1988) used another formula in his work which he indicated that it was originally suggested by C. Jones. The formula is calculated simply as:

$$T = e\,Mf \qquad (2)$$

where $T$ is tolerance, $e$ is the minimum tolerance on the map, and $Mf$ is the derived scale.

Formula (2) obviously neither considers the algorithm nor the source data set.

The next section discusses the development of a proposed formula for data reduction purposes within a scale-dependent context specific to the use of the Douglas-Poiker algorithm. Any formula for data reduction therefore needs to accommodate data, graphic reduction (scale), and algorithm's characteristics being utilised.

5.4.2.2 New proposed formula for scale-dependent data reduction:

Three methods were assessed. They are: the Radical Law (RL), Modified Radical Law 1 (MR1) and 2 (MR2). As discussed in Chapter 2, the Radical Law (Topfer & Pillewizer, 1966) is the only quantitative method describing feature selection during generalisation. The authors indicate that the Radical Law could be applied to predict the number of features (e.g., settlements, place names), line features (e.g., drainage networks), and area features (e.g., lakes). Given the fact that linear data in the digital environment are essentially composed of small objects in the form of line segments connecting digitised points (line segments = number of points -1), it would be reasonable to utilise the concept of this quantitative rule for data reduction purposes. Hence, identifying important items of data (in this sense, critical points) for inclusion on the derived maps can be achieved by selection based on this quantitative method. Justification for its utility is supported by the fact that the number of points is a quantitative element that could be related to scale or resolution (Abraham, 1988). Authors such as Abraham (1988), Whyatt (1991), Joao (1994) and Barber *et al.*, (1995) have all discussed and utilised the Radical Law in the context of digital line generalisation. However, a more elaborate discussion of the appropriateness of this method in relation to scale-dependent data reduction for weeding purposes is presented in the following sections.

It is important to highlight two issues in relation to the use of the Radical Law if it were to be utilised. First, it is essential before conducting scale-dependent data reductions using the Radical Law or any other formula to perform a weeding process once the digitisation process has taken place. This is as discussed in Chapter 3, to weed out the redundant points at the digitising scale so that only critical points representing the lines are retained. If the source or digitised lines were not prepared in this form, application of the Radical Law will be meaningless, since calculation would involve redundant and uncritical points that would have no contribution in the graphical sense, as well as being costly to store and process. The second issue is related to the appropriate application of the Radical Law on different types of line features. Given the fact that straight lines would be depicted as straight lines at any scale, application of this method might well seem unsuitable. Furthermore, regular forms that are composed of straight lines, or in other words represented by limited numbers of critical points, might be beyond the realm of suitable application of the Radical Law's formula. This is because a slight change in the critical points that represent the cartographic form on the map would necessarily lead to a perceptible change or modification of that form. For example, a rectangular shape represented by a finite number of critical points can lose its original shape if the required number of points determined for derived scales by the formula was less (which is expected) than that at the original database. However, if these points were coded as nodes the feature shapes will not be affected by the application of the algorithms, since the simplification algorithms do not remove the points that are coded as nodes representing the start and end points of lines (see Figure 5.5), which in this case raises the question of the validity of using the formula as well as the algorithms. For these reasons, detailed geographic lines (e.g., rivers) or in other words irregular lines or smooth curves that are represented by a relatively large number of points (e.g., hundreds and thousands of points) are the more appropriate domain for application of the Radical Law. This is simply because line details (points) are required to be reduced, and it is expected that a judicious reduction (using appropriate algorithms) would maintain the original shapes of lines, while a lot of their details were removed. In this study, the Radical Law is applied to irregular

irregular and curved features. However, detailed analysis of the Radical Law and its usability on these types of features is presented in the following sections of this chapter.

The Radical Law formula presented here is as discussed in Chapter 2 (section 2.4.3), and given that the symbolic form of the lines is constant at all scales, only its basic form is considered here:

$$Nf = Na\sqrt{Ma/Mf}$$

where

Na = Number of objects on source map with a scale function Ma.

Nf = Number of objects on derived map with a scale function of Mf.

The modified versions (1 and 2) of this expression are proposed in this study and applied in relation to the principle of linear scale and areal scale, respectively. Therefore, the first modification (MR1), reduces the number of points in proportion to the linear scale reduction:

$$Nf = Na * Ma/Mf$$

The second modification, termed as MR2, is expressed in relation to the rate of areal scale change. This modification is theoretically provided just to assess the validity in reducing number of points progressively according to the rate of the areal scale reduction compared to the rate of the linear scale reduction. It should at least indicate, even at the theoretical level, whether it is valid to assume that, since the mapping space for features is progressively reduced line features should be reduced according to this rate. The formula is expressed as follows:

$$Nf = Na * (Ma/Mf)^2$$

In order to appreciate the difference between the three formulae, they are applied to a hypothetical example, where an irregular line feature at its original scale of 1:25,000 has

200 points and is 20 cm long and to be represented at three smaller scales; 1:50,000, 1:250,000, and 1:1,000,000. The results of this example are presented in Table 5.1.

| **Scale** | **RL** (points) | **MR1** (points) | **MR2** (points) | **Line Length to scale** (cm) |
|---|---|---|---|---|
| 1:25,000 (original, 200 points) | 200 | 200 | 200 | 20 |
| 1:50,000 | 141 | 100 | 50 | 10 |
| 1:250,000 | 63 | 20 | 2 | 2 |
| 1:1,000,000 | 32 | 5 | 0.125 | 0.5 |

**Table 5.1: Comparison of the results from the application of the Radical Law and two other modifications for simplifying a line feature with 200 points. (The predicted number of point by MR2 at 1:1,000,000 is only 0.125, whereas the logical choice would be at least 2 points which is the absolute minimum number of points to represent a line).**

As can be seen from Table 5.1, MR2 produces the largest reduction of points, whereas RL produces the least reduction. Considering the length of the line at derived scales, it would be logical to assume that both MR2 and RL are the extreme cases in which the line is either oversimplified or retains superfluous points. For example, if the line at the scale of 1:1,000,000 would appear 0.5 cm long, it would be unreasonable, from a data reduction perspective, to represent it with 35 points, as suggested by RL. Similarly, it would be difficult to represent the line with a '0.125' point as determined by MR2. Irrespective of the detailed complexity of the line, the first option is likely to provide redundant data, whereas the MR2 option provides an unreasonable number of points, and even if it produced two points which is the minimum number of representing a line. As a result, MR1 represents a compromise between the two formulae, and conforms to the rate of linear scale reduction (i.e., data are proportionately reduced to the rate of scale reduction) as well as to the fact that the features concerned are lines. Thus, the formula corresponds to the rate of areal scale (MR2) is most likely to reduce point data excessively, and hence is expected to cause distorted or spiky shapes, which is an undesirable effect from the cartographic point of view (i.e., data reduction, simplification or generalisation). On the other hand, the Radical Law (RL) tends to cause a relatively minimal data reduction, thereby creating redundant data. This corresponds to the finding of Barber *et al.*, (1995) indicating that the "the Radical Law seemed to retain more points than necessary at smaller scale compilations. It also did not account for differences in line complexity. The relationship between scale, line complexity , and how many points to retain require further investigation." (p. 290).

While the above three formulae are employed as a guide for the determination of the required number of points at derived scales, an iterative process has to be performed in deriving those points by the Douglas-Poiker algorithm. Although MR1 is a more sensible choice with comparison to the other two formulae for a data reduction process within the scale-dependent context, it has no implication for the wider issues of any data reduction algorithm, or the effects of graphic reduction. Because the number of points is not a parameter within the simplification algorithms, there is a need for a formula as a tolerance prediction method when using the Douglas-Poiker algorithm. So, for a particular scale, the original number of points is reduced by the algorithm in accordance with the scale reduction, output resolution, data reduction requirements, and the algorithm's behaviour. The considerations that have to be involved in the development of such a formula are therefore mainly concerned with data reduction within the scale-dependent context, and the algorithm's behaviour on different types of lines. Regarding the question of data reduction, such a predictive formula has to accommodate the rate of reduction (or scale change), as well as considering the fact that data reduction has to be appropriately achieved; that is, data have to be reduced but not to reach the stage at which the distortion of shape becomes perceptible at target scales. This implies that so long as the scale-dependent context is concerned the distortion is acceptable if it occurs in the imperceptible realm. Therefore, the requirement for a successful data reduction process at target scales is to produce maximum data reduction accompanied by imperceptible differences between the original feature and its simplified version. The question is, how much is that maximum reduction? Until now, no answer has been presented. This study maintains that unless researchers distinguish between data reduction, cartographic simplification and generalisation, that and other related questions would be impossible to answer. Based on the definition and the scope of the data reduction (or weeding) process provided above, algorithms designed for data reduction purposes, such as the Douglas-Poiker algorithm, are therefore to be used in a way that feature caricatures or distorted shapes of line features at taıget scales are to be avoided. It is assumed, as pointed out above, that scale-dependent data reduction should be performed after correcting digitising errors and after redundant data at the original scale are weeded out. That is, the line at the original or source scale should be represented by a minimum number of digitising points that capture its character, so that when overlaid on its manual or source counterpart at the original scale it should exhibit no perceptible difference.

So far as scale-dependent data reduction for weeding purposes is concerned, it is reasonable to assume that databases of irregular lines have to be reduced with scale. For example, databases of irregular line features containing hundreds, thousands or more points, should be reduced if they were to be displayed at smaller scales. As implied, straight or regular segments or line features which are represented by a limited number of critical points are excluded for reasons discussed above. Accordingly, a *minimum requirement* criterion is proposed for assessing the quality of the data reduction process. This applies to the number of points of line features but not to their lengths. Moreover, the criterion applies to irregular line features or semi-regular or smooth lines that contain a relatively large number of critical points (e.g., databases composed of hundreds or thousands of points), but what is clearly to be excluded are straight lines or features of regular shapes (e.g., town plans and street guides) which are likely to be subject to generalisation processes such as omission and typification, more than to weeding processes. This is important since regular-shaped features are usually represented by a relatively limited number of critical points and further processing of these points may well cause undesirable modifications, since the resulting shapes would appear different across scales.

Thus, based on the discussion of the formulae of RL and MR2, if, for example, a line at the scale of 1:25,000 is to be reduced to 1:50,000, it is proposed that the line should have its database (points) reduced by a minimum percentage of 50 %. That is, according to the above discussion of the Radical Law (RL) and its second modification (MR2) which corresponds to areal scale, it is proposed at this stage it would be safer and more desirable to reduce linear databases in relation to linear scale, in order to avoid excessive or limited data reductions that appeared to be associated with the RL and MR2. If the database was reduced according to the algorithm with no perceptible difference between the simplified database and its original, then it should be seen as an added merit for the algorithm. It is also proposed that for medium scales (for practical purposes defined here as those between 1:50,000 and 1:625,000) data reduction should at least correspond to linear scale. Around 90 per cent data reduction would seem sufficient for small scales (defined here, as smaller than 1:625,000). Based on the results of the previous example (Table 5.1), this is initially regarded as a reasonable general requirement, since beyond this range of data reduction simplified lines would endure extreme data reduction or contain superfluous data. If, however, the algorithm produces larger reductions without causing perceptual shape distortions at target scales, then again it should be attributed to the algorithms' ability. Two significant observations should be considered at this stage. First, given the proposed types of line features and the required number of points, the criterion proposed above is just a general guideline as to how much data should be reduced for a particular scale change, since no answer to this question exists in the literature. Second, it is expected to find that adherence to the criterion does not necessarily yield acceptable graphical results, the criterion is only suggested here and is subject to assessment in this chapter, experimenting with different types of databases.

On the question of the Douglas-Poiker algorithm, it is necessary to first illustrate briefly the description of line features in a typical digital format upon which the Douglas-Poiker algorithm works. Lines in digital forms are represented by points, segments and arcs; as can be demonstrated in Figure 5.5.

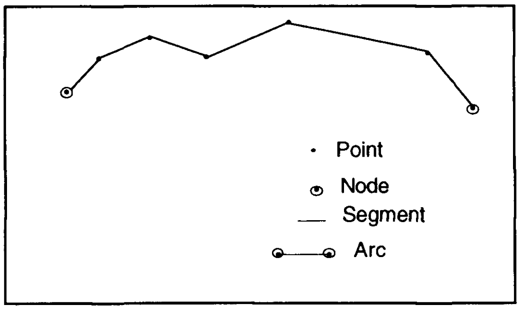



Figure 5.5: Digital encoding of line features.

In Figure 5.5, the line feature is represented by digitising points connected by segments, while the first and last points are coded as nodes marking the start and end of the line. Therefore, the line feature, here, is represented by 7 points, 6 segments, and 1 arc. Obviously, digital lines can be represented by any number of arcs, depending on the process of digitising , type and complexity of the features. With reference to Figure 5.1 (describing the algorithm), the line feature being processed in the figure is represented by one arc, upon which the algorithm is applied. If the line is represented by more than one arc, the algorithm will apply the specified tolerance on each arc at a time; that is arcs representing one line feature are processed independently.

As an already established fact about the algorithm, employing high tolerance values produces high levels of data simplification, whereas small values produce moderate or limited data simplification. Given the observations about the effects of areal scale or graphical reduction on data during the process of data reduction within the scale-dependent context, the question is how to avoid large tolerance values that lead to oversimplification, and similarly how to avoid small values which may lead to a limited data reduction. The following example (Figure 5.6) illustrates the effects of applying different tolerance values on a small data set (165 points), which simultaneously shows the effect of graphical reduction on the resulting simplified lines. This data set represents a small section of a line feature selected from the original data set representing the drainage network in study area 2 for the scale of 1:25,000 (discussed in Chapter 4). In the example, the tolerance values were produced through an iterative process in order to yield the number of points determined by the formula of MR1. So, and given the source scale as 1:25,000, the scale of 1:100,000 required 41 points, which was produced by the tolerance value of 10.00 m, for the scale of 1:250,000, the required number of points was 17, produced by 55.00 m, for the scale of 1:625,000, around 7 points was required, produced by 175.00 m, and for the last scale of 1:1,000,000 the required number of points was only 4 points, produced by 230.00 m. The figure shows that all the resulting simplified lines are first displayed at one fixed scale (1:100,000) in the top row, while these lines are displayed at their representative scales, in the bottom row of line representations. As the figure demonstrates, the lines represented at the scale of 1:100,000 show an increasing effect of data reduction as the tolerance values increase. That is, the difference between original (in Red) and simplified lines (in Green) is increasingly perceptible. On the other hand, the figure also indicates that the magnitude of this effect is markedly reduced, as a result of the effect of graphic reduction (the bottom row of line representations). At the scale of 1:100,000 the difference between the original and simplified line is hardly perceptible, but the Douglas-Poiker algorithm produced progressively spiky or caricatural shapes at the remaining smaller scales. In this sense, the algorithm, apart from the scale of 1:100,000, did not perform a data reduction process for all the remaining three scales; that is, weeding out redundant data at the reduced scales without causing perceptible difference between the original and its simplified version. The 41 points line at the scale of 1:100,000 is, therefore, an ideal; i.e., a large data reduction achieved but no perceptible differences between the original and simplified line generated. On the other hand, the tolerance values used for the remaining scales were undesirable, since perceptible distortions of lines were produced. The example clearly demonstrates that although the effects of the MR1 were a compromise between the two extreme effects of the RL and MR2, the results appeared undesirable, due to the fact that the MR1 does not consider factors related to the algorithm, databases, or quality requirements of data reduction for weeding purposes.

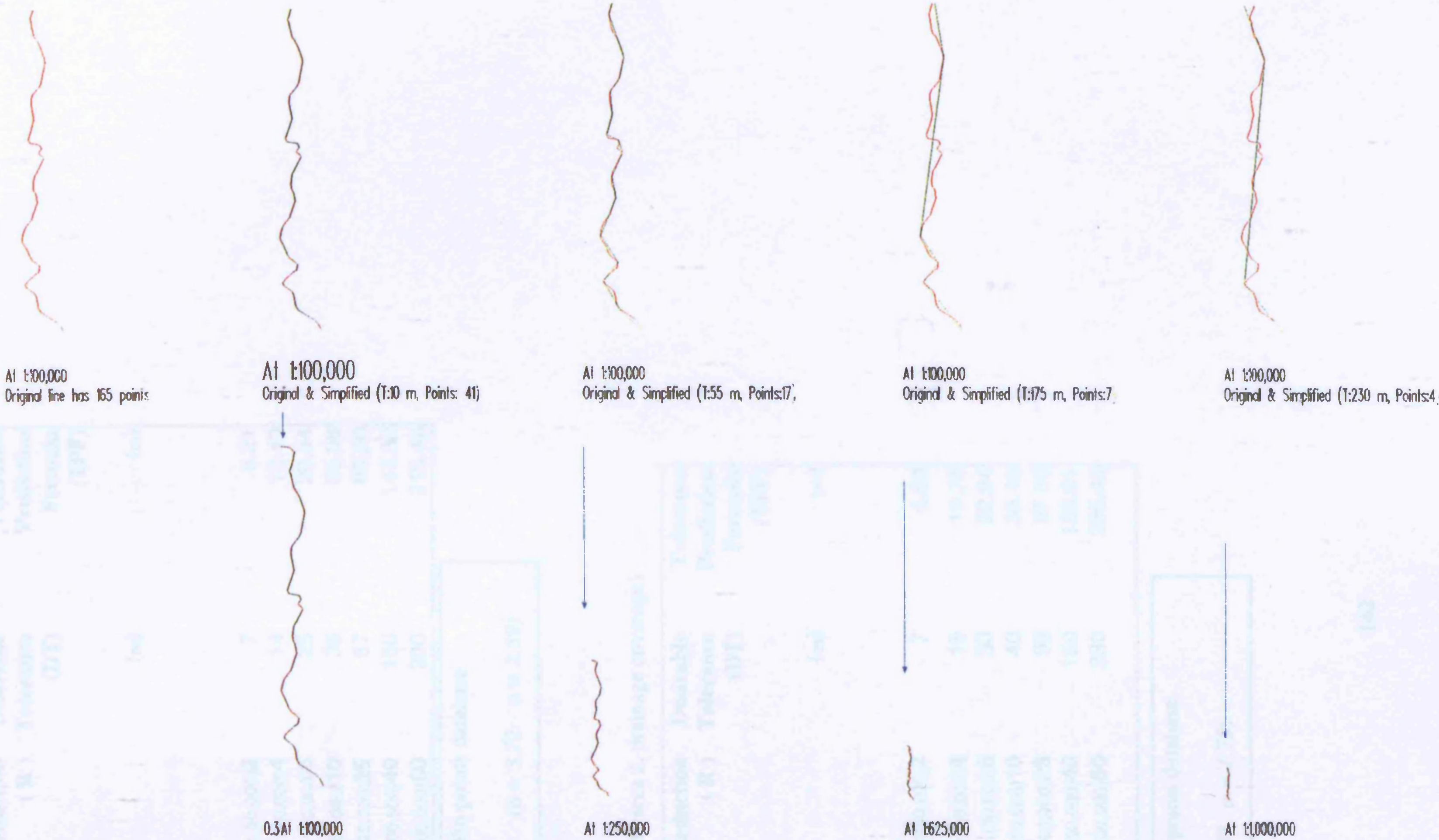



Figure 5.6: The effects of simplification tolerance values by the Douglas--Poiker algorithm on linear data, and the effects of graphic reduction on the simplified data.

**Table 5.2: Relationships between databases, desirable tolerances (DT), reduction (R), and the tolerance prediction formula (TPF).**

***Table 5.2a*** (Database of Study area 1)

| Database (points) | (log of points) | Reduction (R) | Desirable Tolerance (DT) (m) | Tolerance Prediction Formula (TPF) (m) |
|---|---|---|---|---|
| Original scale: 1:25,000 | | | | |
| **4056** | 3.60 | (1:50,000)**2** | 7 | 6.21 |
| 4056 | 3.60 | (1:100,000)**4** | 14 | 13.43 |
| 4056 | 3.60 | (1:150,000)**6** | 25 | 20.64 |
| 4056 | 3.60 | (1:250,000)**10** | 36 | 35.08 |
| 4056 | 3.60 | (1:625,000)**25** | 87 | 89.20 |
| 4056 | 3.60 | (1:1,000,000)**40** | 150 | 143.32 |
| 4056 | 3.60 | (1:1,500,000)**60** | 200 | 215.48 |

R v DT for the 4056 points database
y (DT) = bR + a
(*b = 3.72 a = 2.59*)

***Table 5.2b*** (Database of Study area 2, drainage coverage)

| Database (points) | (log of points) | Reduction (R) | Desirable Tolerance (DT) (m) | Tolerance Prediction Formula (TPF) (m) |
|---|---|---|---|---|
| Original scale: 1:25,000 | | | | |
| **8717** | 3.94 | (1:50,000)**2** | 7 | 6.88 |
| 8717 | 3.94 | (1:100,000)**4** | 16 | 14.76 |
| 8717 | 3.94 | (1:150,000)**6** | 30 | 22.64 |
| 8717 | 3.94 | (1:250,000)**10** | 40 | 38.40 |
| 8717 | 3.94 | (1:625,000)**25** | 98 | 97.50 |
| 8717 | 3.94 | (1:1,000,000)**40** | 150 | 156.61 |
| 8717 | 3.94 | (1:1,500,000)**60** | 230 | 235.42 |

R v DT for the 8717 points database
y (DT) = bR + a
(*b = 3.82 a = 1.79*)

*Table 5.2c* (Database of Study area 2, transport coverage)

| Database (points) | (log of points) | Reduction (R) | Desirable Tolerance (DT) (m) | Tolerance Prediction Formula (TPF) (m) |
|---|---|---|---|---|
| Original scale 1:25,000 **447** | 2.65 | (1:50,000)**2** | 5.5 | 4.30 |
| 447 | 2.65 | (1:100,000)**4** | 12 | 9.60 |
| 447 | 2.65 | (1:150,000)**6** | 20 | 14.90 |
| 447 | 2.65 | (1:250,000)**10** | 26 | 25.50 |
| 447 | 2.65 | (1:625,000)**25** | 70 | 65.25 |
| 447 | 2.65 | (1:1,000,000)**40** | 115 | 105.01 |
| 447 | 2.65 | (1:1,500,000)**60** | 170 | 158.01 |

R v DT for the 447 points database
y (DT) = bR + a
*(b = 2.83 a = 0.27)*

*Table 5.2d* (Database of Study area 4, drainage coverage)

| Database (points) | (log of points) | Reduction (R) | Desirable Tolerance (DT) (m) | Tolerance Prediction Formula (TPF) (m) |
|---|---|---|---|---|
| Original scale 1:25,000 **721** | 2.85 | (1:50,000)**2** | 6 | 4.71 |
| 721 | 2.85 | (1:100,000)**4** | 12 | 10.43 |
| 721 | 2.85 | (1:150,000)**6** | 21 | 16.14 |
| 721 | 2.85 | (1:300,000)**12** | 40 | 33.29 |
| 721 | 2.85 | (1:500,000)**20** | 60 | 56.15 |
| 721 | 2.85 | (1:1,000,000)**40** | 130 | 113.31 |
| 721 | 2.85 | (1:1,500,000)**60** | 150 | 170.47 |

R v DT for the 721 points database
y (DT) = bR + a
*(b = 2.62 a = 5.87)*

*Table 5.2e* (Database of Study area 4, transport coverage)

| Database (points) | (log of points) | Reduction ( R ) | Desirable Tolerance (DT) (m) | Tolerance Prediction Formula (TPF) (m) |
|---|---|---|---|---|
| Original scale 1:25,000 **319** | 2.50 | (1:50,000)**2** | 6 | 4.00 |
| 319 | 2.50 | (1:100,000)**4** | 12 | 9.01 |
| 319 | 2.50 | (1:150,000)**6** | 20 | 14.02 |
| 319 | 2.50 | (1:300,000)**12** | 40 | 29.04 |
| 319 | 2.50 | (1:500,000)**20** | 55 | 49.07 |
| 319 | 2.50 | (1:1,000,000)**40** | 120 | 99.15 |
| 319 | 2.50 | (1:1,500,000)**60** | 175 | 145.22 |

R v DT for the 319 points database
y (DT) = bR + a
(*b* = *2.90* *a* = *1.29)*

**Relationship between R and DT for all the databases:**
**$r^2$ = 0.96**

**Relationship between DT and TPF for all the databases:**
**$r^2$ = 0.98**

Table 5.2 shows the results of experiments on five databases: study area 1 (Table 5.2a - 4056 points), two databases of study area 2 (Tables 5.2b and 5.2c, respectively - drainage coverage: 8717 points; transport coverage: 447 points), and the two databases of study area 4 (Tables 5.2d and 5.2e, respectively - drainage coverage: 721 points; transport coverage: 319 points). These databases are of varying complexity and were subjected to various levels of data reduction (upto 60 fold reduction) corresponding to representative scales. For each database and for each level of data reduction the algorithm was applied experimenting with different tolerance values until a desirable tolerance (DT) was reached for producing desirable graphical results. The value of DT was determined in a manual iterative process, and each entry in the tables may be the outcome of 10 - 15 iterations, or more. The value used was that where no discernible difference occurred between the original and simplified map data at the target scale. As the tables indicate, linear relationships were found between rate of reduction ( R) and desirable tolerances (DT) for all the databases, with an overall coefficient determination equal to 0.96, indicating a significant relationship. The equation of the linear relationship for each database provided a useful insight as to how it is possible to present a formula predicting tolerances which should, at least, approximate the desirable tolerances shown in the results. The equations indicated that substituting *b* with *log* (number of points) in the databases should provide the formula needed, since the results show that the values of *b* for each database exhibit close agreement with *log* (numbers of points) for that database. This formula was called as Tolerance Prediction Formula (TPF) and was, therefore, expressed as:

$$\mathbf{TPF} = k^* ((\log n)^* \mathbf{R}) \qquad \mathbf{(5.1)}$$

where

**TPF** = tolerance required at target reduced scale
***n*** = number of points of original database
***k*** = Dimensioned constant (normally equal to 1m, when units in metre). Its value is dependent on measurement units.
**R** = ratio between derived scale and source scale = Sd/Ss (S is defined as the denominator of scale)

The TPF (Equation 5.1) needed further adjustment. During the preliminary tests, it was found that small changes in tolerance values tend to have a considerable effect on data at especially low levels of simplification, while they tend to have almost no effect at high levels of simplification. Thus, consistent with the data reduction requirements, the formula incorporates a subtraction of ***k*** (in this case 1 metre), in order to avoid over-simplification of data during relatively small changes between source and derived scales. For example, the simplification using the tolerance of 240.00 m for the scale of 1:1,500,000 would not be affected by the process of subtraction of 1.00 m, but a tolerance of 7.00 m for the scale of 1:50,000 would have a noticeable effect on data if it was reduced by (Equation 5.2). That is, the data resulting from either the tolerance of 240.00 m or from the 239.00 m would be the same, whereas the simplified data with the tolerance of 7.00 m is noticeably different (i.e., smaller) than that with a tolerance of 6.00 m. Such an effect is expected since at low levels of data reduction or simplification, minor changes in the tolerance values would necessarily result in changes in the data being processed (i.e., noticeable differences) irrespective of the algorithm used. Whereas at high levels of simplification where the line being process has already undergone a large reduction of data, minor changes in the tolerance values would not produce much difference as opposed to large changes in the tolerance values. This is simply due to the fact that unlike at high levels of simplification, at low levels of simplification the line being processed still retains a lot of the point data, and a slight change in the tolerance value would necessarily incur changes in the data. The TPF is, therefore, preferred as:

$$\mathbf{TPF} = k^* (((\log n)^* \mathbf{R}) - 1) \qquad (5.2)$$

| **Scale** | **TPF without (- 1)** | | **TPF with (-1)** | | **Difference** |
|---|---|---|---|---|---|
| (Original scale is 1:25,000, Database is 165 points) | **Tolerance** | **Points** | **Tolerance** | **Points** | **(number of points)** |
| 1:40,000 | 3.547 | 83 | 2.547 | 102 | 25 |
| 1:50,000 | 4.434 | 75 | 3.434 | 84 | 9 |
| 1:75,000 | 6.65 | 59 | 5.65 | 65 | 6 |
| 1:100,000 | 8.87 | 45 | 7.87 | 49 | 4 |
| 1:250,000 | 22.17 | 29 | 21.17 | 30 | 1 |
| 1:625,000 | 55.44 | 17 | 54.44 | 18 | 1 |
| 1:1,000,000 | 88.70 | 10 | 87.70 | 10 | 0 |

**Table 5.3: The effect of subtraction of *k* (1.00 m) within TPF on a database at low and high levels of data reduction using the Douglas-Poiker algorithm**

Table 5.3 shows the difference in results when the TPF incorporates (- 1) and without (-1) in units of ***k*** (m). For this purpose, the line with a database of 165 points in Figure 5.6 was utilised. Table 5.3 clearly indicates that the difference between the resulting points by both types of the TPF is larger at low levels of data reduction (corresponding to larger scales). On the other hand, this difference becomes increasingly reduced, from 25 to 0 number of points. Graphical results indicated that the results of the TPF without (- 1) appeared less satisfactory at the first two scales (1:40,000, and 1:50,000), since some perceptible difference between the simplified and original lines was detected. It is expected that the effect of both types of the TPF would be larger for larger databases (e.g., with thousands of points); and hence it is expected to find undesirable (excessive) data reduction only at low levels of data reduction. Table 5.2 shows the tolerances from the TPF in Equation 5.2 compared with those of the DT. The results show that there is a very good agreement between the two results with a coefficient determination of 0.98. Most of the graphical results of the TPF are shown later in the chapter.

| **Scale** (Original scale is 1:50,000, Database is 1588 points) | **TPF where ($k$=1)** | | | **TPF where ($k$=2)** | | |
|---|---|---|---|---|---|---|
| | **Tolerance** | **Points** | **Data reduction %** | **Tolerance** | **Points** | **Data reduction %** |
| 1:100,000 | 5.40 | 728 | 54.16 | 10.80 | 492 | 69.02 |
| 1:250,000 | 15.00 | 404 | 74.56 | 30.01 | 260 | 83.63 |
| 1:625,000 | 39.01 | 224 | 85.90 | 78.02 | 151 | 90.50 |
| 1:1,000,000 | 63.02 | 168 | 89.43 | 126.03 | 107 | 93.27 |

**Table 5.4: The effects of increasing tolerance values of the TPF on data reduction when working with a source database at scale smaller than 1:25,000.**

Due to the fact that source databases may be of different resolutions (scales), the TPF was subject to further modification. It is proposed that there is a source scale above which $k$ should be equal to 2, and a scale below of which $k$ = 1. These scales are 1:25,000 and 1:50,000, respectively. For scales between 1:25,000 and 1:50,000 $k$ would vary and for simplicity a linear variation could be used. Given the effect of areal scale (or graphical reduction), this condition is important, since it would ensure that the resulting tolerances would fall within the range of tolerances found to be effective in producing the desirable graphical results and to some extent meet the criterion proposed. Otherwise, the reduction would be limited, as the value of **R** (which is the ratio between scales) would be smaller; hence smaller tolerances. The reason for choosing scales of 1:25,000 and 1:50,000 in this condition, is that a change can be observed to occur in both the numerical and graphical results of Table 5.4, and other preliminary tests with the database of study area 3 (see results of study area 3 in section 5.5.3). An example is provided here to emphasise this point. Table 5.4 shows that the database with 1588 points of the 1:50,000 generalised map of study area 1 was chosen as a source database. The results show that the data reductions resulting from applying the TPF with $k = 2$ are larger than those of the TPF with $k = 1$. However, the graphical results of the TPF with $k = 2$ proved desirable at reduced scales. Obviously, representing the map of this database (1588 points) at the scale of 1:100,000 with 492 points and without detecting a difference between the two maps is more desirable than if the map was represented with 728 points.

The following example (Figure 5.7) demonstrates the application of the TPF (Eqaution 5.2) on the same line (165 points) and for the same scales in Figure 5.6. Given the source scale as 1:25,000, these values and the resulting number of points from their application are as follows: for the scale of 1:100,000, the tolerance value or TPF was 7.87 m, resulting in 49 points; for the scale of 1:250,000, the TPF was 21.17 m, resulting in 30 points; for the scale of 1:625,000, the TPF was 54.44 m, which produced 18 points; and for the last scale of 1:1,000,000, the TPF was 87.70 m, resulting in only 10 points. Figure 5.7 illustrates these results in a similar fashion as in Figure 5.6, where all the resulting simplified lines are first displayed at a one fixed scale (1:100,000) in the top row of line representations, and the same lines are displayed again but at their representative scales below (at 1:100,000, 1:250,000, 1:625,000, 1:1,000,000). Figure 5.7 shows that in the first representation of lines at the scale of 1:100,000, the effects of the simplifications are systematically increased with increasing TPF values. That is, apart from the first simplification for this scale (1:100,000), the simplified lines (in Green) are increasingly distinguishable from the original line (in Red). On the other hand, all the simplified lines appear imperceptible from the original when they are displayed at their representative scales (in the bottom row).

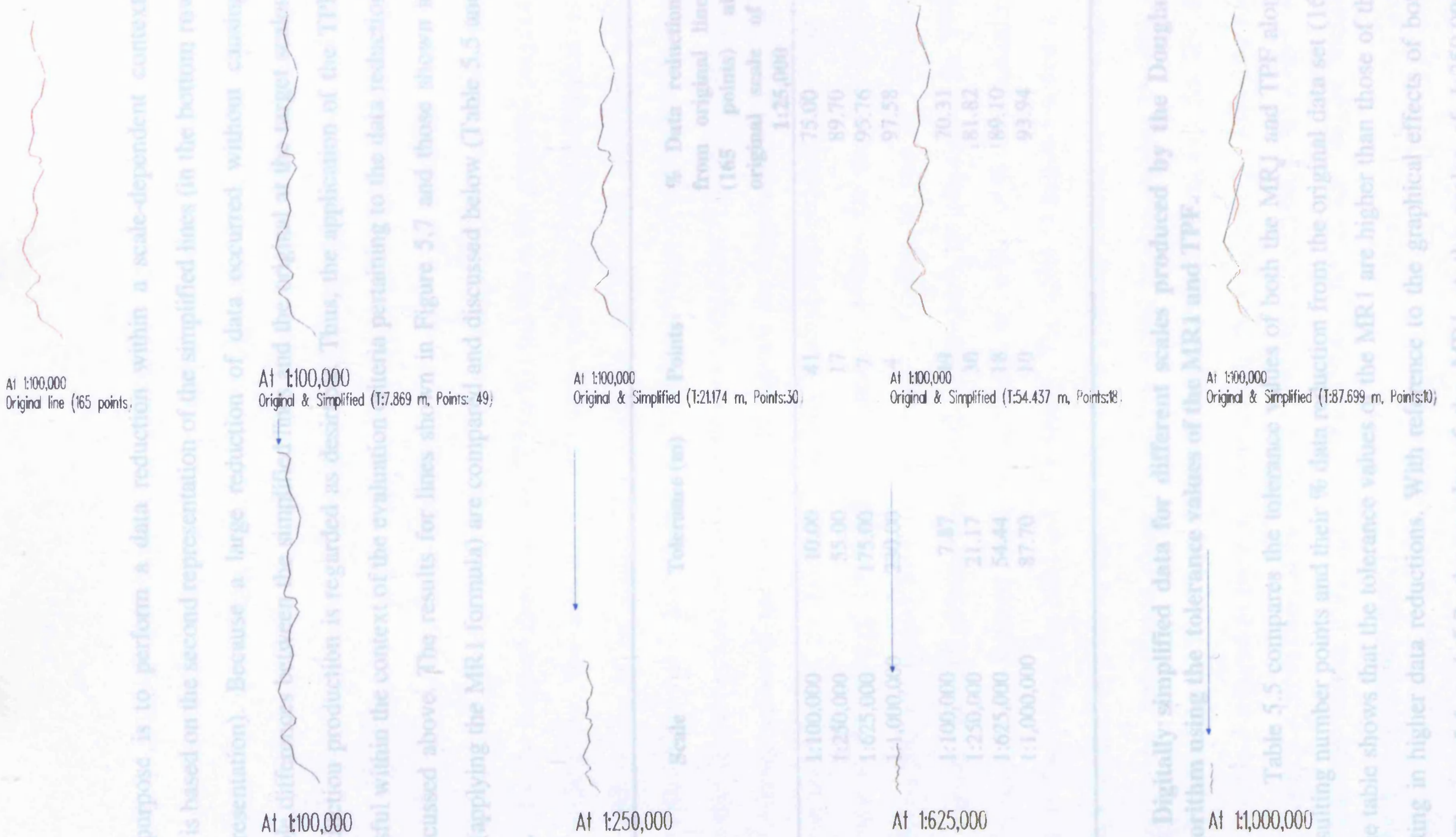



Figure 5.7: The effects of simplification tolerance values of TPF by the Douglas-Poiker algorithm on linear data, and the effects of graphic reduction on the simplified data.

Since the purpose is to perform a data reduction within a scale-dependent context, assessment is based on the second representation of the simplified lines (in the bottom row of the representation). Because a large reduction of data occurred without causing imperceptible differences between the simplified lines and the original at the target scales, the data reduction production is regarded as desirable. Thus, the application of the TPF was successful within the context of the evaluation criteria pertaining to the data reduction process discussed above. The results for lines shown in Figure 5.7 and those shown in Figure 5.6 (applying the MR1 formula) are compared and discussed below (Table 5.5 and Figure 5.8).

| **Formula** | **Scale** | **Tolerance (m)** | **Points** | **% Data reduction from original line (165 points) at original scale of 1:25,000** |
|---|---|---|---|---|
| ***MR1*** | 1:100,000 | 10.00 | 41 | 75.00 |
| | 1:250,000 | 55.00 | 17 | 89.70 |
| | 1:625,000 | 175.00 | 7 | 95.76 |
| | 1:1,000,000 | 230.00 | 4 | 97.58 |
| ***TPF*** | 1:100,000 | 7.87 | 49 | 70.31 |
| | 1:250,000 | 21.17 | 30 | 81.82 |
| | 1:625,000 | 54.44 | 18 | 89.10 |
| | 1:1,000,000 | 87.70 | 10 | 93.94 |

**Table 5.5: Digitally simplified data for different scales produced by the Douglas-Poiker algorithm using the tolerance values of the MR1 and TPF.**

Table 5.5 compares the tolerance values of both the MR1 and TPF along with the resulting number points and their % data reduction from the original data set (165 points). The table shows that the tolerance values of the MR1 are higher than those of the TPF, resulting in higher data reductions. With reference to the graphical effects of both results in Figures 5.6 and 5.7, the results of the MR1 at the scales of 1:250,000, 1:625,000, and 1:1,000,000 are regarded as unacceptable from the data reduction

1:625,000, and 1:1,000,000 are regarded as unacceptable from the data reduction perspective, since the resulting simplified lines are increasingly distorted (Figure 5.6). In contrast, the simplifications resulting from the TPF clearly appear desirable, since large reductions of data occurred whilst the simplified lines show imperceptible changes from the original at all the reduced scales. As mentioned earlier, any differences between the simplified and original lines in the imperceptible realm are regarded cartographically as irrelevant. Such differences can be seen in Figure 5.7 between the simplified lines (in Green) and the original (in Red) in the first row of line representation. The results from the TPF indicates that the range of data reduction (70 to 93) fall within the proposed range of data reduction (50 to 90), provided that the scale range was from 1:25,000 (original) to 1:1,500,000 and smaller target scales. The results of the TPF are, therefore, ideal compared to those of the MR1, since below and beyond this range of data reduction, the lines would either contain redundant data or become oversimplified. Figure 5.8b shows a plot of the tolerance values of the MR1 and TPF against the target scales. The figure shows that the tolerances of the MR1 progressively increase with reduced scale with a marked increase at the scale of 1:625,000. This increase reflects the principle of this formula (MR1) upon which data have to be reduced according to rate of reduction between the original and target reduced scale. On the other hand, the tolerance of the TPF systematically increase with reduced scale. Figure 5.8a shows a plot of % data reduction against target scales for both the MR1 and TPF results. The figure (5.8a) shows that the data reductions resulted by both formulae almost reveal a similar pattern, as can be seen from the two curves. Satisfaction about the curves would be meaningless without a reference to Figures 5.6 and 5.7. Thus, the curve representing the results of the TPF is seen to be desirable compared to the curve representing the results of the MR1. As the figure shows, there is a break point at the scale of 1:250, 000 beyond which the reduction becomes increasingly limited. Full analysis of the results later in the chapter should examine and explain this observation.

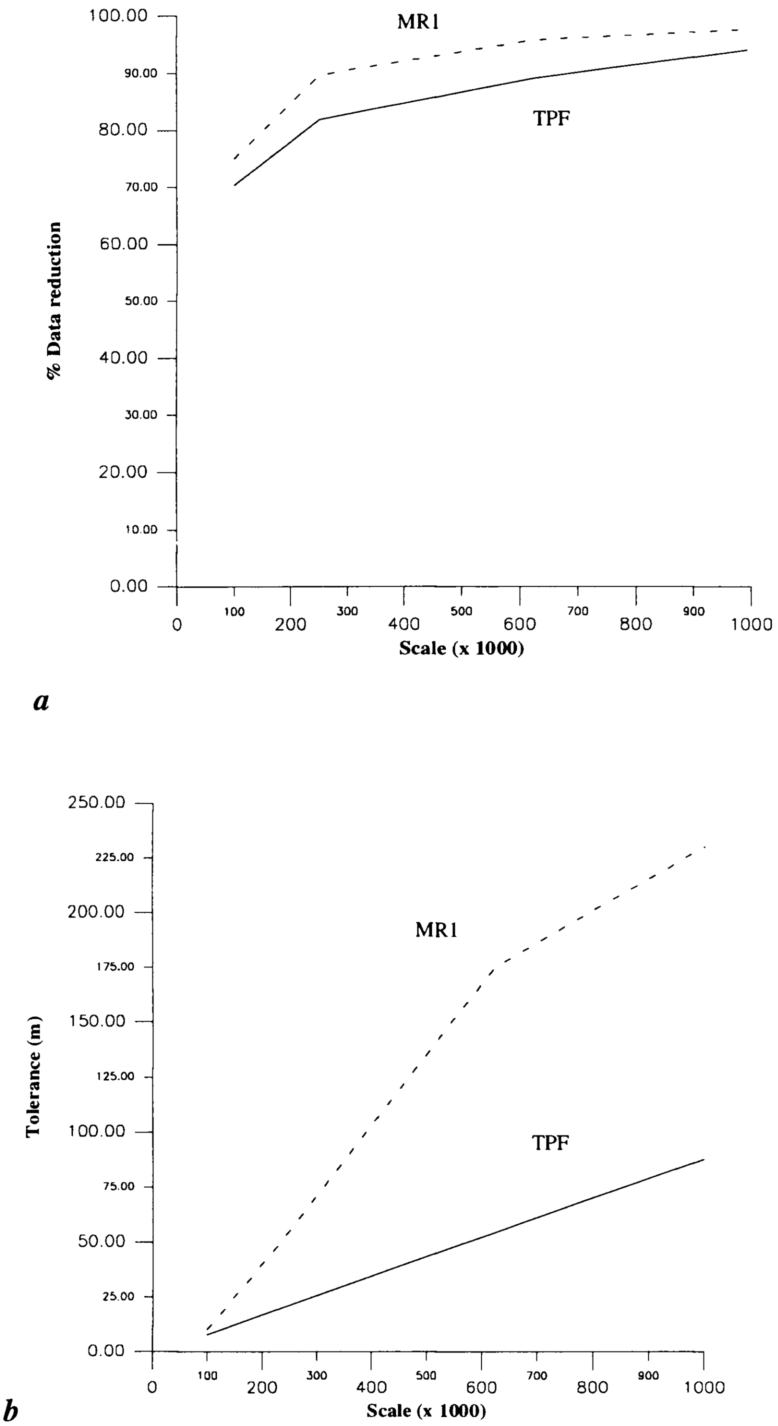



Figure 5.8: Digitally simplified lines from original line (165 points), applying the tolerance values of MR1 and TPF. ***a*** and ***b*** show the relationship between scale, % data reduction, and tolerance values.

These results, especially of the TPF, indicates that the Douglas-Poiker algorithm proved successful in selecting the most critical points of the line in such a way that when the original line was overlaid on its simplified versions at the target scales they produced an imperceptible difference.

Unlike the formula proposed by Abraham (1988) (section 5.4.2.1), the TPF is concerned with the process of data reduction within the scale-dependent context, and considers the algorithm's behaviour (in this case, the Douglas-Poiker algorithm). The previous discussion served as a theoretical background of the development of the TPF which is initially supported by preliminary experiments including the above examples, but the formula (TPF) and the first modification of the Radical Law (MR1), are applied to the data sets selected in this chapter, so as their applicability and validity are further assessed and compared while evaluating the algorithm according to this context (i.e., scale-dependent data reduction for weeding purposes).

*5.4.3 Data sets and methodology:*

The data for this test are the original line sets of all the four study areas, to which their generalised versions are compared in Chapter 4. These lines represent a variety of feature types and complexity, providing a wider context within which the algorithm can be extensively evaluated, applying the two formulae of MR1 and TPF. Once the lines had been digitised, using the Arc/Info program, they underwent a preliminary weeding process for removing redundant points that could have been created during the digitising process (using a very small tolerance of 0.002 inches (1.219 m in ground units) which was the default value in Arc/Info), and correcting digitising errors. Measurement units of all databases are in metre. Other editing was, also, necessary for checking the integrity of the topology of features. The line sets are presented in the form of coverages, some for transport networks and others for drainage networks, except for Study area 1, where all features are presented in one coverage, since most of the lines represent a drainage coverage.

The approach to applying the Douglas-Poiker algorithm within Arc/Info on these coverages is thought to be appropriate, so that the effects of the data reduction process on the topology of the line features are assessed as a whole. The process of data simplification or reduction is performed within a scale-dependent context. All output simplifications are overlaid on their originals at the target scales, and each is drawn in a different colour, supporting an effective visual comparison. Both the originals and their simplified versions are drawn using an output pixel resolution of approximately 0.15 mm. The range of scale is from 1:50,000 to 1:1,500,000, allowing for a reasonable extent of graphic reduction, and hence a thorough assessment. Whilst the application of the proposed formula (TPF) provided direct input tolerance parameters, the other formula (MR1) required iterative processes by the algorithm so that the number of points at target scales can be obtained. The analyses involve perceptual and numerical results. The relationship between the graphic reduction effects and the results of the algorithm application is numerically examined, and an objective method is, therefore, proposed for this purpose, this method is explained in section 5.4.3.1. The results of each study are presented and discussed separately. Throughout these analyses, the term of data reduction is alternately used as data simplification or simplified representations which all imply reducing or weeding out unnecessary data at particular scales.

5.4.3.1 Numeric assessment of graphic reduction effects on data reduction:

Although, the prime aim, here, is to evaluate the Douglas-Poiker algorithm perceptually. However, evaluating the algorithm's approach demands systematic assessments, and most pertinent to this evaluation is the consideration of the significant impact of the process of graphic reduction and output resolution. A quantitative approach to studying this interrelationship is, accordingly, proposed.

The approach is based on the observations stated above about the rate of reduction determined by areal scale. First, the approach involves calculating the total areal displacement of the simplified lines through the overlay process operation supported by the Arc/Info GIS program. This process produces the areal displacement between the original and its simplified versions in the form of *sliver polygons*. Once the areal displacement is calculated by the program, the process of converting the results, which are in ground units, into the map or graphic display units takes place. For example, if the resulting displacement from a simplification for the 1:100,000 scale was 5,809,381 square metres (ground units), it would be about 0.025 square cm (map units). Whilst the resulting displacement is graphically affected by the graphic reduction, the output resolution is invariant at all scales. Second, the approach involves calculating the areal extent of the graphic representation of the actual simplified lines at each scale (in map units) so that they can be compared to the graphical extent at the same scale and using the same map units. This would quantitatively indicate, in general terms, the magnitude of the total areal displacement compared with the graphic representation of the simplified lines, with a constant size of output resolution. Visual inspection will always be essential for effectively evaluating the results of the algorithm, and it is only secondary that such a quantitative approach is regarded as significant at this stage of the evaluation. This is because the desired quantitative results are always based on long-established visual criteria that have, already, been thoroughly tested against the cartographic requirements (i.e., perceptible versus imperceptible change) and hence regarded as standards in the formal realm of cartographic requirements. Data reduction will, therefore, not be acceptable unless there is no deviation between the simplified and original lines; hence it is expected that the quantitative results should support this perceptual result. Theoretically, if the numerical results show that the total areal displacement (in map units) is larger than the total graphical extent of the simplified lines (i.e., thickness of the lines in map units) then it is likely to find that the areal displacement is undesirable (i.e., the displacement is perceptible in the graphical sense). This indicates an excessive reduction of data and the result is therefore rejected. There are, however, some caveats against this assumption; that is, there may be perceptible displacements in some places which clearly can not be acceptable, yet the numerical results may still show that the total areal displacement is far less than the total graphic extent of the simplified lines. This effect is expected with complex details and large coverages. Above all this is a general method and provides a general (total) numerical figure. There is some scope in studying this approach further as to assess the results for both data reduction and simplification perceptually and automatically without the intervention of the cartographer. This can be a research goal which actually implies formalisation of both processes of data reduction and simplification. Thus, this approach can be a valid quantitative measure qualifying digital results in relation to data reduction and cartographic simplification requirements of line features. However, thorough assessment of its validity is yet to be explored which is beyond the limits and scope of this thesis.

## 5.5 Results

### *5.5.1 Study area 1:*

This area is represented by a single coverage comprising different types of features (see section 4.1).

#### 5.5.1.1 Scale, Tolerance, and Data Reduction:

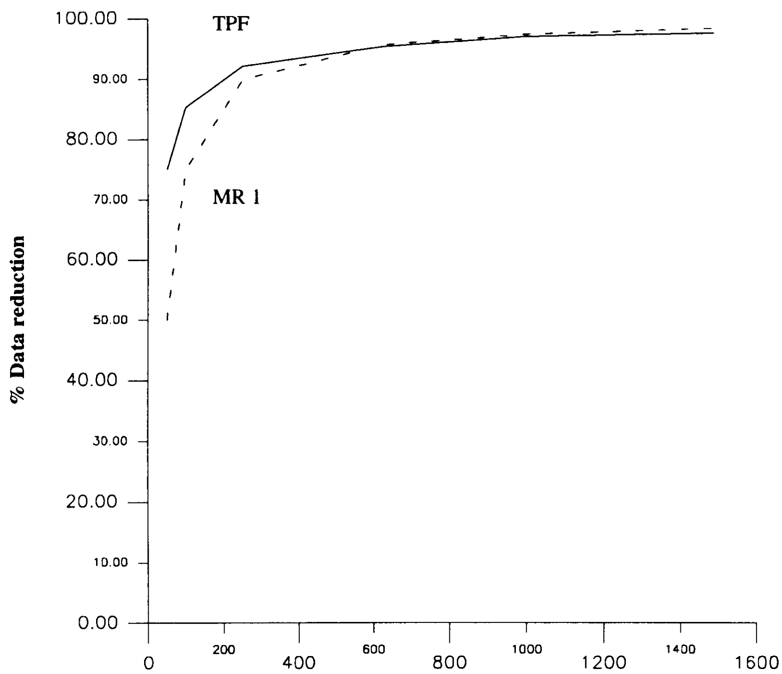
The selected line features in this study area contain 4056 points and 44 arcs, representing the scale of 1:25,000. Figures 5.11, 5.12, 5.13, 5.14, 5.15, and 5.16 show the simplification effects by the algorithm resulted from the application of the MR1 and TPF. The original lines contain 4056 points and the data reduction or simplification was performed for the scales of: 1:50,000, 1:100,000, 1:250,000, 1:625,000, 1:1,000,000, and 1:1,500,000. In the figures , the original line (4056 points) is in Red while the simplified lines are in Green. Visual analysis of the figures indicates that the application of both formulae appear to yield data simplifications which are imperceptible at all scales, except at the scales of 1:1,000,000 and 1:1,500,000. At these two small scales, the application of the MR1 led to high tolerance values, resulting in slight spiky shapes (Figure 5.15). Tables 5.6 and 5.7 show the numeric results of the simplifications from the application of MR1 and TPF, respectively. As the tables and figures indicate, the TPF compared to MR1 allowed for producing smaller tolerance values for small scales, especially for 1:1,000,000 and 1:1,500,000, which resulted in lesser reduction of data; thereby causing almost imperceptible difference between the original and simplified representations. The requirement of minimum reduction, as discussed in section 5.4.2.2, is, therefore, met by the two formulae; however, the TPF appears to produce better results both perceptually at smaller scales (Figure 5.16) and numerically (Tables 5.6 and 5.7). Numerically, Table 5.7 indicates, for example, that the application of the TPF caused the algorithm to reduce data by 74.94 %, for the first scale of 1:50,000, yet no perceptible difference resulted between the original and this simplified representation, as can be seen from Figures 5.12(a) and 5.12(b). Similarly, the TPF reduced data by a 97.50 % for the smallest scale of 1:1,500,000, and yet the simplified lines appear imperceptible from the original. The numeric results are compared in Figure 5.9. The graph in Figure 5.9a, shows that the rate of reduction begins to change to a different pattern at the scale of 1:250,000 at which the reduction starts to become slower towards the smaller scales. Although, subsequent tolerance values were progressively applied, they appear to have little effect on data reduction at these high levels compared to those below the 1:250,000 scale. This limited increase in data reduction beyond this level (around 90 %), reflects the principle upon which the algorithm selects the critical points of the line being processed. Interestingly, the rate of linear scale change at the scale of 1:250,000 is 90 %, and 99 % according to the rate of areal scale change. This is primarily due to the algorithm's approach which is influenced by the position of the points in the original line as opposed to their number; thus the data are not proportionately reduced to the rate of progression of the tolerance values, especially at high levels of data reduction In this, most of the data has already been achieved by the tolerances of both the formulae (24.10 m and 35.32 m) for this scale, and however large might be the tolerance values applied beyond this limit, they only have a limited effect. However, the results show that, even about a 1 % change at the high levels of data reduction (98.27 % by the MR1 and 97.51 % by the TPF) would make a noticeable change in the line shapes, as can be seen at the scale of 1:1,500,000 in Figures 5.15 and 5.16. This emphasises the proposition that reducing data beyond 90 % for small scale representation is questionable from a data reduction point of view, since it might well yield distorted line shapes, whereas a 90 % success is obviously an excellent result itself.

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (4056 points) |
|---|---|---|---|
| 1:50,000 | **2.29** | **2030** | **49.94** |
| 1:100,000 | **6.28** | **1014** | **75.00** |
| 1:250,000 | **24.10** | **409** | **89.92** |
| 1:625,000 | **98.00** | **176** | **95.65** |
| 1:1,000,000 | **192.00** | **109** | **97.31** |
| 1:1,500,000 | **450.00** | **70** | **98.27** |

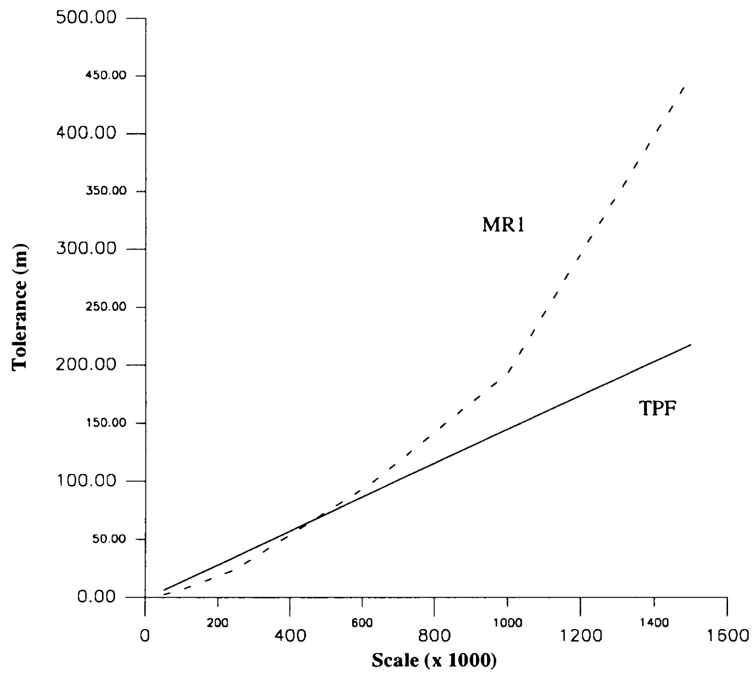
**Table 5.6: Digitally simplified maps of Study area 1 by the Douglas-Poiker algorithm, using the tolerance values of MR1 .**

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (4056 points) |
|---|---|---|---|
| 1:50,000 | **6.264** | **1016** | **74.94** |
| 1:100,000 | **13.53** | **597** | **85.31** |
| 1:250,000 | **35.32** | **320** | **92.11** |
| 1:625,000 | **89.80** | **189** | **95.34** |
| 1:1,000,000 | **144.29** | **123** | **96.97** |
| 1:1,500,000 | **216.94** | **101** | **97.51** |

**Table 5.7: Digitally simplified maps of Study area 1 by the Douglas-Poiker algorithm, using the tolerance values of the TPF.**




*a*




*b*

Figure 5.9: Digitally simplified maps from original map (4056 points, 44 arcs- Study area 1), using two types of tolerance: MR1 and TPF. ***a*** and ***b*** show the relationships between scale, % data reduction and tolerance values.

5.5.1.2 Graphic Reduction and Areal Displacement:

The results of using the TPF with the Douglas-Poiker algorithm, are presented in Tables 5.8 and 5.9, and Figure 5.10. Table 5.8 reports the total areal displacements resulting from the simplification for each scale, while, Table 5.9 compares these displacements in map units with the line areas in the same units at each scale. As mentioned in section 5.4.3.1, the choice of the results of TPF is considered because the MR1 tends to produce perceptible shape distortion at high levels of data reduction, which is not the desired result. In Table 5.8, there is a progressive increase in the displacement with reduced scale, in terms of ground measurements ($m^2$), while there is a progressive decrease in the displacement with reduced scale, in terms of the map units graphic display ($cm^2$). This decrease in the total areal displacements (in map units) is due to the effect of graphic reduction. Table 5.8 also shows that the number of *sliver polygons* (each of which contributes to the total areal displacement) decreases with reduced scale, and the increase in the total displacement. In this sense, the areal extent of the *sliver polygons* increases with reduced scale. At the scale of 1:250,000 (Overlay 3), the total displacement marks a different rate of change, beyond which the displacement becomes progressively large. This reflects the relatively large tolerance value used for the reduction at this scale, compared to the previous two scales of 1:50,000, and 1:100,000 (see Table 5.7). Table 5.9 compares the total areal extent of the graphic representation of the simplified lines at each scale with that of the total areal displacement resulting from the data simplification. The comparison clearly demonstrates that the total areas of lines of the simplified representations, using a pen width of 0.15 mm, are much higher than the sum of areal displacements at target scales, with a small percentage range (from 78 to 82.05). This is to indicate how the increasing displacement is matched by decreasing graphical representations, resulting in almost constant effect, given a constant output resolution. Figure 5.10 shows this effect in a form of the approximately straight line. Obviously, this is an exhaustive quantitative measure describing a general graphic effect, and the question of whether this measure can be utilised to indicate whether the resulting data reduction is acceptable is not tested, since it is a task beyond the scope of this thesis. Consider, for example, the following: if the total areal displacement (in map units) is larger than the graphical areal extent of the simplified lines at a target scale, this might indicate that the data reduction process is not acceptable, since this suggests that the displacement is perceptible. Theoretically, this inference is, to some extent, valid. However, such an example and others are worthy subjects for further research.

In summary, according to the requirements of data reduction suggested above, (section 5.4.2.2) the results showed that the Douglas-Poiker algorithm could successfully produce multiple simplified linear data when applying the proposed formula (TPF). That is, the simplified lines showed imperceptible difference when graphically compared with the original at target scales, especially at the first four scales, even when there was up to 97.51 % data reduction. This clearly emphasises the utility of the Douglas-Poiker algorithm in preserving the most critical points at target scales, while removing less critical ones that are deemed redundant due to the effects of graphic reduction and output resolution.

| Overlays | Number of Sliver Polygons | Sum of Areas ($m^2$) | Sum of Area at target scales ($cm^2$) |
|---|---|---|---|
| Original (1:25,000) | | | |
| | **1452** | 190,849.18 | **.763** (at 1:50,000) |
| **Overlay 1**<br>1:25,000+1:50,000 | | | |
| **Overlay 2**<br>1:25,000+1:100,000 | **820** | 432,883.71 | **.432** (at 1:100,000) |
| **Overlay 3**<br>1:25,000+1:250,000 | **612** | 1,107,158.87 | **.177** (at 1:250,000) |
| **Overlay 4**<br>1:25,000+1:625,000 | **399** | 2,690,000.00 | **.068** (at 1:625,000) |
| **Overlay 5**<br>1:25,000 +<br>1:1,000,000 | **281** | 4,152,155.25 | **.041** (at 1:1,000,000) |
| **Overlay 6**<br>1:25,000 +<br>1:1,500,000 | **254** | 5,809,381.00 | **.025** (at 1:1,500,000) |

**Table 5.8: Areal displacement resulting from overlaying original (4056 points - Study area 1) on digitally simplified maps (using the TPF ).**

| Digitally simplified maps for different scales | Total line length | Area of drawing lines of digitally simplified maps using a Pen Width . 015cm at target scales | Sum of Areal displacement at target scales | Ratio of Difference to Area of Drawing Lines |
|---|---|---|---|---|
| | (m) | ($cm^2$) | ($cm^2$) | % |
| 1:50,000 | 141,665.30 | **4.249** | **.763** (at 1:50,000) | **82.05** |
| 1:100,000 | 140,821.02 | **2.112** | **.432** (at 1:100,000) | **79.55** |
| 1:250,000 | 136,672.86 | **.820** | **.177** (at 1:250,000) | **78.42** |
| 1:625,000 | 133,401.67 | **.320** | **.068** (at 1:625,000) | **78.75** |
| 1:1,000,000 | 128,819.56 | **.193** | **.041** (at 1:1,000,000) | **78.76** |
| 1:1,500,000 | 123,704.39 | **.123** | **.025** (at 1:1,500,000) | **79.68** |

**Table 5.9: Graphic and areal scale effects on the areal displacement resulting from overlaying the 1:25,000 original map (4056 points - Study area 1 ) on digitally simplified maps (using the TPF).**

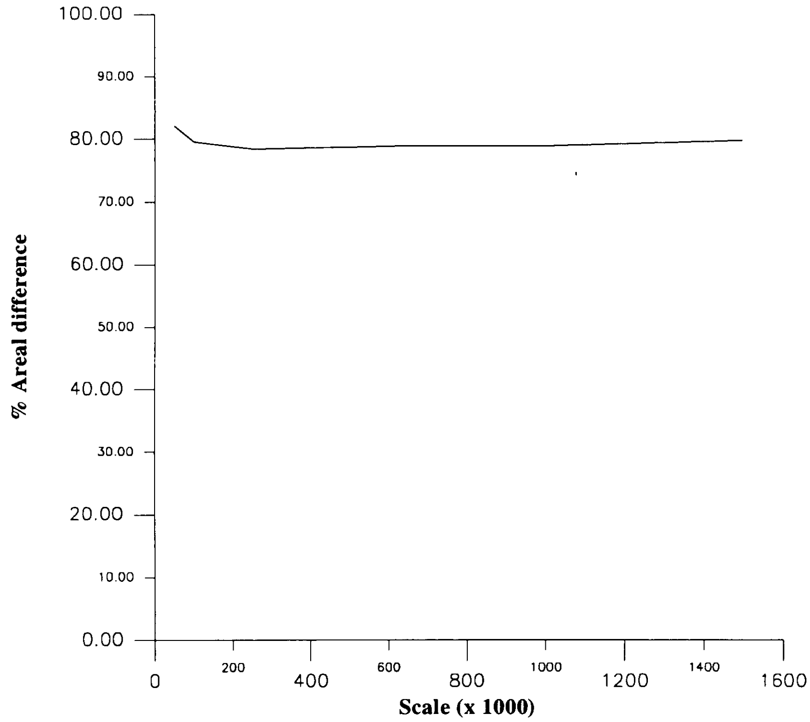



Figure 5.10: Ratio of areal difference (%) to graphic extent of simplified lines. (Difference (in map units) = (total graphic extent of simplified lines) - (total areal displacement). (Study area 1)

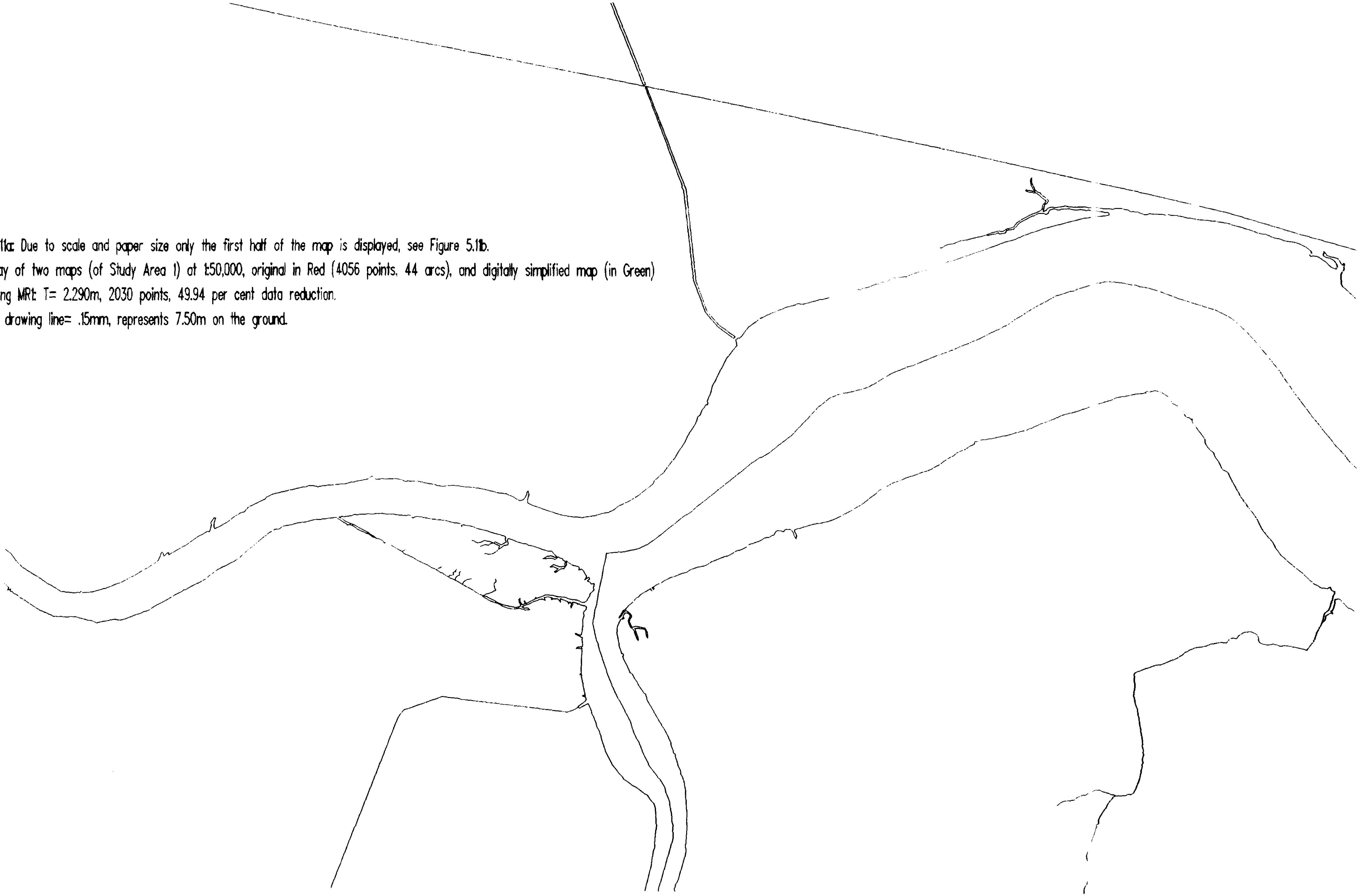


Figure 5.11a: Due to scale and paper size only the first half of the map is displayed, see Figure 5.11b.
An overlay of two maps (of Study Area 1) at 1:50,000, original in Red (4056 points, 44 arcs), and digitally simplified map (in Green) by applying MRt: T= 2.290m, 2030 points, 49.94 per cent data reduction.
Width of drawing line= .15mm, represents 7.50m on the ground.

Figure 5.11b: Due to scale and paper size only the second half of the map is displayed, see Figure 5.11a.
An overlay of two maps (of Study Area 1) at 1:50,000, original in Red (4056 points, 44 arcs), and digitally simplified map (in Green)
by applying MRt: T= 2.290m, 2030 points, 49.94 per cent data reduction.
Width of drawing line= .15mm, represents 7.50m on the ground.

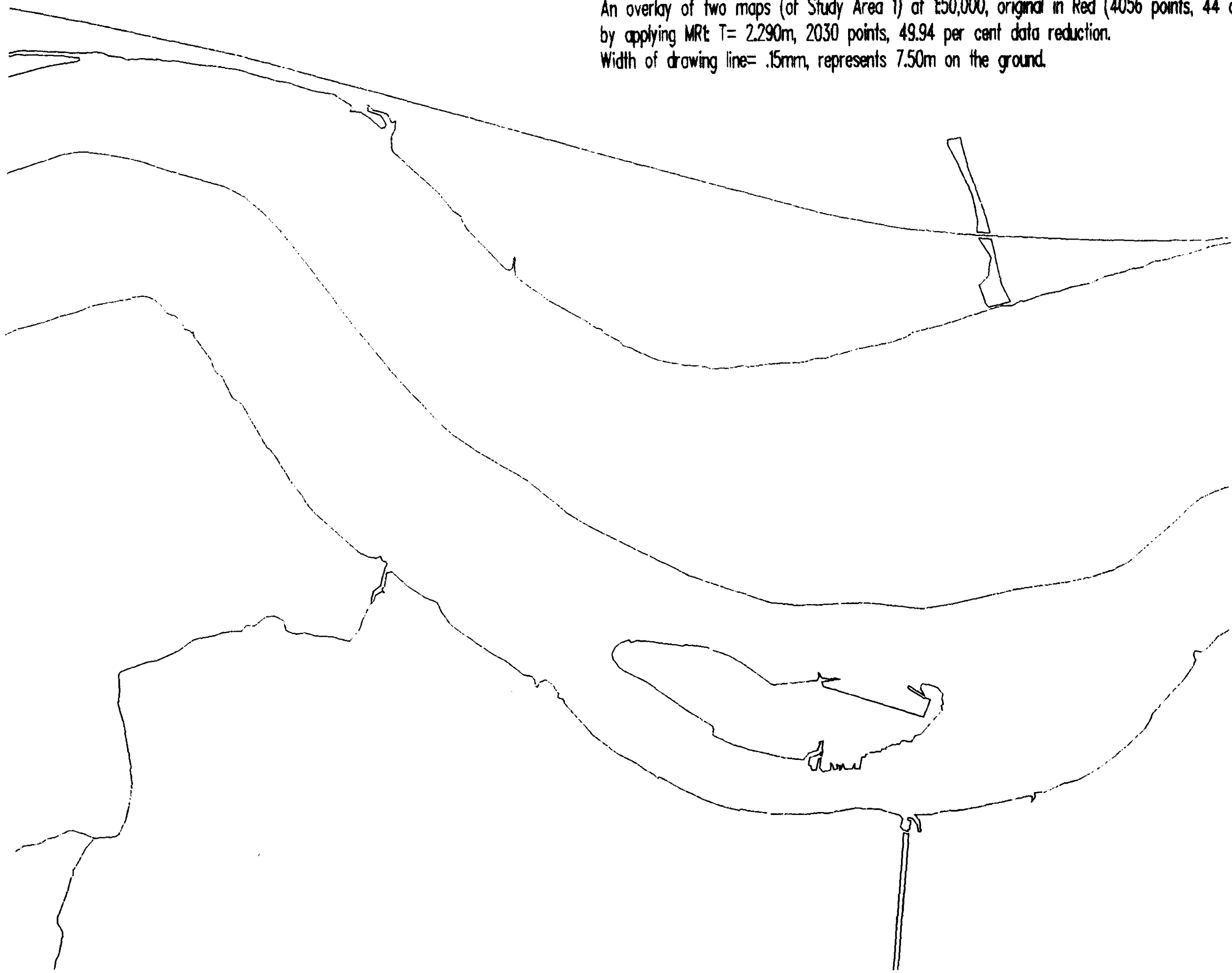


Figure 5.12b: Due to scale and paper size only the second half of the map is displayed, see Figure 5.12a.
An overlay of two maps (of Study Area 1) at 1:50,000, original in Red (4056 points, 44 arcs), and digitally simplified map (in Green)
by applying TPF: T= 6.264 m, 1016 points, 74.94 per cent data reduction.
Width of drawing line= .15 mm, represents 7.50 m on the ground.

Figure 5.12a: Due to scale and paper size only the first half of the map is displayed, see Figure 5.12b.
An overlay of two maps (of Study Area 1) at 1:50,000, original in Red (4056 points, 44 arcs), and digitally simplified map (in Green) by applying TPF: T= 6.264 m, 1016 points, 74.94 per cent data reduction.
Width of drawing line= .15 mm, represents 7.50 m on the ground.

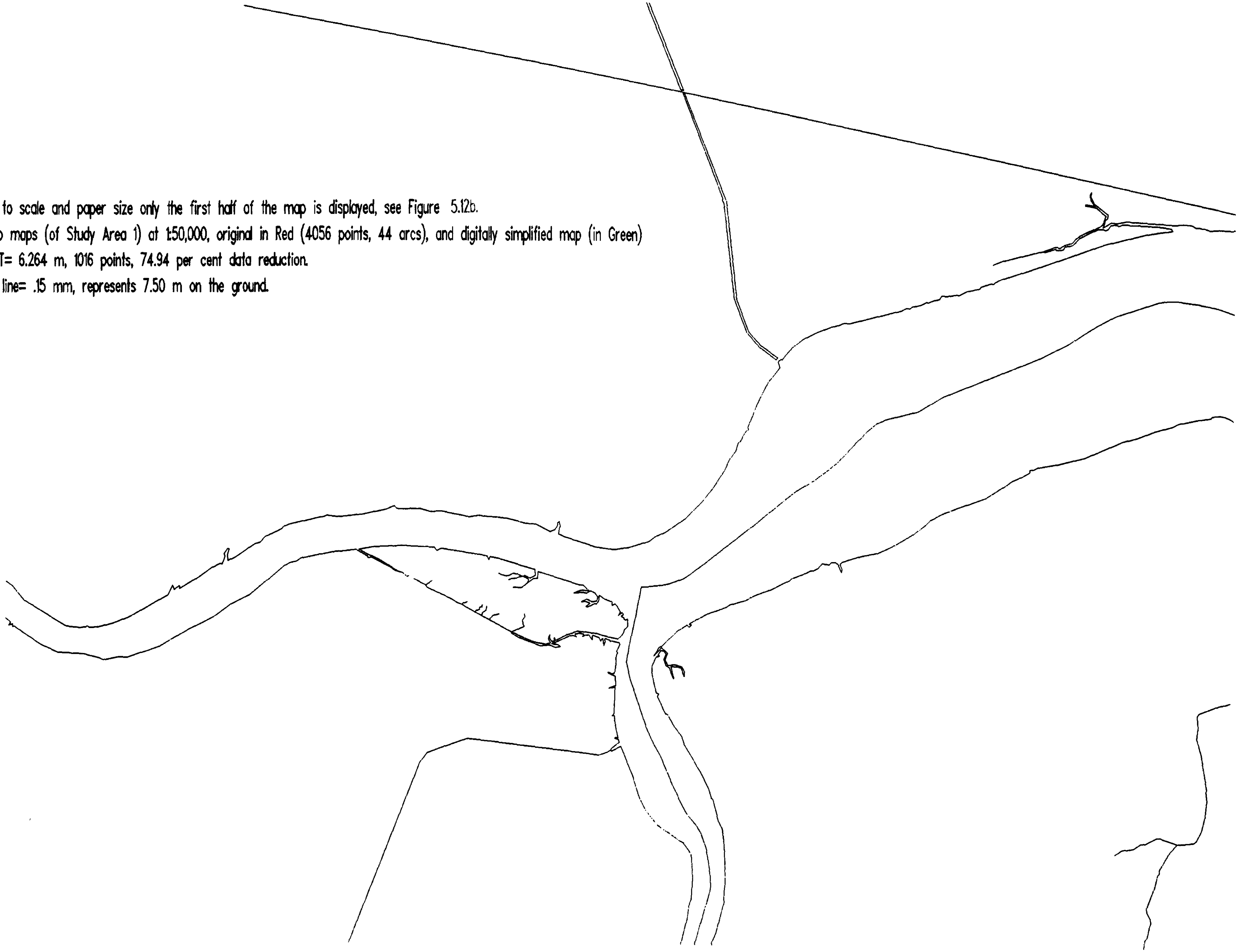
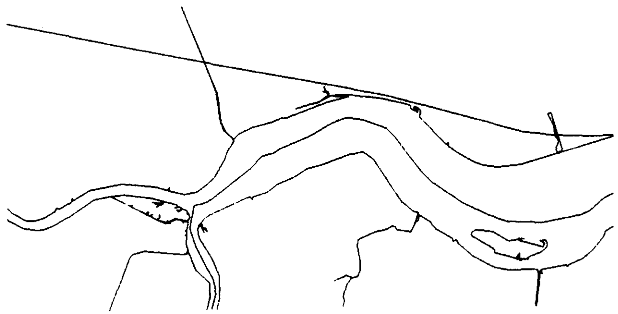


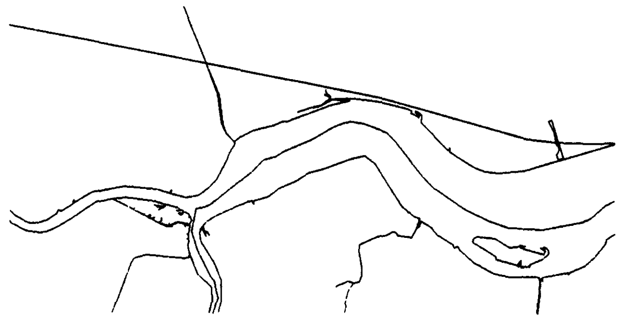
Figure 5.13: An overlay of two maps (of Study Area 1): Original map in Red (4056 points) and, map in Green (1014 points, digitally simplified by applying MR1, T=6.280 m, 75 per cent data reduction).
Width of drawing line: .15mm, represents 15m on the ground.

Figure 5.14: An overlay of two maps (of Study Area 1): Original map in Red (4056 points) and, map in Green (597 points, digitally simplified by applying TPF, T = 13.528 m, 85.31 per cent data reduction).
Width of drawing line: .15 mm, represents 15 m on the ground.

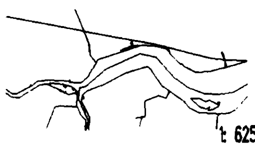
A): Digitally simplified maps in Green (by MR1)

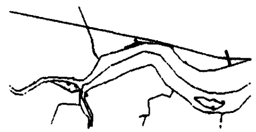
B): Original map in Red (44 arcs, 4056 points) & Simplified maps



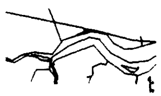


1: 250,000
409 points, 89.92 per cent data reduction
Tolerance value: 24.10 m
Drawing line width: .15mm, represents 37.5m.





1: 625,000
176 points, 95.65 per cent data reduction
Tolerance: 98 m
Drawing line width: .15mm, represnts 93.73 m



1: 1,000,000
109 points, 97.31 per cent data reduction
Tolerance: 192 m
Drawing line width: .15 mm, represents 150 m

1: 1,500,000
70 points, 98.27 per cent data reduction
Tolerance: 450 m
Drawing line width: .15mm, represents 225 m

Figure 5.15: Digitally simplified maps (using MR1) of Study Area 1 overlaid on original at 1:250,000, 1:625000, 1:1,000,000, and 1:1,500,000.

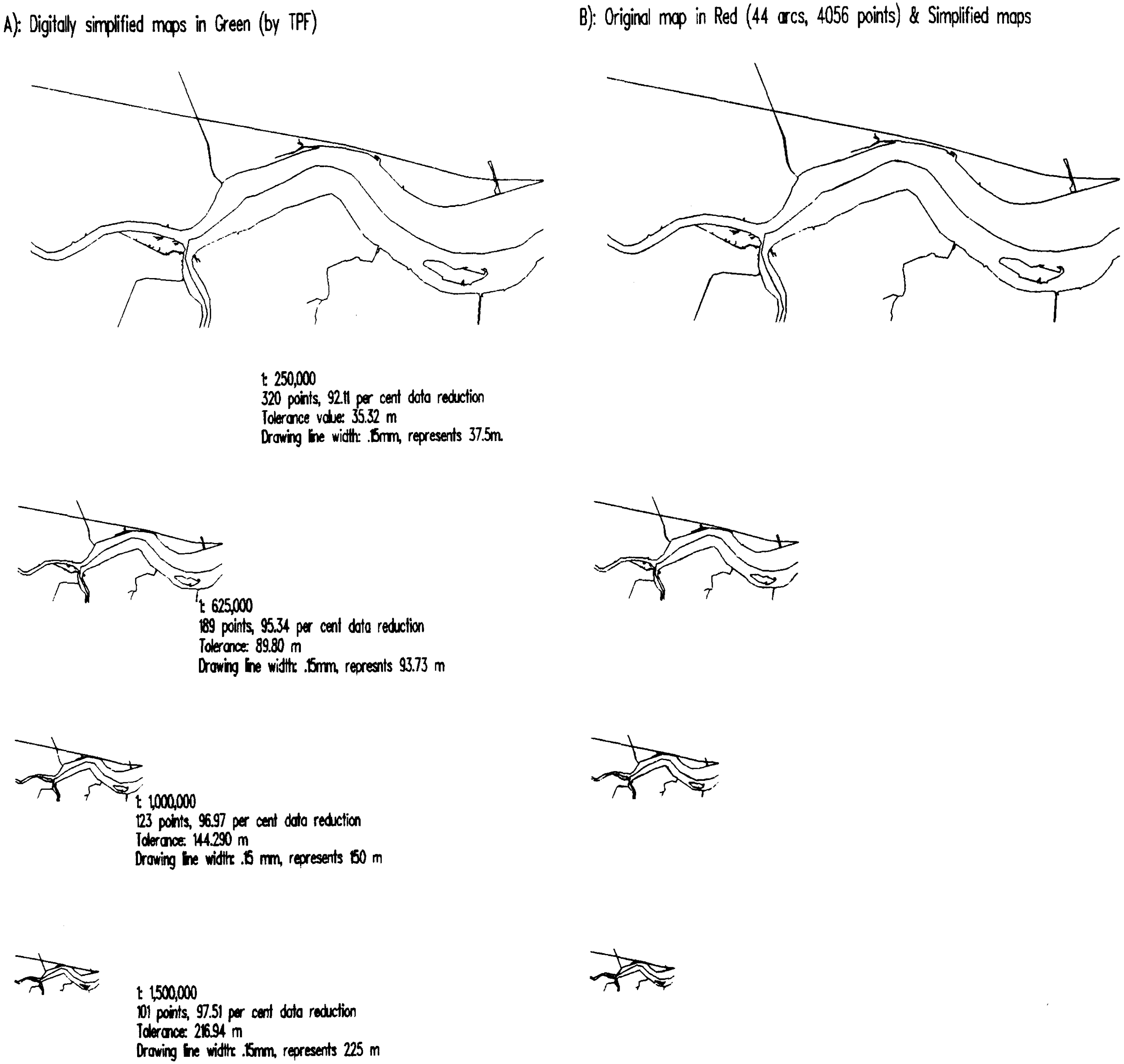



Figure 5.16: Digitally simplified maps (using TPF) of Study Area 1 overlaid on original at 1:250,000, 1:625000, 1:1,000,000, and 1:1,500,000.

*5.5.2 Study area 2:*

This area comprises two types of features: drainage and transport networks (see section 4.1). The results of each network are presented separately.

5.5.2.1 Scale, Tolerance, and Data Reduction (Drainage Coverage):

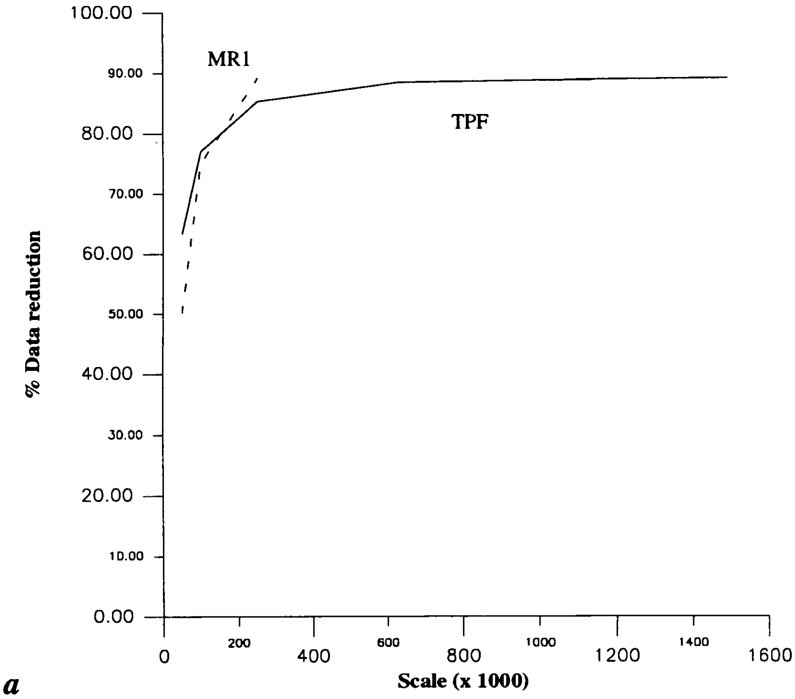
This coverage has a data set of 8717 points and 924 arcs for the scale of 1:25,000. The scales for which the process of data reduction was performed in this coverage include: 1:50,000, 1:100,000, 1:250,000, 1:625,000, 1:1,000,000, and 1:1,500,000. Figures 5.19, 5.20, 5.21, 5.22, and 5.23 show the data reduction effects of the algorithm, using the MR1 and TPF on the original coverage (8717 points), while the numerical results are presented in Tables 5.10 and 5.11. Applying the MR1 caused the algorithm to reach the maximum data reduction at the scale of 1:250,000 (Figure 5.22(A)), where the arcs are all represented by straight lines (original line in Red, and the simplified lines in Green), and so no data reduction was possible for the remaining scales. This maximum reduction was due to using a large tolerance value (500.00 m) which was necessary to produce the required number of points (872) at this scale determined by MR1. At the scales of 1:50,000 and 1:100,000, (Figures 5.19, 4.20, and 5.21) both formulae produced visually acceptable results; as the difference between the original and simplified representations was imperceptible. However, the TPF produced data reductions for all the scales which correspond to the data reduction criteria proposed. That is, the data reductions ranged from over 50 % (63.22) for the scale of 1:50,000 to around 90 % for the scale of 1:1,500,000 (89.05), yet no perceptible differences occurred between the original and simplified lines at all the scales. Tables 5.10 and 5.11 show that the extreme level of data reduction is achieved by MR1 at the scale of 1:250,000 (Figure 5.22(A)), whereas this reduction was approximately arrived at by the TPF at the smallest scale of 1:1,500,000 (Figure 5.23). Figure 5.17 compares the numeric results of both the formulae. In Figure 5.17a , the rate of reduction tends to become slower beyond the scale of 1:250,000. This is due to the reasons provided in section 5.5.1.1. That is, due to the algorithm's approach, application of large tolerance values beyond this scale has limited effect, since data are not proportionately reduced to the rate of progression of the tolerance values, especially at high levels of data reduction. The results, thus, indicate that the application of the TPF was successful in producing the minimum level of data reduction at each scale, according to the evaluation criteria proposed (see section 5.4.2.2).

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (8717 points) |
|---|---|---|---|
| 1:50,000 | **3.971** | **4354** | **50.10** |
| 1:100,000 | **12.85** | **2174** | **75.07** |
| 1:250,000 | **500.00** | **925** | **89.15** |
| 1:625,000 | – | – | – |
| 1:1,000,000 | – | – | – |
| 1:1,500,000 | – | – | – |

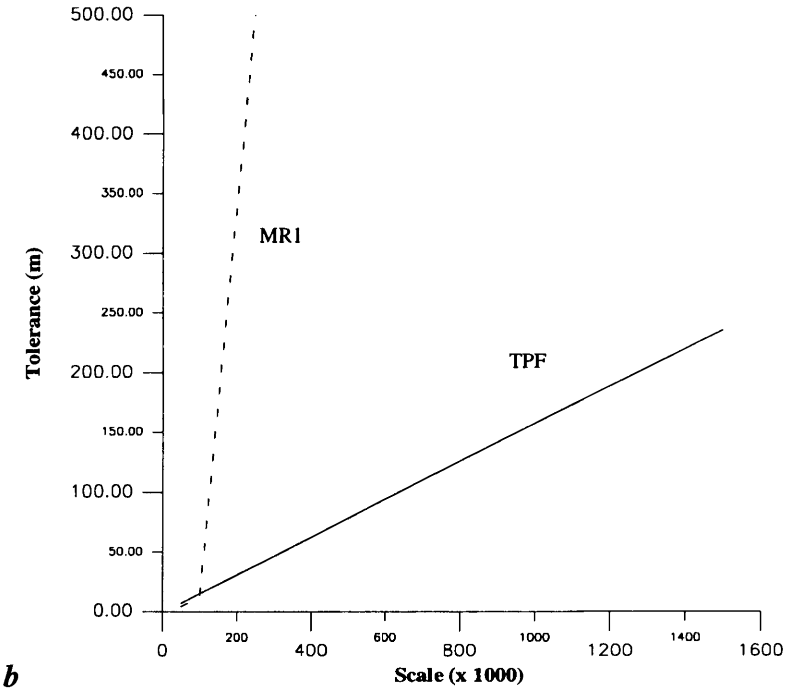
**Table 5.10: Digitally simplified maps from original (Study area 2, Drainage Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the MR1 .**

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (8717 points) |
|---|---|---|---|
| 1:50,000 | **6.87** | **3206** | **63.22** |
| 1:100,000 | **14.76** | **2007** | **76.98** |
| 1:250,000 | **38.40** | **1282** | **85.30** |
| 1:625,000 | **97.50** | **1014** | **88.37** |
| 1:1,000,000 | **156.59** | **980** | **88.76** |
| 1:1,500,000 | **235.39** | **955** | **89.05** |

**Table 5.11: Digitally simplified maps (Study area 2, Drainage Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the TPF.**




*a*




*b*

Figure 5.17: Digitally simplified maps from original map (8717 points, 924 arcs -Study area 2, Drainage Coverage) using two types of tolerance: MR1 and TPF. ***a*** and ***b*** show the relationships between scale, % data reduction and tolerance values.

5.5.2.2 Graphic Reduction and Areal Displacement (Drainage Coverage):

Table 5.12 shows the areal displacements resulting from the data reduction process, using the TPF. The table reveals a systematic increase in these displacements (in ground units) with reduced scale. On the other hand, these displacements are systematically decreased with reduced scale, when converted to the graphic display units. As can be seen from Figures 5.20, 5.21(B), 5.22(B), and 5.23, these displacements are hardly perceptible. Table 5.13 quantitatively verifies this fact, since the total area of the simplified lines resulting from the data reduction process at each scale is larger than the total areal displacements generated by a ratio of areal difference reaches up to 88.21. Figure 5.18 shows that this difference steadily increases at the scale of 1:250,000 towards the smaller scales (1:1,000,000 and 1:1,500,000). This is clearly illustrated, for example, in Figure 5.23 where the areal extent of the lines increasingly becomes more pronounced at smaller scales to a point where it would be impossible to recognise their shapes, let alone the displacement from their original. This observation indicates that the larger the complexity of a coverage the greater is the difference between its areal graphical extent and the areal displacement resulting from a data reduction process, applying the TPF.

| **Overlays** | **Number of Sliver Polygons** | **Sum of Areas ($m^2$)** | **Sum of Area at target scales ($cm^2$)** |
|---|---|---|---|
| Original (1:25,000) | | | |
| **Overlay 1** 1:25,000+1:50,000 | **3242** | 356,034.75 | **1.424** (at 1:50,000) |
| **Overlay 2** 1:25,000+1:100,000 | **2618** | 801,305.81 | **.801** (at 1:100,000) |
| **Overlay 3** 1:25,000+1:250,000 | **1977** | 1,905,649.75 | **.304** (at 1:250,000) |
| **Overlay 4** 1:25,000+1:625,000 | **1598** | 3,735,927.75 | **.095** (at 1:625,000) |
| **Overlay 5** 1:25,000 + 1:1,000,000 | **1487** | 4,719,587.00 | **.047** (at 1:1,000,000) |
| **Overlay 6** 1:25,000 + 1:1,500,000 | **1428** | 5,794,890.00 | **.025** (at 1:1,500,000) |

**Table 5.12 : Areal displacement resulting from overlaying original (8717 points - Study area 2, Drainage Coverage) on digitally simplified maps (using the TPF).**

| Digitally simplified maps for different scales | Total line length | Area of drawing lines of digitally simplified maps using a Pen Width . 015cm at target scales | Sum of Areal displacement at target scales | Ratio of Difference to Area of Drawing Lines |
|---|---|---|---|---|
| | (m) | ($cm^2$) | ($cm^2$) | % |
| 1:50,000 | 232,950.82 | **6.988** | **1.424** (at 1:50,000) | **78.63** |
| 1:100,000 | 229,873.08 | **3.448** | **.801** (at 1:100,000) | **76.77** |
| 1:250,000 | 224,571.80 | **1.347** | **.304** (at 1:250,000) | **77.43** |
| 1:625,000 | 218,218.88 | **.523** | **.095** (at 1:625,000) | **81.84** |
| 1:1,000,000 | 215,009.82 | **.322** | **.047** (at 1:1,000,000) | **85.41** |
| 1:1,500,000 | 212,371.25 | **.212** | **.025** (at 1:1,500,000) | **88.21** |

**Table 5.13: Graphic and areal scale effects on the areal displacement resulting from overlaying the 1:25,000 original map (8717 points -Study area 2, Drainage Coverage ) on digitally simplified maps (using the TPF).**

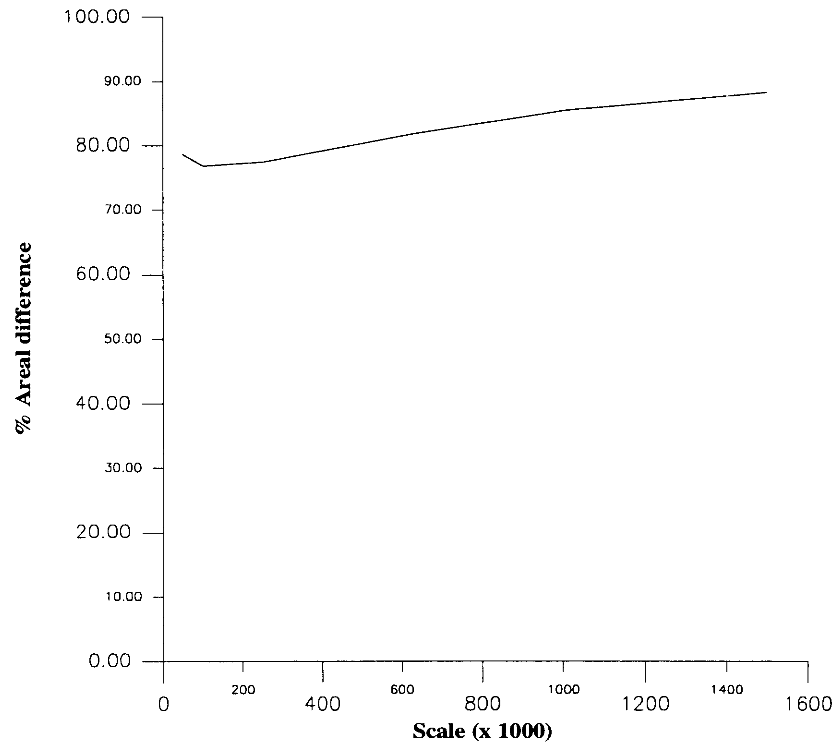



Figure 5.18: Ratio of areal difference (%) to graphic extent of simplified lines. (Difference (in map units) = (total graphic extent of simplified lines) - (total areal displacement). (Study area 2, Drainage Coverage).

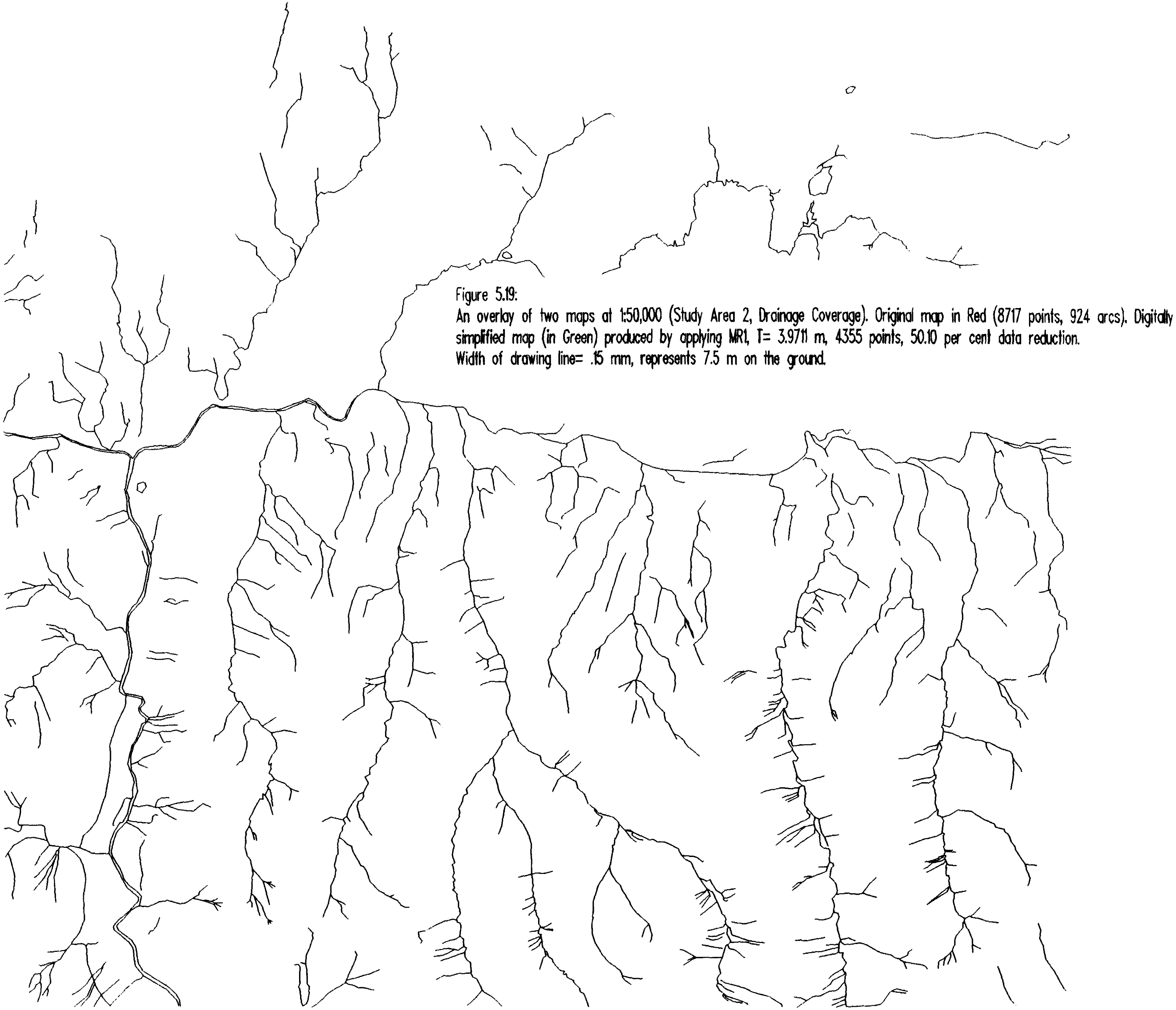


Figure 5.19:
An overlay of two maps at 1:50,000 (Study Area 2, Drainage Coverage). Original map in Red (8717 points, 924 arcs). Digitally simplified map (in Green) produced by applying MR1, T= 3.9711 m, 4355 points, 50.10 per cent data reduction.
Width of drawing line= .15 mm, represents 7.5 m on the ground.

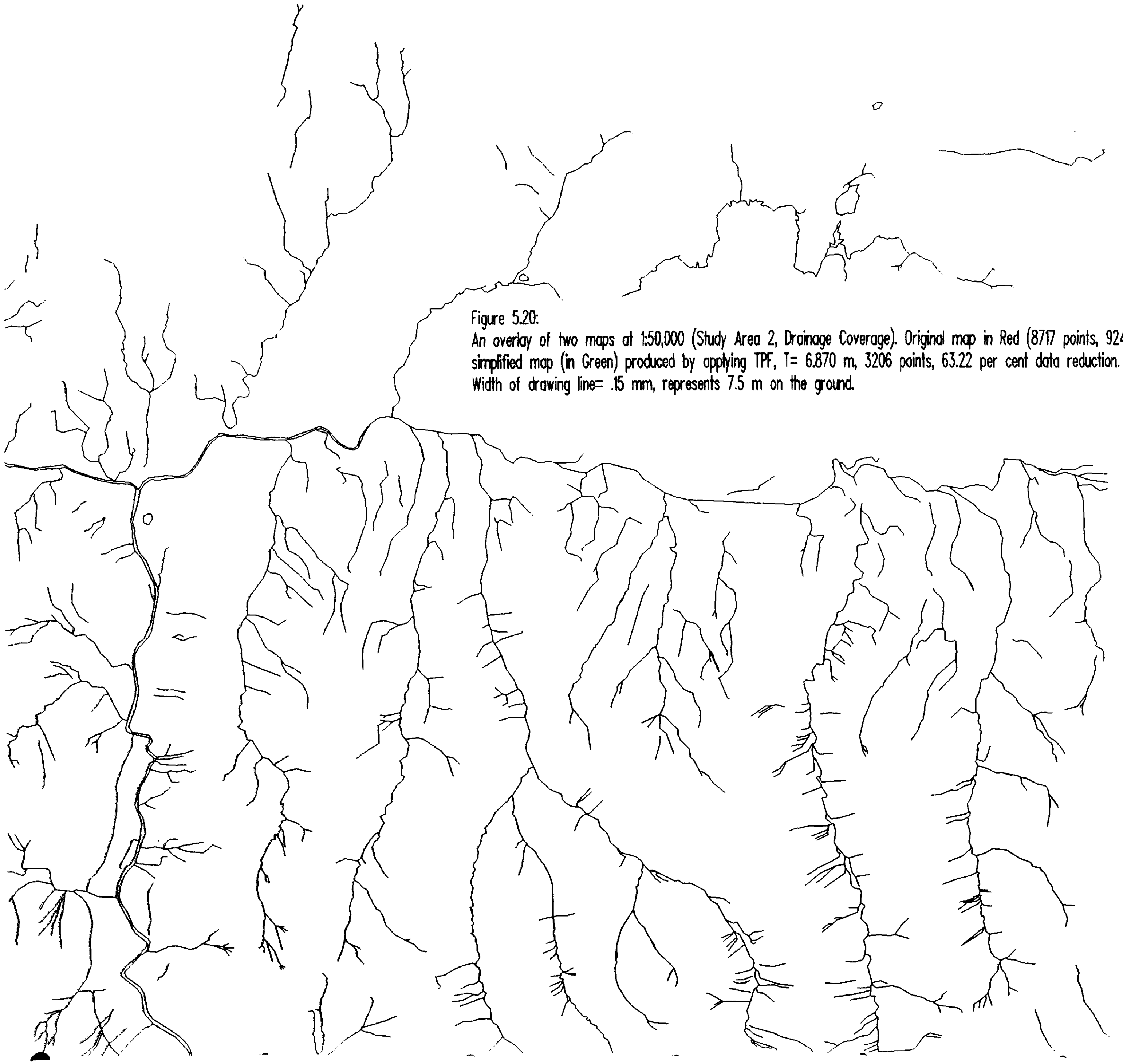
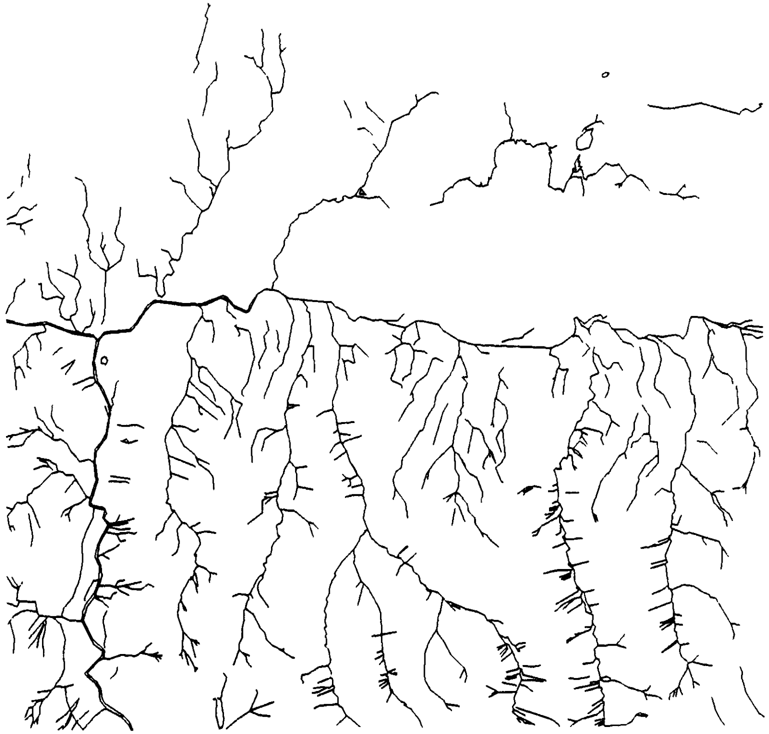
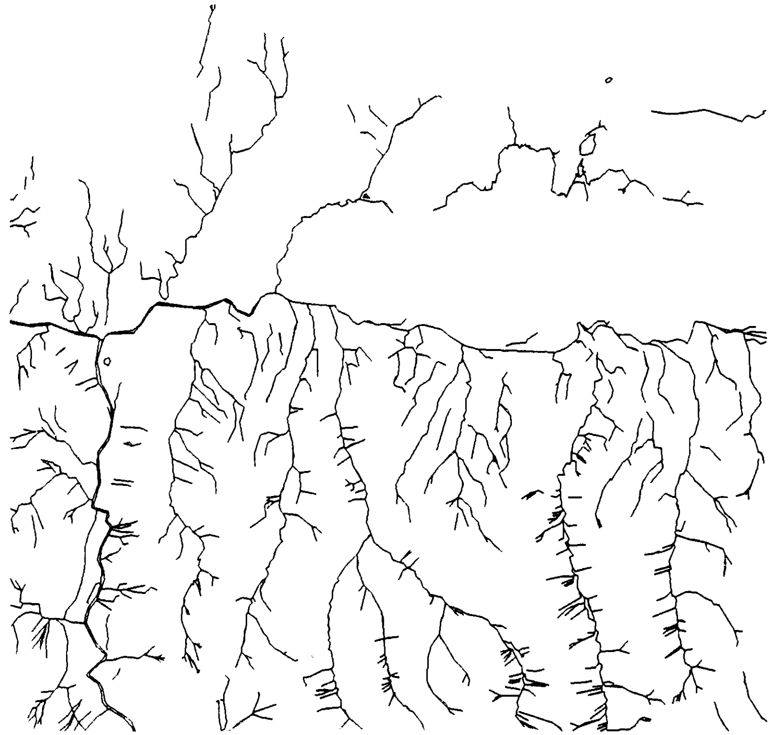


Figure 5.20:
An overlay of two maps at 1:50,000 (Study Area 2, Drainage Coverage). Original map in Red (8717 points, 924 arcs). Digitally simplified map (in Green) produced by applying TPF, T= 6.870 m, 3206 points, 63.22 per cent data reduction.
Width of drawing line= .15 mm, represents 7.5 m on the ground.



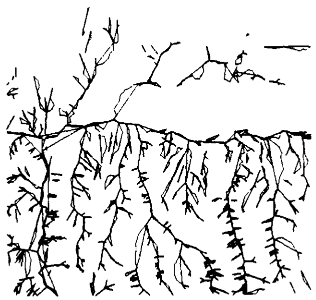
A

Scale: 1:100,000

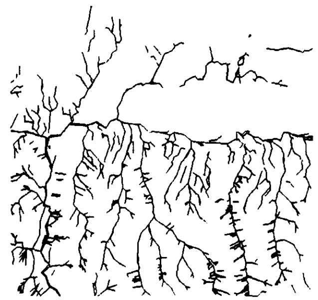


B

Figure 5.21: In (A) an overlay of original -Study Area 2, Drainage Coverage-(in Red, 8717 point) and digitally simplified map (in Green), 2174 points, 75.07 per cent data reduction), produced by a tolerance of 12.850 m (MR1). In (B) an overlay of original (8717 points) and digitally simplified map (in Green), 2007 points, produced by a tolerance of 14,760 m (TPF), 76.98 percent data reduction.
Width of drawing line= .15 mm, represents 15 m on the ground.



A



B

Scale: 1:250,000

Figure 5.22: In (A) an overlay of original (in Red, 8717 point) and digitally simplified map (in Green, 946 points (89.15 per cent data reduction), produced by a tolerance of 500.000 m (MR1). In (B) an overlay of original (8717 points) and, digitally simplified map (in Green), 1282 points, produced by a tolerance of 38.400 m (TPF), 85.30 percent data reduction. Width of drawing line= .15 mm, represents 37.5 m on the ground. (Study Area 2, Drainage Coverage)

1:625 000
Digitally simplified map with 1014 points (88.37 per cent data reduction)
Produced by a tolerance of 97.500 m .
Width of drawing line= .15 mm, represents 93.73 m.

Original map (8717 points), and the digitally simplified map.

1:1000 000
Digitally simplified map has 980 points (88.76 per cent data reduction)
Produced by a tolerance of 156.590 m.
Width of drawing line= .15 mm, represents 150 m.

Original and digitally simplified map.

1:1 500 000
Digitally simplified map has 955 points (89.05 percent data reduction)
Produced by a tolerance of 235.390 m.
Width of drawing line= .15 mm, represents 225 m.

Original and digitally simplified map.

Figure 5.23: Digitally produced maps by only applying TPF, since maximum simplification was achieved by MR1, at which arcs are represented by straight lines (see map for 1:250,000, ). (Study Area 2, Drainage Coverage).

5.5.2.3 Scale, Tolerance and Data Reduction (Transport Coverage):

The line features in this coverage are composed of only 447 points and 21 arcs, representing the scale of 1:25,000. The process of data reduction is performed for the same scales as in the drainage coverage. Figures 5.26, 5.27, 5.28, 5.29, and 5.30 show the simplified lines (in Green) resulted from the process of data reduction for all the scales, while Tables 5.14 and 5.15 report the numerical results. The algorithm, applying the MR1, produced good results only at the scales of 1:50,000 and 1:100,000 (Figures 5.26 and 5.28(A)) compared to the results at the 1:250,000 scale (Figure 5.29(A)); that is, there was no perceptible difference between the original (in Red) and simplified lines (in Green), although the original data were reduced by almost 50 % and 75 %, respectively. Whereas at the scale of 1:625,000, the MR1 produced the least undesirable effect (Figure 5.30(A)), since the simplified lines (in Green) appeared straight indicating the maximum level of reduction where the arc nodes were connected by straight segments. On the other hand, the Douglas-Poiker algorithm produced simplified lines that were imperceptible from the original at all the scales (Figure 5.27, 5.28(B), 5.29(B) and 5.30). Figure 5.24 compares the results from both formulae. As can be seen from the figure (a) and Tables 5.14 and 5.15, over 80 % data reduction occurred within the first three scales, but beyond this range the reduction began to decrease. That is, as in the previous analyses, the 1:250,000 scale marks a break point, reflecting the relatively large tolerance values determined by both formulae. The slow rate of data reduction shown in the results beyond this level, is again due to the algorithm's approach, where application of large tolerance values beyond this level has limited effect, since data are not proportionately reduced to the rate of progression of the tolerance values, especially at high levels of data reduction (see too section 5.5.1.1). Because the TPF produced large data reductions (ranging from 55.71 to 92.62 %), without causing perceptible changes from the original at target scales, and with reference to the data reduction requirements proposed, it can be concluded that the results of the TPF are desirable at all scales.

5.5.2.4 Graphic Reduction and Areal Displacement (Transport Coverage):

Table 5.16 reports the total areal displacements resulting from the simplification process between the original and simplified lines. As the table indicates, data simplification for this coverage resulted in less areal displacement compared to the drainage coverage (Table 5.12). This is due to the relatively small number of features in this transport coverage, and their lack of detailed variations. Table 5.16 shows how the areal displacement systematically increases with higher levels of data simplification, whereas at the highest level (smallest scale) there is an abrupt increase. This is due to the deletion of critical points, which created a large displacement when removed, but the displacement is hardly perceptible at this scale (1:1,500,000 in Figure 5.30). Table 5.17 shows a comparison between the areas of simplified lines and the total areal displacement. As is the case with the previous similar comparisons, this table shows that the graphic representations of the areal displacements are increasingly reduced with reduced scale, and are matched by comparatively large graphic extents of the simplified lines. The comparison shows how this range of differences between the two variables becomes smaller with reduced scale (Figure 5.25). This trend in the relationship emphasises the observations noted above, in which the larger the complexity of a coverage the greater is the difference between its graphical extent and the areal displacement resulting from the data reduction process.

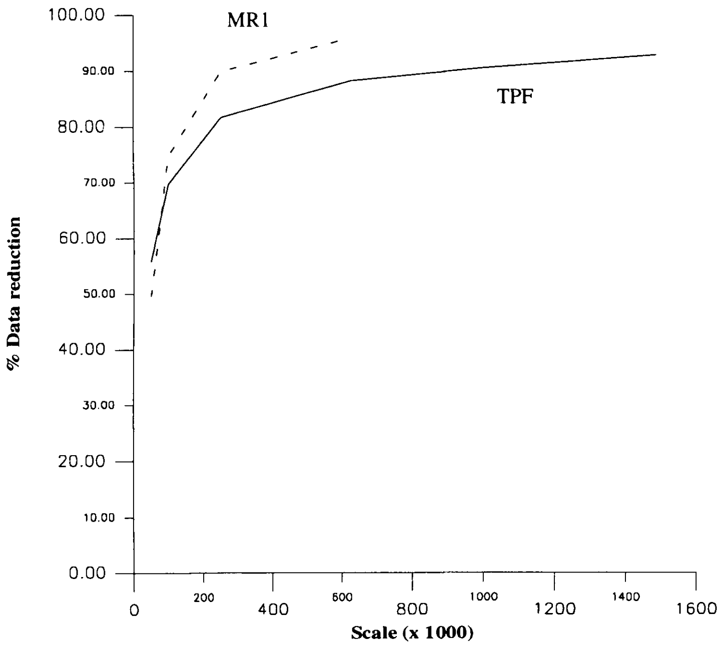
In conclusion, the TPF appeared to present far better results compared to the MR1, both perceptually and numerically in both coverages. Perceptually, the MR1 produced spiky and distorted details, especially beyond the scale of 1:250,000 or the third level of data reduction, whereas the TPF produced results that meets the criteria proposed pertaining to the desirable data reductions. Also, this test showed that the range of the ratio of areal difference to the areal extent of simplified lines is comparable to the range observed in Study area 1. This suggests that the algorithm yields almost similar results within the context of data reduction being examined.

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (447 points) |
|---|---|---|---|
| 1:50,000 | **3.38** | **226** | **49.45** |
| 1:100,000 | **12.80** | **113** | **74.73** |
| 1:250,000 | **83.00** | **46** | **89.83** |
| 1:625,000 | **447.00** | **22** | **95.76** |
| 1:1,000,000 | – | – | – |
| 1:1,500,000 | – | – | – |

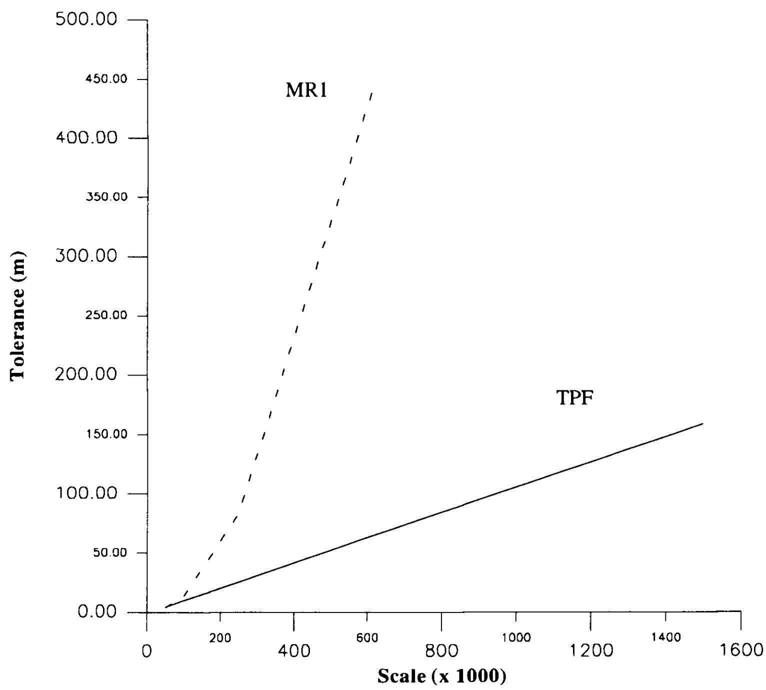
**Table 5.14: Digitally simplified maps from original (Study area 2, Transport Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the MR1.**

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (447 points) |
|---|---|---|---|
| 1:50,000 | **4.291** | **198** | **55.71** |
| 1:100,000 | **9.60** | **136** | **69.58** |
| 1:250,000 | **25.50** | **82** | **81.66** |
| 1:625,000 | **65.25** | **53** | **88.15** |
| 1:1,000,000 | **104.99** | **43** | **90.39** |
| 1:1,500,000 | **157.99** | **36** | **92.62** |

**Table 5.15: Digitally simplified maps (Study area 2, Transport Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the TPF.**




*a*




*b*

Figure 5.24: Digitally simplified maps from original map (447 points, 21 arcs- Study area 2, Transport Coverage), using two types of tolerance: MR1 and TPF. a and b show the relationships between scale, % data reduction and tolerance values.

| Overlays | Number of Sliver Polygons | Sum of Areas ($m^2$) | Sum of Area at target scales ($cm^2$) |
|---|---|---|---|
| Original (1:25,000) | | | |
| **Overlay 1** 1:25,000+1:50,000 | **125** | 31,117.35 | **.124** (at 1:50,000) |
| **Overlay 2** 1:25,000+1:100,000 | **143** | 73,353.74 | **.073** (at 1:100,000) |
| **Overlay 3** 1:25,000+1:250,000 | **107** | 211,272.32 | **.033** (at 1:250,000) |
| **Overlay 4** 1:25,000+1:625,000 | **75** | 591,495.93 | **.015** (at 1:625,000) |
| **Overlay 5** 1:25,000 + 1:1,000,000 | **63** | 889,059.81 | **.008** (at 1:1,000,000) |
| **Overlay 6** 1:25,000 + 1:1,500,000 | **59** | 2,221,373.25 | **.009** (at 1:1,500,000) |

**Table 5.16: Areal displacement resulting from overlaying original (447 points - Study area 2, Transport Coverage) on digitally simplified maps (using TPF) .**

| Digitally simplified maps for different scales | Total line length | Area of drawing lines of digitally simplified maps using a Pen Width . 015cm at target scales | Sum of Areal displacement at target scales | Ratio of Difference to Area of Drawing Lines |
|---|---|---|---|---|
| | (m) | ($cm^2$) | ($cm^2$) | % |
| 1:50,000 | 34,173.67 | **1.025** | **.124** (at 1:50,000) | **87.91** |
| 1:100,000 | 34,147.84 | **.512** | **.073** (at 1:100,000) | **85.75** |
| 1:250,000 | 34,065.85 | **.204** | **.033** (at 1:250,000) | **83.83** |
| 1:625,000 | 33,912.66 | **.081** | **.015** (at 1:625,000) | **81.49** |
| 1:1,000,000 | 33,728.58 | **.050** | **.008** (at 1:1,000,000) | **84.00** |
| 1:1,500,000 | 33,511.45 | **.033** | **.009** (at 1:1,500,000) | **72.73** |

**Table 5.17: Graphic and areal scale effects on the areal displacement resulting from overlaying the 1:25,000 original map (447 points -Study area 2, Transport Coverage) on digitally simplified maps (using the TPF).**

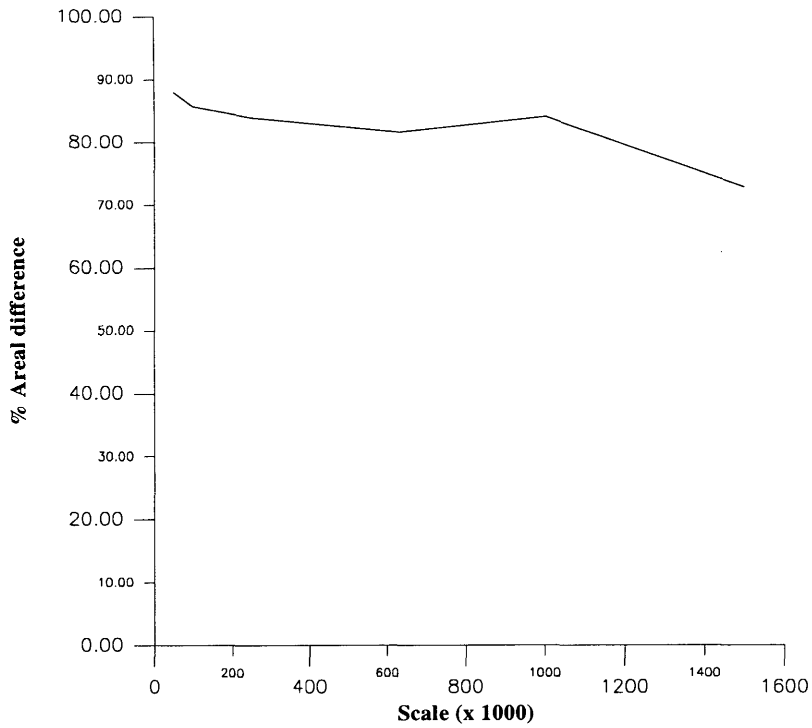



Figure 5.25: Ratio of areal difference (%) to graphic extent of simplified lines. (Difference (in map units) = (total graphic extent of simplified lines) - (total areal displacement). (Study area 2, Transport Coverage).

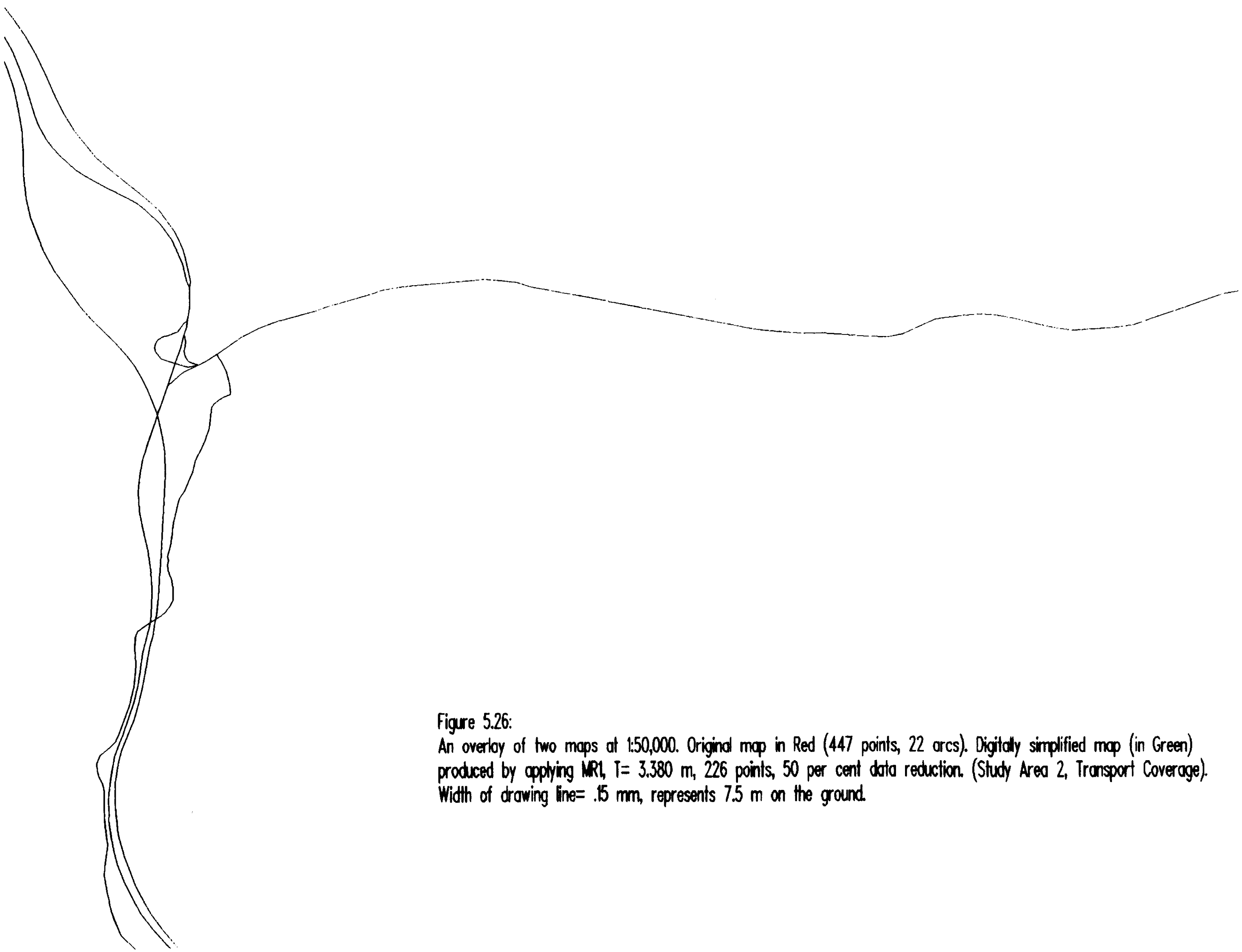


Figure 5.26:
An overlay of two maps at 1:50,000. Original map in Red (447 points, 22 arcs). Digitally simplified map (in Green) produced by applying MR1, T= 3.380 m, 226 points, 50 per cent data reduction. (Study Area 2, Transport Coverage). Width of drawing line= .15 mm, represents 7.5 m on the ground.

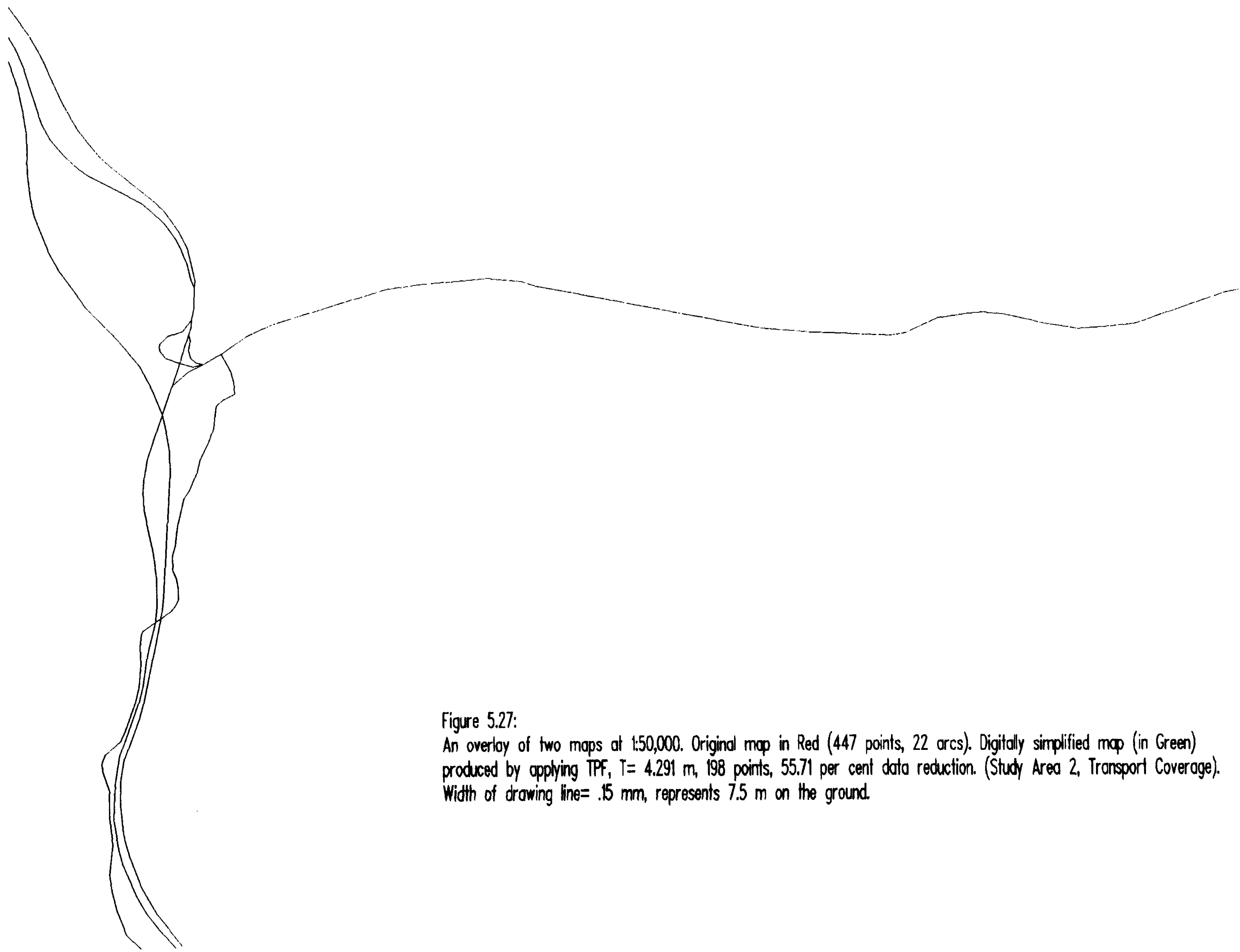
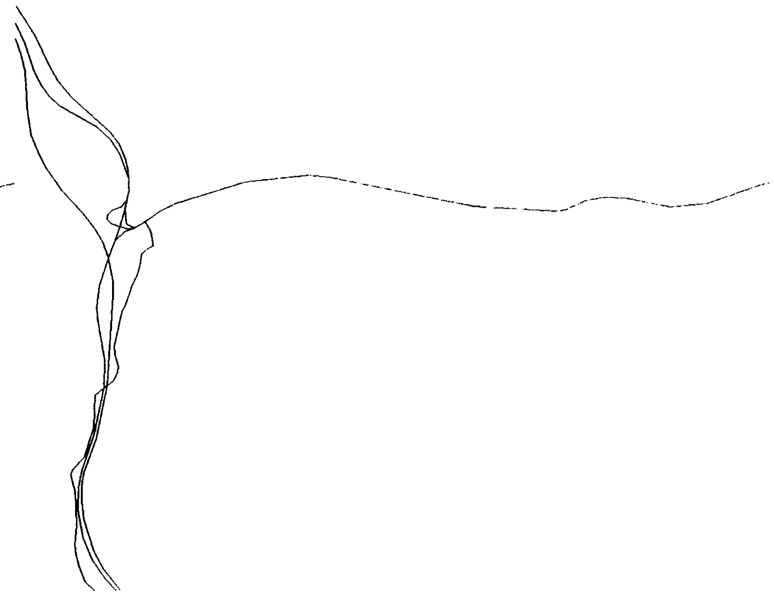


Figure 5.27:
An overlay of two maps at 1:50,000. Original map in Red (447 points, 22 arcs). Digitally simplified map (in Green) produced by applying TPF, T= 4.291 m, 198 points, 55.71 per cent data reduction. (Study Area 2, Transport Coverage). Width of drawing line= .15 mm, represents 7.5 m on the ground.

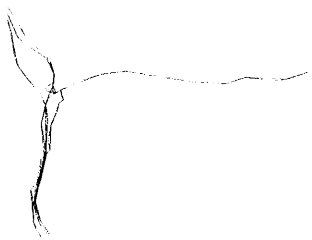
A



B

Scale: 1:100,000

Figure 5.28: In (A) an overlay of original (in Red, 447 points) and digitally simplified map (in Green, 113 points, 74.73 per cent data reduction), produced by a tolerance of 12.800 m (MR1). In (B) an overlay of original (447 points) and, digitally simplified map (in Green), 136 points, produced by a tolerance of 9.600 m (TPF), 69.58 percent data reduction. Width of drawing line= .15 mm, represents 15 m on the ground. (Study Area 2, Transport Coverage).



A

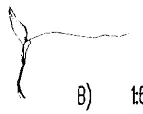
B

Scale: 1:250,000

Figure 5.29: In (A) an overlay of original (in Red, 447 point) and digitally simplified map (in Green, 46 points (89.83 per cent data reduction), produced by a tolerance of 83.000 m (MR1). In (B) an overlay of original (447 points) and, digitally simplified map (in Green), 82 points, produced by a tolerance of 25.500 m (TPF), 81.66 percent data reduction. Width of drawing line= .15 mm, represents 37.5 m on the ground. (Study Area 2, Transport Coverage).

A) 1:625 000

Original (in Red) 447 points, and digitally simplified map (in Green), 23 points (95.76 per cent data reduction), produced by a tolerance of 470.000 m (MR1).
Width of drawing line= .15 mm, represents 93.73 m.



B) 1:625,000

Original map (in Red), 447 points, and the digitally simplified map (in Green).
Digitally simplified map, 53 points (88.15 per cent data reduction)
Produced by a tolerance of 65.250 m (TPF).

1:1000 000
Digitally simplified map has 43 points (90.39 per cent data reduction)
Produced by a tolerance of 104.990 m (TPF).
Width of drawing line= .15 mm, represents 150 m.

Original (447 points) and digitally simplified map.



1:1 500 000
Digitally simplified map has 36 points (92.62 percent data reduction)
Produced by a tolerance of 157.990 m (TPF).
Width of drawing line= .15 mm, represents 225 m.

Original (447 points) and digitally simplified map.

Figure 5.30: Original map (447 points, 21 arcs) and digitally simplified maps. At 1:625,000 (A), the digitally simplified map has reached maximum simplification when using MR1. (Study Area 2, Transport Coverage).

*5.5.3 Study area 3:*

This area is represented by one coverage mainly depicting roads (see section 4.1). The road network is less complex in its detail than the data of study areas 1 and 2. This analysis should further the examination of the algorithm on this type of map (i.e., a network composed of more connected arcs), complexity (i.e., less detail) and scale range (i.e., small range of scale).

5.5.3.1 Scale, Tolerance and Data Reduction:

The original lines in this coverage are represented by 866 points and 103 arcs, representing the scale of 1:158,400, which falls in the range of medium scales. The process data reduction was performed for the scales of: 1:200,000, 1:500,000, 1:730,000, 1:1,000,000, and 1:1,500,000. As it is the case with the previous analyses, the choice of this range of scales is based on that presented in Chapter 4, where this original coverage and its generalised versions were analysed. Figures 5.33, 5.34, 5.35, 5.36, and 5.37 display the simplified lines (in Green) resulted from the data reduction overlaid on the original (in Red), while Tables 5.18 and 5.19 report the numerical results of the process. Figure 5.31 compares these numerical results by the TPF and the MR1. As can be seen from Tables 5.18 and 5.19 and Figure 5.31(a), the TPF produced a smaller data reduction, compared to the MR1. The results by the TPF are shown in Figures 5.34a, 5.34b, 5.35(B), 5.36(B), and 5.37 (B and D). This limited reduction is due to two main factors: 1) the relatively small difference between the scales, and 2) the relatively large number of arcs. The first factor affected the application of the TPF, in the sense that small tolerance values resulted, while the second factor affected the algorithm's approach. Regarding the algorithm's approach, the algorithm is effected by the proportion of the number of arcs to the number of the points in a coverage or line feature. That is, if the coverage or the line consists of 100 points and 50 arcs, this implies that the algorithm will apply the tolerance value independently to each arc (i.e., 50 times). Whereas if the same line is represented by one arc, then the algorithm will apply the tolerance value on the line once. Given the mechanism of the algorithm (discussed in the beginning of this chapter), it is expected that the data reduction that results from the application of the algorithm on the same line 50 times would be far less than the data reduction resulting from one application. This explains the purpose from the process of line segmentation, which was recommended by the researchers in this field, which is aimed at segmenting the line features into different segments or pieces according to their geomorphological structure in order to produce 'balanced simplification'. Another, less significant factor which also contributed to such a small reduction by the algorithm using the TPF is the smaller sinuosity of the lines. The algorithm, applying the MR1, produced acceptable results only for the first scale of 1:200,000 (Figures 5.33a, and 5.33b), and for all others produced increasingly undesirable effects (Figures 5.35(A), 5.36(A), and 5.37(A and C)). As listed in Table 5.18, the tolerance values of the MR1 appear to be the highest compared to the values in the previous analyses produced by either the formulae. This is because the proportion of the number of arcs to the number of points is relatively high, and according to the principle of the formula (MR1), high tolerance values had to be applied in order to acquire the number of points required for each scale. However, the algorithm reached the maximum possible data reduction when the tolerance was 1500.00 m, even before it produced the required number of points (91 points) at the scale of 1:1,500,000 (Figure 5.37(C)). Between these two extreme data reductions by the MR1 and the TPF, an intermediate level can be envisaged to provide a more desirable solution, possibly between 25 and 70 %. However, evaluation of the results by the TPF in this particular analysis, should be considered within other contexts; such as the rate of scale change and the relatively small database of the coverage. Given these, although the results for the TPF did not meet the data reduction criteria initially proposed, the results can be regarded as satisfactory compared to those by the MR1. Hence, if a choice was made between the two sets of results where one was undesirable (e.g., the results by the MR1) and the other was tolerable (e.g., the results of the TPF), it would be reasonable to accept the latter, especially when there were no alternatives. Since both MR1 and TPF perform relatively poorly, it is a reminder of the problem of formalising cartographic knowledge. This example highlights the difficulty in providing formal rules to be uniformly applied to all types of line features. There is a need to consider seriously some compromises between all requirements in order to achieve standardisation of this process.

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (866 points) |
|---|---|---|---|
| 1:200,000 | **20.00** | **688** | **20.56** |
| 1:500,000 | **160.00** | **275** | **68.25** |
| 1:730,000 | **280.00** | **190** | **78.07** |
| 1:1,000,000 | **580.00** | **137** | **84.19** |
| 1:1,500,000 | **1500.00** | **104** | **87.65** |

**Table 5.18: Digitally simplified maps from original ( Study area 3) by the Douglas-Poiker algorithm, using the tolerance values of the MR1.**

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (866 points) |
|---|---|---|---|
| 1:200,000 | **6.414** | **806** | **6.93** |
| 1:500,000 | **17.57** | **709** | **18.13** |
| 1:730,000 | **26.07** | **650** | **24.95** |
| 1:1,000,000 | **36.08** | **592** | **31.64** |
| 1:1,500,000 | **54.63** | **493** | **43.08** |

**Table 5.19: Digitally simplified maps from original (Study area 3) by the Douglas-Poiker algorithm, using the tolerance values of the TPF.**

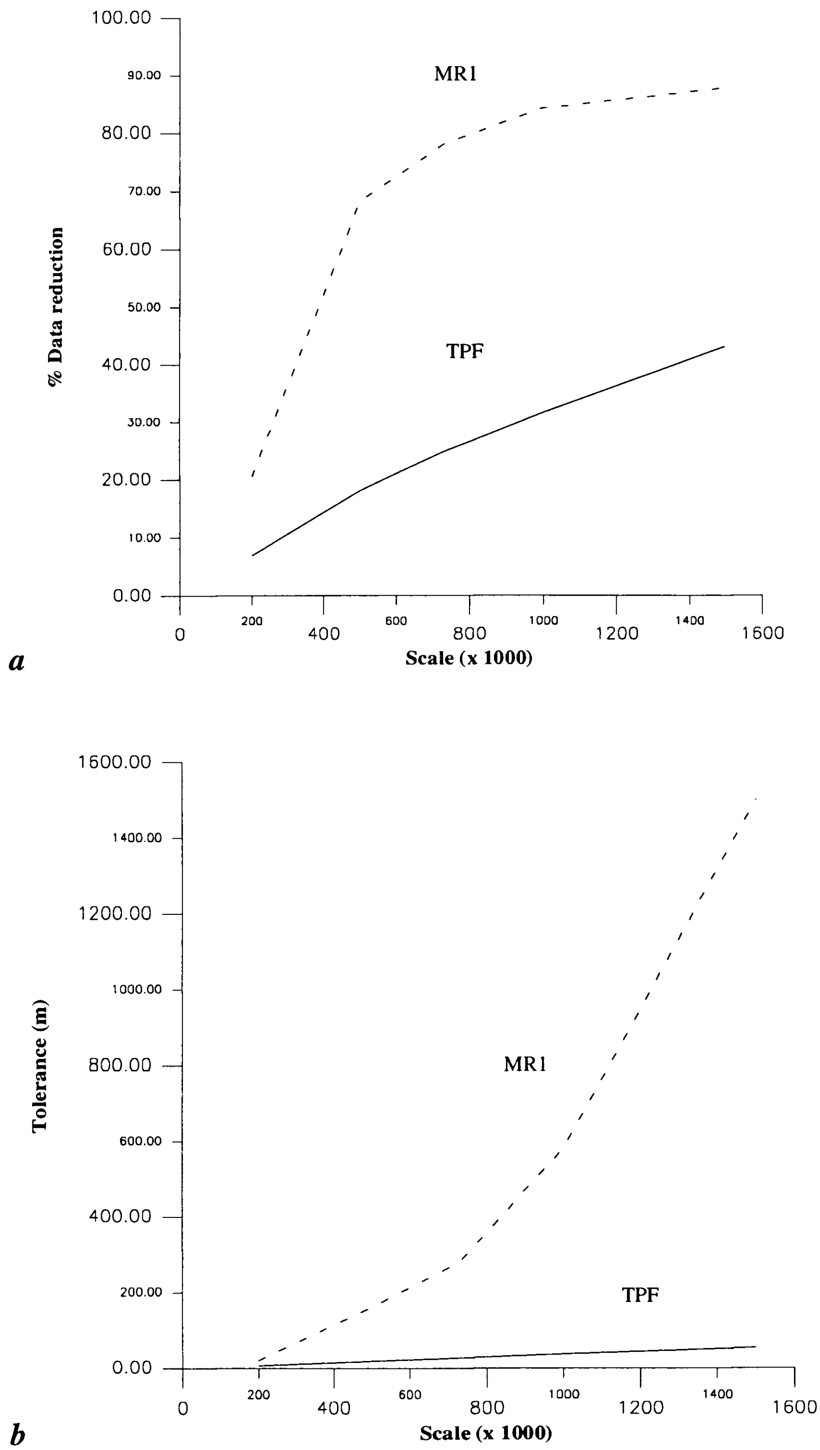



Figure 5.31: Digitally simplified maps from original map (866 points, 103 arcs -Study area 3) using two types of tolerance: MR1 and TPF. ***a*** and ***b*** show the relationship between scale, % data reduction and tolerance values.

5.5.3.2 Graphic Reduction and Areal Displacement:

Table 5.20 reports the total areal displacements resulting from the process of data reduction applying the TPF. The table shows a progressive increase in the areal displacements (in ground units) resulted from the simplification process, from 15,943.92 to 3,237,688.00 square metres. On the other hand, these displacements are systematically decreased with reduced scale when converted to the map units ($cm^2$). As can be seen from Figures 5.34a, 4.34b, 5.35(B), 5.36(B), and 5.37(B and D), these displacements are hardly perceptible on the maps at the target scales. Table 5.21 and Figure 5.32 verifies this fact, since the total area of the simplified lines at each scale is larger than the total areal displacement with a ratio of the difference to the graphic extent of the simplified lines reaches up to 99.86 %.

| **Overlays** | **Number of Sliver Polygons** | **Sum of Areas ($m^2$)** | **Sum of Area at target scales ($cm^2$)** |
|---|---|---|---|
| Original (1:158,400) | | | |
| **Overlay 1** 1:158,400+1:200,000 | **15** | 15,943.92 | **.004** (at 1:200,000) |
| **Overlay 2** 1:158,400+1:500,000 | **99** | 529,258.75 | **.021** (at 1:500,000) |
| **Overlay 3** 1:158,400+1:730,000 | **153** | 938,445.62 | **.017** (at 1:730,000) |
| **Overlay 4** 1:158,400+1:1,000,000 | **203** | 1,726,592.87 | **.017** (at 1:1,000,000) |
| **Overlay 5** 1:158,400+ 1:1,500,000 | **280** | 3,237,688.00 | **.014** (at 1:1,500,000) |

**Table 5.20: Areal displacement resulting from overlaying original (866 points, Study area 3) on digitally simplified maps (using the TPF) .**

| **Digitally simplified maps for different scales** | **Total line length** | **Area of drawing lines of digitally simplified maps using a Pen Width . 015cm at target scales** | **Sum of Areal displacement at target scales** | **Ratio of Difference to Area of Drawing Lines** |
|---|---|---|---|---|
| | **(m)** | **($cm^2$)** | **($cm^2$)** | **%** |
| 1:200,000 | 367,180.59 | **2.753** | **.004** (at 1:200,000) | **99.86** |
| 1:500,000 | 366,814.27 | **1.100** | **.021** (at 1:500,000) | **98.10** |
| 1:730,000 | 366,710.36 | **.753** | **.017** (at 1:730,000) | **97.75** |
| 1:1,000,000 | 366,540.97 | **.549** | **.017** (at 1:1,000,000) | **96.91** |
| 1:1,500,000 | 365,905.56 | **.365** | **.014** (at 1:1,500,000) | **96.17** |

**Table 5.21: Graphic and areal scale effects on the areal displacement resulting from overlaying the 1:158,400 original map (866 points, Study area 3) on digitally simplified maps (using the TPF).**

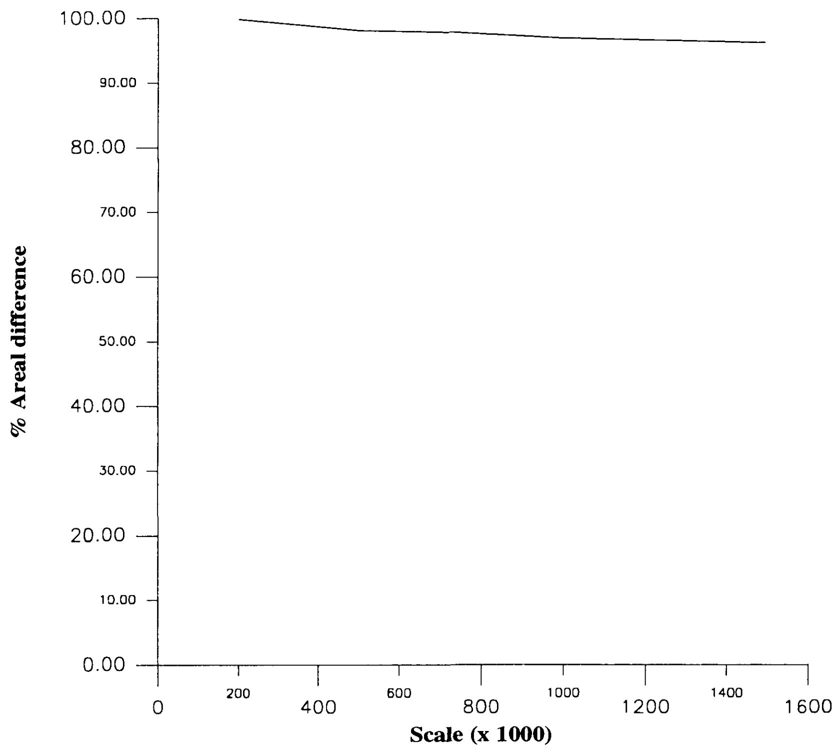
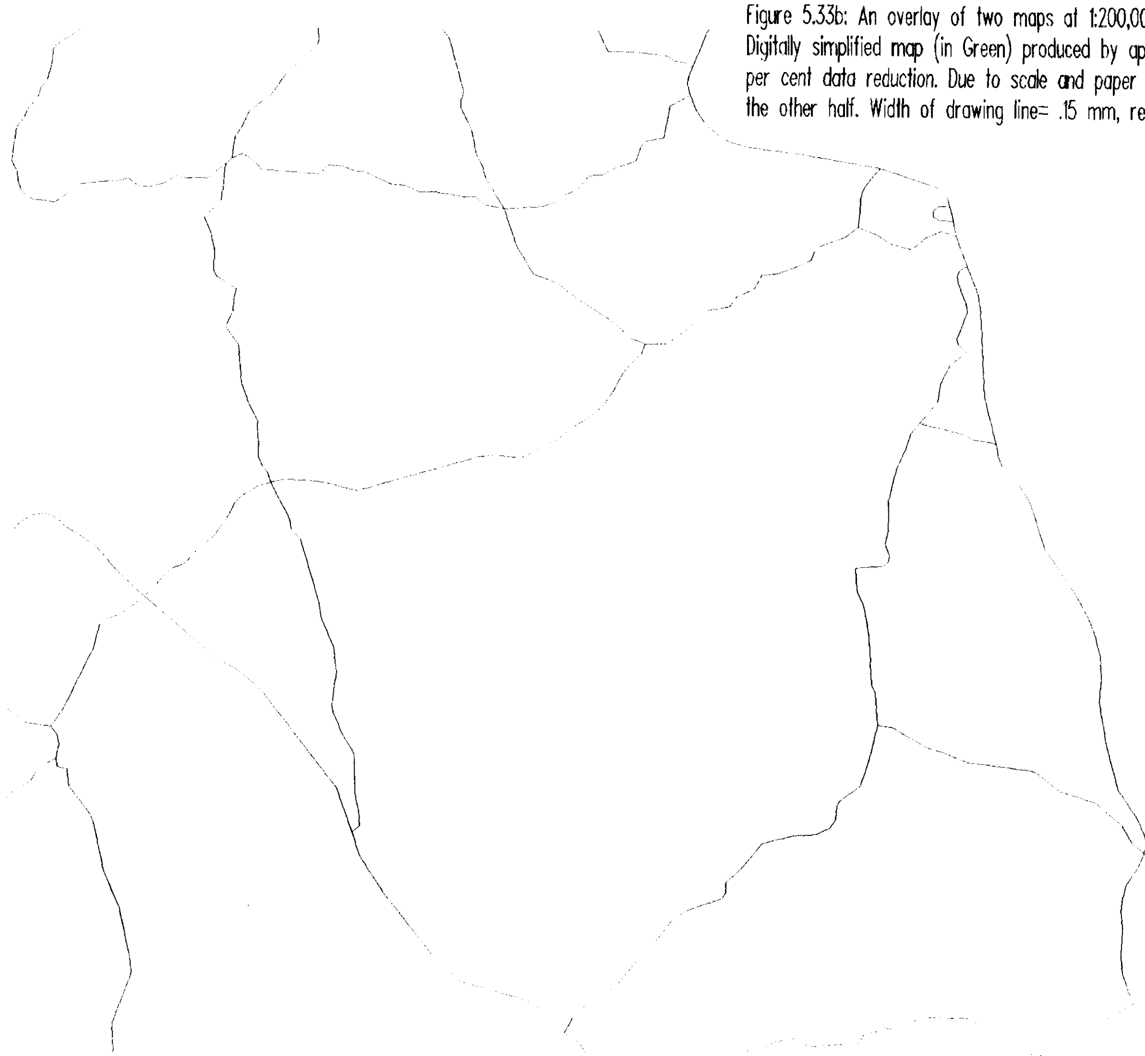



Figure 5.32: Ratio of areal difference (%) to graphic extent of simplified lines. (Difference (in map units) = (total graphic extent of simplified lines) - (total areal displacement). (Study area 3).

Figure 5.33a: An overlay of two maps at 1:200,000 (Study Area 3). Original map in Red (866 points, 103 arcs). Digitally simplified map (in Green) produced by applying MR1, T= 20.000 m, 688 points, 20.56 per cent data reduction.
Due to scale and paper size only half of the map is displayed, Figure 5.33b shows the other half.
Width of drawing line= .15 mm, represents 30 m on the ground.

Figure 5.33b: An overlay of two maps at 1:200,000 (Study Area 3). Original map in Red (866 points, 103 arcs) Digitally simplified map (in Green) produced by applying MR1, T= 20.000 m, 688 points, 20.56 per cent data reduction. Due to scale and paper size only half of the map is displayed, Figure 5.33a shows the other half. Width of drawing line= .15 mm, represents 30 m on the ground.



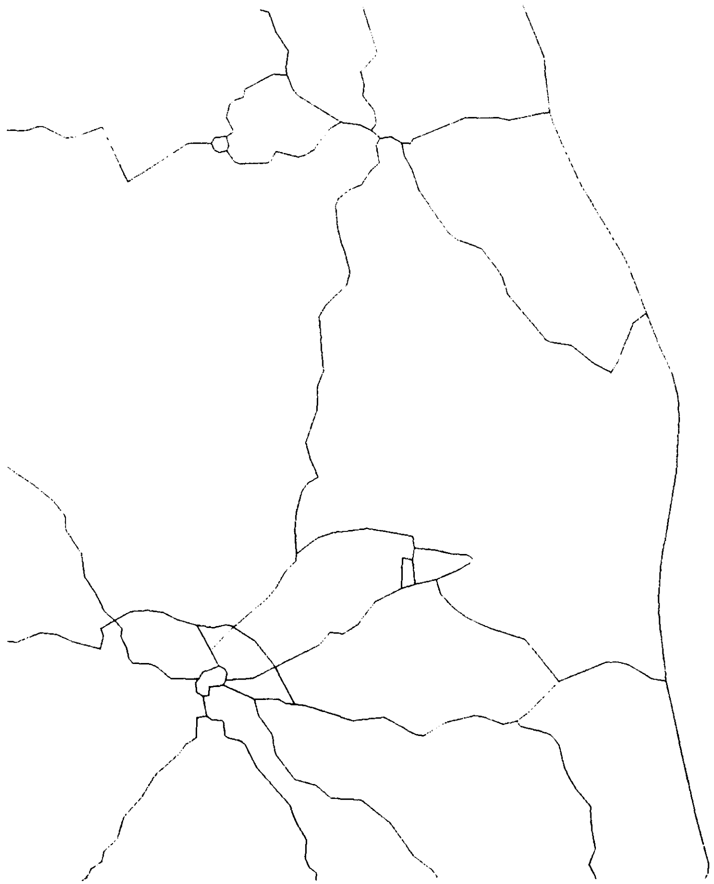


Figure 5.34a: An overlay of two maps at 1:200,000 (Study Area 3). Original map in Red (866 points, 103 arcs). Digitally simplified map (in Green) produced by applying TPF, T= 6.414 m, 806 points, 6.93 per cent data reduction.
Due to scale and paper size only half of the map is displayed, Figure 5.34b shows the other half.
Width of drawing line= .15 mm, represents 30 m on the ground.

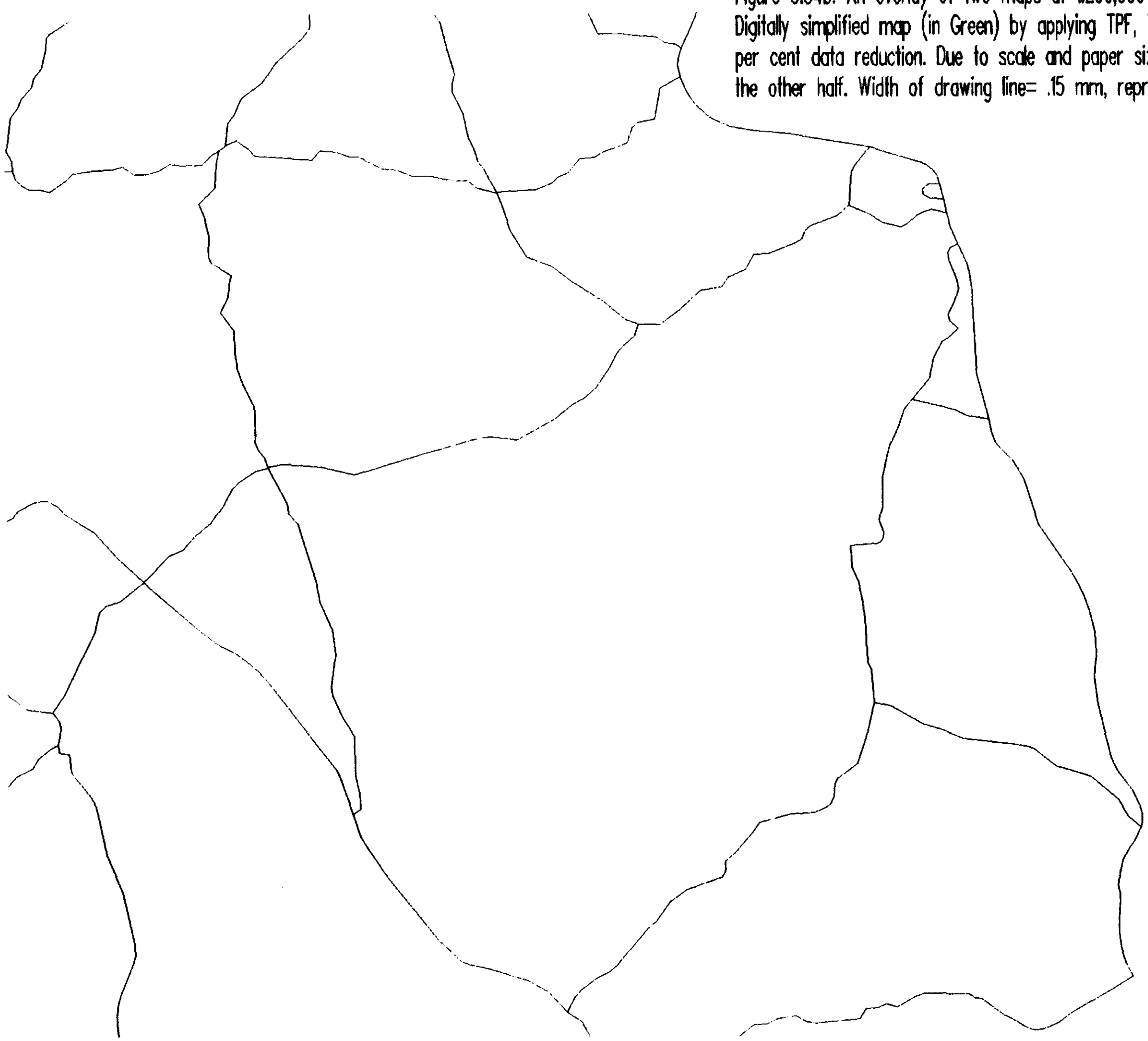
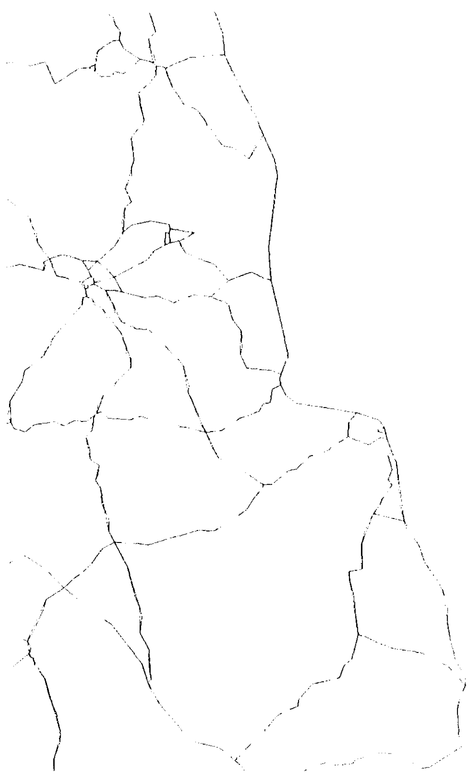


Figure 5.34b: An overlay of two maps at 1:200,000 (Study Area 3). Original map in Red (866 points, 103 arcs). Digitally simplified map (in Green) by applying TPF, T= 6.414 m, 806points, 6.93 per cent data reduction. Due to scale and paper size only half of the mapis displayed, Figure 5.34a shows the other half. Width of drawing line= .15 mm, represents 30 m on the ground.



A Scale: 1:500,000 B

Figure 5.35: In (A) an overlay of original (in Red, 866 point) and digitally simplified map (in Green), 275 points (68.25 per cent data reduction), produced by a tolerance of 160 m (MR1). In (B) an overlay of original (866 points) and, digitally simplified map (in Green), 709 points, produced by a tolerance of 17.571 m (TPF), 18.13 percent data reduction. Width of drawing line= .15 mm, represents 75 m on the ground. (Study Area 3)

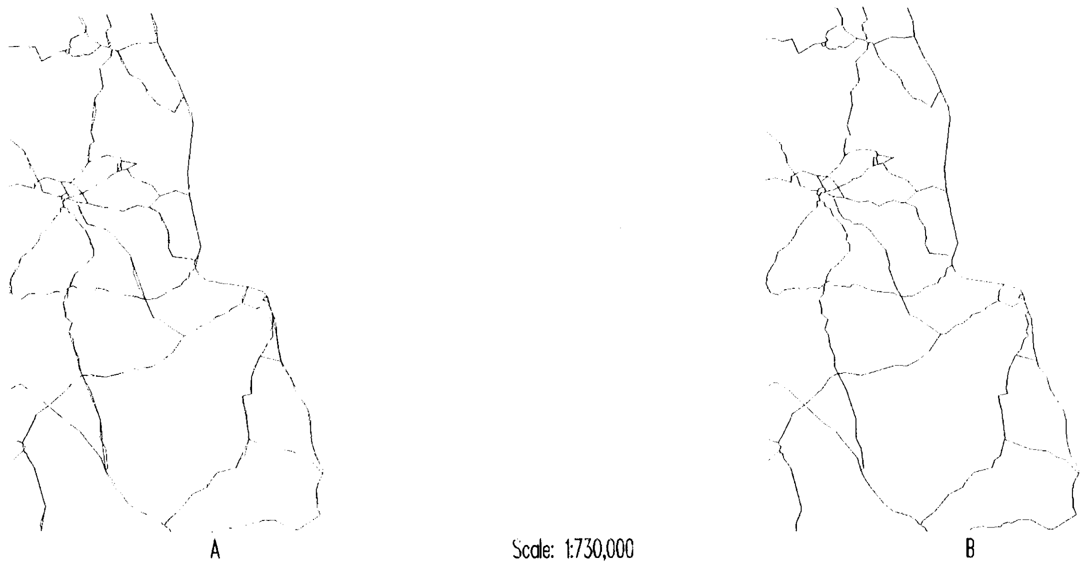
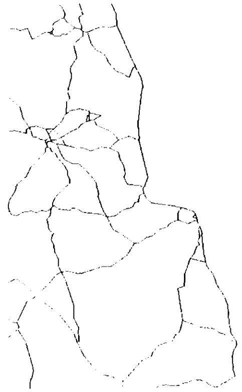



Figure 5.36: In (A) an overlay of original (in Red, 866 point) and digitally simplified map (in Green), 190 points (78.07 per cent data reduction), produced by a tolerance of 280 m (MR1). In (B) an overlay of original (866 points) and, digitally simplified map (in Green), 650 points, produced by a tolerance of 26.070 m (TPF), 24.95 per cent data reduction. Width of drawing line= .15 mm, represents 109.50 m on the ground. (Study Area 3).

A

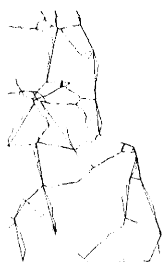
Scale: 1:1,000,000



B

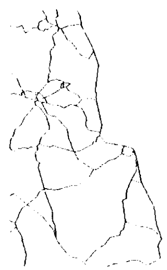
Original (in Red) 866 points, and digitally simplified map (in Green),137 points (84.19 per cent data reduction), produced by a tolerance of 580.000 m (MR1). Width of drawing line= .15 mm, represents 150 m.

Digitally simplified map, 592 points (31.64 per cent data reduction) Produced by a tolerance of 36.083 m (TPF). Original map (in Red), 866 points, and the digitally simplified map (in Green).



C

Scale: 1:1,500,000



D

Original (in Red) 866 points, and digitally simplified map (in Green),104 points (87.65 per cent data reduction), produced by a tolerance of 1500.000 m (MR1). Width of drawing line= .15 mm, represents 225 m.

Digitally simplified map, 493 points (43.08 per cent data reduction) Produced by a tolerance of 54.625 m (TPF). Original map (in Red), 866 points, and the digitally simplified map (in Green).

Figure 5.37: Original map (866 points, 103 arcs) and digitally simplified maps at 1:1,000,000, and 1:1,500,000 (Study Area 3).

*5.5.4 Study area 4:*

This area is comprised of two types of line features: drainage network and transport network (see section 4.1). The maps of this study area have relatively less line details and complexity compared to those of the previous study areas. Each network is represented by a separate coverage, and the results and discussion are presented accordingly.

5.5.4.1 Scale, Tolerance and Data Reduction (Drainage Coverage):

This coverage consists of 721 points and 4 arcs, digitised at a scale of 1:25,000, as the original scale. The scales for which the data reduction process was performed are: 1:50,000, 1:100,000, 1:300,000, 1:500,000, and 1:1,000,000, and 1:1,500,000. As for the previous analyses, the choice of the scale range in this analysis for both the drainage coverage and transport coverage is based on the range of scales studied in Chapter 4. The results from the algorithm using the TPF and MR1 are shown in Figures 5.40, 5.41, 5.42, 5.43, and 5.44, and the numerical results are presented in Tables 5.22, and 5.23. In perceptual terms, application of the TPF allowed the Douglas-Poiker algorithm to produce simplified line shapes (in Green) that did not differ from the original (in Red) (Figures 5.41, 5.42(B), 5.43(B), and 5.44). Numerically, the TPF led to large reductions of data, up to 94.87 % (Table 5.23), and yet, as mentioned above, the figures reveal imperceptible deviations at the target scales between the simplified and original lines. Comparatively, applying the MR1 increasingly caused the algorithm to effect spiky shapes at the scales of 1:300,000 (Figure 5.43(A)), 1:500,000, 1:1,000,000, and 1:1,500,000 (Figure 5.44). Figure 5.38 compares the results of the algorithm by the two formulae. Figure 5.38(a) and Tables 5.22 and 5.23 show that an almost 90 % data reduction was achieved by the two formulae for the scale of 1:300,000 (i.e., the third level of data reduction). This is due to the relatively large tolerance values applied compared to those applied for the previous scales (1:50,000 and 1:100,000). The reason for the limited effect of the application of larger tolerances beyond this level of data reduction is explained in section 5.5.1.1. As a result, the application of the TPF as opposed to the MR1 appears successful in producing the data reduction which was proposed to be desirable in beginning of this chapter. That is, the TPF could produce data reductions ranging from over 60 % to over 90 %, whilst no deviation was detected between the simplified lines and original at all the target scales.

5.5.4.2 Graphic Reduction and Areal Displacement (Drainage Coverage):

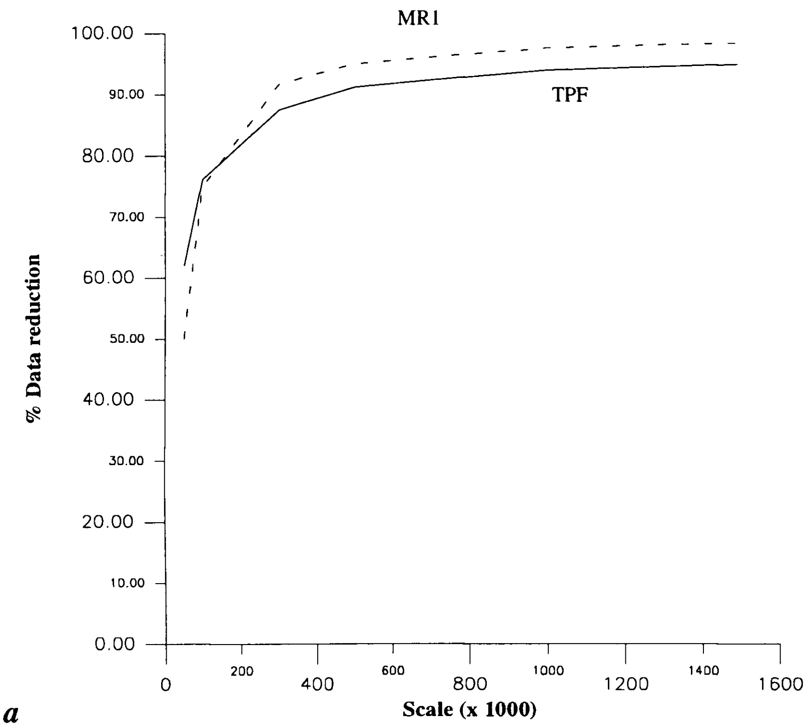
Table 5.24 shows the total areal displacements resulting from the application of the TPF between the simplified and original lines for each scale. The table shows a steady increase in the areal displacements (in ground units) with reduced scales. In terms of map units ($cm^2$), these displacements are systematically decreased with reduced scale. Table 5.25 and Figure 5.39 indicate that the total area of the simplified lines at each scale is larger than the total areal displacement with a ratio of the difference to the graphic extent of the simplified lines reaches up to 87.83 (at 1:50,000) but not less than 78.58 % (at 1:500,000). As the tables and the figure show, the results are, again, comparable to those of the previous analyses. That is, the graphic representation of the simplified lines is larger than that of the areal displacements caused by the data reduction process by a relatively high percentage (not less than 75 %). This emphasises the perceptual observation about the resulting simplification using the TPF, in which the difference between the simplified and original lines is imperceptible at the target reduced scales. However, and as mentioned before, this might not imply that no perceptible difference in one particular place or another exists.

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (721 points) |
|---|---|---|---|
| 1:50,000 | **2.947** | **361** | **49.94** |
| 1:100,000 | **9.85** | **180** | **75.04** |
| 1:300,000 | **67.00** | **60** | **91.68** |
| 1:500,000 | **176.00** | **36** | **95.01** |
| 1:1,000,000 | **420.00** | **18** | **97.51** |
| 1:1,500,000 | **840.00** | **12** | **98.34** |

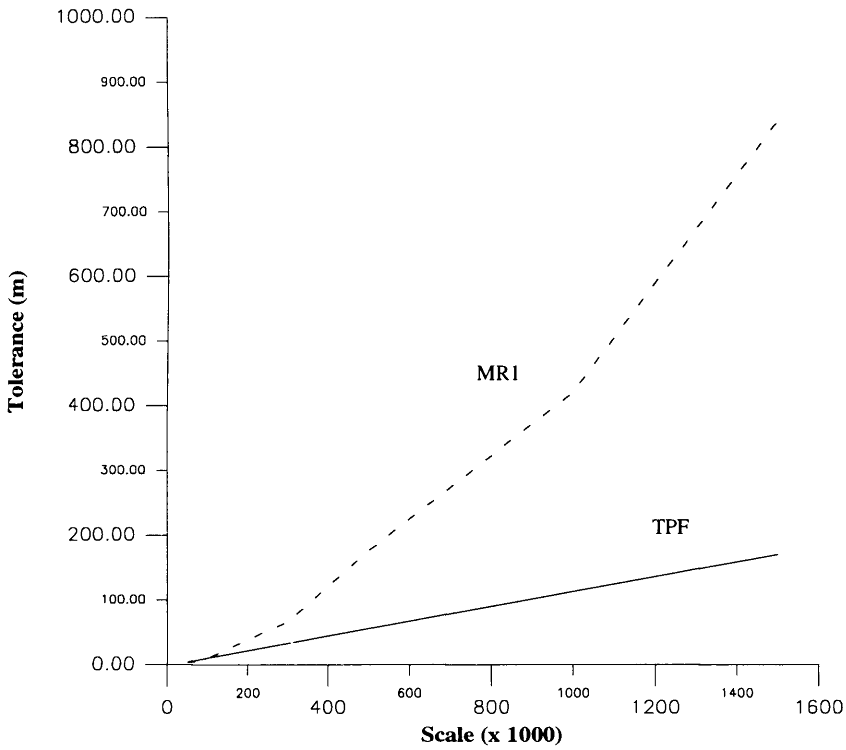
**Table 5.22: Digitally simplified maps from original (Study area 4, Drainage Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the MR1.**

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (721 points) |
|---|---|---|---|
| 1:50,000 | **4.699** | **274** | **62.00** |
| 1:100,000 | **10.40** | **172** | **76.15** |
| 1:300,000 | **33.20** | **90** | **87.52** |
| 1:500,000 | **56.10** | **64** | **91.13** |
| 1:1,000,000 | **113.31** | **44** | **93.90** |
| 1:1,500,000 | **170.47** | **37** | **94.87** |

**Table 5.23: Digitally simplified maps from original (Study area 4, Drainage Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the TPF.**




*a*




*b*

Figure 5.38: Digitally simplified maps from original map (721 points, 4 arcs -Study area 4, Drainage Coverage) using two type of tolerance: MR1 and TPF. ***a*** and ***b*** show the relationships between scale, % data reduction and tolerance values.

| Overlays | Number of Sliver Polygons | Sum of Areas ($m^2$) | Sum of Area at target scales ($cm^2$) |
|---|---|---|---|
| Original (1:25,000) | | | |
| **Overlay 1** 1:25,000+1:50,000 | **218** | 34,774.30 | **.139** (at 1:50,000) |
| **Overlay 2** 1:25,000+1:100,000 | **219** | 93,393.68 | **.093** (at 1:100,000) |
| **Overlay 3** 1:25,000+1:300,000 | **146** | 331,007.00 | **.036** (at 1:300,000) |
| **Overlay 4** 1:25,000+1:500,000 | **104** | 614,909.25 | **.024** (at 1:500,000) |
| **Overlay 5** 1:25,000 + 1:1,000,000 | **96** | 1,041,936.43 | **.010** (at 1:1,000,000) |
| **Overlay 6** 1:25,000 + 1:1,500,000 | **92** | 1,516,382.12 | **.006** (at 1:1,500,000) |

**Table 5.24: Areal displacement resulting from overlaying original (721 points, Study area 4, Drainage Coverage) on digitally simplified maps (using the TPF) .**

| Digitally simplified maps for different scales | Total line length | Area of drawing lines of digitally simplified maps using a Pen Width . 015cm at target scales | Sum of Areal displacement at target scales | Ratio of Difference to Area Drawing Lines |
|---|---|---|---|---|
| | (m) | ($cm^2$) | ($cm^2$) | % |
| 1:50,000 | 38,098.72 | **1.1420** | **.139** (at 1:50,000) | **87.83** |
| 1:100,000 | 38,020.62 | **.570** | **.093** (at 1:100,000) | **83.69** |
| 1:300,000 | 37,835.83 | **.189** | **.036** (at 1:300,000) | **80.96** |
| 1:500,000 | 37,665.36 | **.112** | **.024** (at 1:500,000) | **78.58** |
| 1:1,000,000 | 37,348.80 | **.056** | **.010** (at 1:1,000,000) | **82.15** |
| 1:1,500,000 | 37,089.86 | **.037** | **.006** (at 1:1,500,000) | **83.79** |

**Table 5.25: Graphic and areal scale effects on the areal displacement resulting from overlaying the 1:25,000 original map (721 points, Study area 4, Drainage Coverage ) on digitally simplified maps (using the TPF).**

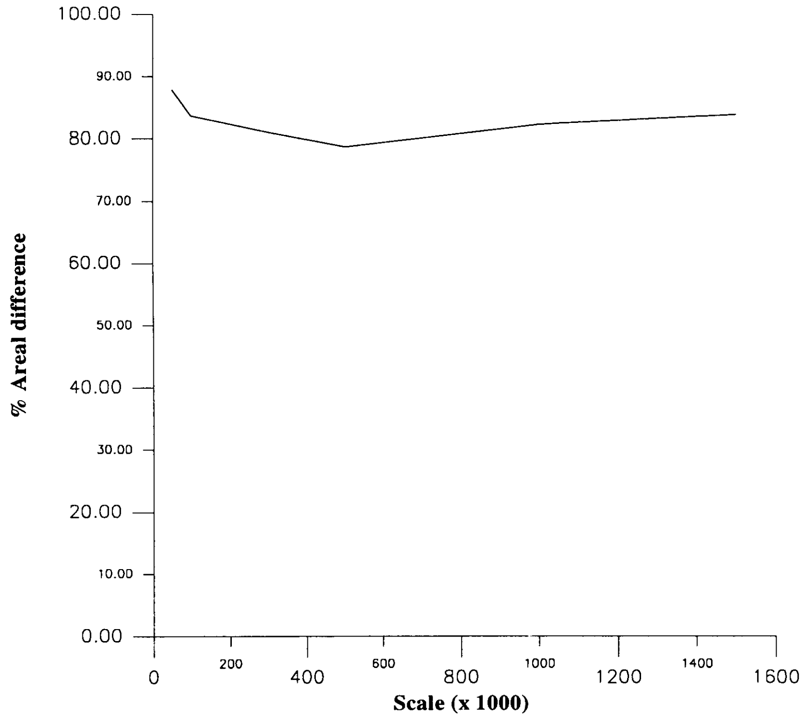



Figure 5.39: Ratio of areal difference (%) to graphic extent of simplified lines. (Difference (in map units) = (total graphic extent of simplified lines) - (total areal displacement). (Study area 4 - Drainage Coverage).

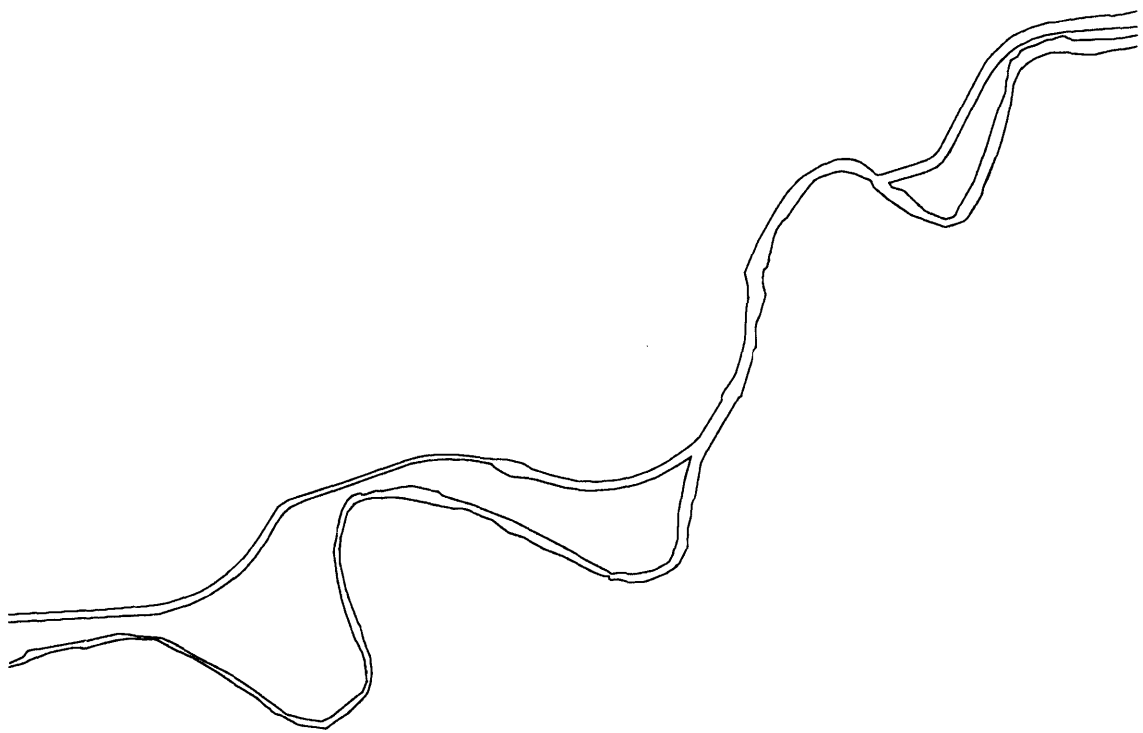


Figure 5.40:
An overlay of two maps at 1:50,000 (Study Area 4, Drainage Coverage). Original map in Red (721 points, 4 arcs). Digitally simplified map (in Green) produced by applying MR1, T= 2.947 m, 361 points, 49.94 per cent data reduction.
Width of drawing line= .15 mm, represents 7.5 m on the ground.

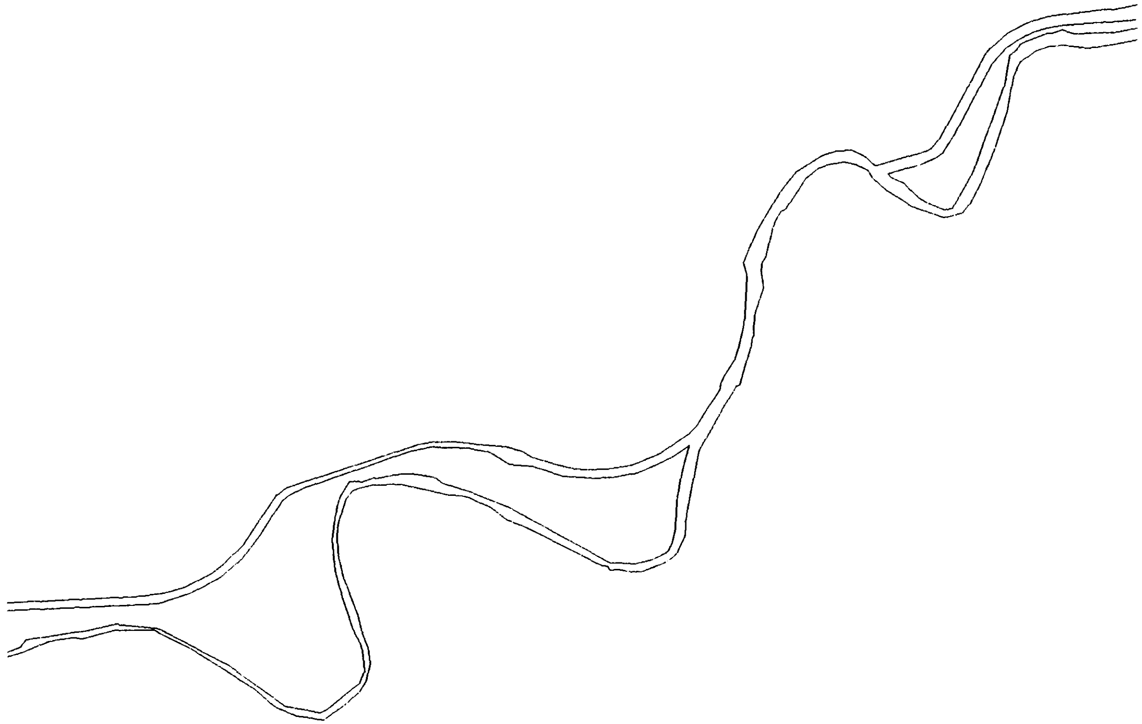
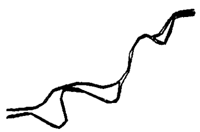


Figure 5.41:
An overlay of two maps at 1:50,000 (Study Area 4, Drainage Coverage). Original map in Red (721 points, 4 arcs). Digitally simplified map (in Green) produced by applying TPF, T= 4.699 m, 274 points, 62.00 per cent data reduction.
Width of drawing line= .15 mm, represents 7.5 m on the ground.

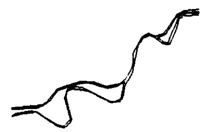
A B

Scale: 1:100,000

Figure 5.42: In (A) an overlay of original (in Red, 721 points) and digitally simplified map (in Green, 180 points, 75.04 per cent data reduction), produced by a tolerance of 9.850 m (MR1). In (B) an overlay of original (721 points) and, digitally simplified map (in Green), 172 points, produced by a tolerance of 10.400 m (TPF), 76.15 percent data reduction. Width of drawing line= .15 mm, represents 15 m on the ground.(Study Area 4, Drainage Coverage)



A



B

Scale: 1:300,000

Figure 5.43: In (A) an overlay of original (in Red, 721 points) and digitally simplified map (in Green, 60 points (91.68 per cent data reduction), produced by a tolerance of 67.000 m (MR1). In (B) an overlay of origial (721 points) and, digitally simplified map (in Green, 90 points, produced by a tolerance of 33.200 m (TPF), 87.52 per cent data reduction. Width of drawing line= .15 mm, represents 45 m on the ground. (Study Area 4, Drainage Coverage).

1:500 000
Width of drawing line= .15 mm, represents 75 m.

An overlay of two maps: digitally simplified map in Green, 36 points, 95.01 per cent data reduction, produced by a tolerance of 176.000 m (MR1), and original in Red 721 points.

An overlay of two maps: digitally simplified map in Green, 64 points, 91.13 per cent data reduction, produced by a tolerance of 56.100 m (TPF), and original in Red 721 points.

1:1000 000
Width of drawing line= .15 mm, represents 150 m.

An overlay of two maps: digitally simplified map in Green, 18 points, 97.51 per cent data reduction, produced by a tolerance of 420.000 m (MR1), and original in Red 721 points.

An overlay of two maps: digitally simplified map in Green, 44 points, 93.90 per cent data reduction, produced by a tolerance of 113.310 m (TPF), and original in Red 721 points.

1:1 500 000
Width of drawing line= .15 mm, represents 225 m.

An overlay of two maps: digitally simplified map in Green, 12 points, 98.34 per cent data reduction, produced by a tolerance of 840.000 m (MR1), and original in Red 721 points.

An overlay of two maps: digitally simplified map in Green, 37 points, 94.87 per cent data reduction, produced by a tolerance of 170.470 m (TPF), and original in Red 721 points.

Figure 5.44: Original map ( in Red, 721 points) and digitally simplified maps (in Green) at 1:500,000, 1:1,000,000, and 1:1,500,000, using tolernaces of MR1 and TPF. (Study Area 4, Drainage Coverage).

5.5.4.3 Scale, Tolerance and Data Reduction (Transport Coverage):

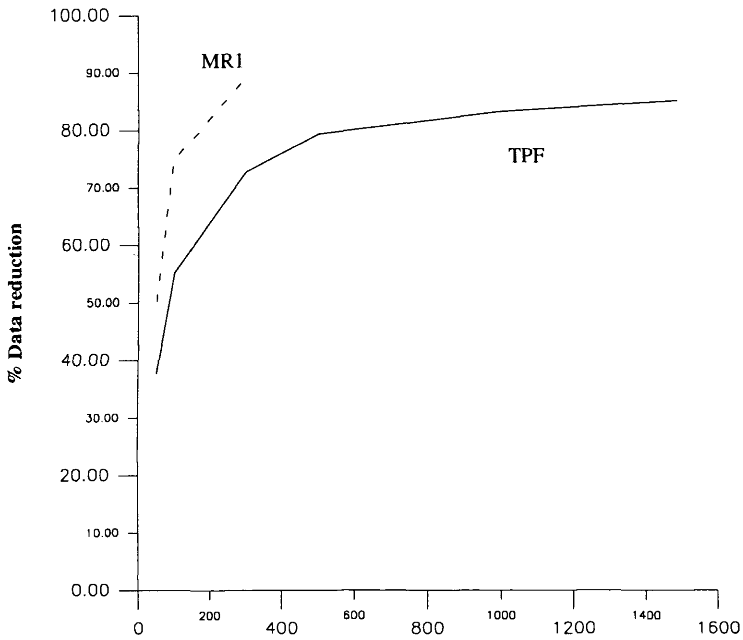
The transport coverage depicts a relatively small number of points (319), although it has a larger number of arcs (33) compared to the drainage coverage (4). The line features in this coverage are also less detailed. Figures 5.47, 5.48, 5.49, 5.50, and 5.51 display the simplified lines (in Green) produced by the algorithm overlaid on their original (in Red). The numerical results are presented in Tables 5.26 and 5.27. As the tables report, the algorithm when the MR1 is used produced better results, in terms of data reduction and perceptual analysis, for the scale of 1:50,000 (Figure 5.47) compared to those by the TPF (Figure 5.48). That is about 50.16 % data reduction was achieved, yet no perceptible difference is detectable between the simplified and original lines. On the other hand, the TPF produced better results at the remaining smaller scales compared to those by the MR1 (Figures 5.49, 5.50, and 5.51). In order to produce the required number of points (27) for the scale 1:300,000, the application the MR1 caused the algorithm to reach the maximum data reduction where arcs are represented by straight lines as can be seen in Figure 5.50(A); hence no further data reduction was possible for the remaining scales. The results by the TPF show that there is a minimum increase of data reduction beyond the scale of 1:300,000 (Table 5.27, and Figures 5.45(a), 5.50(B), and 5.50. For the same reasons explained in the previous analyses, this limited increase is due to the process by which the algorithm selects critical points in the line during the simplification, in which beyond a large reduction of data (or points), in this case 72.73 %, the algorithm does not necessarily reduce the data in proportion to the progressive application of tolerance values. This is, and as explained before, partly to the algorithm's behaviour which is affected by the position of the points, but not by their number (cf. section 5.5.1.1).

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (319 points) |
|---|---|---|---|
| 1:50,000 | **6.80** | **159** | **50.16** |
| 1:100,000 | **32.25** | **80** | **74.93** |
| 1:300,000 | **1500.00** | **34** | **89.35** |
| 1:500,000 | – | – | – |
| 1:1,000,000 | – | – | – |
| 1:1,500,000 | – | – | – |

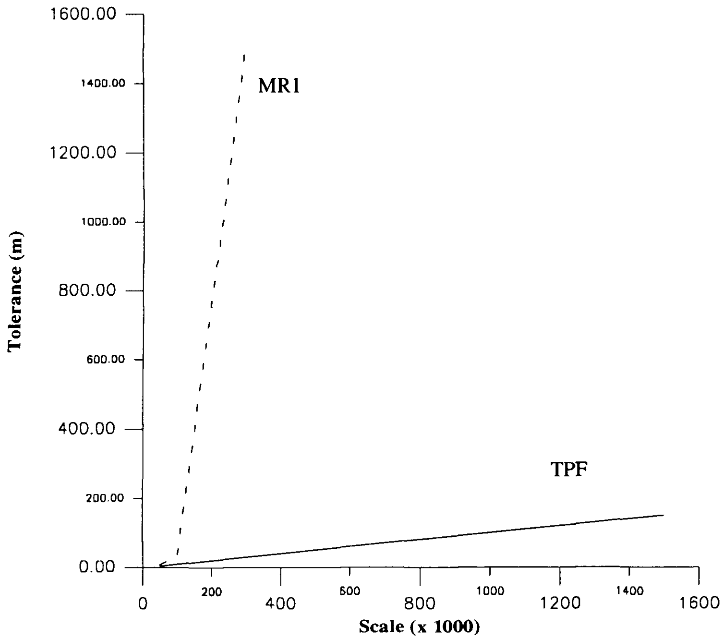
**Table 5.26: Digitally simplified maps from original (Study area 4, Transport Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the MR1.**

| Digitally simplified maps for different scales | Tolerance (m) | Number of points | % Data reduction from original (319 points) |
|---|---|---|---|
| 1:50,000 | **4.00** | **199** | **37.62** |
| 1:100,000 | **9.00** | **143** | **55.18** |
| 1:300,000 | **29.00** | **87** | **72.73** |
| 1:500,000 | **49.00** | **66** | **79.32** |
| 1:1,000,000 | **99.15** | **54** | **83.08** |
| 1:1,500,000 | **149.22** | **48** | **84.96** |

**Table 5.27: Digitally simplified maps from original (Study area 4, Transport Coverage) by the Douglas-Poiker algorithm, using the tolerance values of the TPF.**




*a*




*b*

Figure 5.45: Digitally simplified maps from original map (319 points, 33 arcs, Study area 4, Transport Coverage) using two types of tolerance: MR1 and TPF. ***a*** and ***b*** show the relationships between scale, % data reduction and tolerance values.

As in the case with the previous analyses, especially Study area 3, the topological pattern in this coverage (i.e., the relatively large number of arcs compared to the number of points) appeared to have constrained the behaviour of the algorithm in such a way that the maximum level of data reduction was quickly reached at the third scale (1:300,000), using a large tolerance value of 1500.00 m. On the other hand, the TPF produced smaller tolerance values (within the same range as the previous TPF values), resulting in data reductions for all scales; hence avoiding the extreme levels of reduction. However, the relatively low data reduction resulting from the Douglas-Poiker algorithm using the TPF values, is an effect partly caused by the presence of relatively fine details in the line features; hence a small number of points is present, and largely caused by the proportion of the number of arcs to the number of the points. In order to further demonstrate the effect of segmenting line features on restricting the effect of the algorithm, an example is provided where the drainage coverage in this study (721 points, and 4 arcs) were subdivided into 64 arcs total, and a range of tolerances were applied before and after the segmentation (Table 5.28). The table shows that the coverage represented by the fewer arcs (4) underwent larger data reductions compared to the coverage with more arcs (63) using the same tolerance values which ranged from 0.1 to 500 m. As the table reads, and as expected, the higher the tolerance value the higher the data reduction resulting from both the databases of the coverages; i.e., the smaller the number of points. However, the data reductions resulting from the coverage with fewer arcs are larger than those resulted from the coverage with more arcs. The difference between the two types of data reduction increasingly, yet systematically, appear larger at low levels of data reduction, using smaller tolerances, and at higher levels of data reduction, using larger tolerances. As marked in the table, the base line below and beyond which this difference systematically increase, is the tolerance of 30 m where the level of data reduction reached 86.97 % for both coverages. Beyond this level of reduction the application of the algorithm using progressively larger tolerances appears of limited effect. The example in this table emphasises that the tolerance value (or around it), and about this level of data reduction, are shown throughout the analyses in this chapter to represent the limit beyond which large tolerance values have little effect in producing correspondingly large data reductions. The previous analyses showed that this finding proved valid irrespective of the type of line features (see for example the figures comparing the data reduction results from the MR1 and TPF). As explained in sections 5.4.2.2 (explaining the parameters of the TPF) and 5.5.1.1, this critical limit of data reduction using such a tolerance value is due to the effect of the algorithm's approach which is based on the principle of critical point selection. According to this principle, the algorithm's behaviour is influenced by the positions of individual points on the source line but not their numbers, as shown in the beginning of the chapter, and explained in the sections referred to above. In regards to the difference between the two types of data reduction in Table 5.28, the increase below and beyond the tolerance of 30 m appears to be systematic, but an inconsistency is associated with tolerances of 3 and 4m. The data reduction, here, is expected to be consistent with increase in the tolerance values, (the difference should be smaller when using the tolerance of 4 m) but it is not. Instead, the difference is smaller (20 points) compared to that of the tolerance of 3 m (26 points). This indicates that such a difference is unpredictable, and it is mainly due to the reason pointed out above; that is the effect of the position of individual points in the lines. Although the example discussed here supports some observations found about the algorithm throughout the whole analyses, it is important to remember the purpose behind its presentation. The example clearly demonstrated that the algorithm is constrained by the existence of a large number of arcs in the line features being simplified.

| Tolerance (m) | Resulting number of points from the coverage with more arcs (63), 784 points | Resulting number of points from the coverage with few arcs (4), 721 points | Difference (number of points) |
|---|---|---|---|
| | | | |
| 0.1 | 718 | 674 | 44 |
| 0.2 | 701 | 658 | 43 |
| 0.5 | 642 | 603 | 39 |
| 1.0 | 574 | 539 | 35 |
| 2.0 | 448 | 416 | 32 |
| 3.0 | 378 | 358 | 20 |
| 4.0 | 313 | 287 | 26 |
| 10.0 | 187 | 178 | 10 |
| 15.0 | 149 | 139 | 10 |
| 30.0 | 102 (86.99 % data reduction) | 94 (86.97 % data reduction) | 8 |
| 50.0 | 77 | 66 | 11 |
| 70.0 | 73 | 59 | 14 |
| 100.0 | 67 | 45 | 22 |
| 150.0 | 66 | 40 | 26 |
| 500.0 | 64 (maximum reduction) | 16 | 48 |
| - | - | 5 (maximum reduction) | 59 |

**Table 5.28: The effects of line segmentation on the process of data reduction by the Douglas-Poiker algorithm.**

Given the above observations, it can be concluded that, although smaller scales require higher levels of data reduction, using large tolerance values, in most cases, does not necessarily yield acceptable database simplification of linear features at reduced target scales, especially if the database is in the form of the database under study; i.e., contains a relatively large number of arcs. Evaluation of the success of the TPF should consider the context of the data reduction requirements and the linear data themselves. Therefore, one should first consider whether the maximum or extreme reduction is an option. According to the data reduction requirements (explained in section 5.4.2.2), it is assumed that the maximum reduction resulting from the MR1 should be avoided by the TPF at any scale level, since otherwise it would produce undesirable line shapes. Given the limited database of this coverage, the fact that the range of data reduction resulting was not far from the level of data reduction that was proposed to be desirable (50 to 90 %), and the effect of applying large tolerance values, it can be concluded that the TPF produced desirable data reduction for all the scales in this coverage.

5.5.4.4 Graphic Reduction and Areal Displacement (Transport Coverage):

Unlike the drainage coverage of study area 4, the transport coverage suffered less areal displacement, using the TPF (Table 5.29). This is due to the relatively small data reduction. As Table 5.29 shows, there is a systematic increase in the sum of areal displacements with reduced scale (in ground units). In contrast, these displacements, in terms of map units, show a systematic decrease with reduced scale. Table 5.30 shows that the total areal extents of graphic representations of the simplified lines are larger than the total graphic representations of these displacements with a ratio of the difference to the graphic extent of the simplified lines reaches up to 93.41 %. The range of difference between the scales is about 15.41 %, (Table 5.30, and Figure 5.46). Again, these results are comparable to the previous analyses.

In conclusion, the Douglas-Poiker algorithm's behaviour on both coverages, using the TPF, produced perceptually and numerically acceptable results, compared to the MR1, according to the requirements of data reduction within the scale-dependent context (section 5.4.2.2). In terms of the relationship between the resulting areal displacement and the process of graphic reduction, the results appeared comparable to those of the previous study areas. Generally, and based also on the previous analyses, the rate of data reduction appeared to have been restricted by the number of arcs. It can, thus, be concluded that the larger the number of arcs within a line, the less data reduction can be achieved. Furthermore, this effect can well be intensified if the line already has a limited number of points. Thus, long and unsegmented line features such as coastlines, contours, and long and meandering streams, are highly likely to undergo large data reductions by the Douglas-Poiker algorithm (using the tolerances of either the MR1 or TPF) compared to man-made or cultural line features such as transport networks, and the like.

| Overlays | Number of Sliver Polygons | Sum of Areas ($m^2$) | Sum of Area at target scales ($cm^2$) |
|---|---|---|---|
| Original (1:25,000) | | | |
| **Overlay 1** 1:25,000+1:50,000 | **84** | 16,073.88 | **.064** (at 1:50,000) |
| **Overlay 2** 1:25,000+1:100,000 | **110** | 121,670.82 | **.121** (at 1:100,000) |
| **Overlay 3** 1:25,000+1:300,000 | **97** | 180,241.93 | **.020** (at 1:300,000) |
| **Overlay 4** 1:25,000+1:500,000 | **89** | 377,189.33 | **.015** (at 1:500,000) |
| **Overlay 5** 1:25,000 + 1:1,000,000 | **73** | 687,579.31 | **.006** (at 1:1,000,000) |
| **Overlay 6** 1:25,000 + 1:1,500,000 | **63** | 1,014,396.56 | **.004** (at 1:1,500,000) |

**Table 5.29: Areal displacement resulting from overlaying original (319 points, Study area 4, Transport Coverage) on digitally simplified maps (using the TPF) .**

| Digitally simplified maps for different scales | Total line length (m) | Area of drawing lines of digitally simplified maps using a Pen Width . 015cm at target scales ($cm^2$) | Sum of Areal displacement at target scales ($cm^2$) | Ratio of Difference to Area of Drawing Lines % |
|---|---|---|---|---|
| 1:50,000 | 32,350.99 | **.970** | **.064** (at 1:50,000) | **93.41** |
| 1:100,000 | 32,317.39 | **.484** | **.121** (at 1:100,000) | **75.00** |
| 1:300,000 | 32,168.90 | **.160** | **.020** (at 1:300,000) | **87.50** |
| 1:500,000 | 32,008.28 | **.096** | **.014** (at 1:500,000) | **84.38** |
| 1:1,000,000 | 31,799.80 | **.047** | **.006** (at 1:1,000,000) | **87.24** |
| 1:1,500,000 | 31,639.06 | **.031** | **.004** (at 1:1,500,000) | **87.10** |

**Table 5.30: Graphic and areal scale effects on the areal displacement resulting from overlaying the 1:25,000 original map (319 points, Study area 4, Transport Coverage) on digitally simplified maps (using the TPF).**

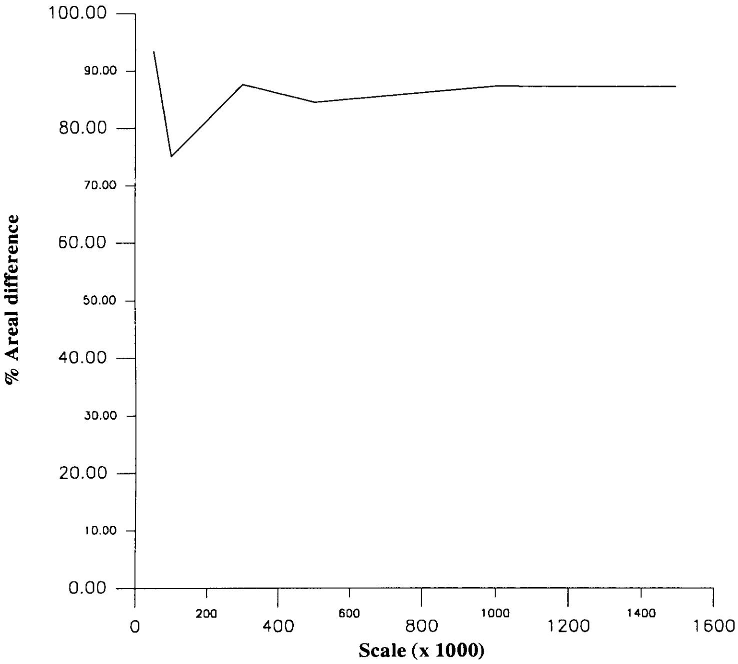



Figure 5.46: Ratio of areal difference (%) to graphic extent of simplified lines. (Difference (in map units) = (total graphic extent of simplified lines) - (total areal displacement). (Study area 4 - Transport Coverage).

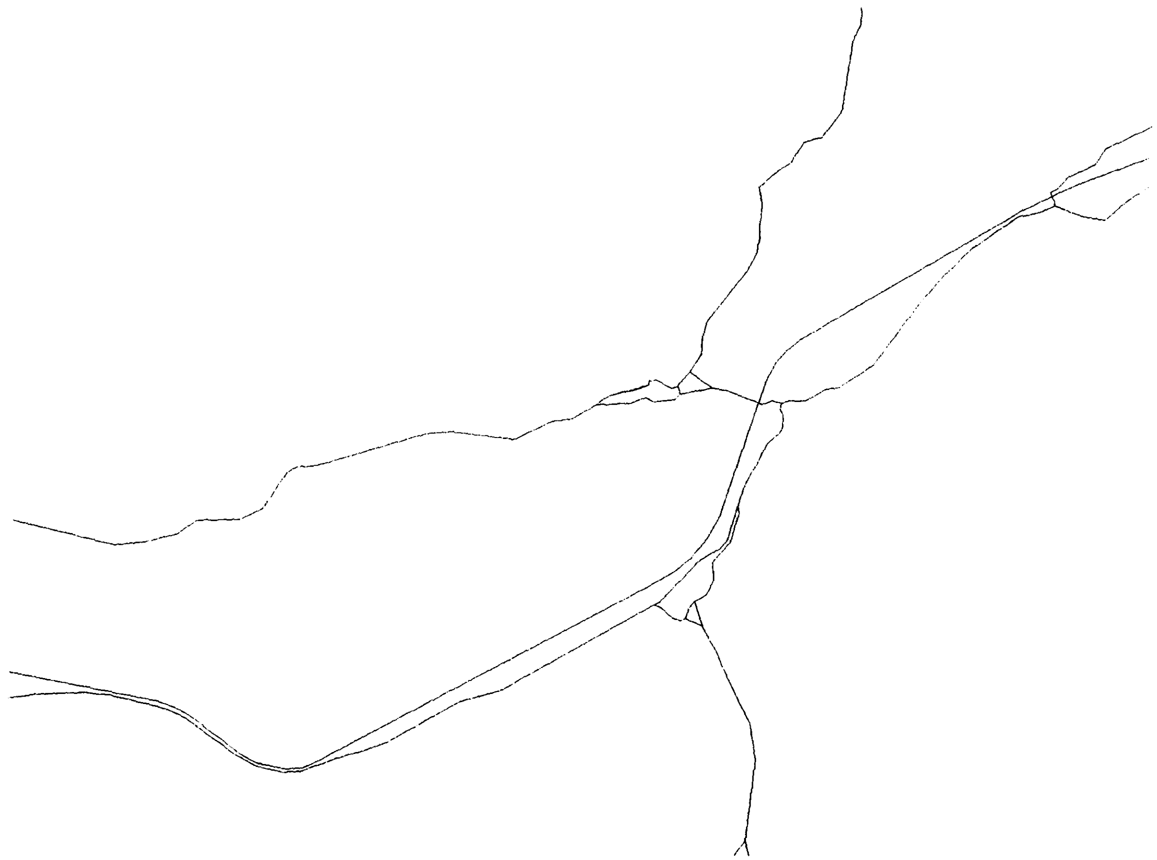


Figure 5.47:
An overlay of two maps at 1:50,000 (Study Area 4, Transport Coverage). Original map in Red (319 points, 33 arcs). Digitally simplified map (in Green) produced by applying MR1, T= 6.800 m, 159 points, 50.16 per cent data reduction.
Width of drawing line= .15 mm, represents 7.5 m on the ground.

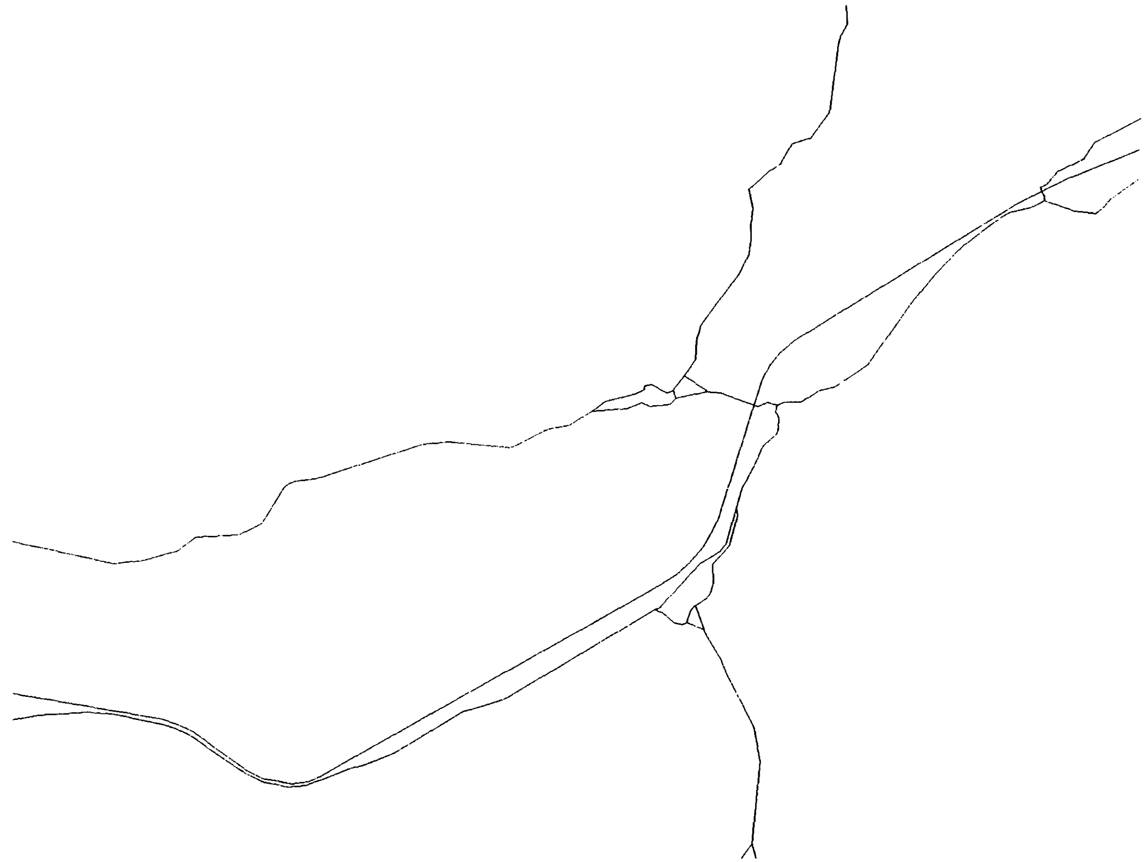
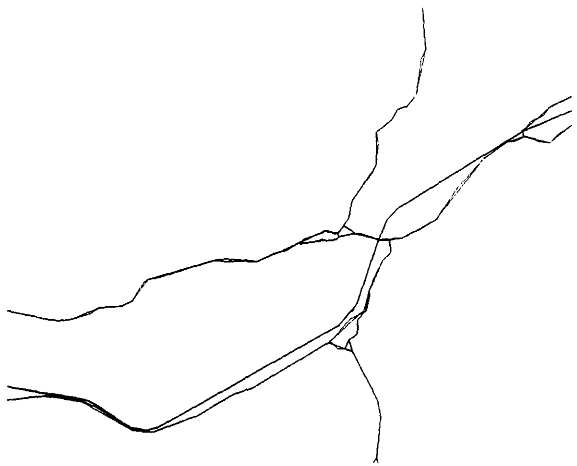
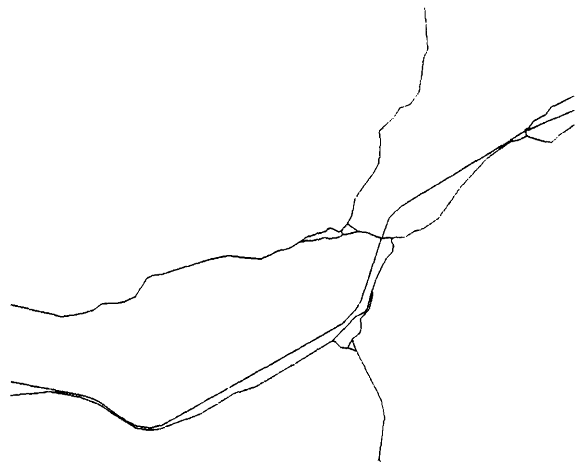


Figure 5.48:
An overlay of two maps at 1:50,000 (Study Area 4, Transport Coverage). Original map in Red (319 points, 33 arcs). Digitally simplified map (in Green) produced by applying TPF, T= 4.000 m, 199 points, 37.62 per cent data reduction.
Width of drawing line= .15 mm, represents 7.5 m on the ground.



A



B

Scale: 1:100,000

Figure 5.49: In (A) an overlay of original (in Red, 319 points) and digitally simplified map (in Green, 80 points, 74.93 per cent data reduction), produced by a tolerance of 32.250 m (MR1). In (B) an overlay of original (319 points) and, digitally simplified map (in Green), 143 points, produced by a tolerance of 9.000 m (TPF), 55.18 per cent data reduction. Width of drawing line= .15 mm, represents 15 m on the ground. (Study Area 4, Transport Coverage).

A

B

Scale: 1:300,000

Figure 5.50: In (A) an overlay of original (in Red, 319 points) and digitally simplified map (in Green, 34 points (89.35 per cent data reduction), produced by a tolerance of 1500.000 m (MR1). In (B) an overlay of original (319 points) and, digitally simplified map (in Green, 87 points, produced by a tolerance of29.000 m (TPF), 72.73 per cent data reduction. Width of drawing line= .15 mm, represents 45 m on the ground. (Study Area 4, Transport Coverage).

1:500 000

Digitally simplified map, 66 points (79.32 per cent data reduction)
Produced by a tolerance of 49.000 m (TPF).
Width of drawing line= .15 mm, represents 75 m.

Original (in Red, 319 points), and digitally simplified map.

1:1000 000
Digitally simplified map has 54 points (83.08 per cent data reduction)
Produced by a tolerance of 99.150 m (TPF).
Width of drawing line= .15 mm, represents 150 m.

Original (319 points) and digitally simplified map.

1:1 500 000
Digitally simplified map has 48 points (84.96 per cent data reduction)
Produced by a tolerance of 149.220 m (TPF).
Width of drawing line= .15 mm, represents 225 m.

Original (319 points) and digitally simplified map.

Figure 5.51: Original map ( in Red, 319 points) and digitally simplified maps (in Green) at 1:500,000, 1:1,000,000, and 1:1,500,000. Using tolerance values of TPF. (Study Area 4, Transport Coverage)

## 5.6 Summary

This chapter presented an evaluation of the Douglas-Poiker algorithm on the basis of data reduction within a scale-dependent context. Unlike the previous evaluations of the algorithm, this evaluation studied the algorithm according to the main purpose of its design; that is data reduction. Most of the previous implementations and evaluations have not adhered to this principle, and hence misuse of the algorithm was resulted. For example, the algorithm has been used for cartographic line simplification which is, as seen in Chapter 4, a complex and context-dependent process. Some have gone further and used the algorithm for generalisation purposes which is more complex than simplification. In the evaluation presented here, a definition of the process of data reduction for weeding purposes within the scale-dependent context was proposed, since it has never been formally presented. It was proposed that the process of data reduction within this context primarily demands that the integrity of the line shapes should be maintained, so that there should be no perceptible shape distortion, and so far as a cartographic representation is concerned, imperceptible distortion is irrelevant. This is the first criterion. It was also suggested that only irregular line features or features that are not straight lines or regular geometrical shapes (e.g., town plans and street guides) can be subject to this type of data reduction. This is important since regular-shaped features are usually represented by a relatively limited number of critical points and further processing of these points may cause undesirable modifications, since the resulting shapes would appear different across scales.

The second criterion is that, given the effects of graphic reductions (or areal scale) on line feature representation, a desirable range of data reduction is from approximately 50 % to 90 %, if the scale range was from 1:25,000 or larger (as the original) to 1:1,500,000 or smaller (as the reduced scale). Both criteria are applied to databases (points) of line features but not to their lengths. Furthermore, only databases that are composed of hunderds or thousands of points upon which the two criteria apply. Since there are no solutions as to how to achieve this range of data reductions objectively, preliminary trials were performed to test the Douglas-Poiker algorithm based on these two criteria, so that a formula can be developed to produce direct input parameters as tolerance values to be used by the algorithm. For a given target scale the formula can be utilised to derive the required tolerance value to be applied by the Douglas-Poiker algorithm. A formula was proposed and termed as Tolerance Prediction Formula (TPF) (section 5.4.2.2). Initial tests and the theoretical justification of the TPF proved promising. Values derived from the formula were then tested using larger data sets of different line features.

The original formula of the Radical Law (RL), discussed in section 5.4.2.2, was utilised and proved to be unsuitable as it tends to results in redundant data. On the other hand, employing the MR2 which was modified from the RL in order to consider the effects of graphical reduction (or areal scale) appeared to be undesirable because reducing databases in relation to areal scale appeared to be excessive. The MR1 was seen as a comprise between the two formulae for RL and MR2, where the data reduction related to linear scale. The TPF and MR1 were both applied to give tolerances for scale-dependent data reductions in the test data sets. Unlike the TPF, MR1 is basically a method of prediction of the required data reduction in the form of the number of points, but not in the form of tolerance values. This implies that the user has to perform a series of iterations until the required number of points, determined by the formula, is reached.

The evaluation showed that the TPF results compared to those by the MR1 generally conformed well to the criteria and the evaluation measure proposed above; data were reduced (around and beyond 90 % for the smallest scales, and about or beyond the minimum of 50 %), but no perceptible differences resulted between the simplified and original lines at the target scales. However, the evaluation revealed that the numerically desirable results do not necessarily coincide with the graphically desirable results throughout the analyses. This was due to factors such as the proportion of the number of arcs to the number of points (or, the degree of segmentation of the lines), rate of scale change, and complexity of the line details. For example, due to the effects of line segmentation, and line details (less details), the Douglas-Poiker algorithm produced data reduction of less than 50 % for the coverages of study area 3, and the transport coverage in study area 4 (see Table 5.28 and associated discussion). It can, therefore, be concluded that the greater the severity of the effects of these factors the less data reduction occurs. It is, thus, important to bear in mind that while the criteria were intially proposed to provide general evaluation guides or measures, they are like any general rules that do not necessarily apply to every individual case.

On the one hand, this unconformity between the numerical and graphical results might be regarded as reasonable especially when large tolerances would lead to extreme data reductions which commonly lead to distorted line shapes that can be perceptible at target scales; see for example, the effects of point reduction advised by the MR1. This indicates that so far as these factors or constraints are concerned, employing large tolerances does not necessarily lead to graphically desirable or acceptable results from the Douglas-Poiker algorithm when used with the TPF. This supports the criterion proposed that limiting the range of data reduction within the scale-dependent context provides generally desirable graphical results, backed by reasonably acceptable numerical results. The evaluation indicated that the Douglas-Poiker algorithm, using either of MR1 or TPF, could produce data reduction of around 75 % for the scale of the 1:50,000, but no deviations between the simplified and original lines were perceptible at the target scale. This result was always obtained when the algorithm's performance is unaffected by the constraints mentioned above; for example, see the results of study area 1. The evaluation also showed that the Douglas-Poker algorithm appeared to reach a level of data reduction, usually between 75 and 90 %, beyond which the effects of the increased tolerance values were limited, irrespective of the line complexity. This is primarily due to the fact that the algorithm's behaviour is affected by the position of individual points but not by their numbers. Hence, data reduction is largely achieved by applying certain tolerance values, while beyond this large reduction a progressive increase in the tolerance values does not necessarily lead to a progressive data reduction (see sections, 5.4.2.2, 5.5.1.1, and 5.5.4.3). It was observed that this level of reduction approximately corresponds to a range of tolerance values from 20 to 30 m. Thus, the algorithm's performance on different types of line features was invariant. This poses the question whether applying a single tolerance value on a segment of line or lines is valid from the data reduction perspective.

Buttenfield (1985) among others argues that in order to produce balanced results, that are consistent with the geomorphological shapes of line features, the lines have to be segmented so that for each segment a suitable tolerance value might be applied. While this proposition of segmentation might well be understandable in terms of generalisation, (although, a single simplification process can not solve or substitute generalisation), it would be much harder to implement and possibly cartographically untenable, in terms of data reduction within the scale-dependent context. Cartographically, it would lead to an inconsistent data-reduction in the imperceptible realm. This is because, all types and complexity of lines stored in a digital file are a geometrical problem according to which the algorithm presents a consistent treatment; hence result. On the other hand, the coverages, as can be seen from the analyses, are essentially compiled from segmented lines. Economically, processing every line segment in a coverage is time consuming and costly. However, the results indicate that the application of the algorithm, using TPF did achieve consistent data reductions. The significance of the proposed formula (TPF) is that it is a practical solution for the user of the Douglas-Poiker algorithm for performing data reduction process within a scale-dependent context from a single detailed database, provided that the databases as well as the TPF values are in meter units.

In the analyses, the relationship between the output resolution, graphic reduction and the graphic representation of the reduced data was addressed. It is a quantification approach for such a relationship. The method is a general numerical description of the relationship between the total areal displacement resulting from the data reduction and the total areal graphic extent of that reduction to the respective scale. Given the size of the output resolution, all the four analyses showed that the total areal extents of the graphical representations were larger than the total graphical extents of the total areal displacement at target scales by a percentage extending up to 99.51 %. This result conformed to the perceptual results, in which the displacements were almost imperceptible. This measure should help indicate whether the total areal displacement is or is not perceptible at the target scale. However, results showed that there might be perceptible deviations between the simplified and original lines whereas the numerical results indicate that the total areal extent of the simplified lines is larger than the total areal displacement between the simplified and original lines; and hence the numerical results can not be relied upon for assessing the quality of data reduction performed.

As the results demonstrated, the algorithm was capable of selecting the critical points of a line feature in such a way that when the original and simplified lines were overlaid they exhibit imperceptible differences at the target scale. Thus, and contrary to the decision reached by Whyatt (1991), it is concluded that the tests demonstrated the ability of the algorithm for weeding purposes when used in conjunction with the TPF. It is, therefore, the purpose of what follows in this thesis to establish whether this property has a cartographic value in the process of digital cartographic line simplification during a cartographic production process.

# CHAPTER SIX: CARTOGRAPHIC PRODUCTION AND DIGITAL LINE SIMPLIFICATION

The aim of this chapter is to explore the scope of digital line simplification as it is implemented in a digital cartographic system in cartographic production for producing an acceptable quality. Unlike the previous studies, this study adopts the definition and role of simplification within generalisation. For this purpose, two digital simplification algorithms are tested; namely, the Douglas-Poiker algorithm, and the Cubic Spline smoothing routine. This experiment specifically explores the potential of the integration of the two algorithms, based on the requirements of cartographic line simplification within a typical digital cartographic production. In this chapter, previous attempts at digital generalisation are first referred to, and then a proposed model for cartographic line simplification within the digital generalisation paradigm is presented. The results and discussion of the application of the model on a selected set of line features are finally presented.

## 6.1 Digital Cartographic Generalisation

The review of the cartographic generalisation processes in Chapters 2 and 4 provided an insight as to how generalisation is commonly conceived and practised in the traditional realm of mapping. There has been little effort expended in examining the traditional definitions of these processes and how they relate to the digital context. McMaster and Shea (1992) point out key areas in which digital generalisation differs from manual generalisation. First, manual generalisation is labour-intensive compared to the digital processes which are fast computer-driven operations. Second, the manual process is characterised as a highly subjective process. In contrast, digital generalisation is necessarily objective. Third, the process of manual generalisation is known to be "holistic in its perception and execution", while digital generalisation works in serial mode in which the processes are performed independently.

The subjective elements associated with manual generalisation are not only based on the skill and experience of the cartographer, but, also, on his understanding of geographical characteristics of the features being processed in order to create a representation consistent with existing geographical knowledge. This special requirement makes it almost an impossible process to automate (McMaster and Shea, 1992). Several efforts in automated generalisation have used the manual process for guidance, although some researchers have stressed that, automated generalisation should not be a direct translation of manual methods (Rhind, 1973). Research work in automation shows two different trends. Most of it has tended toward oversimplification (e.g., McMaster, 1986, 1987; Li and Openshaw, 1992, 1993; Abraham, 1988; Zoraster *et al.*, 1984). Such efforts have looked to the process of automated generalisation as a problem that can be solved by single algorithms such as those for simplification. In contrast, others (e.g., Buttenfield, 1991; Shea, 1991) have pointed out that the process of generalisation can be decomposed into specific rules so that the process of generalisation can be incorporated into expert systems. For example, a rule might be:

IF *an overlap of two features is below the minimum spacing requirement*

THEN *displace the overlapping features to meet the minimum spacing requirement.*

The first part of the rule might invoke proximity detection algorithm, while the second part might invoke a displacement algorithm. As Shea (1991) points out, these rules (called production rules) are acknowledged to be readable and easy to understand. Second, they behave much like independent pieces of knowledge and as such, rules in the knowledge base can be independently modified with little direct effort on other rules. Finally, being an ordered list, the rules fired in the decision-making process help to explain the line of reasoning or to justify any conclusions reached (Shea, 1991). The disadvantages of these rules are that "rules in a production system lack topology. The strong modularity of the [rules] results in inefficiencies of program execution, and makes rule maintenance unwieldy as the number of rules increases. Several other knowledge representation schemes address these deficiencies. Production rules may be expressed in several ways: as logic, structured in a semantic network or they may be gathered together in a frame" (Shea, 1991, p.11).

The logical methods can be represented by conditional and combinatorial expressions (such as AND, OR, IF-THEN-ELSE). These schemes for knowledge representation are useful for representing facts about objects and how those objects relate to each other (Shea, 1991). As Shea (1991) indicates, that the advantages of logic representation are ": 1) logic represents a natural way to express an intuitive understanding of a domain; 2) is precise and consistent in the expression of the formalisation; 3) offers flexibility in that the representation of fact is not tied to the use of that fact; and 4) includes logical assertions which are modular and independent from one another" (Shea, 1991, p.11). Robinson and Zaltash (1989) report a study in which the generalisation rules used by the Ordnance Survey, Great Britain in generalising from large (1:250 and 1:2500) to medium scale (1:10,000) were analysed and put into a rule-based expert system advisor called OSGEN (OSGENeralisation). The aim of their study was to investigate the nature of rules, interaction among them and how complete they were and whether they could be translated into a computer-compatible form (Robinson and Zaltash, 1989). The result of the study indicated that most of the rules were too specific and could not be applied as stated in many cases. Also, the rules were not complete, and according to the authors that was due to the fact that a great deal of cartographic knowledge is held in the cartographer's head. Such a valuable practical approach is what is actually needed, and, only, the results of these approaches which would matter in the current and future efforts in the field of formalising cartographic knowledge.

A program called CHANGE was developed at the University of Hanover, and is a combination of procedural steps which might come close to the idea of fully automated generalisation (Powitz and Schmidt, 1992; Gruenreich, 1993, 1995). The goal for developing this program is to generalise automatically some feature classes of German topographic maps from 1:5,000 to 1:25,000. However, the program encounters two main drawbacks which are: 1) it generalises features (mainly topographic) for a limited ranges of scales, and 2) it performs only 50 to 60 % of the work (Muller *et al.*, 1995b) .

Nickerson (1988) developed a system for automated generalisation of topologically structured cartographic data. The program supports processes such as feature elimination, simplification and interference detection. The system is only intended to generalise within the scale range of 1:24,000 to 1:250,000 (Nickerson, 1988).

Some researchers have expressed doubts about the cartographic validity of the expert systems (e.g., Fisher and Mackaness, 1987). Beard (1991) effectively outlines three main factors which "impede" the formalisation of rules for generalisation.

"First, generalisation has traditionally been practised as an individual artistic skill and therefore incorporates subjective components which do not readily decompose into logical rules. Formalising the subjective elements is difficult and tends to sacrifice unique and creative aspects of map making. A second impediment arises in tailoring generalisation for a specific map purpose. Rules effective for the generalisation of one map type may not be effective for another. Rules for the generalisation of a soil map, for example, may not be transferable to the generalisation of nautical charts. Development of a common rule base therefore potentially loses sensitivity to requirements of a particular purpose or application. A third difficulty arises in responding to variation in the spatial and non-spatial characteristics of the geography being represented. Rules should be responsive to local context, considering the spatial and attribute relationships of neighbourhoods of objects and not simply objects in isolation. Spatial and attribute relations among objects, however, can be very diverse, with each variation requiring a slightly different generalisation decision or rule" (Beard, 1991a, p. 121).

A more pragmatic approach to automated generalisation is considered by many digital mapping users and researchers, which is the interactive approach where the generalisation problem can be tackled by a joint human-machine task. In this approach, the computer executes some tasks where it is superior to the human but relies on the user's control and knowledge. Such an approach is suggested by Weibel (1991) and is termed the *amplified intelligence* approach. "The present trend is to use the interactive environment made available through work stations, PCs and powerful interfaces" (Muller *et al.*, 1995b, p. 9). "In principle, the interactive approach to automated generalisation offers more flexibility; it allows the application of procedural knowledge closer to the needs of the user. Its drawback could be that it may require a large period of interactive operation" (Muller *et al.*, 1995b, p.11).

This call for some user control is, in fact, supported by early works by researchers like Brophy (1973) and Brassel (1985). Brophy (1973) identified the main subjective element as map purpose for which users should have some control over parameters used in the generalisation process. In this direction, the system MGE Map Generalizer was developed by INTEGRAPH and is a significant contribution towards digital generalisation (Lee, 1993, 1995). This system will be briefly presented in section 6.1.2.1.

### *6.1.1 Models*

Attempts have been made at producing comprehensive conceptual models for generalisation. Some of these efforts are discussed below with much emphasis on the work presented recently by McMaster and Shea (1992).

Ratajski (1967) proposed one of the first formal models of generalisation. He identifies two fundamental components of the generalisation processes: quantitative and qualitative. Generally, the quantitative approach involves the gradual reduction in map information which depends on scale change, while the qualitative approach is concerned with the transformation of symbolic forms to more generalised ones. Morrison (1974) presented a model in which the relationship between the four main processes of generalisation (simplification, classification, symbolisation, and induction) were formalised based on set theory. Rhind (1973) argues that automated generalisation should be based on the manual approaches, but not be a direct translation of them. He lists the essential components of automated generalisation as line sinuosity reduction, feature transposition, within and between category amalgamation, feature or category elimination, and graphic coding change. In Nickerson and Freeman (1986) a model was proposed on which an operational system has been based and published by Nickerson (1988). The model consists of three components: 1 ) A *source map* with known scale (denoted 1 million) symbolism and area $w * h$. At this level of generalisation processes of deletion, simplification, combination, and type conversion are applied. 2) An *Intermediate map,* at which the symbol size is enlarged to $Ka$, where $K$ is a factor greater than unity. At this level, both feature relocation and symbol placement occurs at the intermediate scale. 3) The *target map,* which is produced from scale reduction and subsequent map displacement.

A further model is developed by Brassel and Weibel (1988) which is regarded as conceptually more detailed. The authors propose five separate processes of generalisation in a digital environment. These processes are: 1) *structure recognition*; 2) *process recognition*; 3) *process modelling*; 4) *process execution*; and 5) *data display*. *Structure recognition* involves identifying specific cartographic objects or aggregates of objects, spatial relations and measures of importance. This is followed by process *recognition*, which is concerned with what type of generalisation operation is needed. Process modelling is concerned with compiling rules and procedures from the process library. The last process is *data display*, which transforms the target data to the target map. The authors present what they call statistical and cartographic generalisations. Statistical generalisation is defined as a filtering process, where the focus is on data manipulation, whereas cartographic generalisation is related to the issues of graphic representation of the generalised map.

Brassel (1985) proposes a linkage between generalisation operations or functions (e.g., expand, displace, eliminate and smooth) and geographic objects. These objects are categorised to include point, line, composite, area, and volume features. For example, the operations that might be required to generalise line features are as followed: 1) expand/shrink/select/eliminate; 2) reduction of sinuosity; or change topology in linear networks; 4) displace; and 5) classify.

There are, also, other models for automated generalisation (e.g., Geraint *et al.*, 1995; van Oosterom, 1989, 1991, 1993). McMaster and Shea (1988) were the first to propose a comprehensive, conceptual generalisation model, based on a philosophy of digital generalisation. In a book published in 1992, devoted entirely to generalisation in digital cartography, the authors state that

"existing definitions of the generalisation processes are inadequate when considering the unique and complex processing paradigm of a digital environment. What is lacking thus far is a precise and carefully conceived definition of the nature of generalisation which embraces a digital generalisation philosophy that focuses on numerically-based manipulations. We would like to address this issue by presenting a definition of digital generalisation, not offered as a replacement for the existing body of work that has viewed generalisation strictly as a manual endeavour, but as a definition that is more clearly aligned with the unique and current needs of digital processing environments. Although this definition focuses on the generalisation of digital data, it remains applicable to manual generalisation processes" (McMaster and Shea, 1992, p. 3).

As McMaster and Shea (1992) indicate, "digital generalisation can be defined as the process of deriving, from a data source, a symbolically or digitally-encoded cartographic data set through the application of spatial and attribute transformation" (p. 3).

In their model, McMaster and Shea (1992) divide the generalisation process into "three operational areas: 1) a consideration of the philosophical objectives of **why** to generalise; 2) a cartometric evaluation of the conditions which indicated **when** to generalise; and 3) the selection of the appropriate spatial and attribute transformations which provided the techniques on **how** to generalise" (p. 27). Figure 6.1 illustrates this model, in which the three main components are further decomposed into other elements. With reference to the figure, the authors indicate that a systematic organisation of the **when** and **how** of generalisation, in the form of operators, algorithms, or tolerances, can help to form a "complete approach to digital generalisation". The authors conclude by stating that the generalisation process is now "matured, and the potential for appropriate computing resources is available" (p. 12). Certainly, this is not the case, since there is an evident lack of understanding of the cartographic generalisation processes. The confusion of terms describing generalisation operations, as between line simplification and line generalisation, and the absence of a fully automated generalisation system or program are clear evidence of this lack of understanding.

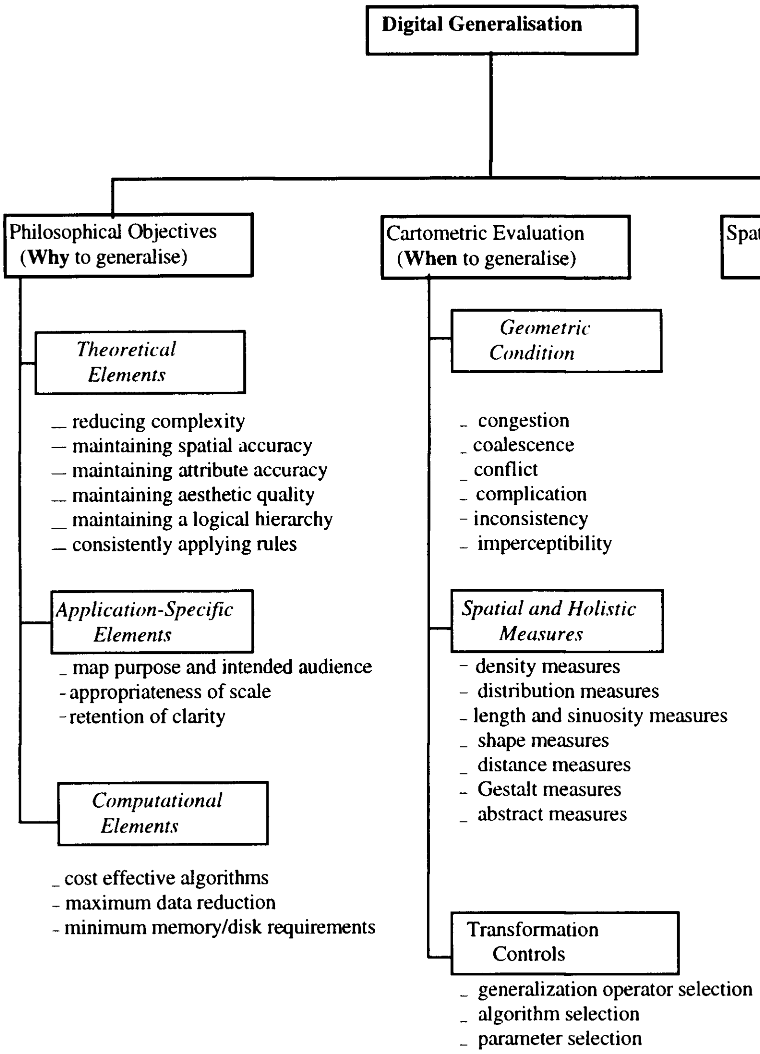



Figure 6.1: The Digital Generalisation Model proposed by McMaster and Shea, 1992.

*6.1.2 Interactive GIS-based generalisation:*

As noted above, the interactive approach to generalisation seems to be a pragmatic step towards total automation of generalisation. The advantage of this approach lies in the fact that it helps aid in gaining an insight about how cartographic generalisation is perceived and executed by cartographers. It has recently been recognised that Geographic Information Systems (GIS) appear to provide a new context for considering problems of map compilation and generalisation. Morehouse (1995) points out that GIS-based cartography separates the mapping process into three parts to which generalisation issues are relevant. These parts include: 1) compilation of geographic database; 2) processing tools, which allow for translation of geographic information from one form to another; and 3) rendering tools, which transform objects in the geographic database into symbolic forms on a map. Morehouse graphically illustrates how GIS supports a wide variety of graphic displays (maps) that can be generated from one single geographic database (Figure 6.2).

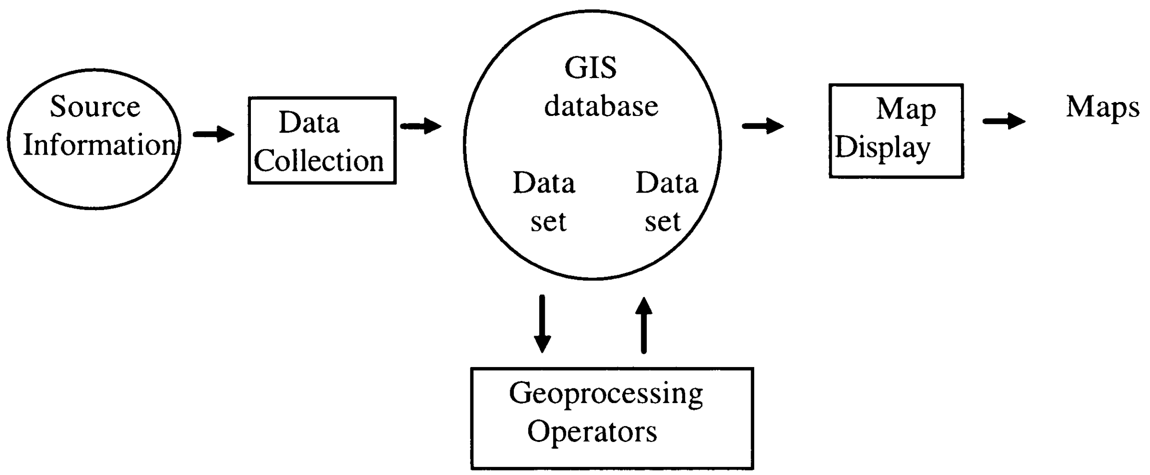



Figure 6.2: GIS-based cartographic process (after Morehouse, 1995).

Spiess (1995) emphasises that searching for a complete batch solution to generalisation is an "illusion". He, first highlights the need for generalisation in a GIS environment, since visualisation of GIS-based products in the form of paper will continue to be a demand, "as not everybody has access to a workstation or the necessary background to use it" (Spiess, 1995, p. 37). Spiess points out that the interactivity aspect provided by GIS tools leads to generalisation modules that are clearly defined. "The result of these modules must be evaluated visually step by step and not only the final image. This allows us to gain experience and expertise in the effects produced by certain parameters or algorithms on specific occurrence of data" (p. 45).

6.1.2.1 MGE Map Generalizer (MGMG):

The powerful potential mapping capabilities of GIS have certainly encouraged progress in incorporating generalisation operators within these systems. The Intergraph GIS presents a significant contribution to interactive generalisation (Delucia and Black, 1987; Lee, 1993, 1995). As Lee (1995) explains "the fundamental task in developing such a system is to translate the conventional (manual) experience of map generalisation to a computer process. This can not be achieved without a complete understanding of how the cartographer handles the art and science of generalisation" (Lee, 1995, p. 219). Algorithms that are supported by this system are based on the belief that they should "emulate the human approach to pattern recognition through the principles of perceptual organisation described by psychologists and used to advantage in communication by the cartoonist and somewhat less by the cartographer" (Delucia and Black, 1987, p.169).

As Delucia and Black (1987) point out, the proposed approach to automated generalisation through this system, is based upon "the cartographic reality that all points in space are not equal in "information value". Therefore, generalisation algorithms should determine the hierarchy of both importance "between features" and "within features".

Delucia and Black (1987) state that "the fruitful path to a working automated generalisation lies in the blending of mathematical and heuristic algorithms to produce results that more closely approach those that can now be produced by the experienced and able cartographer" (p. 170).

Being an interactive system, MGMG provides the user with "the most flexibility and control on what to generalise, how to generalise, and how much to generalise" (Lee, 1995, p.219). The formalisation of the generalisation process in the system is implemented as a learning process performed between the user and the system (Lee, 1995). There are nine generalisation operations within MGMG, which are thought to represent the process of manual generalisation. They are:

| | |
|---|---|
| 1- selection / elimination | 6- aggregation |
| 2- simplification | 7- collapse |
| 3- typification / refinement | 8- exaggeration |
| 4 - classification | 9- displacement |
| 5- aesthetic refinement | |

Some of these functions have been implemented. Lee (1995) outlines some aspects that should be considered when formalising a generalisation work flow using the product with the available functions and tools:

1- *feature class dependency:*

Each feature class (e.g., hydrography, landuse) is characterised differently and, therefore, requires a unique generalisation sequence.

2- *feature dimension dependency*

There are different dimensions of features (i.e., point, line, and area) for almost every feature class. According to the system, most of its generalisation operators are designed for certain feature dimensions. For example, simplification is for linear features, and aggregation is for points and areas. So, the work flow for a feature class can be subdivided based on feature dimensions.

3- *operational logic*

It would be logical and efficient to start a generalisation process with the operations of selection or elimination. An iteration process is often necessary with other generalisation operations.

4- *minimising spatial accuracy reduction*

Although the process of digital generalisation causes displacement of data points from their original positions, this displacement can be minimised. For example, displacement can partly be minimised by carefully choosing algorithms and parameters, and executing operations in a certain order.

5- *maximising process efficiency*

When processing efficiency becomes important in certain cases (e.g., large data, or large scale change), the workflow should be customised for maximising the process efficiency throughout.

6- *minimising the need for subsequent readjustment*

It is highly recommended to avoid unnecessary repetitions of the same operation. For example, in generalising a cluster of small areas, aggregation followed by smoothing would be better than smoothing followed by aggregation, because the latter may require further smoothing of the aggregated boundaries.

These and other guidelines are most helpful in the context of interactive generalisation, and can be incorporated in the system. Another property of this system, is that all the parameters can be "saved and further adjustments could be carried out to refine the process, improve the result, and optimise the workflow" (Lee, 1995, p.234).

Although this system may still reflect the subjective nature of generalisation, the processing time and consistency are considered significantly superior to manual generalisation. However, work is still going on to improve the system. Muller *et al.*, (1995b), recognise the advantages of the MGMG system as being more flexible and versatile than a batch program, "...but the question is whether it is practical in large production environments. The provision of logging and script capabilities which can 'remember' the values of the parameter used and the scenarios that were adopted for similar map situations may give a partial answer"(p. 10). However, MGMG is very helpful for acquisition of knowledge relating to generalisation (Lee, 1995).

This thesis maintains that the interactive approach using GIS capabilities provides an excellent way forward for cartographic generalisation which can be closely explored and analysed, so as to be able to contrive ways for formalising this knowledge. In the remainder of this chapter, the role of line simplification within a digital context using the Arc/Info GIS is assessed, but first there is a review of the previous efforts in digital generalisation as specifically applied to line features.

## 6.2 Digital line generalisation

Previous discussion was focused on the general issue of generalisation and how geographic features (i.e., point, line, and area features) can be generalised in the digital environment, in either vector or raster format. For reasons discussed earlier, line features, in vector form, have been the focus of digital automation efforts more than have point and areal features. Many workers proposed simplification algorithms as a single solution to generalisation. On the other hand, some have presented conceptual and operational models for digital line generalisation (Abraham, 1988; Nickerson and Freeman, 1986, 1988; and McMaster,1989, 1995). These models are based on the cartographic fact that line features should be generalised by more than a single cartographic step such as simplification.

Nickerson and Freeman (1988) specifically suggest four processes that might be involved in a line generalisation process. These processes include: 1) feature deletion; 2) feature combination, where line features can be combined in a single feature; 3) feature simplification; and 4) interference detection. There are two processes involved in interference detection: a) determining if two line features will collide; and b) the subsequent displacement of the features. Zoraster *et al.,* (1984) propose that manual line generalisation can be replicated by three digital processes which are: simplification, smoothing, and displacement. McMaster (1989) proposes that a sequence of operators is necessary for performing an effective generalisation of digital line data. He proposes the application of two integrated operations which are simplification and smoothing (Figure 6.3). The application involves two separate stages: "database preparation and mapping". The first stage is concerned with preparation of digital data for storage within the database through cleaning (i.e., removing digitising errors), preliminary smoothing (using routines such as weighted-averaging or distances averaging algorithms), and finally simplification (using a sequential approach such as Lang's algorithm). The second stage is required for "mapping or cartographic display", where there is a scale change. McMaster advocates the use of global simplification algorithms such as the Douglas-Poiker algorithm (where the concern is about critical point selection) followed by a smoothing process for aesthetic purposes. McMaster points out that these two stages correspond to the terms applied by Brassel and Weibel (1988) which are "statistical and cartographic generalisation".

**Proposed Generalisation Procedure**
For Generalising Linear Data

**A. Database generalisation:**

*1. Clean digitised data*
- *eliminate duplicate co-ordinate pairs*
- *clean digitising errors*
  - *spikes*
  - *switchbacks*
  - *polygonal knots*

*2. initial smoothing*
- *apply weighted-moving average to eliminate effects of digitising*

*3. Database simplification*
- *apply sequential method for initial weeding of co-ordinate density previous to storage of data*
- *Lang algorithm*
- *Reumann-Witkam algorithm*
- *Opheim algorithm*

**B. Mapping - Cartographic Generalisation**

*4. Mapping simplification*
- *apply Douglas global algorithm*
- *primary concern with critical points*

*5. Secondary smoothing*
- *smooth lines for aesthetic quality*
- *apply weighted averages / splines*

Figure 6.3: McMaster's Model for the sequential processing of digital line data (after McMaster 1989).

Based on this model McMaster (1989) finds that the application of a smoothing routine before simplifying digitised lines increases the efficiency based on two geometric measures; 1) the total areal displacement between the original line -both smoothed and unsmoothed- and the simplified line; and 2) the areal displacement index (McMaster, 1989, p.114): the smaller the values of these two measures the greater the efficiency or the quality. Clearly, the proposed model by McMaster (1989) does not adequately address the wider context of digital line generalisation. This is due to the complexity of the generalisation processes that might well be involved in line generalisation. The significance of this work might be related to the fact that McMaster presents a study about the interrelationship between two algorithms together applied to digital line data, although he provides no clues as to what tolerance values might be utilised.

Weibel (1996) emphasises that line simplification itself is not a simple task, which can not be solved by existing algorithms as they are based on very restricted criteria. He discusses some constraints on line simplification, specifically the simplification of polygonal subdivisions. These constraints include: metric, topologic, semantic, and Gestalt constraints. Weibel indicates that while the first three constraint types relate to the basic aspects of data modelling and can be formally defined, the fourth (Gestalt constraints) is dictated by aesthetic and perceptual criteria and thus more difficult to operationalise. Based on the discussions of simplification and generalisation in Chapters 2 and 4, and in this chapter of this thesis, it would appear that Weibel implicitly refers to the process of line generalisation, not line simplification. Weibel suggests that the process of line simplification that considers all these constraints should be called an "extended line simplification" but not line generalisation. One is therefore left with the question of: what is line generalisation? Unlike this framework proposed by Weibel, the current thesis maintains that line simplification is only a part of the comprehensive process of line generalisation. This is important, because although the distinction between the two processes is difficult to define, formalising it would be more difficult if a single simplification algorithm was to achieve the intricate task of generalisation.

## 6.3 Digital Cartographic Line Simplification

The following sections explore the role of cartographic line simplification within the context of digital cartographic generalisation according to its perceived role as a sub-process in cartographic generalisation. For this purpose, a model is proposed using two simplification algorithms, namely, the Douglas-Poiker algorithm and the Cubic Spline smoothing routine. These are independently implemented in the Arc/Info software, although their uses are not formally defined whether for data reduction (weeding process), simplification, or generalisation. In short, their application is determined by the user.

### *6.3.1 Theoretical bases:*

Cartographic line simplification is perceived to be a process by which minor details along lines are removed while the larger ones are retained (see Chapters 2 and 4). As noted in Chapter 2, specific rules or guidelines as to how much details should be removed or retained is a question not only determined by the scale and purpose of the intended map, but, also, by the subjective decisions made by the cartographer. This particularly makes any attempt at standardisation of the simplification process in the digital realm very difficult. Whilst this chapter is principally concerned with this very issue of formalism, such a difficulty has to be investigated in order to find some understanding or solutions as to how cartographic line simplification can be formalised in the digital context, and what are the criteria for judging the quality of the product? It is by no means a simple task. Previous analyses in Chapters 3, 4, and 5 suggest that some formalism can be presented based on important observations, and these are discussed in detail below.

With reference to the discussions in Chapters 2 and 4, the first consideration, therefore, should be related to the assumed role of line simplification as a *cartographic* process within cartographic line generalisation in the digital environment. Simplifying a line feature implies a transformation in both the shape and data of the line, depending on the feature complexity. In the traditional practice of generalisation, simplification is concerned with manipulation of minor details in accordance with the map scale and purpose. Generally, larger details of line features are subject to other generalisation processes such as exaggeration, and typification. Also, the interaction between simplification and other processes during generalising line features for different mapping purposes and scales is almost unknown. It is therefore assumed that line simplification should be *a* process in digital generalisation of line features. That is, a digital line simplification algorithm should not replace other digital processes required for cartographic line generalisation. Digital cartographic line simplification is, therefore, a process to be viewed cartographically within the wider context of cartographic line generalisation.

The second consideration is related to the digital concept and mechanism upon which multiple cartographic productions are to be based. The question is now how can an acceptable and multiple cartographic simplification might be achieved digitally? Given the lack of guidelines or well defined rules for this task, this study proposes a conceptual model for cartographic line simplification, within which formal elements are involved. The model is principally based on an important observation in Chapter 5 and supported by the existing literature. This observation is related to the fact that multiple production from single detailed databases can be achieved by hierarchical algorithms by which data at high hierarchical levels are only subsets of the original or low levels of the hierarchy. That is, the database is designed hierarchically around the concept of critical points selection. The concept of hierarchy is utilised for the process of multiple cartographic productions such as those proposed by Abraham (1988) and McMaster (1989), but they were inappropriately based on the concept of generalisation, although the work by Abraham does include data reduction. Digital algorithms with such a property (i.e., hierarchical) are called global (see Chapter 3), e.g., the Douglas-Poiker algorithm. Whilst some authors have understandably argued that points along generalised lines are not always a subset of the original in traditional practice, this property is important according to the digital concept upon which digital generalisation processes should be to process line features. In Chapter 5, the Douglas-Poiker algorithm was shown to be capable of critical points selection. For this reason, the algorithm is recommended for use as a digital method that meets the requirement of scale-dependent generation of multiple cartographic product from single databases. Having established this basic requirement, there is a need for an objective method as to how multiple production can be achieved.

The third and final consideration is also related to the digital mode to which the process of digital generalisation has to adhere; that is the sequential approach of the digital mode. Generalisation processes can be modelled as a sequence of digital algorithms which can be integrated to produce the required cartographic product. Given the fact that cartographic line simplification should incur a change in both the shape and data of the line, depending on the line complexity, it is necessary to consider how such a result might be realised in the digital mode? It is, therefore, proposed that two principal digital operations can be integrated to perform a single process; namely, a simplification algorithm that deals with the selection of critical points, and a smoothing algorithm which primarily deals with the aesthetic aspects of the resulting shape. The Douglas-Poiker algorithm is therefore used to perform the first task, and the Cubic Spline smoothing algorithm to be utilised for the second task. Although these two types of digital routines have already been suggested and used, there has been no proposal as to how they or other algorithms might be utilised *objectively* and in conformity with *digital cartographic* simplification, and *digital cartographic* generalisation (c.f., section 6.2). Given the sequencing in the digital process, the combination of the two digital routines should, therefore, achieve a *minimum manipulation* of line feature details. This is obviously due to the fact that further transformation or manipulation of line details, specifically larger ones, should be another task that has to take place during further application of other generalisation operations such as exaggeration, displacement, or symbolisation. Based on the above three observations, a model is proposed by which the definition of this minimum manipulation is further discussed in the next sections.

*6.3.2 Design objectives and quality requirements of the simplification model:*

Design objectives of the proposed model can therefore be summarised as follows:

1) Producing multiple, repeatable, and consistent cartographic line simplification consistent with the concept and practice of digital generalisation for a typical topographic and thematic mapping.

2) Unlike the traditional processing of line details during simplification, the proposed simplification should aim for error reduction and consistent simplification.

3) Producing an interactive system with which the user can interact and decide when and what to simplify among a given set of line features during cartographic production.

4) Providing an empirical framework to test the proposed simplification in the realm of digital cartographic production, so that knowledge about line simplification can be attained.

6.3.2.1 Quality measures:

The difficulty faced in the development of this simplification model is that no recommended quality measures are available. However, this was not seen as a hindrance, since research has to provide proposals and search for solutions. This is so because, like any established measures in cartography, solutions are first proposed and subsequently proved useful through testing and experience. Similarly, it is suggested in this study that once the results of the proposed model are approved visually through extensive testing, the formulation itself is therefore regarded to be a measure by which this and other products can be assessed. In this respect, the formulation process is concerned with the *minimum elimination* of minor line details, and mathematically states how much elimination occurs. The assessment is based on qualitative measures; i.e., of perception.

In a digital line generalisation, there has to be a sequence starting with selection, and simplification of some minor details. This process is, therefore, referred to, here, as *preliminary simplification*. This is due to the observations outlined above, that during a generalisation process manipulating feature details is a context dependent process; hence the remaining details have to be processed accordingly, whereas this preliminary process serves as a digital method for selecting of critical points and is accompanied by some elimination of minor details. The small details that have to be removed at this stage are those that are regarded as redundant from both cartographic and economic perspectives; cartographically, they do not bear information value, and economically, they demand large computer resources for processing.

Due to the complexity of generalisation, it is reasonable to assume that further removal of feature details might well be needed. It is, therefore, expected that simplification of different processes might be required using different types of algorithms and parameters. However, the two simplification algorithms referred to above are considered to be appropriate for this first stage. Thus, it is necessary to note that the whole role of simplification at different scales for various mapping purposes is yet to be explored. Regarding the question of what determines a good or acceptable preliminary simplification is judged against the proposed requirements outlined below, which are essentially based on requirements of cartographic <u>perception</u> . They are as follows:

IF the line feature has been selected and has complex details (perceptually judged by the cartographer), and the difference between original and derived scale is equal or larger than 50 %;

THEN minimal removal of line details is required.

And / or

IF the line feature has been selected and has rather smooth or small details
or, the details are deemed important according to the mapping context (e.g., thematic mapping);

THEN no perceptible simplification is required, only a weeding process is to be performed. (This process is referred to as imperceptible or database simplification, and is applied at any target reduced scale).

Thus, there are two types of simplification: *perceptible simplification* which entails a perceptual change in the feature details, most importantly at large and medium scales, and *imperceptible simplification* which does not incur a change. The model is referred to as Preliminary Cartographic Line Simplification (PCLS). Its formulation and implementation are further illustrated in Figures 6.4 and 6.5 below.

6.3.2.2 Preliminary Cartographic Line Simplification (PCLS):

Figure 6.4 graphically illustrates the first type of simplification according to the design objectives of the model using the two simplification algorithms mentioned above. The two algorithms are included in the Arc/Info GIS, where the model was implemented. The figure illustrates how the two digital routines can be utilised to achieve the two simplification types. The figure schematically shows the Douglas-Poiker algorithm being used first for selecting the critical points in the line, employing large tolerance values (a), followed by the smoothing process where the Cubic Spline smoothing routine is applied on the selected points of the line in (b). The last process involves the use of the Douglas-Poiker algorithm again but only for weeding purposes, in which most of the newly added points are removed, using small tolerance values (c). Given the effects of graphic reduction, it is anticipated that the overall simplification might well be imperceptible at the target scale, especially at small scales. Such an effect is, however, implied and valid according to the requirements of the model, since the result will be read by other digital generalisation processes which are concerned with processing the remaining detail.

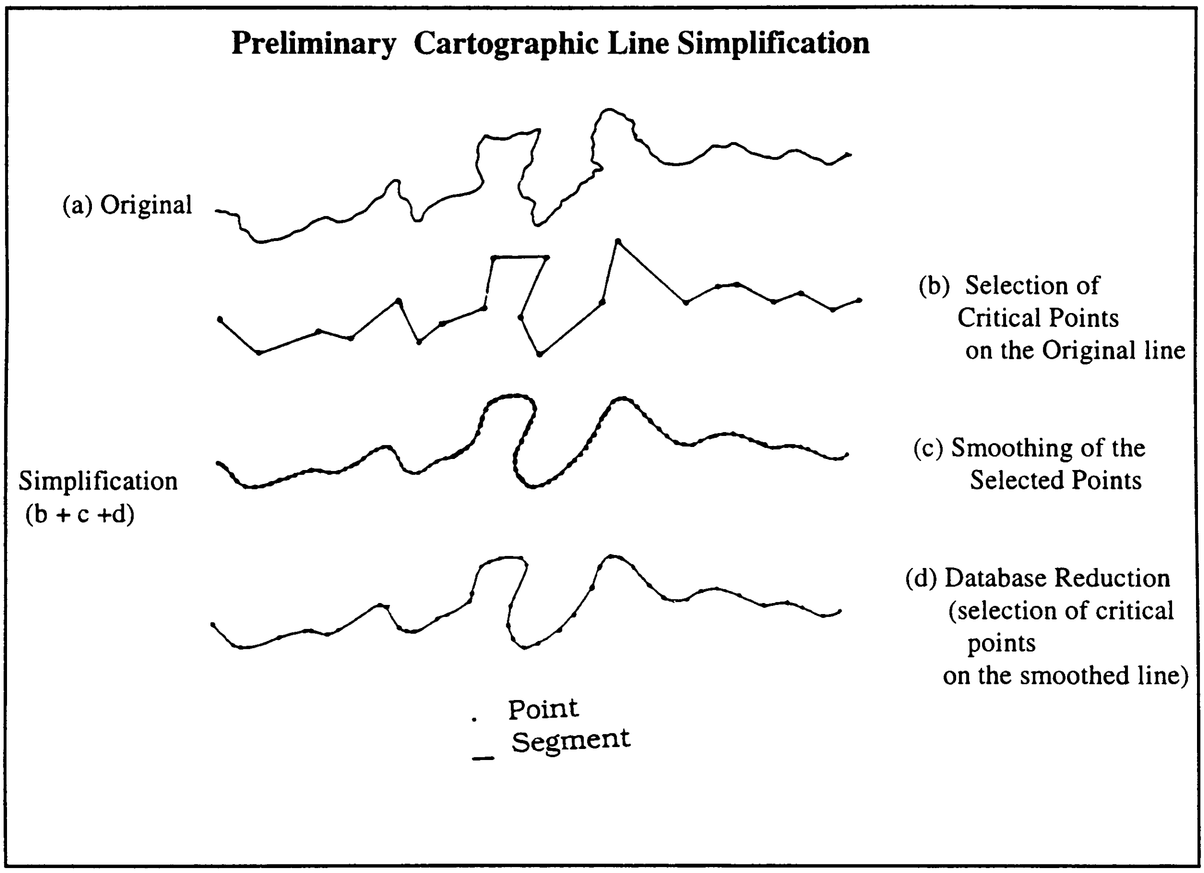



Figure 6.4: Implementation of the proposed simplification model (PCLS).

In order to meet the objectives of minimum manipulation defined for the model to be operational, three formulae are proposed for the three stages shown in Figure 6.4. As can be seen from that figure, the first stage is concerned with selecting critical points of the line so that major details and some of its sinuosity are retained. Given the success of the Douglas-Poiker algorithm in selecting critical points, based on the formula (TPF) proposed in Chapter 5 for multiple data reduction, the algorithm, as explained in the previous section, is utilised for this task. As pointed out in Chapter 5, the TPF was designed to produce tolerance values in accordance with: original data, difference between original and derived scale, and the requirements of the process of weeding (reduction) or simplification of the original database. Because larger tolerance values will produce distorted shapes or produce unwarranted caricatural shapes, this formula was chosen as a first step, but in order to produce perceptible simplification in the line details, lines will require a further process.

The second stage in Figure 6.4 shows one further manipulation of the resulting shape, using the Cubic Spline routine. The focus, now, is how to perform a multiple smoothing process consistent with the results of the previous stage, and with the requirement of incurring some perceptible reduction of detail, if needed, at the target scale, especially at large and medium scales (e.g., 1:50,000 to 1:150,000). Both algorithms are supported by Arc/Info, but their implementation within the system has to be understood, before applying them as a single integrated process. Whilst the first step has already been discussed, the second process of Cubic Spline smoothing, described in Arc/Info as a Cubic Polynomial Spline, is designed to smooth the shape of arcs using a user-defined tolerance. The tolerance here is in the form of an inter-vertex distance (i.e., the distance between points), a parameter input by the user. Figure 6.5 shows that arcs will be smoother and have more vertices if the tolerance is smaller than the current distance between vertices (b), whereas arcs will be more "generalised" if the tolerance is larger than the current distance.

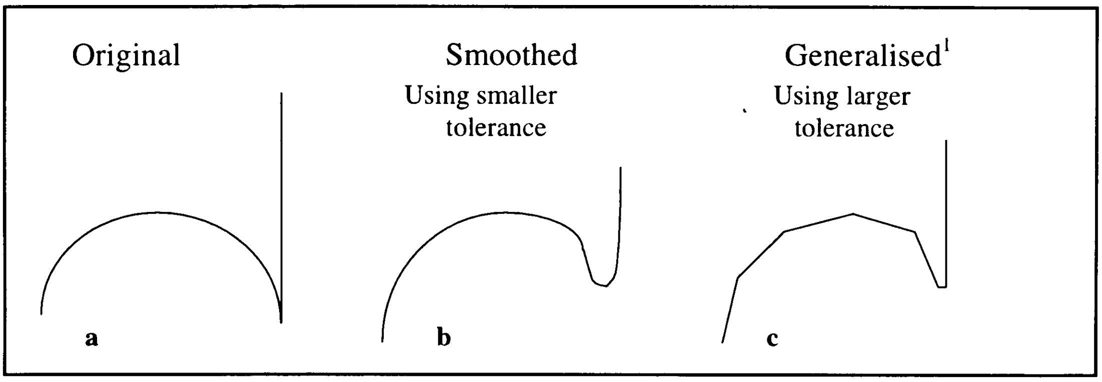



Figure 6.5: Implementation of the Cubic Polynomial Spline smoothing routine as shown in the Arc/Info Manual (source: ESRI 1982-1996). ([1] Note the misuse of terms as shown in the ESRI document).

In order to produce an input tolerance for the Cubic Spline function that can be used directly to smooth the line resulting from the previous step, it is considered that the following variables should be considered: the number of segments in the line, the average segment length, and the ratio between the number of segments in the original and simplified line resulting from the first step. Theoretically, in order to add points on a simplified line so that the undesirable effects such as spikes can be smoothed out, it would be logical to at least add points equivalent in the number to those of the original line. It is, in fact, a simple procedure of subdividing a simplified line in relation to the subdivision in its original, and since both lines are defined in terms of segments is simply related to the points of the lines (number of points = number of segments + 1) it is, therefore, an easy task to add points on the simplified line in relation to the number of points or segments on the original. For this purpose, and in order to provide a general result, the average segment length of the simplified line is calculated and then subdivided in relation to the ratio between the number of segments on the simplified and original lines. Accordingly, and with reference to Figure 6.5, the resulting value is used as the vertex distance that can directly be applied to the already simplified line This method should render the simplified line smoother; and hence aesthetically pleasing, as can be seen in Figure 6.4 (c). Initial testing proved that such a method was valid, since a smaller number of points tended to produce unsatisfactory results such as distorted or spiky line shapes, whereas a larger number of points retained redundant information. Thus, in order to provide a consistent relationship between the first two stages in the simplification model implemented within the same GIS system (Arc/Info), the second formula should provide direct input parameters in the form of *vertex distances* (VDs). The formula is, therefore, expressed as:

$$VD = SL * \left(\frac{NSd}{NSs}\right) \qquad \textbf{(6.1)}$$

where

**VD** = vertex distance (as input parameter)
**SL** = average segment length of the simplified line from the previous stage
(i.e., resulted by the Douglas-Poiker algorithm)
= total line length / number of segments
**NS** = number of segments of original or source line feature (s), and on the simplified (d).
(number of segment = number of points - 1)

Since further manipulation (though minimal) of minor details is required by the model, Equation (6.1) requires further consideration. Given the implementation of the Cubic Spline routine within Arc/Info, this manipulation is in the form of a further reduction and exaggeration. So, in order to enable a noticeable simplification, the resultant parameter from Equation (6.1) has to be *slightly* increased, otherwise line details will be oversimplified if the increase was large. Such an increase is thought of as a factor by which the parameter is constantly and modestly increased.

| Original Databases | Scale Reduction (original scale 1:25,000) | VD (m) | Log(VD) | D. VD (m) | PVD = (log(VD) * VD) (m) |
|---|---|---|---|---|---|
| 8717 | | | | | |
| | 2 (1:50,000) | 26.72 | 1.426 | 35 | *38.12* |
| | 4 (1:100,000) | 26.37 | 1.421 | 43 | *37.47* |
| | 6 (1:150,000) | 26.12 | 1.416 | 45 | *37.01* |
| | 10 (1:250,000) | 25.76 | 1.411 | 44 | *36.34* |
| | 25 (1:625,000) | 25.03 | 1.398 | 42 | *35.00* |
| | 40(1:1,000,000) | 24.66 | 1.391 | 43 | *34.32* |
| 493 | | | | | |
| | 2 | 35.93 | 1.555 | 48 | *55.89* |
| | 4 | 35.77 | 1.553 | 47 | *55.57* |
| | 6 | 35.66 | 1.552 | 50 | *55.35* |
| | 10 | 35.41 | 1.549 | 55 | *54.85* |
| | 25 | 34.76 | 1.541 | 44 | *53.57* |
| | 40 | 34.16 | 1.533 | 45 | *52.39* |
| 165 | | | | | |
| | 2 | 38.34 | 1.583 | 55 | *60.72* |
| | 4 | 38.2 | 1.582 | 57 | *60.43* |
| | 6 | 38.12 | 1.581 | 65 | *60.27* |
| | 10 | 37.97 | 1.579 | 65 | *59.97* |
| | 25 | 37.25 | 1.571 | 63 | *58.52* |
| | 40 | 36.39 | 1.561 | 67 | *56.80* |
| | | | **Overall average** : 1.511 | | |

**Linear Relationship between VD and DVD for all databases:**
Intercept (a) at 0:
y (**DVD**) = 1.535 * **VD** $R^2$ = 0 56

Intercept at 6.63 (default):
y (**DVD**) = 1.339 * **VD** + 6.63 $R^2$ = 0.57

**Linear Relationship between DVD and PVD for all databases:**
Intercept at 0:
y (**PVD**) = 0.98 * **DVD** ≈ y (**PVD**) = DVD
$R^2$ = 0.56
Intercept at 8.09 (default)
y (**DVD**) = 0.82 * **DVD**
$R^2$ = 0.58

**Table 6.1: Relationships between Vertex Distance (VD) values and Desirable VD (DVD) values for smoothing the resulting simplifications from three original databases (8717, 493, and 165 points), and Predicted VD values.**

Table 6.1 shows the results of experiments on the effect of increasing VD in order to achieve the desirable simplification suggested above. The databases are: the original database of study area 2 -drainage coverage: 8717 points, and two other smaller databases selected from this database: 493 points and 165 points. These three databases are of varying complexity and were subjected to various levels of simplification (upto 40 fold reduction, given the original scale as 1:25,000) corresponding to representative scales. For each database and for each level of simplification, the first stage in the simplification (applying the Douglas-Poiker algorithm) was performed followed by application of the Cubic Spline algorithm experimenting with different Vertex Distance (VD) values until a Desirable VD (DVD) was reached for producing desirable graphical results; i.e., desirable simplification. The value of DVD was determined through an iterative process, and each entry in the table required many iterations. The databases and VDs are in metre units. As the table indicates, the best relationship found between all the values of VD and DVD was linear ($R^2 = 0.56$) with the intercept being set to 0, indicating a positive, but not very significant, relationship. The equation of the linear relationship between the two variables (VD and DVD) for all the three databases provided a useful insight as to how it is possible to present a formula predicting new VD values which should, at least, approximate the Desirable VD values shown in the results. The linear equation of VD and DVD for all the databases is expressed as follows when the intercept (*a*) was set to 0:

$$y = b\text{VD} + a \quad (\text{i.e., } \mathbf{DVD} = 1.535 * \mathbf{VD}) \qquad \textbf{(6.2)}$$

Equation (6.2) shows that the value of *b* shows close agreement with the *log* (VD) for all the databases (Table 6.1), specifically compare the overall average of *log* (VD) (1.511) with the value of *b* in Equation (6.2). Equation (6.2), therefore, indicates that substituting *b* with *log* (VD) in the databases would generally provide the formula needed. Thus, the required formula providing the Predicted Vertex Distance (PVD) is expressed as:

$$\mathbf{PVD} = (\log \mathbf{VD}) * \mathbf{VD} \qquad \textbf{(6.3)}$$

The values resulting from application of Equation (6.3) are presented in the table (6.1). The appropriateness of this formula can be shown through examining the relationship between the DVD and PVD values for all databases. As Table 6.1 shows, the linear relationship between the two types of values indicates that the value of *b* is almost close to 1, especially when the intercept was set to 0. This is to imply that the PVD values are similar to DVD values; (PVD $\approx$ DVD).

With reference to Equation (6.1) and given the above discussion regarding the required increase in VD values as suggested by Equation (6.3), VD is, therefore, finally re-expressed as:

$$PVD = VD * (\log VD) \qquad \textbf{(6.4)}$$

where

***PVD*** = Vertex Distance (as input parameter)

$$VD = SL * \left(\frac{NSd}{NSs}\right)$$

SL = average segment length of the simplified line from the previous stage (i.e., resulted by the Douglas-Poiker algorithm)
= total line length / number of segments
(number of segment = number of points - 1)

NS = number of segments of original or source line feature (s), and number of segments of the simplified (d).

Since the second stage produces a great deal of redundant data (Figure 6.4 c), it is, therefore, essential to utilise the Douglas-Poiker algorithm as a third stage in the simplification for weeding purposes (Figure 6.4d). This requires production of an input parameter for multiple simplification of the resulting database from the second stage. In order to provide the required input tolerances, reference to Chapter 5 about the Douglas-Poiker algorithm as well as understanding of the nature the second stage in the simplification model are required. As explained in Chapter 5, the algorithm is influenced by the position of the points as opposed to their numbers. So, because the second stage in the model results in a large number of points most of which are located away from the most critical points (i.e., the positions of marked changes in the line) to which the algorithm is sensitive, small tolerance values would appear effective in removing most of such redundant data (Figure 6.4 c). Although all the results in Chapter 5 contribute to the above fact about the algorithm, specific example can be seen as a more appropriate for reference here where minor changes of tolerance values at low level data reductions or simplifications incur large changes on the database (Table 5.3). This specific example demonstrates how the database was largely reduced by the first few tolerance values, as well as how slight changes in those values resulted in large reduction of data. Accordingly, the formula is simply and practically expressed as the function of the difference between the source and derived scale, and is, therefore, expressed as follows:

$$T_2 = j * \left(\frac{Sd}{Ss}\right) \qquad \textbf{(6.5)}$$

where

**$T_2$** = Input tolerance value
*j* = Constant normally equai to 1 m, when units in metre. Its value is dependent on measurement units.
**Sd, Ss** = scale of derived scale, and source scale, respectively (**S** is the scale denominator).

Given the results and discussion of the algorithm in Chapter 5, and the nature of the data reduction required at this stage, if the difference between the two scales is greater than 90 %; then *j* should be equal to 0.5. This requirement is designed to ensure production of constrained (smaller tolerance values), especially for small scales, in order to avoid shape distortion of the smoothed line. Experimenting with the line, for example, in Figure 6.6 suggested that if *j* was not reduced by at least half of its value at the simplification for the scale of 1:250,000, the simplified line would have resulted in a spiky shape.

Figure 6.6 illustrates how the three processes of the model were implemented. The original line experimented on here is the same as is used in Figure 5.7, which underwent four levels of simplification. These levels are all displayed at a constant scale (1:100,000), showing how the simplified lines resulting from all the three steps as in

Figure 6.4. As Figure 6.6 shows, the tolerance values (T1) determined by the first step (using TPF) are progressively increased. As discussed above, this first step in the simplification increasingly removes redundant points whilst selecting the critical points in the line. For example, the first level of simplification (first row of lines) shows that the simplified line required a tolerance of 7.87 m, whereas the fourth line, representing the highest level of simplification, required a larger tolerance (87.70 m). On the other hand, whilst the tolerance values (T1) are progressively increased, the predicted vertex distances (shown in the figure and later in the chapter as VD) for Cubic Spline are almost similar. The relationship between the two processes here indicates that while the first step in the simplification (applying the Douglas-Poiker algorithm) increasingly removes points, the second step (applying the Cubic Spline routine) smoothes out the resulting simplified line, through generation of new points. For example, the first level of simplification (first row of lines) shows that the simplified line required a tolerance value of 7.87 m and a vertex distance of 60.51 m, whereas the fourth line, representing the highest level of simplification, required a larger tolerance (87.70 m) but a vertex distance slightly smaller than the previous VD value (56.89 m). This relationship between the tow parameters ensures that removal of points by the first step in the simplification is matched by a reciprocal and systematic addition of points by the second step in the simplification (smoothing process) for each level of simplification. The relationship is based on the number of points in the original line, in order to enable a systematic calculation of points for each level of simplification. The consistency in the calculation is reflected in the similarity of VD values (only between 56.89 m and 60.51 m), indicating that applying the VD values on the resulting simplified line from the first stage in the simplification enables a systematic subdivision or segmentation of the line (thereby generating new points) for each level of simplification in relation to the original number of points of the line. It is important, however, to note that the actual number of points resulting immediately after applying the VD values for each simplification level is similar to the original number of points in this example; i.e., 165 points. The figure shows only the final number of points after applying the third stage of the simplification; that is the weeding process using Equation (6.4).

The VD values used in Figure 6.6 are those resulting from Equation (6.3). In order to appreciate the difference between the results from Equations (6.1) and (6.3), Equation (6.1) would produce smaller VD values than those by Equation (6.3) which when applied by the smoothing routine would only add more points on the simplified lines (i.e., more subdivisions of the segments of the simplified lines), but would not lead to the perceptible simplification required. On the other hand, VD values resulting from Equation (6.3) would produce larger values -as shown in Table 6.1- which according to the implementation of the smoothing routine in Figure 6.5, could cause a minimum perceptible simplification of line details at the target scales, as Figure 6.6 shows in the bottom row of line representations. That is, Equation (6.3) produces large *vertex distances*; hence less segments (or points) on the simplified line so that the simplified line will appear smoother, accompanied by a slight exaggeration of the retained details (due to changes in positions of the added points) and reduction of minor details. The figure, also, shows the effect of the third step; that is using the Douglas-Poiker algorithm for weeding purposes. Tolerance values, termed as T2 in the figure (Equation 6.4), were used for this purpose, after the lines had been smoothed. For example, in the first simplification, the T2 was 4.00 m, whereas it became larger as the level of simplification increased.

The figure also reveals that the resulting line shapes from the simplification appeared consistent with the requirement of minimum manipulation of minor details. Hence, the approach successfully produced minimal multiple reductions of line details from a single detailed database. Whilst this example served as a preliminary illustration of the implementation of the proposed simplification, full testing of this implementation on sets of line features is presented in the remainder of this chapter.

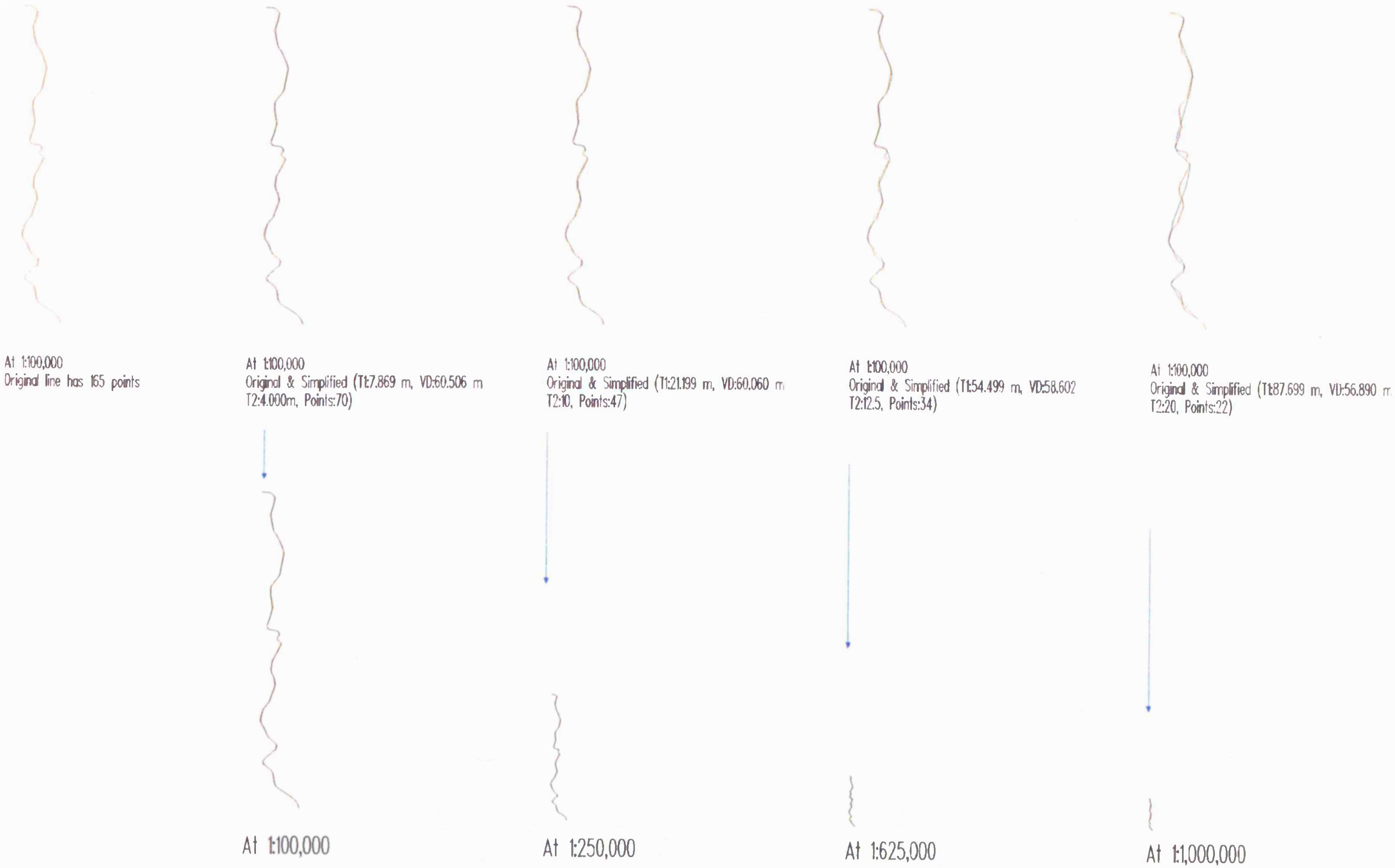



Figure 6.6: Implementation of the proposed formulation of the integration between the Douglas-Poiker algorithm and the Cubic Spline algorithm within the simplification model.

Figure 6.7 illustrates the complete design construction of the model (PCLS). The sequence of processes is assumed to be used in an interactive mapping process, and by an expert cartographer. For this purpose, a program was written within the Arc/Info system for implementing the processes of feature selection and simplification process. Details of the program are given in Appendix A. The program in the Arc Macro language (AML) and allows interaction with the user (i.e., cartographer) through pull-down menus and graphical interfaces (Appendix A.3). The user is first prompted to choose the database and input the original target scale. Through this interface, the user's attention is directed by the READ FIRST button which activates a pull-down screen explaining guidelines as to what and when to simplify. According to these guidelines (as explained in section 6.3.2.1), the user has to exercise his/her own cartographic sense and experience as to whether the lines to be processed do require either shape/data or database simplification. The interfaces to either of the two simplifications prompt the user to select either the whole or a selection of the lines for processing. Once the choice has been made, the user has to click on a button activating the automatic simplification. The results can be viewed and compared with the original at any scale defined interactively by the user, and finally the results can be saved. The approach illustrated in Figure 6.7 assumes that the digital data file is already prepared; that is the file is free from digitising errors, and redundant digitising points. The significance of ensuring that redundant points are removed, is that the process of simplification can work on the remaining points which are assumed to be the most critical points that capture the character of the line feature at the source level of resolution. Otherwise, the process will work on redundant data that are regarded as a waste of computing resources, and hence, produce meaningless results (see Chapter 5). The following sections are dedicated to the results and discussion of the application of the PCLS on different sets of digital lines.

**Preliminary Digital line Simplification Model**
**For Cartographic Production**

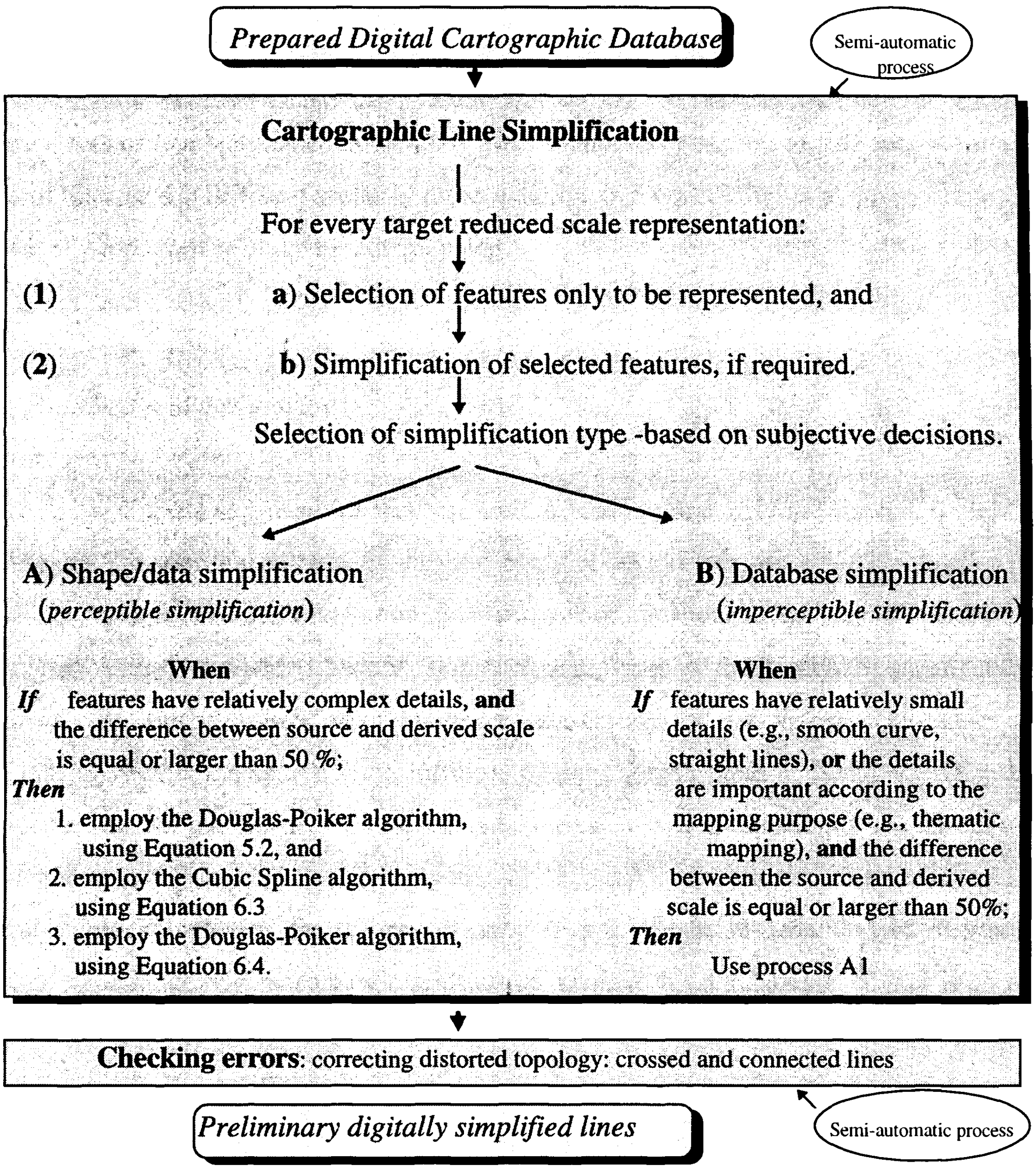



Figure 6.7: Complete implementation steps of the proposed simplification model (Preliminary Cartographic Line Simplification: PCLS) for a typical cartographic production process.

## 6.4 Tests:

The following sections describe tests designed to evaluate the quality of the proposed line simplification. This is to examine whether the application of the model on different sets of line features could actually yield minimum reduction of line details for all types of lines, and at different scales of representation, in a typical digital mapping context, specifically topographic and thematic mapping.

### *6.4.1 Data sets and methodology:*

Data sets used for this test are those used previously in Chapters 4 and 5, specifically, the source databases of the four study areas (see section 4.1). They differ in complexity and detail sufficiently enough to support the purpose of the test. Considering their previous use, the added benefit from their use, here, is familiarity. Being digitally prepared, the four data sets meet the prerequisite stage in the model. The scale of 1:25,000 represents the source scales of three of them (study areas 1, 2 and 4), whereas the source scale of the fourth is 1:158,400 (study area 3). The application of the model on these area tests will be as follows: study areas 1, 2 and 4, will be tested in a topographic mapping context, whereas the testing of study area 3 will be in a context of thematic mapping. The range of representation scales is the same as used in Chapter 5; that is up to the scale of 1:1,500,000 (60 fold reduction). Such a range is thought to be large enough that the resulting simplifications can be extensively assessed. The process of feature selection for different scales within this range is performed accordingly. Although it may not be the ideal choice, it serves as a typical exercise during digital cartographic production. Throughout the analysis, data/shape simplification is referred to as the *first simplification*, whereas database simplification is referred to as the *second simplification*. The focus of the analysis will be perceptual changes at each scale, based on the requirement of the model. Results of each test are first presented, and discussion of all follows. Since the proposed simplification deals with minor details of line features, it is essential to exercise a close visual examination of the simplified lines.

*6.4.2* Study area 1 (44 arcs, 4056 points):

Table 6.1 shows the numerical results of the simplification of study area 1, reporting the process of feature selection, tolerance values, number of points, and per cent data reduction. For the 1:50,000 scale, 38 arcs were selected for both types of simplification, whilst the remaining arcs were either deleted or underwent neither of the simplifications, such as the railway (see original map in section 4.1). At this scale, 37 arcs (with 3563 points) underwent the first simplification (data/shape simplification) resulting in 1274 points, a 64.25 % data reduction from the relative database (3563 points), 70.30 % of the original database (4056 points). The remaining arc (424 points), representing a district boundary, underwent the second simplification (database simplification), resulting in 102 points, a 75.95 % data reduction from the relative database (424 points), 97.14 % data reduction of the original (4056 points). Both simplifications resulted in a 66.08 % data reduction from the original database (4056 points). Figures 6.9 to 6.11 graphically illustrate the effects of the simplifications at this scale (1:50,000). Figures 6.9.1 and 6.9.2 show the effects of the first simplification. The overlay of the original (in Red) and simplified lines (in Green) is shown in Figures 6.10.1 and 6.10.2, whereas Figures 6.11.1 and 6.11.2 show the overlay of the original line and its simplified version using the second simplification. Figures 6.10.1 and 6.10.2 indicate that a minimal reduction resulted from the first simplification. Given the original line details, and the target scale, the result appears to be consistent with the model requirements, since further reduction would inappropriately render the lines much smoother. Some shape distortion occurred at four places (A1, 2, 3, and 4), and are all perceptible at the scale (Figure 6.10.1, 6.10.2 and 6.10.3). They were all caused by the smoothing process (i.e., the Cubic Spline algorithm) in which lines that closely lie within the specified tolerance (VD) are likely to connect. On the other hand, Figure 6.11 reveals an imperceptible difference between the original line and its simplified version, using the second simplification. This is a desirable effect, since the character of the line requires no further processing except simplification of its database.

The arcs that were selected for the 1:50,000 scale were, also, selected for representation at 1:100,000; i.e., 38 arcs (with 3987 points). The two simplifications were, also, applied on these arcs. That is, 37 arcs underwent data/shape simplification, resulting in 860 points, a 75.87 % data reduction from the relative database (3563 points), 79.50 % of the original (4056 points). The remaining arc (424 points) underwent simplification, resulting only in 53 points, an 87.50 % data reduction from the relative database (424 points), 98.52 % of the original (4056 points). Data reduction from both simplifications constitutes a 77.50 % reduction from the original database (4056 points). Figures 6.12 and 6.13 display the graphical effects of the two simplifications, respectively. Figure 6.12.1 shows the simplified (in Green) and original (in Red) lines. In the figure, the overlay of both maps (below) indicates a barely perceptible difference between the two, in spite of the 75.87 % data reduction. However, close visual examination reveals that the slight perceptible difference between the two maps is caused by the process of exaggeration more than by the reduction in minor details. This effect is cartographically desirable, since small details are either eliminated or exaggerated. Since smoothing out all or most of the details shown at the scale does not constitute acceptable simplification at this early stage, the simplification process is generally regarded as acceptable. Ideally, it would be desirable if minor details in the form of single dots with a size equal to the line symbol are removed. On the other hand, the simplified arc, using the second simplification, reveals an imperceptible difference from the original at the scale (Figure 6.13). This is a desirable and required effect. The first simplification produced topology distortion at five places, three of them are imperceptible (B1,2, and 3), and two are perceptible (A1, and 2) (Figures 6.12.1 and 6.12.2). They were all caused by the smoothing process. This is due to the same factors as discussed for the previous scale.

At the scale of 1:250,000, 18 arcs (with 3348 points) were selected for representation. Table 6.1 indicates that when the first simplification was applied to these arcs 396 points resulted, an 88.18 % data reduction from the relative database (3348 points), 90.24 % of the original (4056 points). As can be seen from Figures 6.14.1, the simplified lines (in Green) appear with some accentuating of the remaining minor details compared to the original (in Red). Many small details such as those shown at 1:50,000 were reduced, but due to the effect of graphic reduction this is imperceptible. Because the simplification did not cause total removal of line details, yielded an 88.18 % data reduction from the relative database, and was accompanied by a limited exaggeration of the retained details, the simplification at this scale is regarded as acceptable. However, there was some distortion in three places (B1, 2, and 3), and all are imperceptible at scale (Figures 6.14.1 and 6.14.2). The distortion at B2 and B3 (Figure 6.14.2) were caused by the Douglas-Poiker algorithm, whereas the distortion at B1 was caused by the smoothing process.

At the scale of 1:625,000, only 16 arcs (with 3133 points) were selected. First simplification was used, resulting in 265 points, a 91.55 % data reduction from the relative database (3133 points), 93.46 % of the original. Close visual examination of Figure 6.15.1 shows that the simplified lines (in Green) appear with some details being reduced, and with some being exaggerated (e.g., the island). Although there are small details still shown on the simplified lines, their removal, according to the model's requirements, is undesirable, even if they appear in the size of the line symbol at this scale. This is due to the fact that their presence at this small scale emphasises their geographical significance, in terms of magnitude, as they were not removed in spite of the large reduction of data (91.55 %). Although, it is difficult to discern perceptible changes, this is not a concern in the preliminary simplification, and so the result is regarded as acceptable. From these results, it is important to make two observations at small scales: the presence of small details at small scales indicates that they are significant in terms of size; and although a great deal of data reduction occurs most of it is hardly perceptible at the scale. There was some distortion of topology at three places, which are all imperceptible (Figure

6.15.2, (B1, B2, and B3)). Distortion at B3 was caused by the Douglas-Poiker algorithm, whereas the distortions at B1 and 2 were caused by the smoothing process.

At the scale of 1:1,000,000, only 10 arcs (with 3099 points) were selected. The first simplification produced 170 points, a 94.52 % data reduction of the relative database (3099 points), 95.81 % of the original (Table 6.2). Close visual examination of Figure 6.16 should reveal that the simplified lines (in Green) appears smoother compared to the original (in Red), accompanied by some exaggeration in some parts (e.g., the island). With reference to the model requirements and the discussions provided above for the previous scales and especially the last two observations for the scale of 1:625,000, the result of simplification for this scale is regarded as acceptable. However, shape distortion occurred at two places, but both are imperceptible, (Figures 6.16.1 and 6.16.2 (B1 and B2)). The distortion at B1 was caused by the smoothing process, whereas the second was caused by the Douglas-Poiker algorithm.

At the scale of 1:1,500,000, only 8 arcs (with 2491) were selected. They underwent the first simplification, resulting in 94 points, a 96.23 % data reduction from the relative database (2491 points), 97.68 % of the original (Table 6.2). In Figure 6.16.1, the simplified lines (in Green) shows a slight difference from the original (in Red). This small difference is largely due to the effect of graphic reduction (a 60 fold reduction), although there was a large reduction of data. Again, with reference to the previous results discussed for the previous scales, such an effect is consistent with the acceptance of the result as consistent with model requirements. Distortion of topology occurred at two places, and both are imperceptible at the scale (Figures 6.16.1 and 6.16.3, (B1) and (B2)). The first distortion was caused by the smoothing process, whereas the second was caused by the Douglas-Poiker algorithm.

| Maps after features Selection Process for different scales | T1 (m) | VD (m) | T2 (m) | Resulting Points | % Data reduction | |
|---|---|---|---|---|---|---|
| Original map before feature selection (**4056** points, **44** arcs) | | | | | from relative database | from original database (4056 points) |
| 1:50,000 **a:** (**3563** points, **37** arcs) | **6.10** | **33.61** | **2.00** | **1274** | **64.25** | **70.30** |
| 1:50,000 **b:** (**424** points, **1** arc) | **4.254** | _ | _ | **102** | **75.95** | **97.14** |
| **Total:** 38 arcs, 3987 points | | | | **1376** | | **66.08** |
| 1:100,000 **a:** (**3563** points, **37** arcs) | **13.20** | **32.96** | **4.00** | **860** | **75.87** | **79.50** |
| 1:100,000 **b:** (**424** points, **1** arc) | **9.51** | _ | _ | **53** | **87.50** | **98.52** |
| **Total:**38 arcs, 3987 points | | | | **913** | | **77.50** |
| 1:250,000 (**3348** points, **18** arcs) | **34.29** | **33.28** | **10.00** | **396** | **88.18** | **90.24** |
| 1:625,000 (**3133** points, **16** arcs) | **86.50** | **31.98** | **12.50** | **265** | **91.55** | **93.47** |
| 1:1,000,000 (**3099** points, **10** arcs) | **138.81** | **30.70** | **20.00** | **170** | **94.52** | **95.81** |
| 1:1,500,000 (**2491** points, **8** arcs) | **203.01** | **32.01** | **30.00** | **94** | **96.23** | **97.68** |

**Table 6.2: Digitally simplified maps from their 1:25,000 source databases (relative) after features have been selected from original (4056 points, 44 arcs -study area 1) according to the PCLS.**

Figure 6.8 shows the processes of selection and simplification effect in terms of data reduction from the original database (4056 points) at all the scales. The graph shows that the rate of data reduction decreases beyond the scale of 1:250,000. As can be seen from the figure, around 90 % of data reduction occurred within the first three scales. This is primarily due to the effect of the process of feature selection, and the effect of the Douglas-Poiker algorithm's behaviour (see Chapter 5), and is partly due to the fact that the resulting data reductions (both from the processes of simplification and feature selection) produced for the remaining scales (1:625,00, 1:1,000,000, and 1:1,500,000) represent a relatively small percentage range (90.24 - 97.68) in relation to the original database.

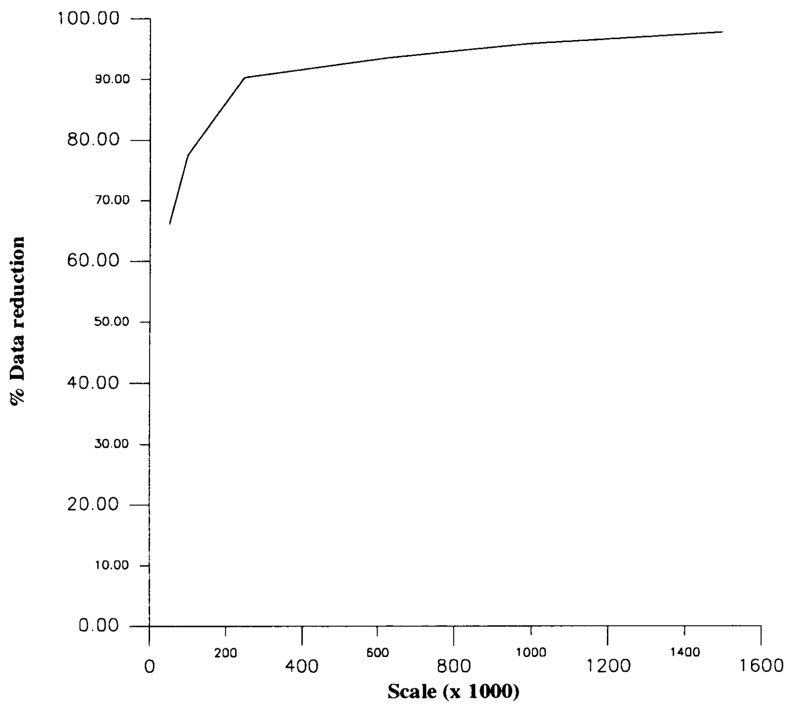



Figure 6.8: Scale against % Data reduction for digitally simplified maps (using the PCLS) from the original database (4056 points, 44 arcs -study area 1).

Given the requirements outlined for the model, stating that complete elimination of the line detail is regarded as unacceptable at this stage in the digital generalisation, and given the fact that the lines in this study area are relatively less detailed, the formulation proposed practically and generally proved acceptable. That is, some details were removed or reduced while some were slightly exaggerated. Additionally, the larger forms or details of the lines were not affected. However, some shape distortions were caused by both the algorithms, but in only a few places and were largely imperceptible at the target scales. Such a distortion has to be corrected if necessary before applying other generalisation processes. Thus, for each production scale the preliminary simplified lines are regarded as suitable and ready for further generalisation operations, and the retained details can be processed according to the mapping purpose, which is here topographic. Furthermore, and given the relatively smaller details within the lines in this study area, further removal of the retained details is not recommended, as otherwise the character of the original lines will be lost. In this respect, these details might well be further exaggerated but not removed.

Figure 6.9.1: Due to scale and paper size only half of the map is displayed, see Figure 6.9.2.
Simplified map at 1:50,000, 1274 points, 64.25 per cent data reduction from 3563 points database and,
70.30 per cent from 4056 points database. 1) Feature Selection, 2) T1= 6.10 m, 3) VD= 33.61 m, 4) T2= 2.00 m.

Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 1).

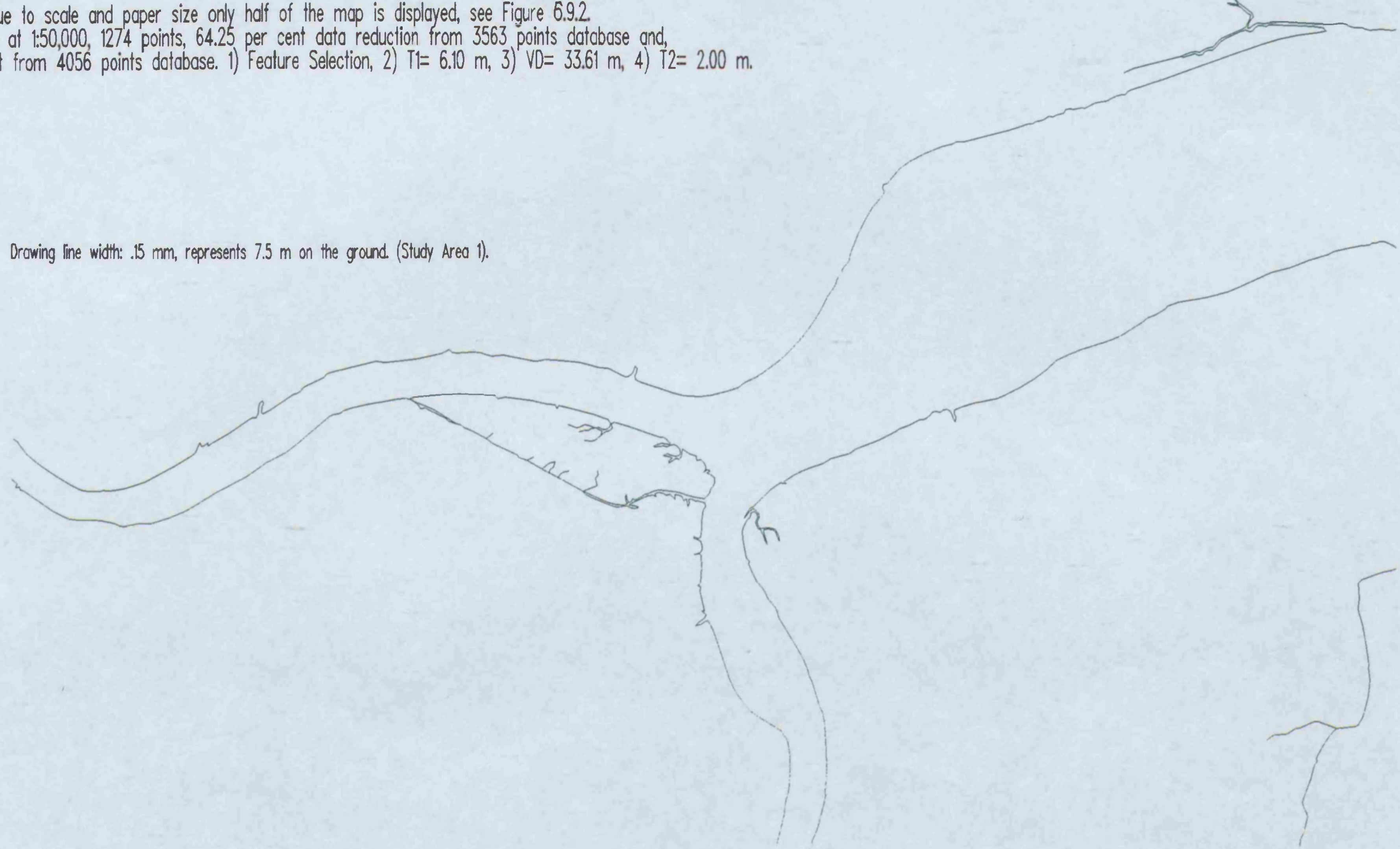


Figure 6.9.2: Due to scale and paper size only half of the map is displayed, see Figure 6.9.1.
Simplified map at 1:50,000, 1274 points, 64.25 per cent data reduction from 3563 points database and,
70.30 per cent from 4056 points database. 1) Feature Selection, 2) T1= 6.10 m, 3) VD= 33.61 m, 4) T2= 2.00 m.
Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 1).

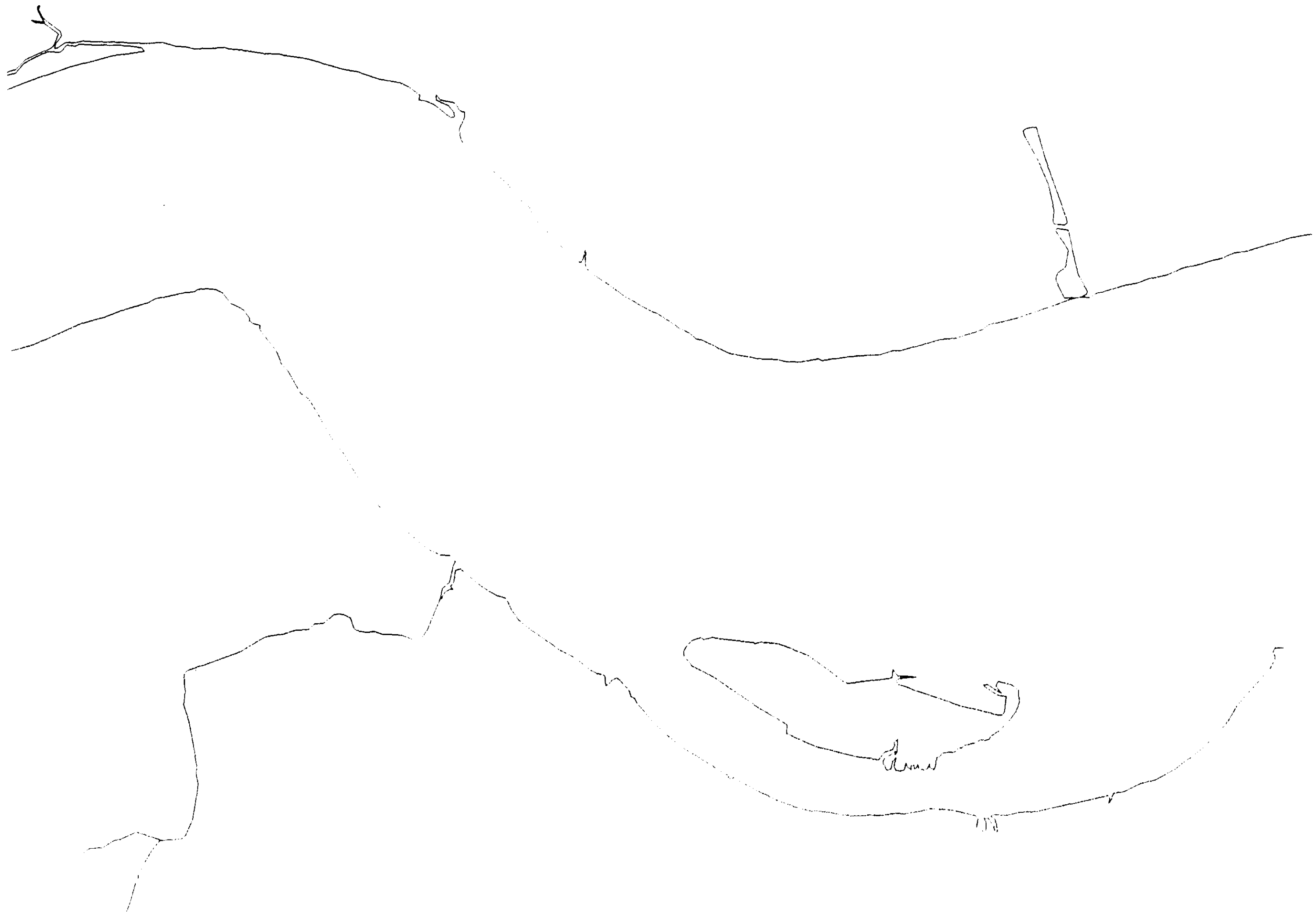


Figure 6.9.2: Due to scale and paper size only half of the map is displayed, see Figure 6.9.1.
Simplified map at 1:50,000, 1274 points, 64.25 per cent data reduction from 3563 points database and, 70.30 per cent from 4056 points database. 1) Feature Selection, 2) T1= 6.10 m, 3) VD= 33.61 m, 4) T2= 2.00 m.
Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 1).

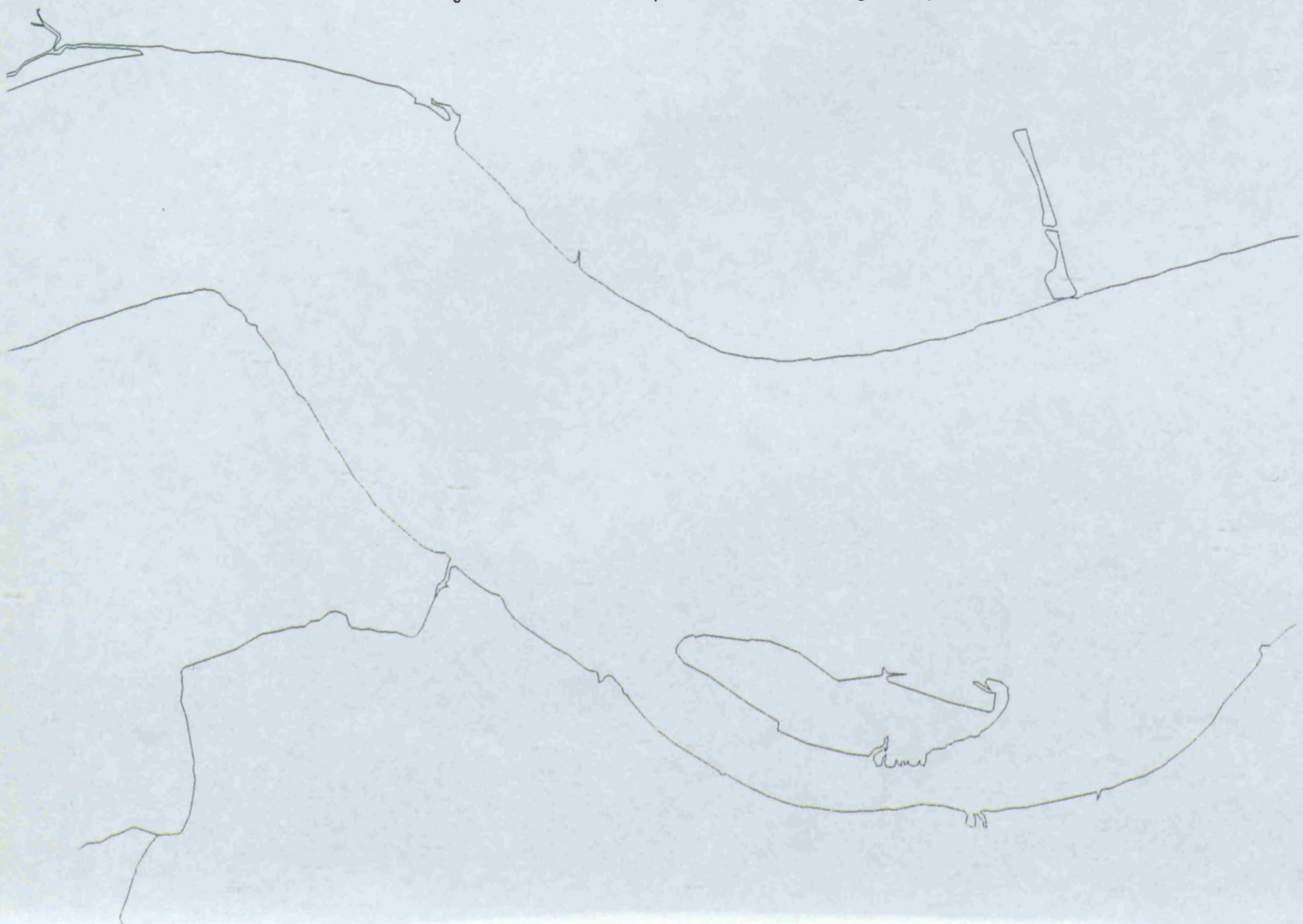


Figure 6.10.1: Two maps at 1:50,000. Original (in Red) after selection of the features that need to be simplified, with 3563 points and 37 arcs. Simplified map (in Green), 1274 points, 64.25 per cent data reduction from 3563 points database and, 70.30 per cent from 4056 points database. 1) Feature Selection, 2) T1= 6.10 m, 3) VD= 33.61 m, 4) T2= 2.00 m. (Study Area 1, see Figure 6.10.2 for the other half of the map).
Drawing line width: .15 mm, represents 7.5 m on the ground.

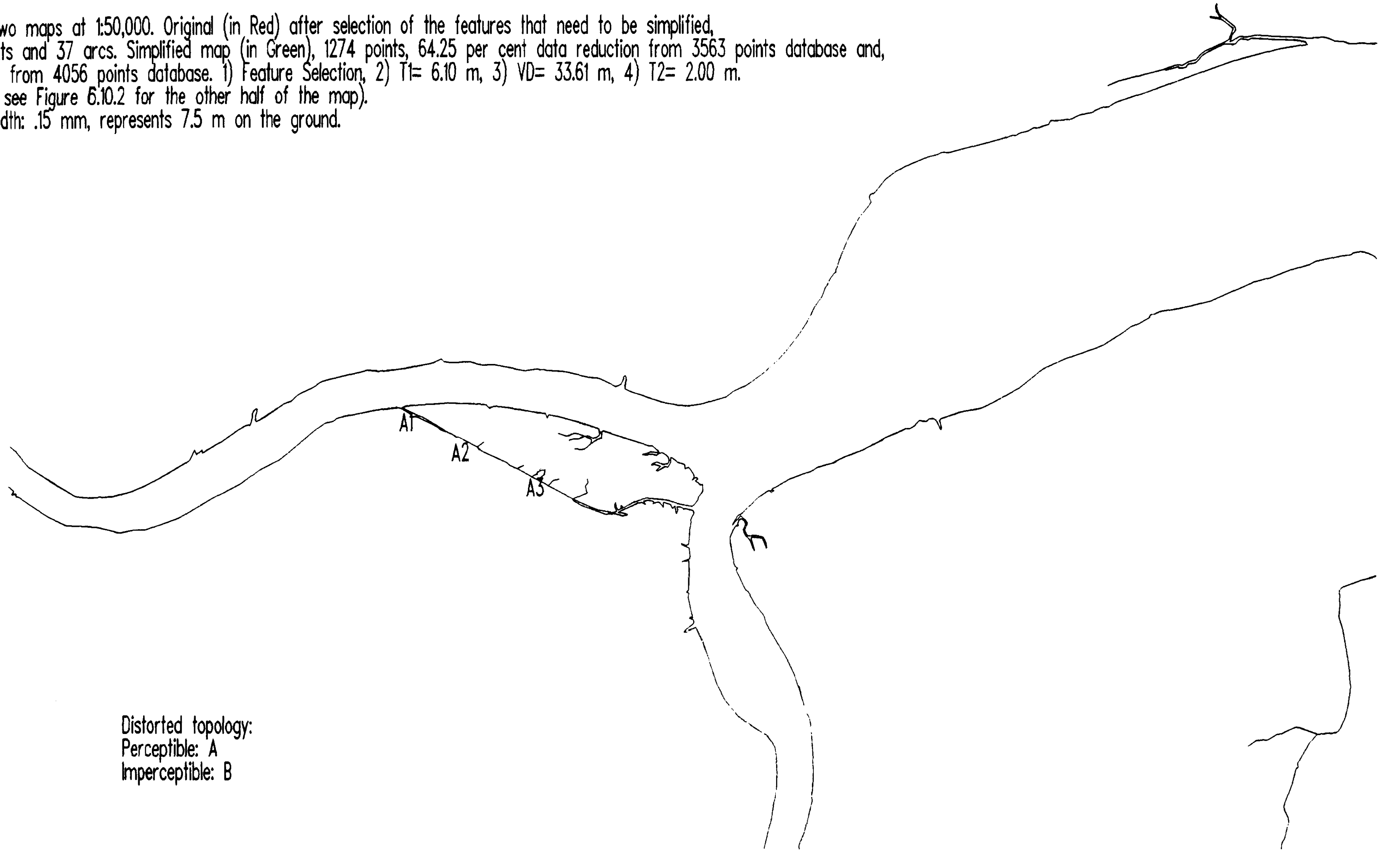
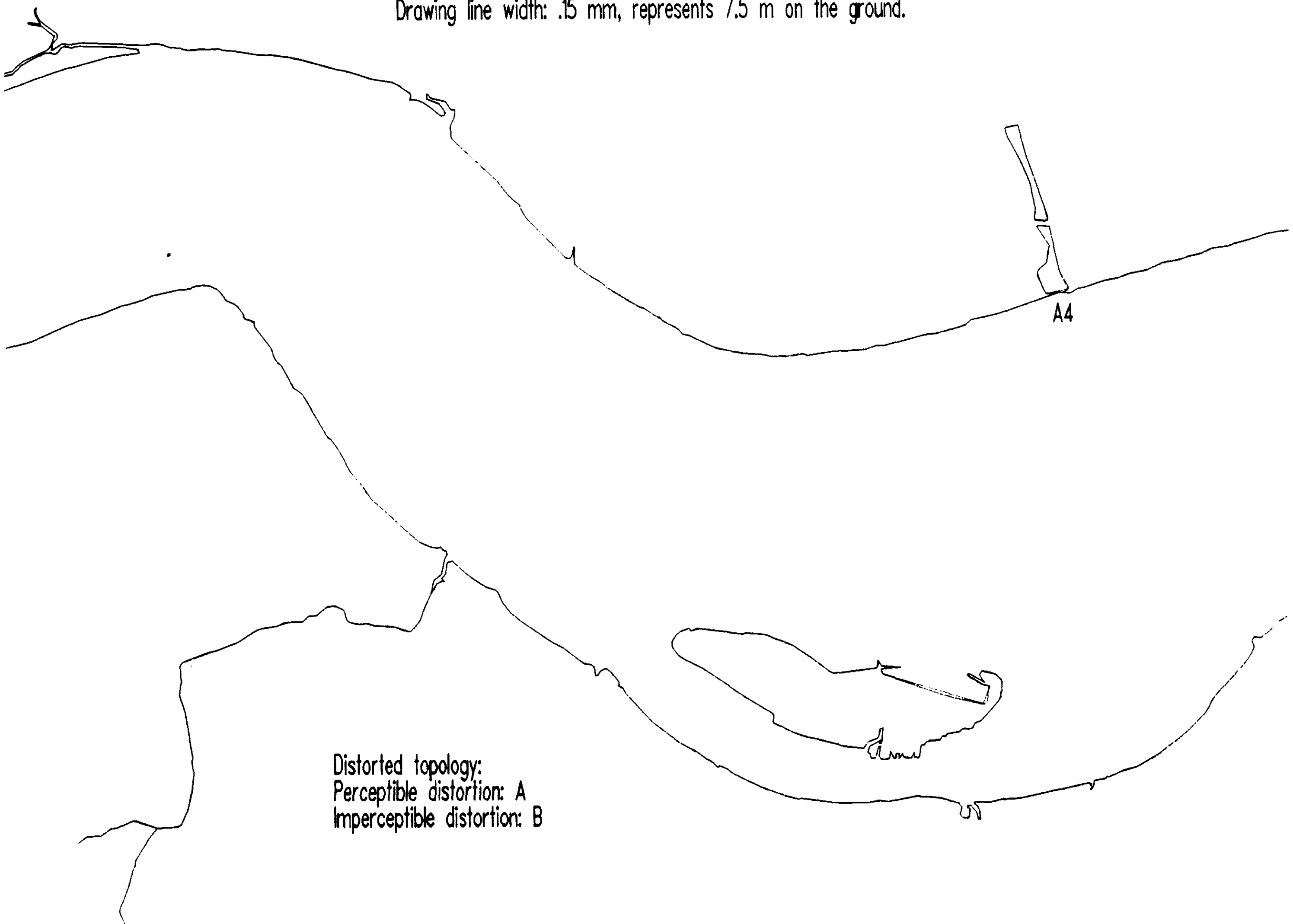


Figure 6.10.2: Two maps at 1:50,000. Original (in Red) after selection of the features that need to be simplified, with 3563 points and 37 arcs. Simplified map (in Green), 1274 points, 64.25 per cent data reduction from 3563 points database and, 70.30 per cent from 4056 points database. 1) Feature Selection, 2) T1= 6.10 m, 3) VD= 33.61 m, 4) T2= 2.00 m.
(Study Area 1, see Figure 6.10.1 for the other half of the map).
Drawing line width: .15 mm, represents 7.5 m on the ground.



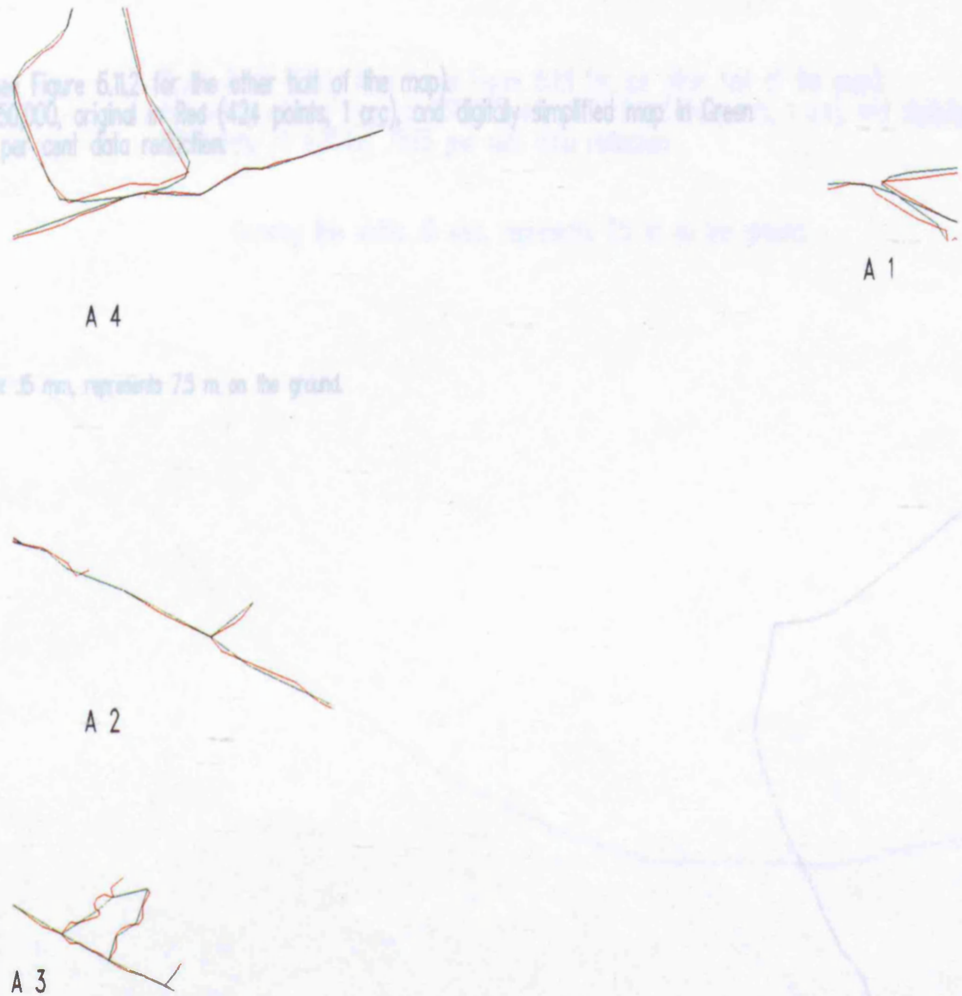



Figure 6.10.3: Distorted topology resulted from the simplification produced for the 1:50,000 scale
The distortion is perceptible (A) at the scale of 1:50,000, while it is enlarged in this figure to the scale 1:10,000.
Original in Red and digitally simplified map in Green.

Figure 6.11.1: (Study Area 1, see Figure 6.11.2 for the other half of the map).
An overlay of two maps at 1:50,000, original in Red (424 points, 1 arc), and digitally simplified map in Green
102 points, T1: 4.254m, 75.95 per cent data reduction.

Drawing line width: .15 mm, represents 7.5 m on the ground.

Figure 6.11.2: (Study Area 1, see Figure 6.11.1 for the other half of the map).
An overlay of two maps at 1:50,000, original in Red (424 points, 1 arc), and digitally simplified map in Green 102 points, T1: 4.254m, 75.95 per cent data reduction.

Drawing line width: .15 mm, represents 7.5 m on the ground.

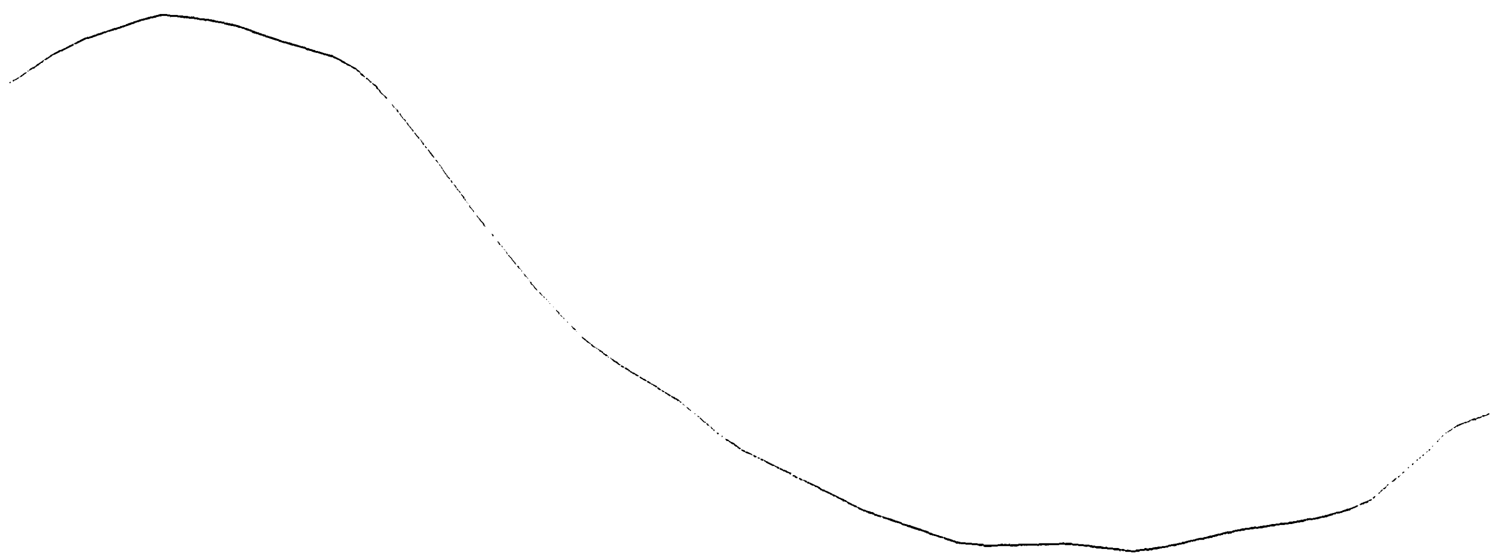


Figure 6.12.1: Two maps at 1:100,000. Original map in Red (3563 points), and the simplified map in Green (860 points, 1) Features Selection, 2) T1= 13.20 m, 3) VD= 32.96 m, 4) T2= 4.00 m. 75.25 per cent data reduction from 3563 points database and, 79.50 per cent from 4056 points database. (Study Area 1).

Drawing line width: .15 mm, represents 15 m.

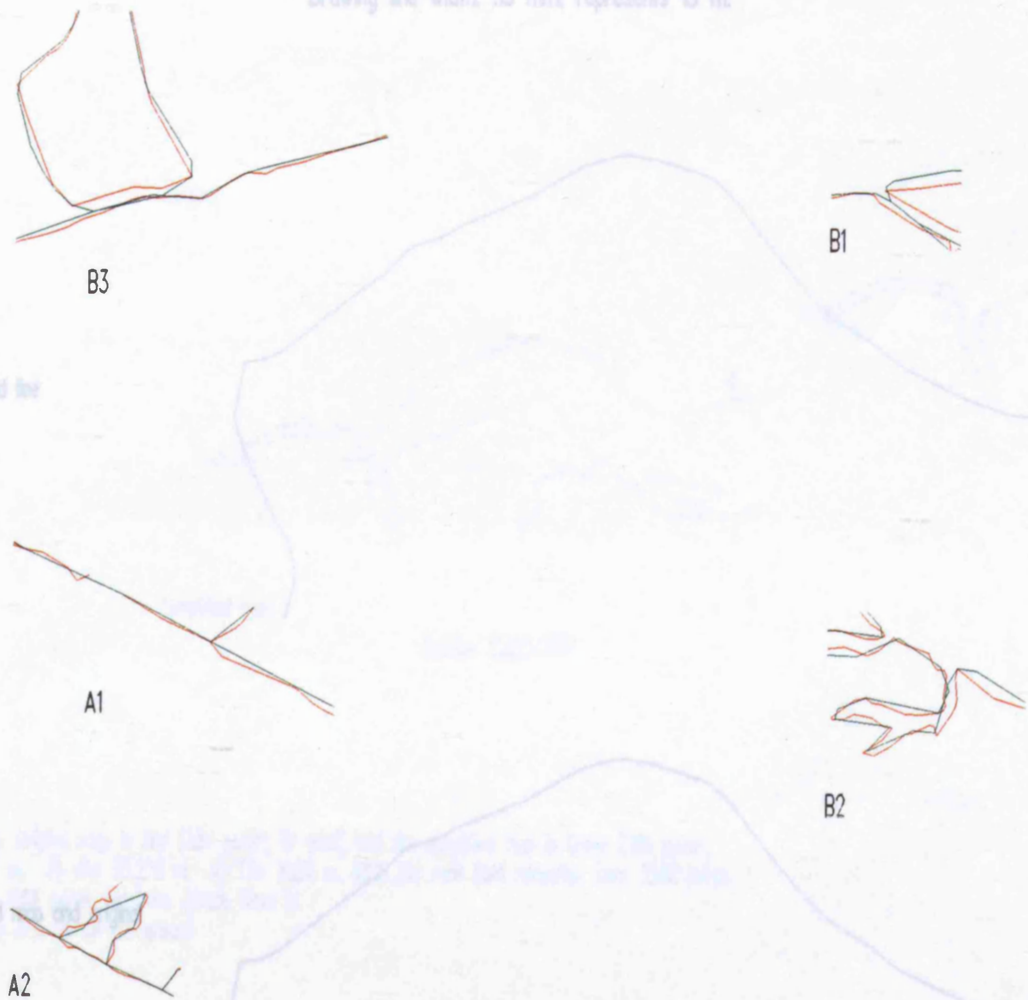
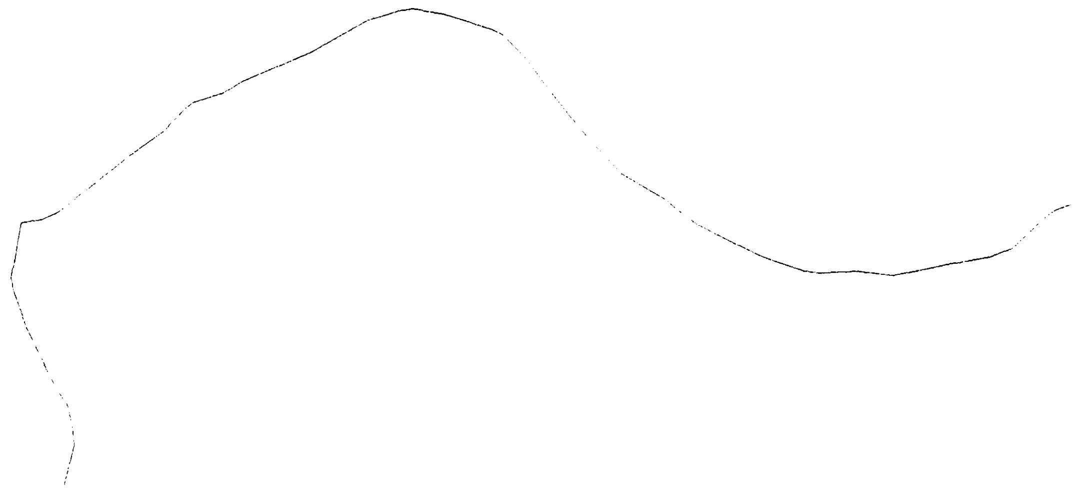


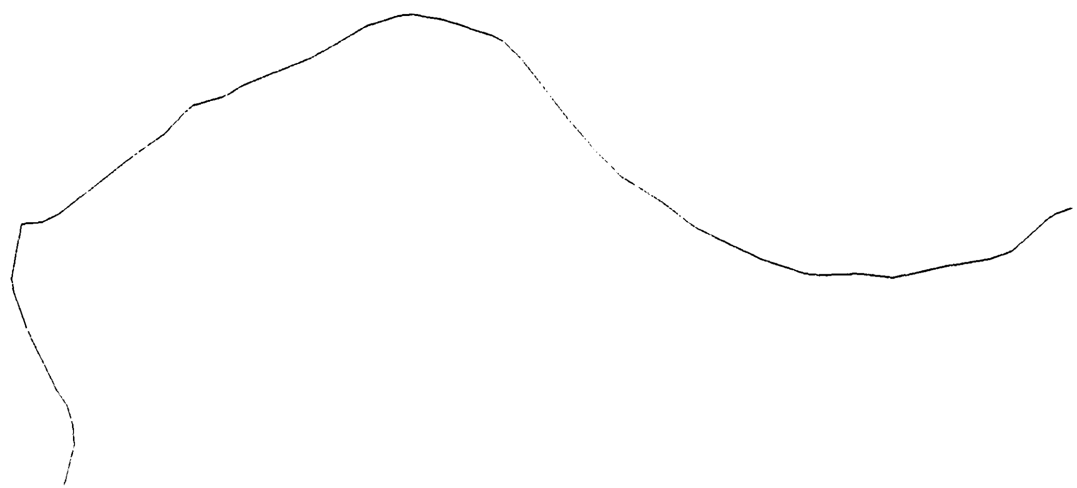

Figure 6.12.2: Distorted topology resulted from the simplification produced for the 1:100,000 scale
Some of the distortions are perceptible (A) and some are imperceptible (B) at the scale of 1:100,000
while they are enlarged in this figure to the scale of 1:10,000.
Original in Red and digitally simplified map in Green (see figure 6.12.1)

Figure 6.13: Two maps at 1:100,000. Original map in Red (424 points, 1 arc), and the simplified map in Green (53 points, T1: 9.509 m, 87.50 per cent data reduction. (Study Area 1).
Drawing line width: .15 mm, represents 15 m.

Simplified line



Simplified map and original



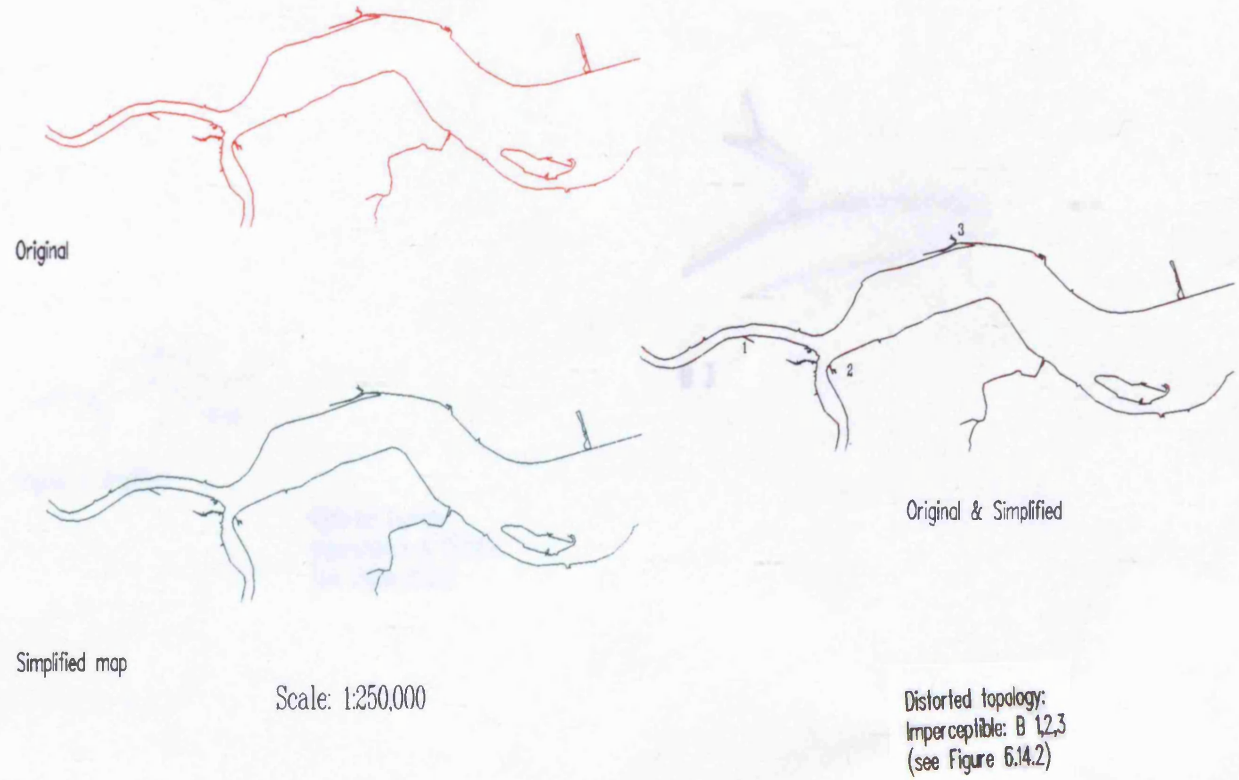



Figure 6.14.1: Two maps at 1:250,000. Original map in Red (334 points, 18 arcs), and the simplified map in Green (396 points, 1) Features Selection, 2) T1= 34.287 m, 3) VD= 33.278 m, 4) T2= 10.00 m. 88.18 per cent data reduction from 3348 points database and, 90.24 per cent from 4056 points database. (Study Area 1).
Drawing line width: .15mm, represents 37.5 m. on the ground

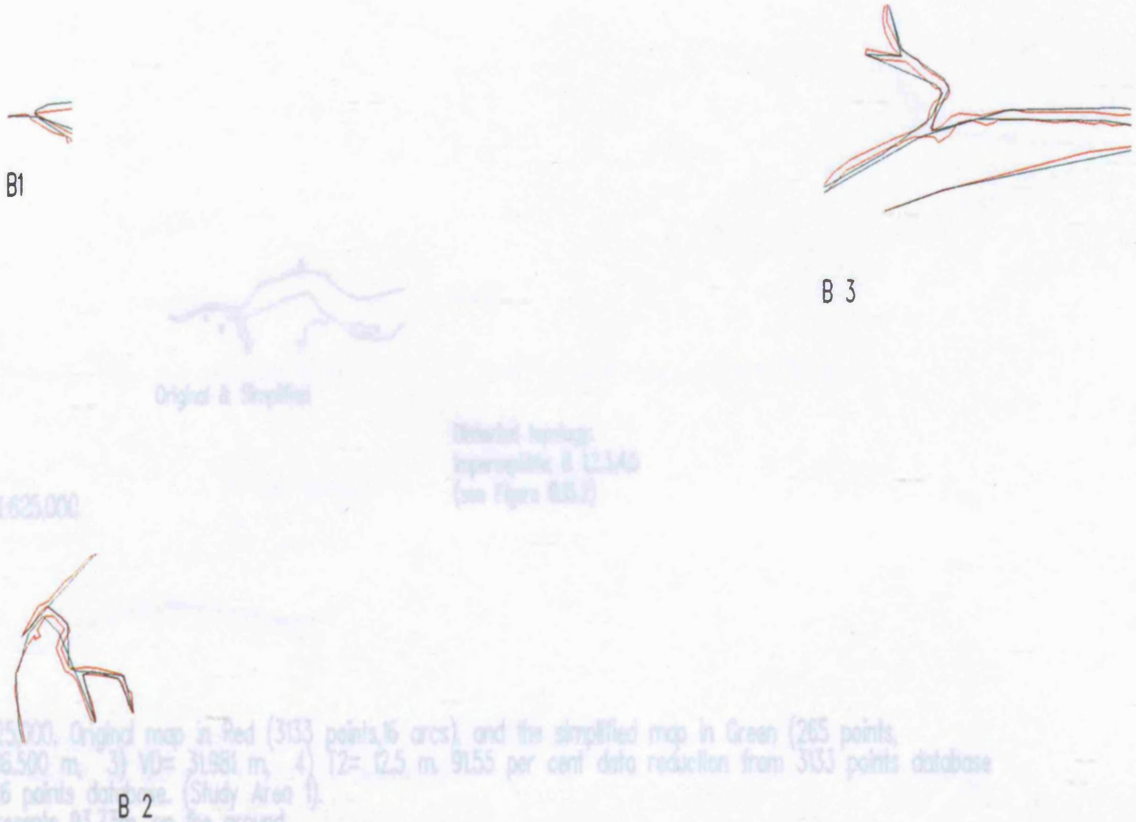



Figure 6.14.2: Distorted topology resulted from the simplification produced for the 1:250,000 scale
The distortion is imperceptible (B) at the scale of 1:250,000, while it is enlarged in this figure to the scale of 1:20,000.
Original in Red and the simplified in Green.


Original

Simplified

Original & Simplified

Distorted topology:
Imperceptible: B 1,2,3,4,5
(see Figure 6.15.2)

Scale 1:625,000

Figure 6.15.1: Two maps at 1:625,000. Original map in Red (3133 points,16 arcs), and the simplified map in Green (265 points, 1) Features Selection, 2) T1= 86.500 m, 3) VD= 31.981 m, 4) T2= 12.5 m. 91.55 per cent data reduction from 3133 points database and, 93.47 per cent from 4056 points database. (Study Area 1).
Drawing line width: .15mm, represents 93.73m. on the ground.

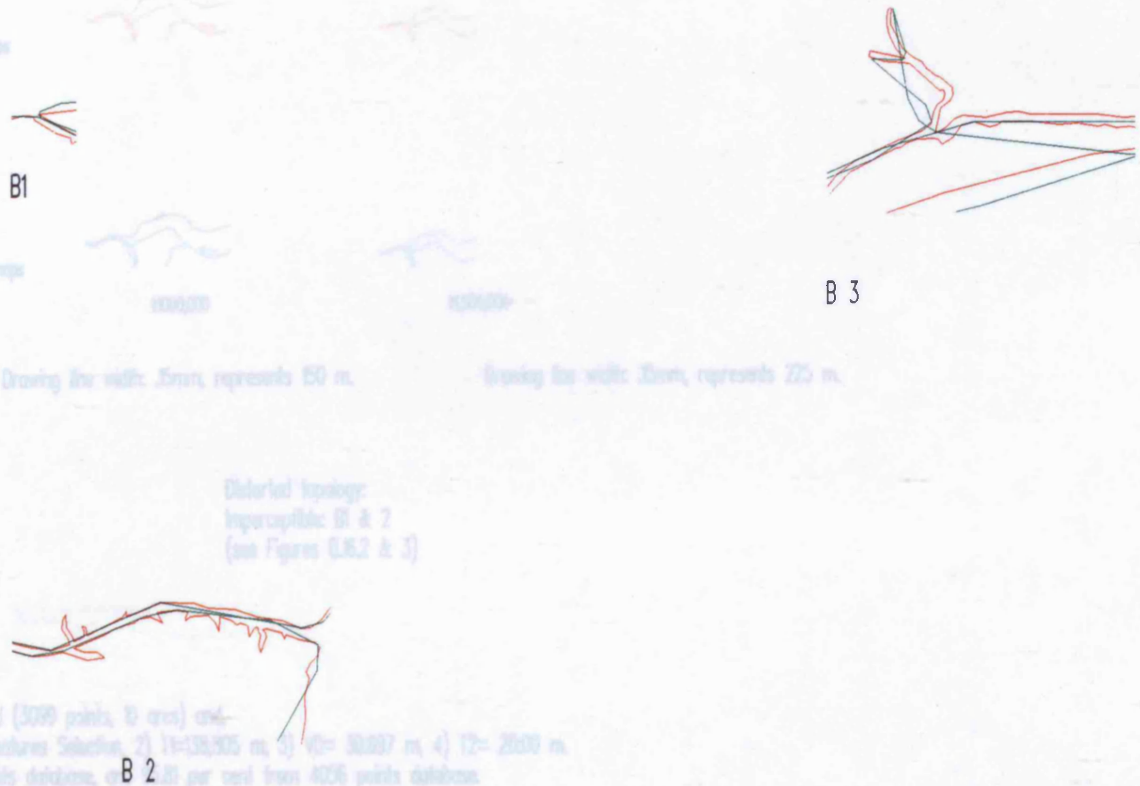
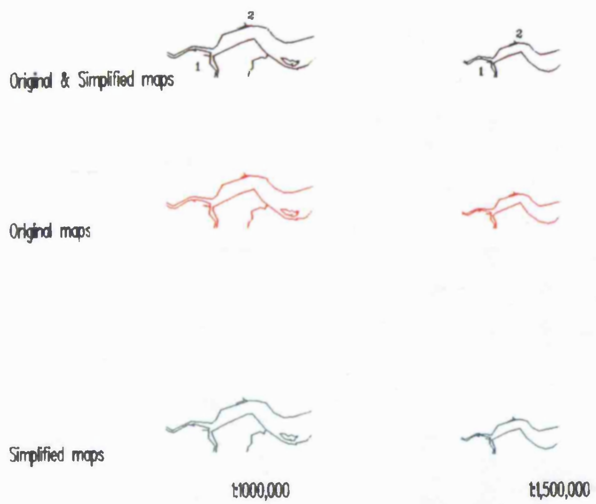



Figure 6.15.2: Distorted topology resulted from the simplification produced for the 1:625,000 scale
The distortion is imperceptible (B) at the scale of 1:625,000, while it is enlarged in this figure to the scale of 1:20,000.
Original in Red and the simplified in Green.



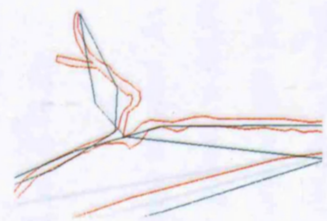

Drawing line width: .15mm, represents 150 m.

Drawing line width: .15mm, represents 225 m.

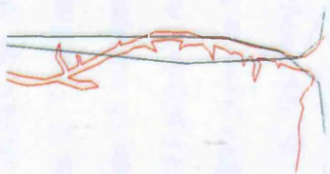
Distorted topology:
Imperceptible: B1 & 2
(see Figures 6.16.2 & 3)

Figure 6.16.1: (Study Area 1).
At the scale of 1:1000,000, original map in Red (3099 points, 10 arcs) and,
the simplified map in Green (170 points), 1) Features Selection, 2) T1=138,805 m, 3) VD= 30.697 m, 4) T2= 20.00 m.
94.52 per cent data reduction from 3099 points database, and 95.81 per cent from 4056 points database.

At the scale of 1:1,500,000, original map in Red (2491 points, 8 arcs), and
the simplified map in Green (94 points), 1) Feature Selection, 2) T1= 203.011 m, 3) VD= 32.006 m, 4) T2= 30.00 m.
96.23 per cent data reduction from 2491 points database, and 97.69 per cent from 4056 points database.



B 2



B 1

Figure 6.16.2: Distorted topology resulted from the simplification produced for the 1:1000,000 scale
The distortion is imperceptible (B) at the scale of 1:1000,000, while it is enlarged in this figure to the scale of 1:20,000.
Original in Red and the simplified in Green.

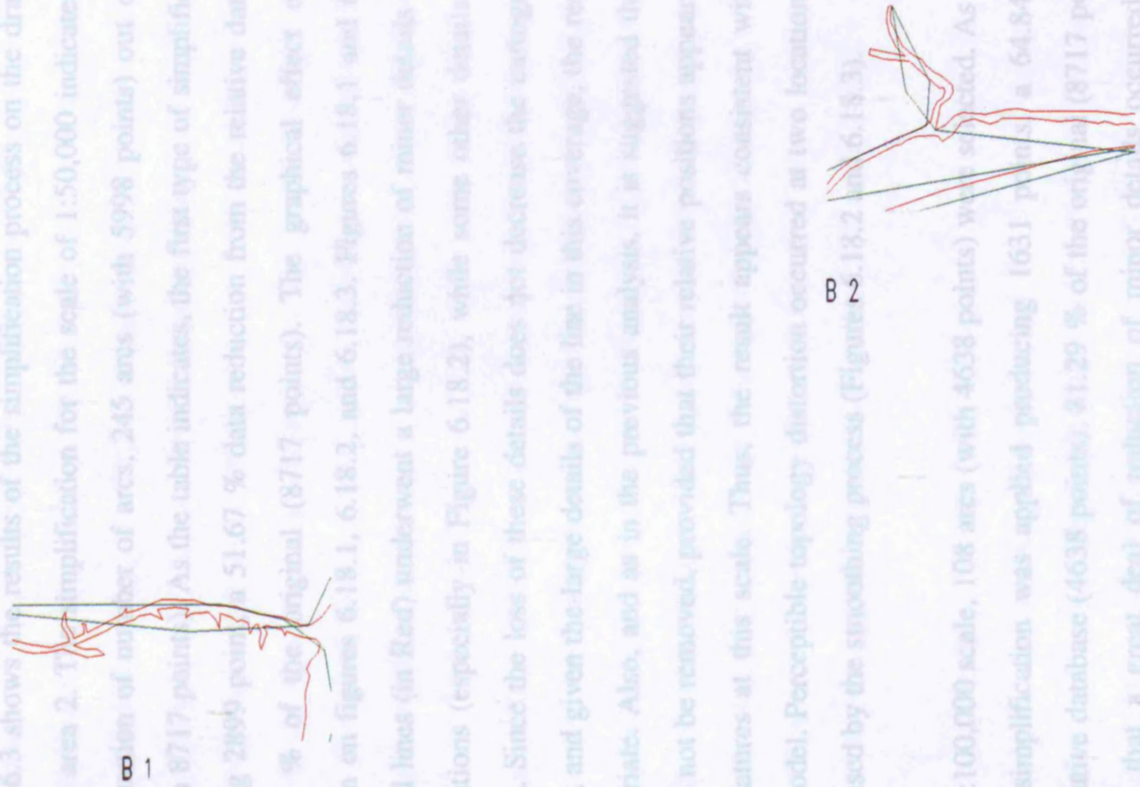



Figure 6.16.3: Distorted topology resulted from the simplification produced for the 1:1,500,000 scale
The distortion is imperceptible (B) at the scale of 1:1,500,000, while it is enlarged in this figure to the scale of 1:20,000.
Original in Red and the simplified in Green.

*6.4.3 Study area 2:*

6.4.3.1 Drainage Coverage (924 arcs, 8717 points):

Table 6.3 shows the results of the simplification process on the drainage coverage of the study area 2. The simplification for the scale of 1:50,000 indicates that there was a large reduction of number of arcs, 245 arcs (with 5998 points) out of the original 924 arcs (with 8717 points). As the table indicates, the first type of simplification was applied, producing 2899 point, a 51.67 % data reduction from the relative database (5998 points), 66.75 % of the original (8717 points). The graphical effect of the simplification is shown on figures 6.18.1, 6.18.2, and 6.18.3. Figures 6.18.1 and 6.18.2 reveal that the original lines (in Red) underwent a large reduction of minor details in the form of minor fluctuations (especially in Figure 6.18.2), while some other details were minimally accentuated. Since the loss of these details does not decrease the cartographic usefulness at the scale, and given the large details of the line in this coverage, the removal is regarded as appropriate. Also, and as in the previous analysis, it is suggested that the retained details should not be removed, provided that their relative positions appear in no conflict with other features at this scale. Thus, the result appears consistent with the requirements of the model. Perceptible topology distortion occurred at two locations (A1 and 2) which were caused by the smoothing process (Figures 6.18.2 and 6.18.3).

At the 1:100,000 scale, 108 arcs (with 4638 points) were selected. As Table 6.3 shows, the first simplification was applied producing 1631 points, a 64.84 data reduction from the relative database (4638 points), 81.29 % of the original (8717 points). Figure 6.19.1 reveals that a great deal of reduction of minor details occurred. The discussion provided for the previous scale about the performance of the simplification is also applies to this scale. Moreover, and given the effects of graphic reduction, the simplification, here, is particularly successful. Figure 6.19.1 also shows that there are two places (B1 and 2) where shape distortion resulted, but both are imperceptible at the scale. As Figure 6.19.2 suggests, both distortions were caused by the smoothing process.

At the scale of 1:250,000, only 25 arcs (with 2247 points) were selected. Following the first type of simplification, 348 points were kept, a 84.59 % data reduction from the relative database (2257 points), 96.01 % of the original. Figure 6.20 shows that the simplified lines (in Green) appear less detailed compared to their original (in Red), and in some places the lines appear to have spiky shapes which were largely caused by the effect of graphic reduction. Furthermore, most of the effect of the large reduction of data resulting from the simplification is imperceptible, due to the effect of graphic reduction. Although the resulting spiky shapes are not desirable, the arguments provided for the previous scales are also valid for this scale, and since minor fluctuations in the original lines were removed, while some were accentuated, the simplified lines still appear appropriately detailed. The result of the simplification is therefore consistent with the objectives and requirements of the model in achieving a minimal reduction of small details.

At the scale of 1:625,000, only 20 arcs (with 1974 points) were selected. They all underwent the first simplification, producing only 184 points, a 90.68 % data reduction of the relative database (1974 points), 97.89 % of the original (Table 6.3). As shown in Figure 6.21, the simplified lines (in Green) appear less detailed compared to the original lines (in Red). Although there was a large reduction of data, the figure graphically demonstrates that the effect is only slightly perceptible at the scale. This is, again, due to the effect of graphic reduction. Again, the result is, therefore, consistent with the model requirements.

At the scale of 1:1,000,000, only 11 arcs (with 1316 points) were selected. They underwent the first simplification, resulting in 88 points, a 93.32 % data reduction from the relative database (1316 points), 99.00 % of the original (Table 6.3). Close visual analysis of Figure 6.22 should reveal that the simplified lines (in Green) appear less detailed, in terms of minor details, compared to their originals (in Red). However, this reduction is imperceptible, as it is due to graphic reduction. Although this has further increased at this scale, the simplification achieved some perceptible reduction of minor, redundant details. Thus, the simplification is regarded as acceptable.

Only 3 arcs (with 493 points) were selected for the 1:1,500,000 scale. The first simplification was applied on them, producing only 35 points, a 92.92 % data reduction from the relative database (493 points), 99.60 % of the original database (8717 points). As can be seen from Figure 6.23, the simplified lines (in Green) appear with no little difference from their original, whilst, as Table 6.3 indicates, there was a large reduction of data. Given the previous arguments about the relationship between the model requirements in terms of minor reduction of line details, and the effect of graphic reduction, and given the relatively large details of the original lines, the simplification result for this scale is, again, regarded as satisfactory.

The results of the simplification for all the scales showed that large data reduction achieved, ranging from 51.67 to 92.92 % from the immediate or relative databases. At larger scales, especially at 1:50,000, 1:100,000, and 1:250,000, some effects of these reductions were perceptible at the scales. On the other hand, the effects of data reductions at the smaller scales (i.e., the remaining scales) are increasingly became imperceptible; the simplified lines largely revealed imperceptible differences from their originals. As pointed out in the previous analysis (study area 1), the effect at small scales does not negate the principle upon which the proposed model was based; namely, manipulating feature details during generalisation is exercised with reference to the mapping context, and only those minor details that are cartographically and economically redundant should be lost. Accordingly, line details should not necessarily undergo excessive reduction in order to incur perceptible graphic results at these scales, as this would entail a removal of important details which may be appropriate to the mapping context. For example, in order to induce a perceptible difference between the simplified and original lines at the scale of 1:1,500,000, the simplified lines would need to be almost straight, but this is, obviously, unacceptable. Database simplification alone would not be a reasonable option, because the original lines contain large details and by applying the first simplification it is, therefore, more likely to incur perceptible changes on them than if the lines were less detailed or smoothed.

| Maps after features Selection Process for different scales | T1 (m) | VD (m) | T2 (m) | Resulting Points | % Data reduction | |
|---|---|---|---|---|---|---|
| Original map before feature selection (**8717** points, **924** arcs) | | | | | from relative database | from original database (8717 points) |
| 1:50,000 (**5998** points, **245**arcs) | **6.55** | **42.03** | **2.00** | **2899** | **51.67** | **66.75** |
| 1:100,000 (**4638** points, **108**arcs) | **13.67** | **40.39** | **4.00** | **1631** | **64.84** | **81.29** |
| 1:250,000 (**2247** points, **25** arcs) | **32.57** | **37.25** | **10.00** | **348** | **84.59** | **96.01** |
| 1:625,000 (**1974** points, **20** arcs) | **81.48** | **35.57** | **12.50** | **184** | **90.68** | **97.89** |
| 1:1,000,000 (**1316** points, **11**arcs) | **123.91** | **38.37** | **20.00** | **88** | **93.32** | **99.00** |
| 1:1,500,000 (**493** points, **3** arcs) | **160.81** | **51.24** | **30.00** | **35** | **92.92** | **99.60** |

**Table 6.3: Digitally simplified maps from their 1:25,000 source databases (relative) after features have been selected from original (8717 points, 924 arcs, study area 2 - Drainage Coverage) according to the PCLS.**

Figure 6.17 graphically illustrates the effect of the processes of selection and simplification in terms of data reduction from the original database (8717 points). The graph shows that the 1:250,000 scale represents a breaking point below which the rate of reduction is comparatively higher than the one beyond this point. As can be seen, more than 90 % of the data was removed by the first three scales changes. This is primarily due to the effect of the process of feature selection, and the effect of the Douglas-Poiker algorithm's behaviour (see Chapter 5).

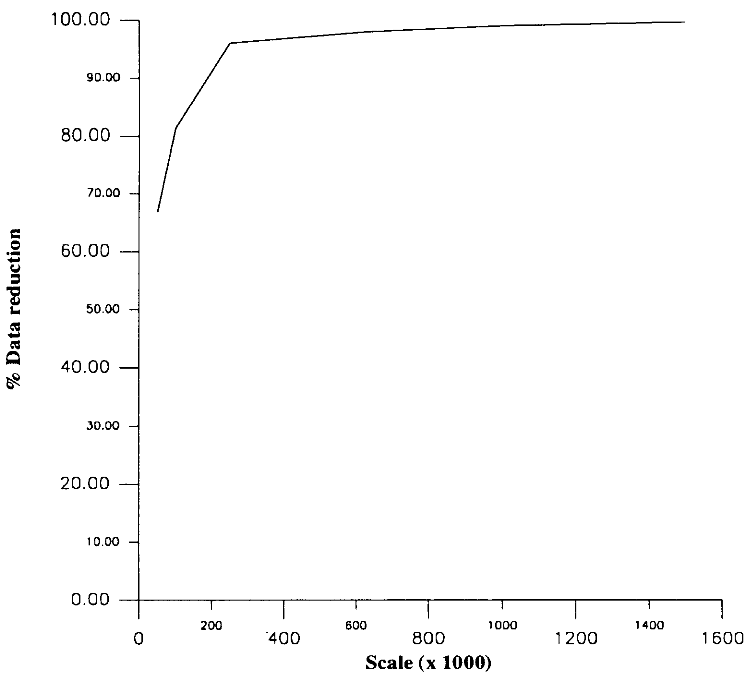



Figure 6.17: Scale against % Data reduction for digitally simplified maps (using the PCLS) from the original map (8717 points, 924 arcs, study area 2 -Drainage Coverage).

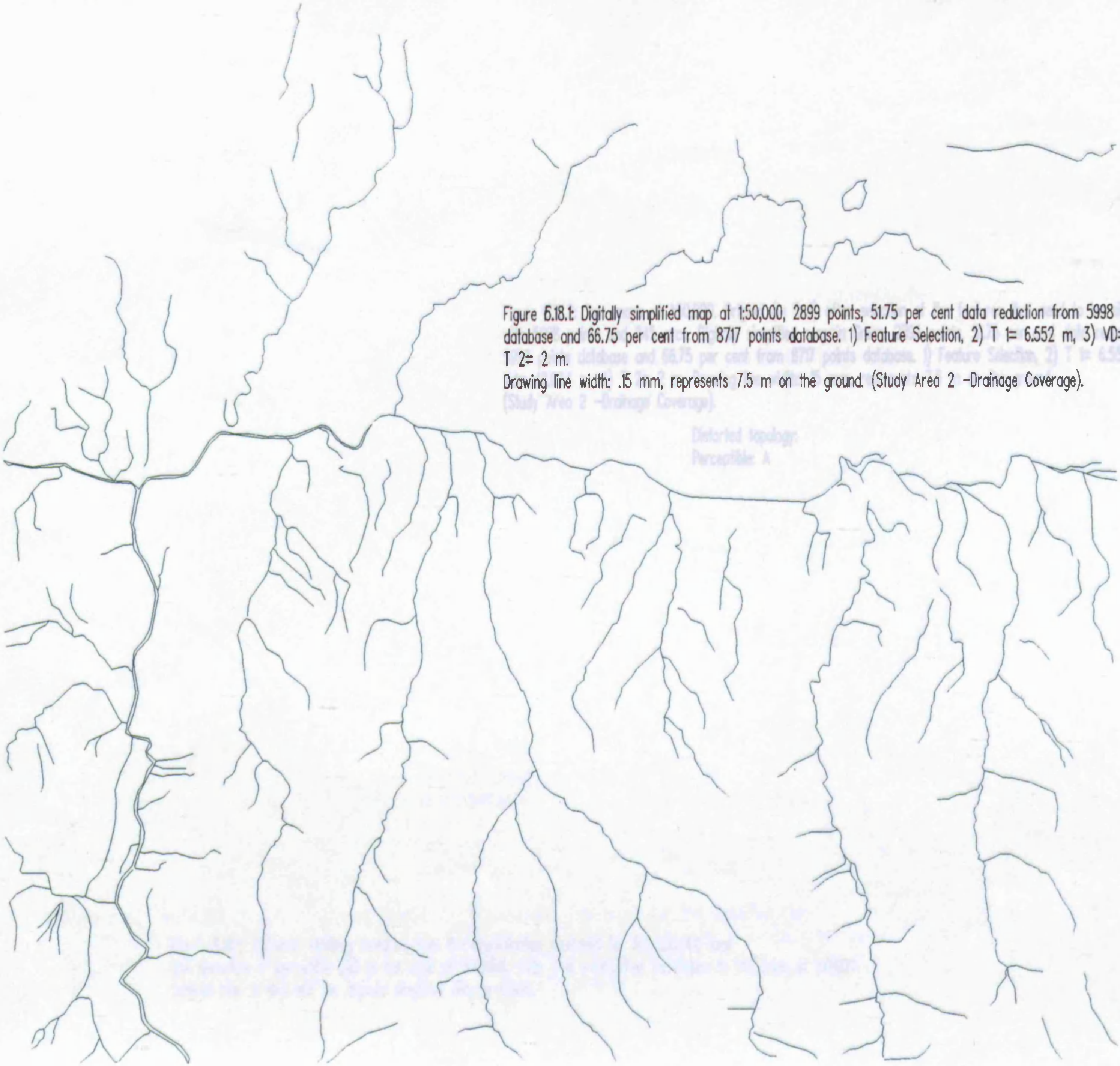


Figure 6.18.1: Digitally simplified map at 1:50,000, 2899 points, 51.75 per cent data reduction from 5998 points database and 66.75 per cent from 8717 points database. 1) Feature Selection, 2) T 1= 6.552 m, 3) VD= 42.034 m, 4) T 2= 2 m.
Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 2 -Drainage Coverage).

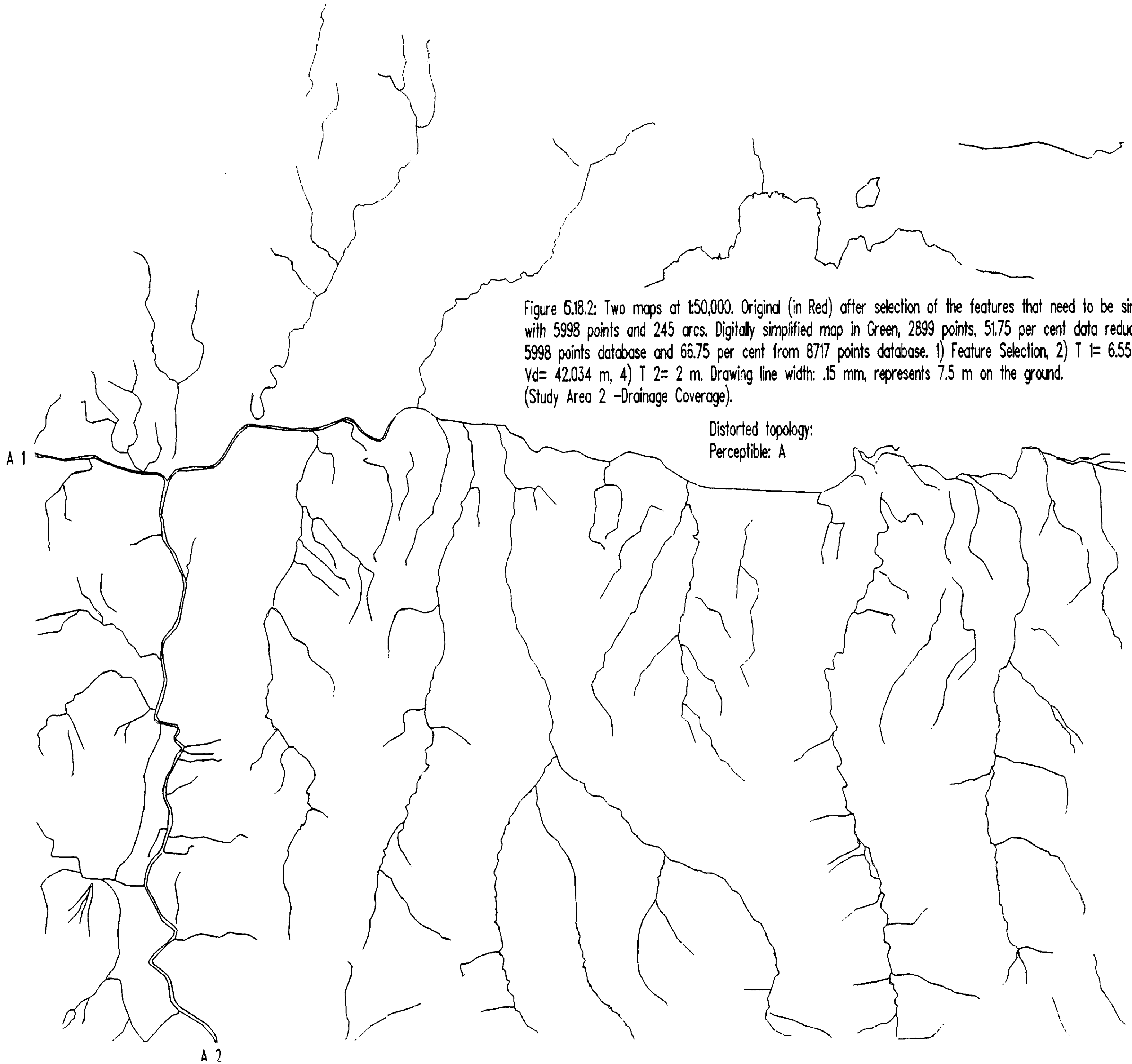



Figure 6.18.2: Two maps at 1:50,000. Original (in Red) after selection of the features that need to be simplified, with 5998 points and 245 arcs. Digitally simplified map in Green, 2899 points, 51.75 per cent data reduction from 5998 points database and 66.75 per cent from 8717 points database. 1) Feature Selection, 2) T 1= 6.552 m, 3) Vd= 42.034 m, 4) T 2= 2 m. Drawing line width: .15 mm, represents 7.5 m on the ground.
(Study Area 2 -Drainage Coverage).

Distorted topology:
Perceptible: A

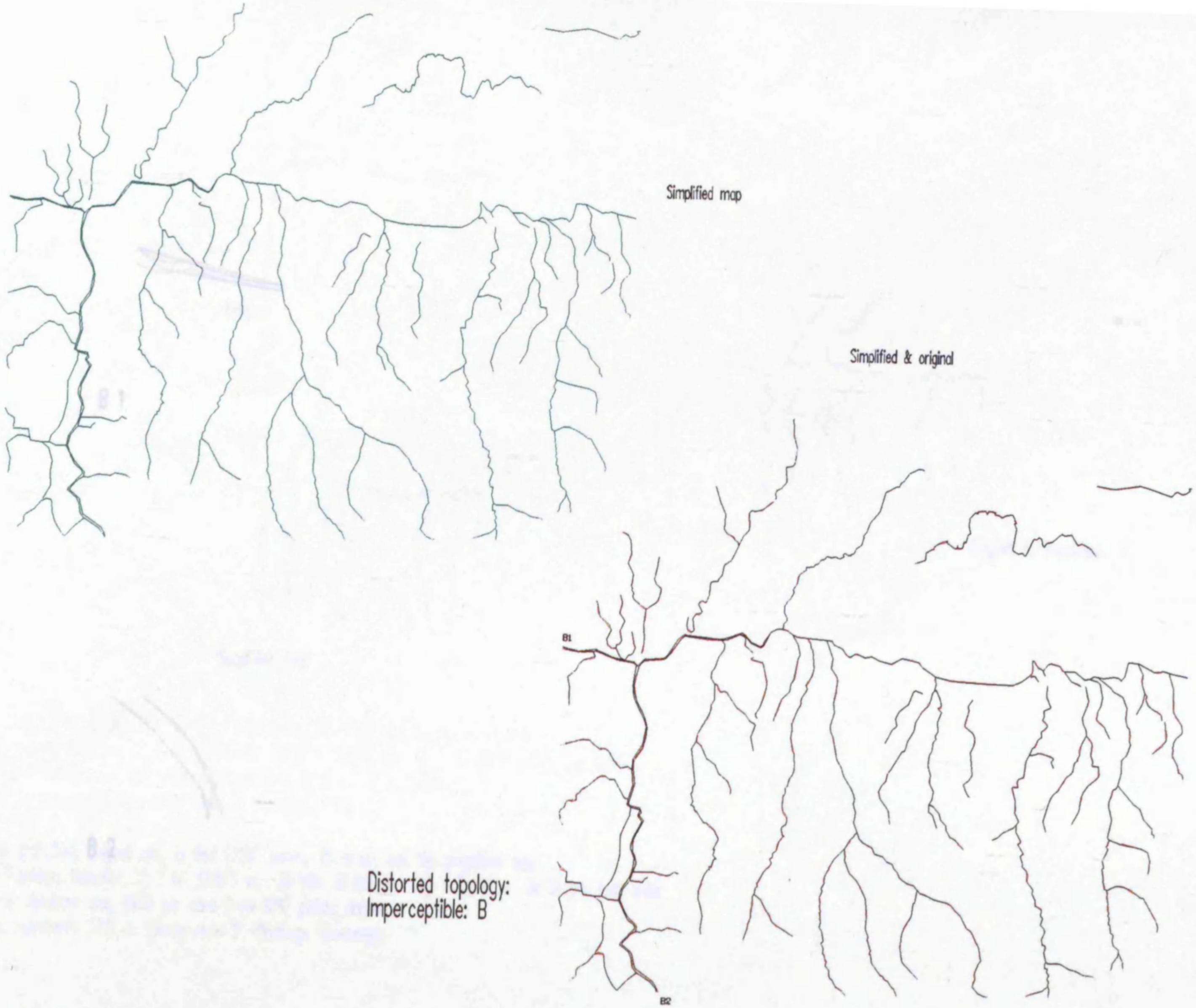



Figure 6.19.1: Two maps at 1:100,000. Original map in Red (4638 points, 108 arcs), and the simplified map in Green (1631 points). 1) Features Selection, 2) T 1= 13.665 m, 3) VD= 40.391 m, 4) T 2= 4 m. 64.84 per cent data reduction from 4638 points database and, 81.29 per cent from 8717 points database. Drawing line width: .15 mm, represents 15 m. (Study Area 2 -Drainage Coverage).

B 1

B 2

Figure 6.19.2: Distorted topology resulted from the simplification produced for the 1:100,000 map
The distortion is imperceptible (B) at the scale of 1:100,000, while it is enlarged in this figure to the scale of 1:10,000
Original map in Red and the digitally simplified in Green.

Original

Simplified map

Original & Simplified

Figure 6.20: Two maps at 1:250,000. Original map in Red (2257 points, 25 arcs), and the simplified map in Green (348 points). 1) Features Selection, 2) T 1= 32.573 m, 3) VD= 37.248 m, 4) T 2= 10 m. 84.59 per cent data reduction from 2257 points database and, 96.01 per cent from 8717 points database.
Drawing line width: .15mm, represents 37.5 m. (Study Area 2 -Drainage Coverage).


Figure 6.21: Two maps at 1:625,000. Original map in Red (1974 points, 20 arcs), and the simplified map in Green (184 points). 1) Features Selection, 2) T 1= 81.476 m, 3) VD= 35.570 m, 4) T 2= 12.5m. 90.68 per cent data reduction from 1974 points database, and 97.89 per cent from 8717 points database.
Drawing line width: .15mm, represents 93.73m. (Study Area 2 –Drainage Coverage).


Figure 6.22: Two maps at 1:1,000,000. Original map in Red (1316 points, 11 arcs), and the simplified map in Green (88 points). 1) Features Selection, 2) T 1= 123.910 m, 3) VD= 38.368 m, 4) T 2= 20 m. 93.32 per cent data reduction from 1316 points database, and 99.00 per cent from 8717 points database. Drawing line width: .15mm, represents 93.73m. (Study Area 2 -Drainage Coverage).

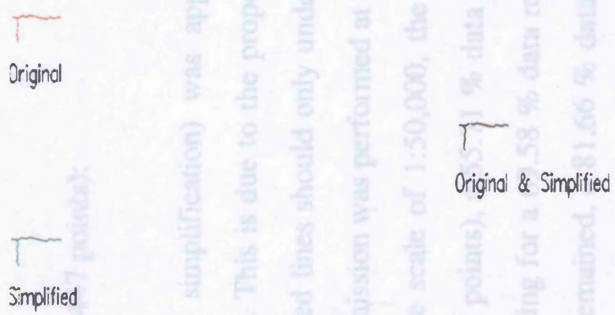



Figure 6.23: Two maps at 1:1,500,000. Original map in Red (494 points, 3 arcs), and the simplified map in Green (35 points). 1) Features Selection, 2) T 1= 160.805 m, 3) VD= 51.239 m, 4) T 2= 30m. 92.92 per cent data reduction from 494 points database, and 99.60 per cent from 8717 points database.
Drawing line width: .15mm, represents 93.73m. (Study Area 2 -Drainage Coverage).

6.4.3.2 Study area 2 - Transport coverage (21 arcs, 447 points):

The second simplification (database simplification) was applied on this coverage at all scales of representation (Table 6.4). This is due to the proposition of the model, which suggests that smoothed or less detailed lines should only undergo this type of simplification. Table 6.4 shows that no feature omission was performed at the first three scales (1:50,000, 1:100,000, and 1:25,000). At the scale of 1:50,000, the result of the simplification was 198 points from the original (447 points), a 55.71 % data reduction. At the scale of 1:100,000, 136 points resulted, accounting for a 69.58 % data reduction from the original. At the scale of 1:250,000, 82 points remained, a 81.66 % data reduction of the original. Visual inspection of the figures showing the graphical effects of these results for the three scales reveal imperceptible differences between the original lines and their simplified versions (Figures 6.25, 6.26, and 6.27).

The process of feature selection was applied at the remaining scales; namely the 1:625,000, 1:1,000,000, and 1:1,500,000. At the scale of 1:625,000, 12 arcs (with 394 points) were selected. The process of simplification resulted in 42 points, representing a 88.44 % data reduction from the relative database (394 points), and a 90.16 % from the original (447 points). At the scale of 1:1,000,000, the arcs and points selected for the previous scale were also selected. The simplification process produced 32 points, representing a 91.88 % data reduction from the relative database (394 points), and a 92.86 % from the original. At the scale of 1:1,500,000, only 6 arcs (with 291 points) were selected. The process of simplification produced only 16 points, representing a 94.51 data reduction from the relative database (291 points), and a 96.43 % from the original (447 points). Figure 6.28 shows the graphical effect of the simplification, where there is imperceptible difference between the original lines and their simplified versions. The results for all the scales showed that large reductions of data (upto 94.51 %) from the relative databases were achieved, without causing perceptible changes between the simplified and original lines at the target scales. The results of the database simplifications at all scales are, therefore, acceptable.

| **Maps after features Selection Process for different scales** | **T1 (m)** | **VD (m)** | **T2 (m)** | **Resulting Points** | **% Data reduction** | |
|---|---|---|---|---|---|---|
| Original map before feature selection (**447** points, **21** arcs) | | | | | from relative database | from original database (447 points) |
| 1:50,000 (**447** points, **21** arcs) | **4.291** | _ | _ | **198** | **55.71** | **55.71** |
| 1:100,000 (**447** points, **21** arcs) | **9.60** | _ | _ | **136** | **69.58** | **69.58** |
| 1:250,000 (**447** points, **21** arcs) | **25.50** | _ | _ | **82** | **81.66** | **81.66** |
| 1:625,000 (**394** points, **12** arcs) | **63.89** | _ | _ | **42** | **88.84** | **90.16** |
| 1:1,000,000 (**394** points, **12** arcs) | **102.89** | _ | _ | **32** | **91.88** | **92.85** |
| 1:1,500,000 (**291** points, **6** arcs) | **146.83** | _ | _ | **16** | **94.51** | **96.43** |

**Table 6.4: Digitally simplified maps from their 1:25,000 source databases (relative) after features have been selected from original (447 points, 21 arcs, study area 2 - Transport Coverage) according to the PCLS.**

Figure 6.24 shows the effects of the selection and simplification processes in terms of data reduction from the original database (447 points). The graph reflects how the process produced a large data reduction well over 80 % within the first three scales. This is primarily due to the effect of the Douglas-Poiker algorithm. On the other hand, the graph shows a lesser rate of progression beyond the 1:250,000 scale, although feature selection and simplification contribute to the pattern. This is due to the fact pointed out in the discussions of this effect, observed in Figure 6.18.

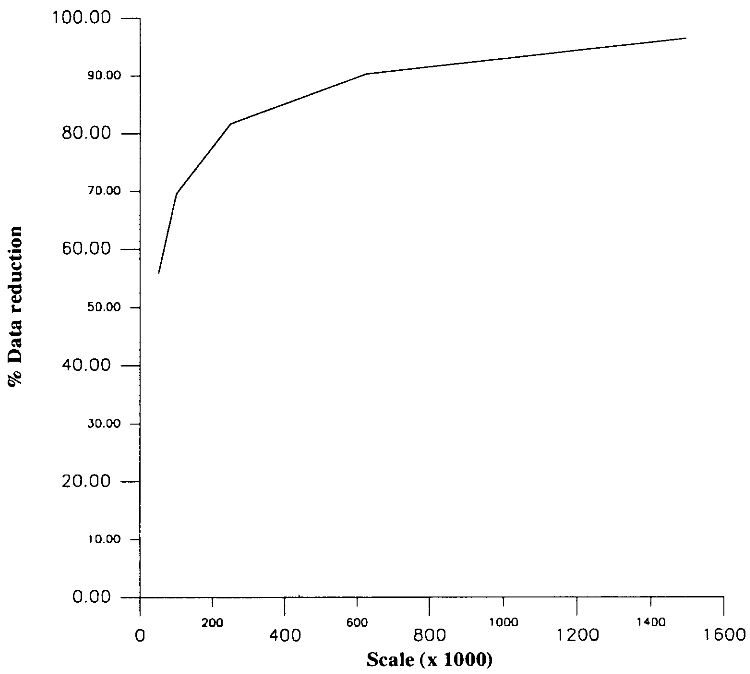



Figure 6.24: Scale against % Data Reduction for digitally simplified maps (using the PCLS) from the original map (21 arcs, 447 points, study area 2 -Transport Coverage).

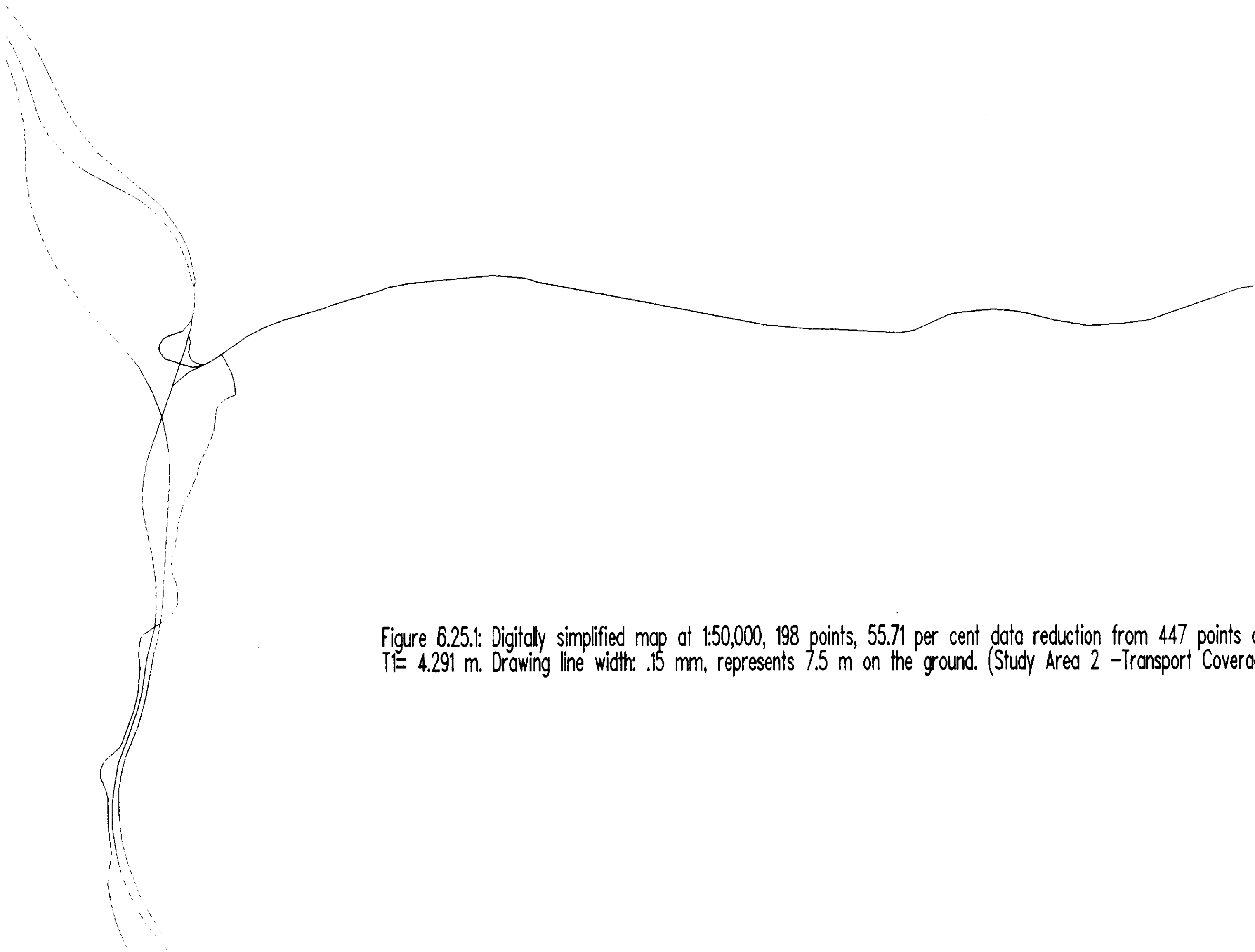


Figure 6.25.1: Digitally simplified map at 1:50,000, 198 points, 55.71 per cent data reduction from 447 points original database. T1= 4.291 m. Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 2 -Transport Coverage)

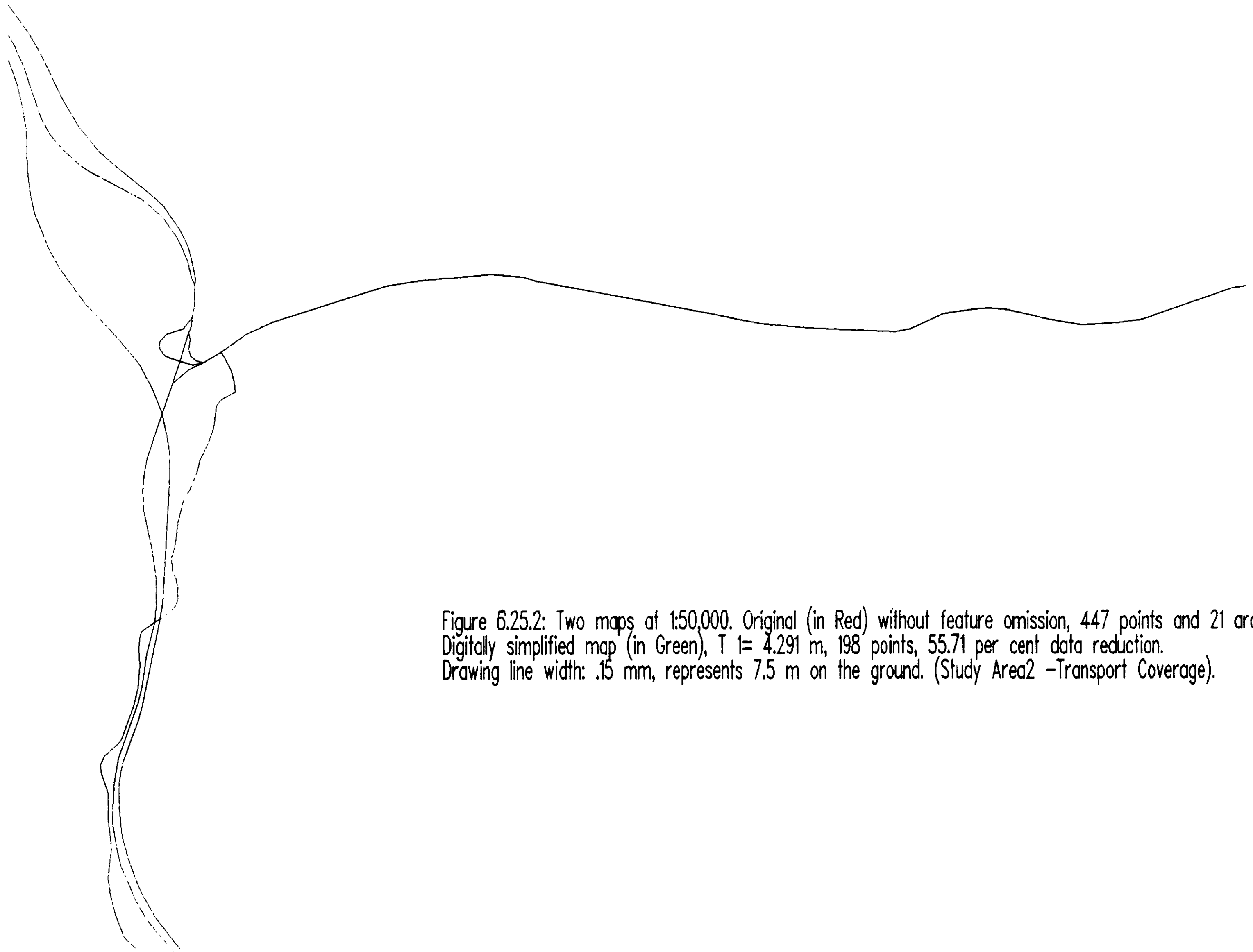
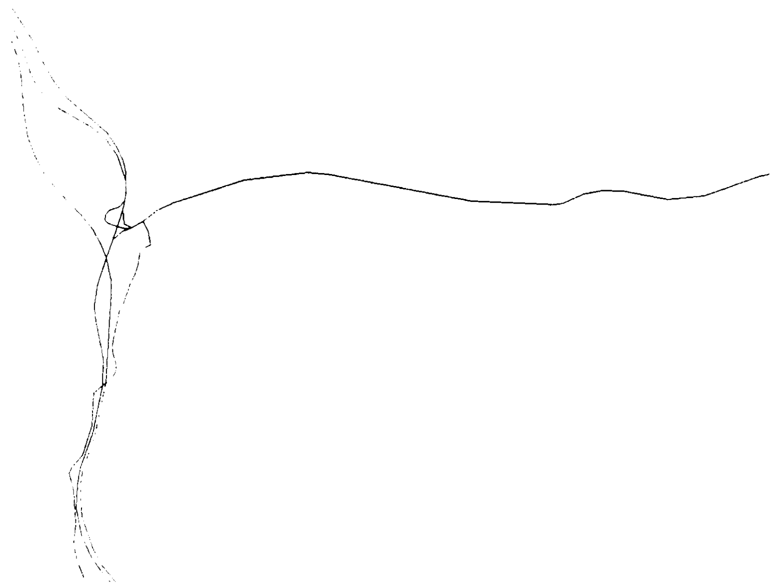
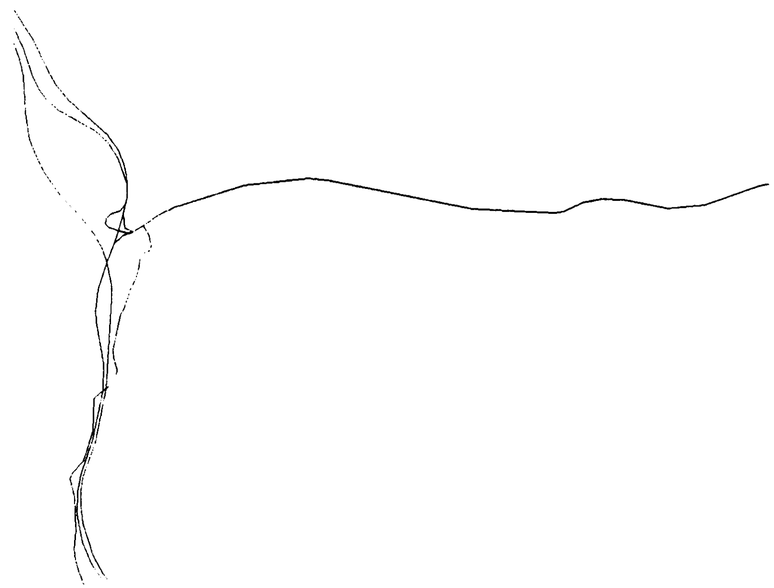


Figure 6.25.2: Two maps at 1:50,000. Original (in Red) without feature omission, 447 points and 21 arcs
Digitally simplified map (in Green), T 1= 4.291 m, 198 points, 55.71 per cent data reduction.
Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area2 -Transport Coverage).



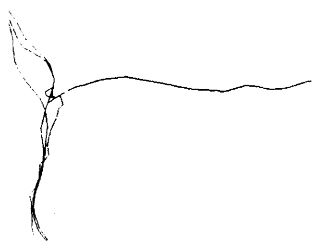
Simplified



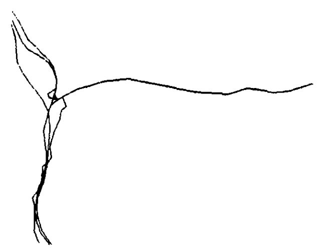
Simplified & Original

Scale: 1:100 000

Figure 6.26: Two maps at 1:100,000. Original (in Red) without feature omission, 447 points and 21 arcs
Digitally simplified map (in Green), T 1= 9.600 m, 136 points, 69.58 per cent data reduction.
Width of drawing line= .15 mm, represents 15 m on the ground. (Study Area 2 -Transport Coverage).



Simplified



Simplified & Original

Scale: 1:250,000

Figure 6.27: Two maps at 1:250,000. Original (in Red) without feature omission, 447 points and 21 arcs
Digitally simplified map (in Green), T 1= 25.500 m, 82 points, 81.66 per cent data reduction.
Width of drawing line= .15 mm, represents 37.5 m on the ground. (Study Area 2 -Transport Coverage).

1:625 000

Digitally simplified map (in Green), T 1= 63.887 m, 42 points (88.84 per cent data reduction
from 394 points database, and 90.16 per cent from 447 points database.
Width of drawing line= .15 mm, represents 93.73 m.

Digitally simplified map (in Green), and original map, after feature selection, (in Red, 394 points, 12 arcs).

1:1000 000
Digitally simplified map (in Green), T 1= 102.891 m, 32 points (91.88 per cent data reduction
from 394 points database, and 92.88 per cent from 447 points database.
Width of drawing line= .15 mm, represents 150 m.

Digitally simplified map (in Green), and original map after feature selection(in Red, 394 points, 12 arcs).

1:1 500 000
Digitally simplified map (in Green), T 1= 146.833 m, 16 points (94.51 per cent data reduction
from 291 points database, and 96.43 per cent from 447 points database.
Width of drawing line= .15 mm, represents 225 m.

Digitally simplified map (in Green), and original map after feature selection (in Red, 291 points, 6 arcs).

Figure 6.28: Digitally simplified maps of Study Area 2, Transport Coverage, using the Simplification Model, of 1:625,000, 1:1,000,000, and 1:1,500,000.

*6.4.4 Study area 3 (103 arcs, 866 points):*

The original data set of this test has 103 arcs with 866 points. It represents a road network (see section 4.1) at the scale of 1:158,400. Five reproduction scales were derived from it. Table 6.5 shows that no features were omitted for the first three scales. As it is the case in the previous analyses, manually generalised maps and requirements of road mapping were the basis for decisions regarding the feature selection and simplification processes. In this case, the basis for such decisions were based on two main factors. First, is the fact that road mapping like any thematic mapping process, should emphasise the subject matter of the map, i.e., its graphic details in contrast to other subordinate themes of the map. The second factor relates to the source database which was from medium scale maps as opposed to a large scale (e.g., 1:10,000, or 1:25,000), and given the significance of the first factor, it is therefore important not to perform a perceptible simplification of the feature details. Another factor, is related to the nature of the details of the lines, which are less complex here than are those in the previous two test areas. As a sequence, the second type of simplification (database reduction) was applied for the first three scales: 1:200,000, 1:500,000, and 1:730,000. At the scale of 1:200,000, the resulting points were 806 points, representing only a 6.93 % data reduction from the original database (866 points). At the scale of 1:500,000, there were 709 points resulted, representing only an 18.13 % data reduction from the original. At the third scale of 1:730,000, the resulting points were 650 points, being only a 24.95 % data reduction. The graphical effects of the simplifications at these three scales are shown in Figures 6.30, 6.31, 6.32, and 6.33. The figures display imperceptible deviation between the original and the simplified lines. Although there was some reduction of data, previous analysis of the Douglas-Poiker algorithm in Chapter 5 suggests that it would have been possible to reduce the data by larger percentages and maintain the imperceptible difference between the original and the simplified lines.

Table 6.5 indicates that the processes of feature selection and the first simplification were both applied for the remaining scales; namely the 1:1,000,000, and

1:1,500,000 scale. This is due to one main factor: the differences between the source scale and these derived scales are relatively large compared to that of previous scale changes. At such scales (i.e., 1:1million and smaller) cartographic features primarily undergo symbolisation, since their representation to scale becomes impossible. It was therefore necessary to apply the first simplification. So, 91 arcs (with 847 points) from the original 103 arcs were selected for both scales of 1:1,000,000, and 1:1,500,000. At the scale of 1:1,000,000, 294 points were retained, representing a 65.29 % data reduction from the relative database (847 points), 66.06 % of the original (866 points). At the last scale of 1:1,500,000, the resulting points were 284 points, representing a 66.47 % data reduction from the relative database (847 points), 67.21 % of the original. Figure 6.34 graphically illustrates how the simplification effects at these last two scales are clearly perceptible, compared to those at the previous larger scales. In this figure, the simplified lines (in Green) at both scales appear smoother as a result of the reduction of much of the small details. The relatively small difference between these two data reductions (or details) for the last two scales is regarded desirable, since a larger increase for the scale of 1:1,500,000 would actually cause spiky or distorted shapes.

The simplifications for both the scales show a uniform reduction of details for all segments of the lines. This emphasises the desirable property of consistency inherent within digital processes as opposed to manual approaches. Close visual examination of every single line segment at both the scales would reveal that the simplification results adhere to the model requirements. Given the significance of line details for this type of mapping and given the requirements of the model in terms of minimum reduction of details, some oversimplification of some details, could have been better avoided, especially for the scale of 1:1,000,000. These undesirable effects can be identified more clearly on the overlays (Figure 6.34) at locations where the difference between the original and simplified lines becomes much pronounced.

With comparison of the VD values in the previous results, the increase in the values, here, was determined by two factors: first, the total length of the lines which are, here, longer than those in the previous study areas, and second, by the relatively small database. Based also on the observation about the algorithm's approach, it can be concluded that large line coverages in the form of networks (i.e., large numbers of arcs) and with small databases have to undergo database simplification but not shape/data simplification for the large and medium scales. This is to avoid unnecessary reduction of line details. If such line coverages are composed of large databases, or complex line details, then performing the shape/data simplification process would be possible. Therefore, the simplification results for the last two scales in this analysis are desirable, as at these scales some line details are regarded as redundant and have to be removed. At this stage in the analyses, it proved evident that both the development and evaluation of digital cartographic solutions have to be performed within a context-dependent mapping processes. As a result, the complexity of cartographic processes such as line simplification demands that the applicability of a given digital solution or formulation to a particular cartographic problem is limited. For example, during digital generalisation line simplification algorithms have to adhere to specific mapping contexts.

| Maps after features Selection Process for different scales | T1 (m) | VD (m) | T2 (m) | Resulting Points | % Data reduction | |
|---|---|---|---|---|---|---|
| Original map before feature selection (**866** points, **103**arcs) | | | | | from relative database | from original database (866 points) |
| 1:200,000 (**866** points, **103** arcs) | **6.41** | _ | _ | **806** | **6.93** | **6.93** |
| 1:500,000 (**866** points, **103**arcs) | **17.57** | _ | _ | **709** | **18.13** | **18.13** |
| 1:730,000 (**866** points, **103** arcs) | **26.07** | _ | _ | **650** | **24.95** | **24.95** |
| 1:1,000,000 (**866** points, **91** arcs) | **36.08** | **1132.50** | **6.31** | **294** | **65.29** | **66.06** |
| 1:1,500,000 (**866** points, **91** arcs) | **54.51** | **1130.35** | **9.47** | **284** | **66.47** | **67.21** |

**Table 6.5: Digitally simplified maps from their 1:158,400 source databases (relative) after features have been selected from original (866 points, 103arcs) according to the PCLS. (study area 3).**

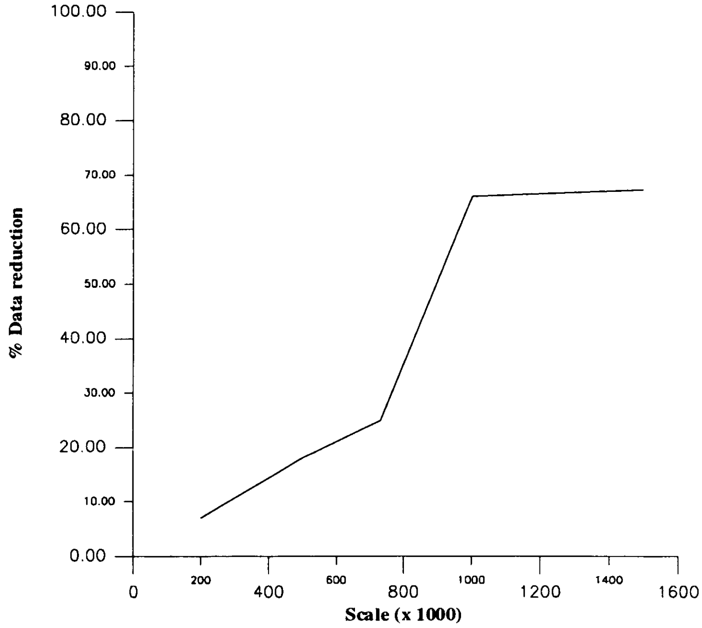



Figure 6.29: Scale against % Data reduction for digitally simplified maps (using the PCLS) from the original map (866 points, 103 arcs, study area 3).

Figure 6.29 shows the effect of feature selection and the simplification of both types in terms of the amount of data reduction at all the scales. In the figure, the rate of reduction steadily increases with reduced scale, but is abruptly increased at the scale of 1:1,000,000, though it continues to follow the same pattern. This marked increase was primarily caused by the smoothing process, where a large tolerance (VD) was applied, according to which line details underwent further reduction and exaggeration.

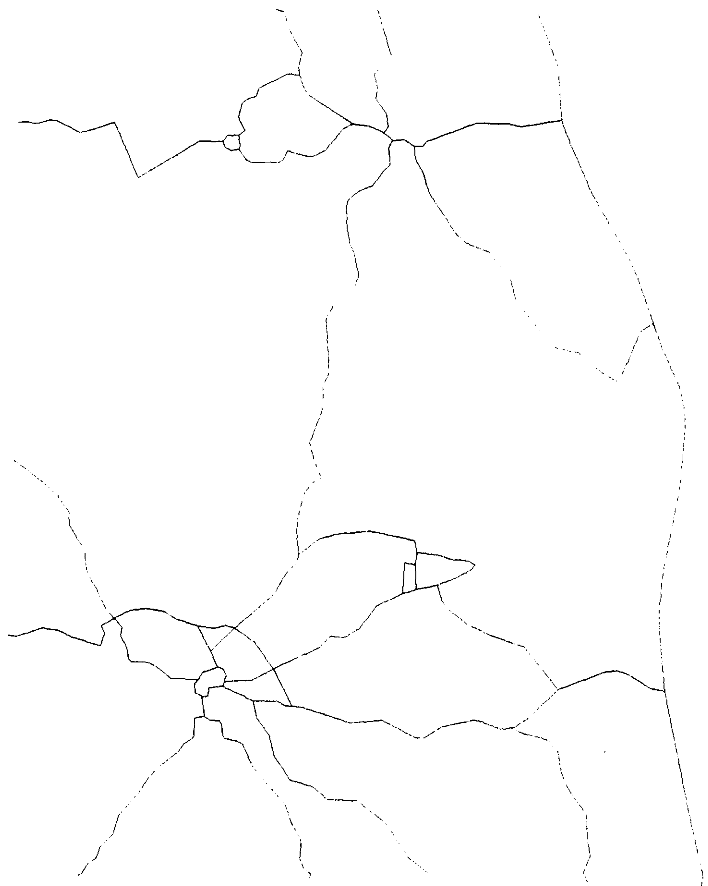


Figure 6.30.1: Digitally simplified map at 1:200,000, 806 points, 6.93 per cent data reduction from 866 points original database, Tt= 6.414 m. (Study Area 3). Drawing line width: .15 mm, represents 30 m on the ground.
Due to scale and paper size only half of the map is displayed, Figure 5.30.2 shows the other half.

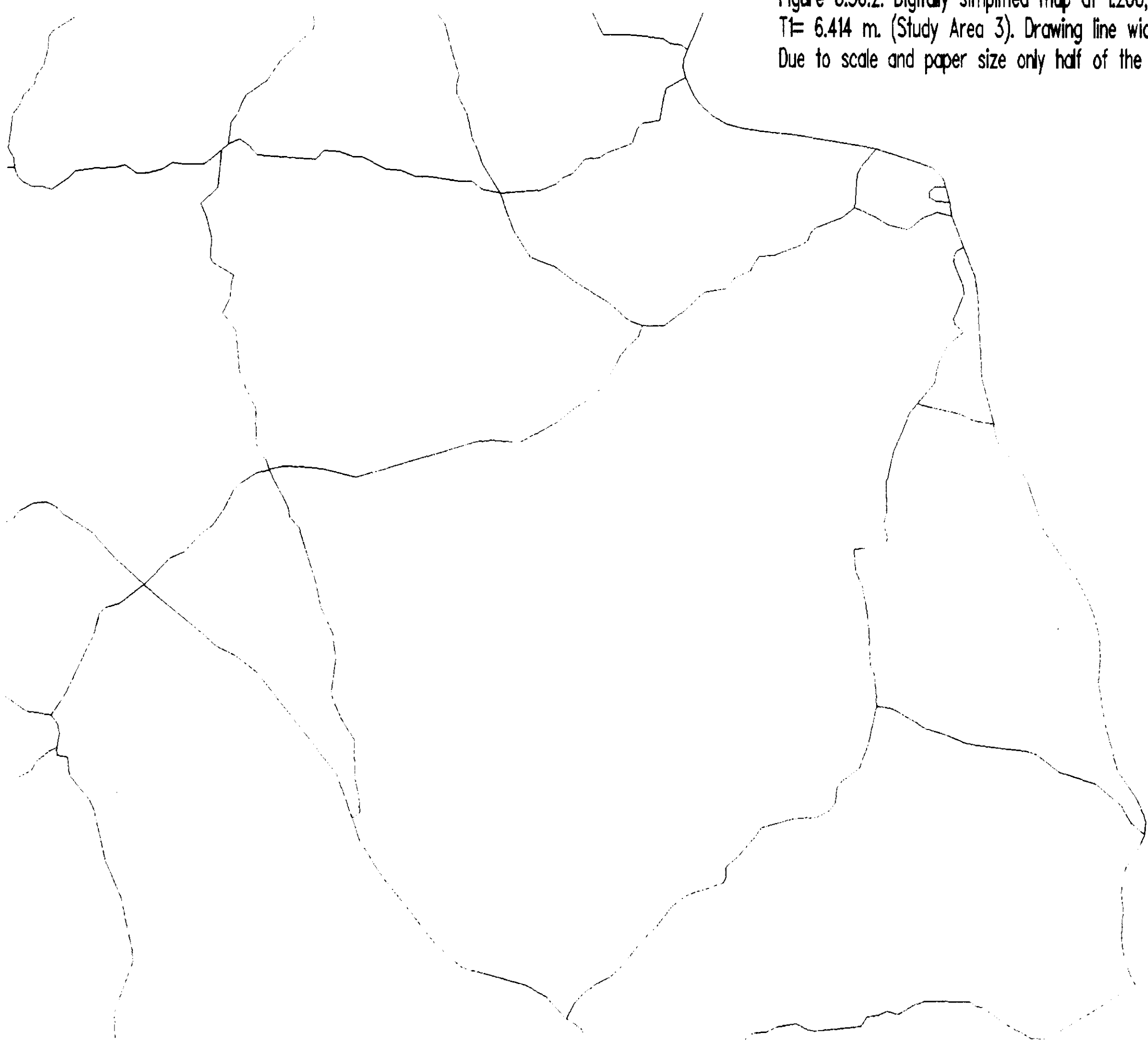


Figure 5.30.2: Digitally simplified map at 1:200,000, 806 points, 6.93 per cent data reduction from 866 points original database, T1= 6.414 m. (Study Area 3). Drawing line width: .15 mm, represents 30 m on the ground.
Due to scale and paper size only half of the map is displayed, Figure 5.30.1 shows the other half.

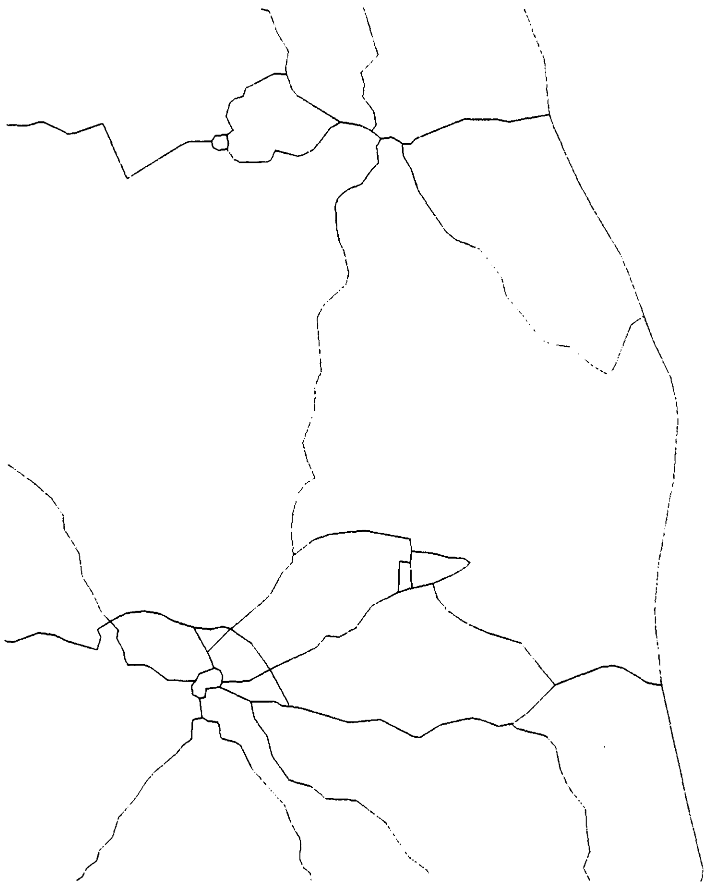


Figure 6.31.1: Two maps at 1:200,000. Original (in Red) without feature omission, 866 points, 103 arcs.
Digitally simplified map (in Green), T 1= 6.414 m, 806 points, 6.93 per cent data reduction.
Due to scale and paper size only half of the map is displayed, Figure 5.31.2 shows the other half.
Width of drawing line= .15 mm, represents 30 m on the ground. (Study Area 3).

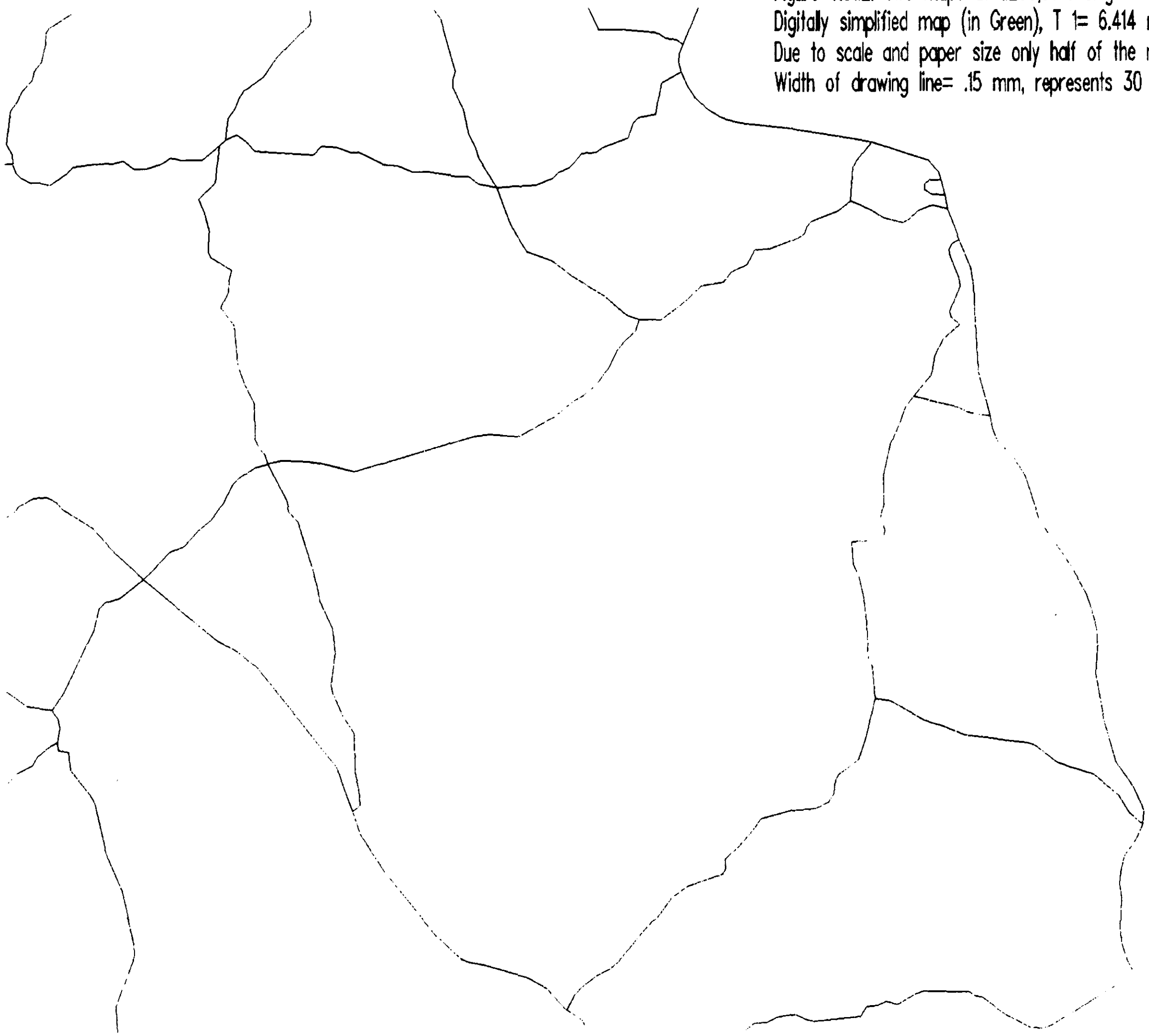
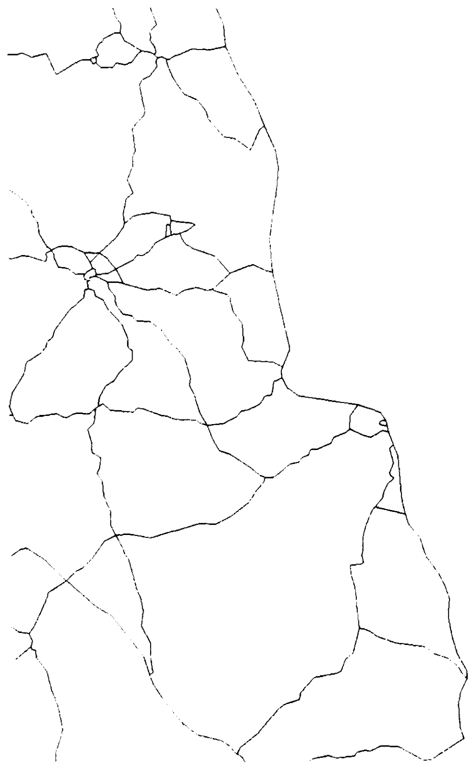


Figure 6.31.2: Two maps at 1:200,000. Original (in Red) without feature omission, 866 po ints, 103 arcs
Digitally simplified map (in Green), T 1= 6.414 m, 806 points, 6.93 per cent data reduction.
Due to scale and paper size only half of the map is displayed, Figure 5.31.1 shows the other half.
Width of drawing line= .15 mm, represents 30 m on the ground. (Study Area 3).



Scale: 1:500,000

Simplified

Simplified & Original

Figure 6.32: Two maps at 1:500,000. Original (in Red) without feature omission, 866 points Digitally simplified map (in Green), T 1= 17.571 m, 709 points, 18.13 per cent data reduction. Width of drawing line= .15 mm, represents 75 m on the ground. (Study Area 3).

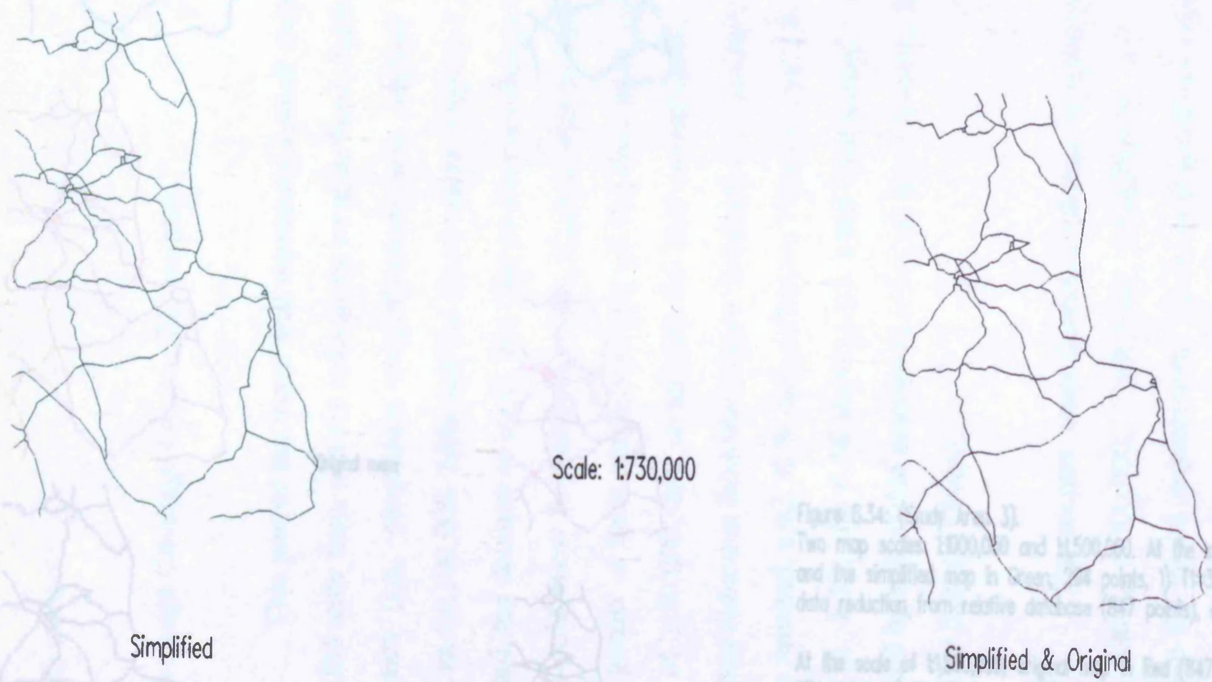



Figure 6.33: Two maps at 1:730,000. Original (in Red) without feature omission, 866 points
Digitally simplified map (in Green), T 1= 26.070 m, 650 points, 24.95 per cent data reduction.
Width of drawing line= .15 mm, represents 109.50 m on the ground. (Study Area 3).

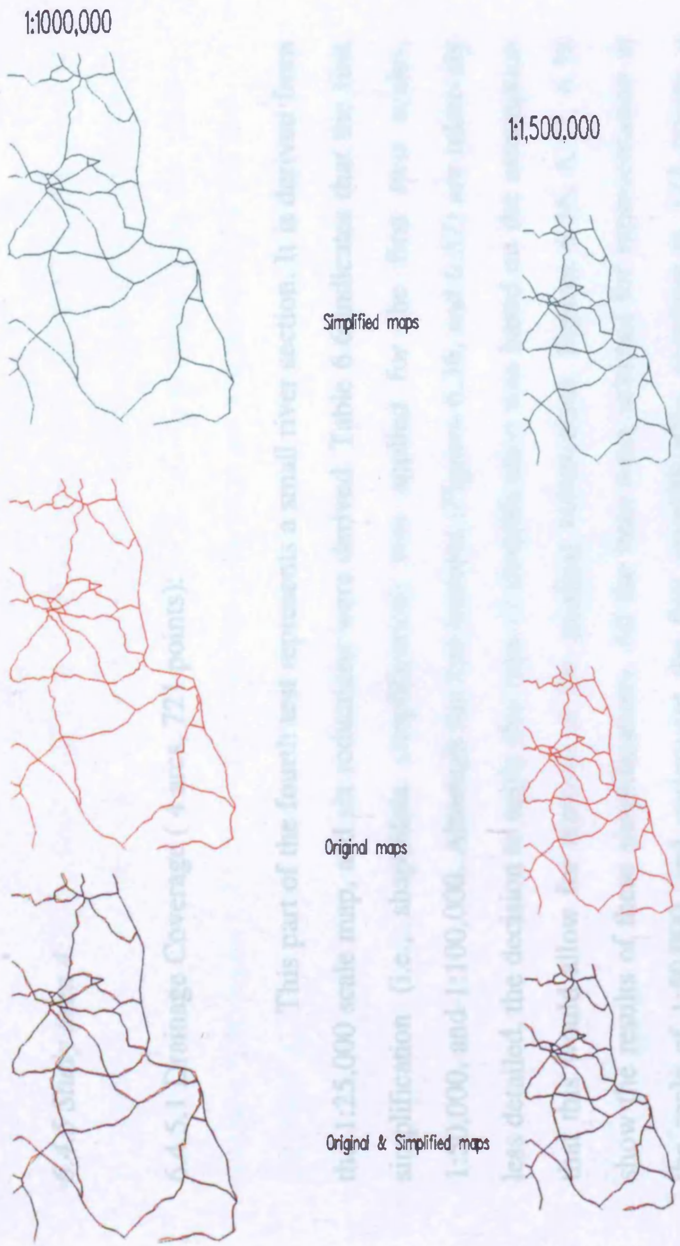



Figure 6.34: (Study Area 3).
Two map scales: 1:1000,000 and 1:1,500,000. At the scale of 1:1000,000, original map in Red (847 points, 91 arcs), and the simplified map in Green, 294 points, 1) T1=36.009 m, 2)VD= 1132.500 m, 3) T2= 6.313 m. 65.29 per cent data reduction from relative database (847 points), and 66.06. from original database (866 points).

At the scale of 1:1,500,000, original map in Red (847 points, 91 arcs), and simplified map in Green, 284 points, 1) T1= 54.514 m, 2) VD=1130.348 m, 3) T2=9.469 m. 66.47 per cent data reduction from realtive database (847 points), and 67.21 per cent from original database (866 points).

*6.4.5 Study area 4:*

6.4.5.1 Drainage Coverage ( 4 arcs, 721 points):

This part of the fourth test represents a small river section. It is derived from the 1:25,000 scale map, and six reductions were derived. Table 6.6 indicates that the first simplification (i.e., shape/data simplification) was applied for the first two scales, 1:50,000, and 1:100,000. Although the line features (Figures 6.36, and 6.37) are relatively less detailed, the decision to apply this type of simplification was based on the assumption that this would allow for removal of the smallest indentations. Figures 6.36, 6.37, 6.38 show the results of these simplifications. All the lines were selected for representation at the scale of 1:50,000, and underwent the first simplification, resulting in 323 points, a 55.21 % data reduction from the original database (721 points). At the scale of 1:100,000, 269 points resulted, a 62.70 % data reduction from the original database (Figure 6.38). As Figures 6.36, 6.37, and 6.38 show, such small indentations were successfully removed accompanied by a desirable smoothing effect of the remaining details, as can be seen from the simplified lines (in Green).

The second simplification (database simplification) was applied for the remaining scales, 1:300,000, 1:500,000, 1:1,000,000, and 1:1,500,000 (Table 6.6). According to the model requirements, this is an appropriate choice for two reasons. First, there is already a limited amount of fine details. Second, the effect of graphic reduction masks out small details such as those that were required to be removed at the previous scales; hence the application of the first simplification would be meaningless. At the scale of 1:300,000, there were 90 points, being a data reduction of 87.52 % from the original database (721 points). At the scale of 1:500,000, a selective omission process of features was applied which resulted unusually in 5 arcs (with 356 points) compared to the original 4 arcs. This increase in the number of arcs resulted from the process of the selective omission, in which the river was represented by a single line instead of double lines. The simplification resulted in only 35 points, representing a 90.17 % data reduction from the relative database (356 points), 95.15 % of the original (721 points). At the scale of 1:1,000,000, only one arc (with 256 points) was selected. The process of simplification resulted in only 16 points, representing a 93.75 % data reduction from the relative database (256 points), 97.79 % of the original. The same arc was selected for the scale of 1:1,500,000. The simplification resulted in only 12 points, representing a 95.32 % data reduction from the relative database (256 points), and a 98.34 % from the original. Figures 6.39, and 6.40 graphically show the simplification effects at these four scales. The figures reveal imperceptible differences between the original and simplified lines. The results of both simplifications are, therefore, consistent with the model requirements.

| **Maps after features Selection Process for different scales** | **T1 (m)** | **VD (m)** | **T2 (m)** | **Resulting Points** | **% Data reduction** | |
|---|---|---|---|---|---|---|
| Original map before feature selection (**721** points, **4** arcs) | | | | | from relative database | from original database (721 points) |
| 1:50,000 (**721** points, **4** arcs) | **4.67** | **89.48** | **2.00** | **323** | **55.21** | **55.21** |
| 1:100,000 (**721** points, **4** arcs) | **10.40** | **90.95** | **4.00** | **269** | **62.70** | **62.70** |
| 1:300,000 (**721** points, **4** arcs) | **33.20** | _ | _ | **90** | **87.52** | **87.52** |
| 1:500,000 (**356** points, **5** arcs) | **50.03** | _ | _ | **35** | **90.17** | **95.15** |
| 1:1,000,000 (**256** points, **1** arc) | **95.33** | _ | _ | **16** | **93.75** | **97.79** |
| 1:1,500,000 (**256** points, **1** arc) | **143.49** | _ | _ | **12** | **95.32** | **98.34** |

**Table 6.6: Digitally simplified maps from their 1:25,000 source databases (relative) after features have been selected from original (721 points, 4 arcs) according to the PCLS. (study area 4 -Drainage Coverage).**

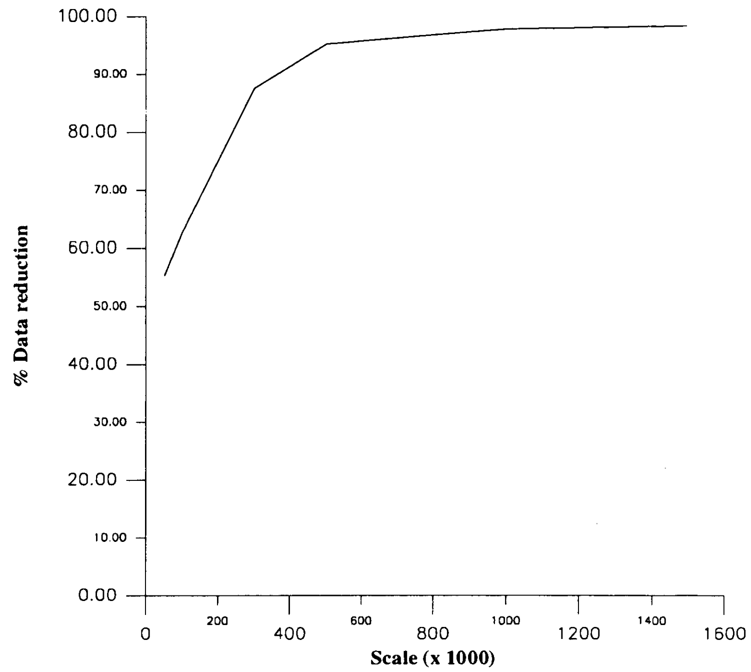



Figure 6.35: Scale against % Data reduction for digitally simplified maps (using the PCLS) from the original map (721 points, 4 arcs, study area 4 -Drainage Coverage).

Figure 6.35 graphically shows the effects of feature selection and both simplifications in terms of data reduction from the original database (721 points). The figure indicates that well below 90 % of data reduction is produced at the first three scales. As the figure shows, the rate of reduction gradually increases with reduced scale and noticeably increases at the scale of 1:300,000, whereas it generally resumes the same pace of increase but more slowly. This effect is caused by the relatively large reduction of data at the 1:300,000 scale, resulted from the relatively large tolerance used for this scale compared to the previous two larger scales (Table 6.6). Although both processes of feature selection and simplification contributed to data reduction at the last three scales, the amount of reduction is not comparable to that which occurred at the previous scales. As explained in the previous analyses, this is due to the fact that although the resulting data reductions at these small scales are all high, all fall within a small percentage in relation to the original database (87.52 to 95.32 %).


Figure 6.36: Digitally simplified map for the scale of 1:50,000 according to the Simplification Model (323 points,
1) Features Selection, 2) T 1= 4.699 m, 3) VD= 89.480 m, 4) T 2= 2 m. 55.21 per cent data reduction from
721 points original database (in Red). (Study Area 4 -Drainage Coverage).
Drawing line width: .15 mm, represents 7.5 m on the ground.

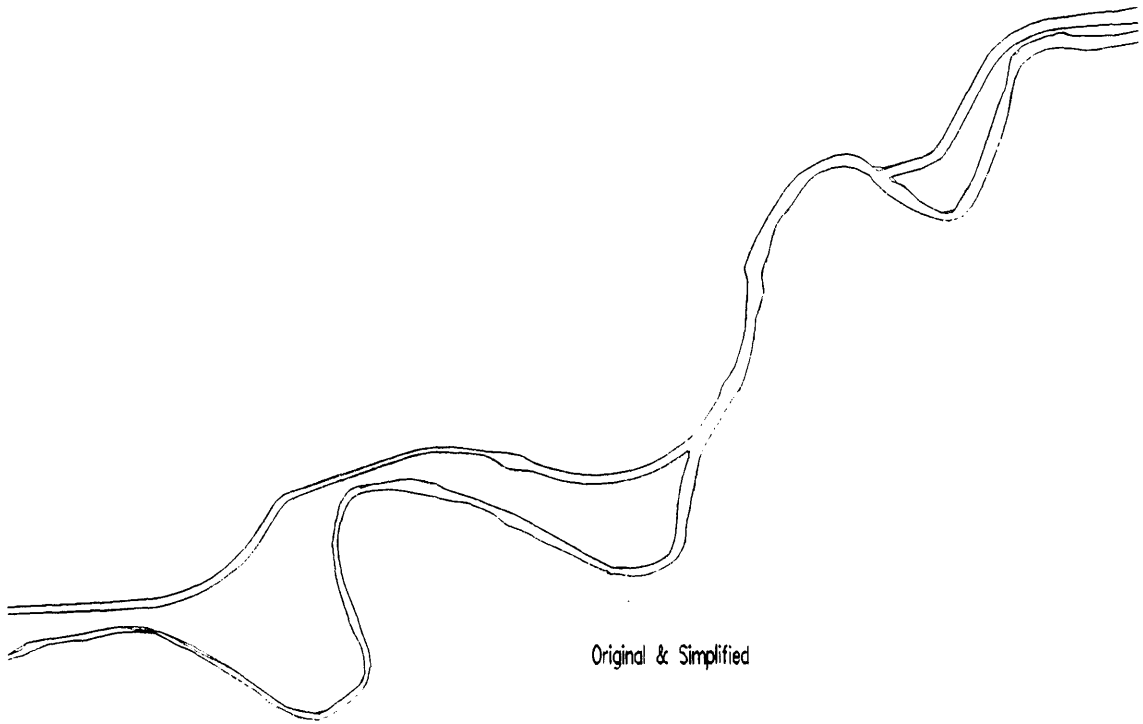



Figure 6.37: An overlay of original (in Red) and digitally simplified map (in Green) at 1:50,000
1) Feature Selection, 2) T 1= 4.699 m, 3) VD= 89.480 m, 4) T 2= 2 m, 323 points, 55.21 per cent data reduction from 721 points original database.

Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 4 -Drainage Coverage).

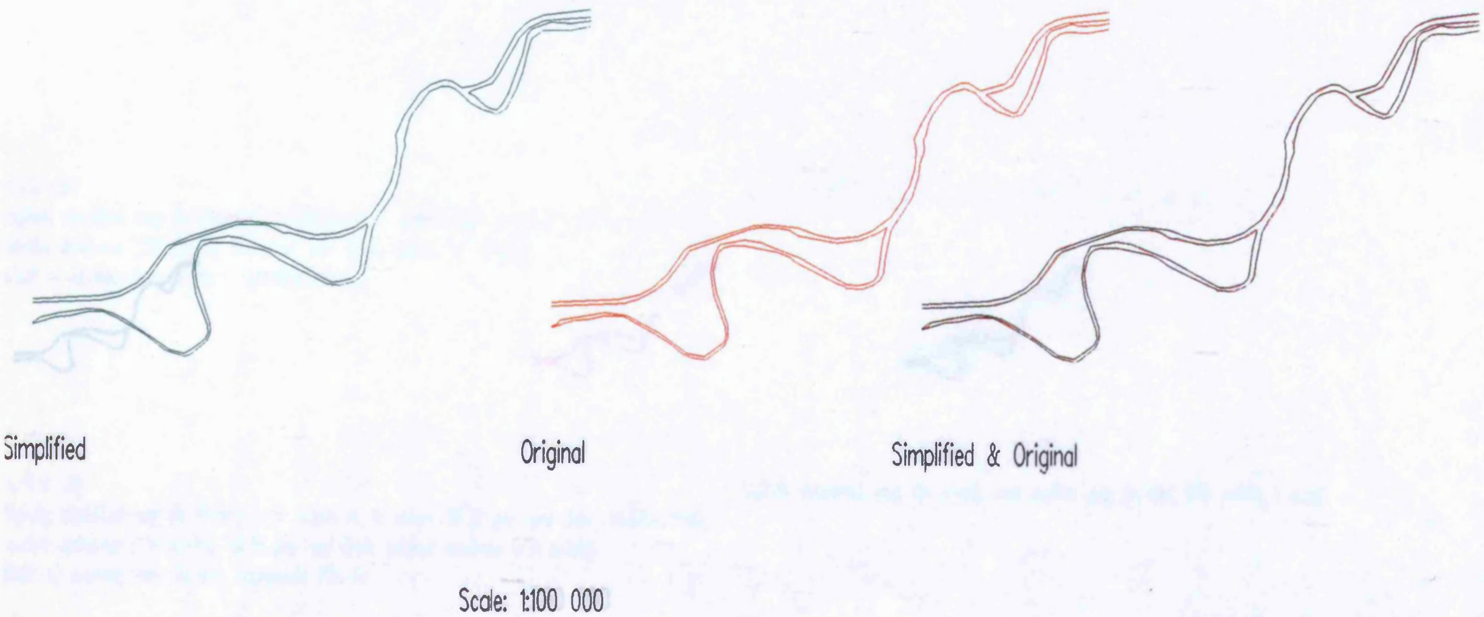



Figure 6.38: Two maps at 1:100,000. Original map in Red (721 points), and the simplified map (in Green), 269 points, 1) Feature Selection, 2) T1= 10.400 m, 2) VD= 90.950 m 3) T=4 m. 62.70 per cent data reduction.
Width of drawing line= .15 mm, represents 15 m on the ground. (Study Area 4 -Drainage Coverage).

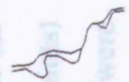

Figure 6.39: Two maps at 1:300,000. Original map in Red (721 points), and the simplified map (in Green), 90 points, 87.52 per cent data reduction.
Width of drawing line= .15 mm, represents 45 m on the ground. (Study Area 4 -Drainage Coverage).

1:500 000

Digitally simplified map (in Green), T 1= 50.029 m, 53 points, 90.17 per cent data reduction from redtive database (356 points), 95.15 per cent data reduction from original database (721 points).
Width of drawing line= .15 mm, represents 75 m.

1:1000 000
Digitally simplified map (in Green), T 1= 95.329 m, 16 points, 93.75 per cent data reduction from relative database (256 points), 97.79 per cent from original (721 points).
Width of drawing line= .15 mm, represents 150 m.

1:1 500 000
Digitally simplified map (in Green), T 1= 143.49 m, 12 points, 95.32 per cent data reduction from relative database (256 points), 98.34 per cent from original database (721 points).
Width of drawing line= .15 mm, represents 225 m.



Digitally simplified map (in Green), and relative map (in Red, 356 points, 5 arcs).

Digitally simplified map (in Green), and relative map (in Red, 256 points, 1 arc).

Digitally simplified map (in Green), and relative map (in Red, 256 points, 1 arcs).

Figure 6.40: Digitally simplified maps at 1:500,000, 1:1,000,000, and 1:1,500,000, according to the Simplification Model. (Study Area 4 - Drainage Coverage).

6.4.5.2 Transport Coverage (33 arcs, 319 points):

Six reductions were derived from the 1:25,000 source data for the transport coverage. According to the model requirements, the second simplification was applied on this coverage, since the database represents smooth and straight curves. As Table 6.7 reveals, no features were omitted for the first four scales. At the scale of 1:50,000, 199 points resulted from the simplification, a 37.62 % data reduction from the original database (319 points). At the 1:100,000 scale, 143 points resulted from the simplification, representing a 55.18 % data reduction. At the scale of 1:300,000, 87 points resulted, representing a 72.73 data reduction. At the fourth scale of 1:500,000, 66 points were retained, a 79.32 % data reduction. Figures 6.42, 6.43, 6.44, and 6.45 all show that there is no perceptible deviation between the simplified and original lines.

Table 6.6 indicates that the process of feature selection was performed at the remaining two scales of 1:1,000,000 and 1:1,500,000. Only two arcs (with 143 points) were selected for these two small scales. At the scale of 1:1,000,000, only 30 points were retained, representing a 79.03 % data reduction from the relative database (143 points), and a 90.60 % from the original (319 points). At the last scale of 1:1,500,000, only 26 points resulted, representing an 89.90 % data reduction from the relative database (143 points) and a 91.85 % from the original. Figure 6.46 displays the effect of the simplification at these two scales. The figure reveals an imperceptible difference between the original (in Red) and the simplified (in Green) lines. The results are, therefore, consistent with the model requirements, in the sense that the line details here do not require perceptible simplification, since they are essentially less detailed. However, there was some distortion of topology as a result of the simplification at the scales of 1:1,000,000, and 1:1,500,000, respectively (Figure 6.46.1, 6.46.2, and 6.46.3). These distortions were caused by the Douglas-Poiker algorithm.

| Maps after features Selection Process for different scales | T1 (m) | VD (m) | T2 (m) | Resulting Points | % Data reduction | |
|---|---|---|---|---|---|---|
| Original map before feature selection (319 points, 33 arcs) | | | | | from relative database | from original database (319 points) |
| 1:50,000 (319 points, 33 arcs) | 4.00 | _ | _ | 199 | 37.62 | 37.62 |
| 1:100,000 (319 points, 33 arcs) | 9.00 | _ | _ | 143 | 55.18 | 55.18 |
| 1:300,000 (319 points, 33 arcs) | 29.00 | _ | _ | 87 | 72.73 | 72.73 |
| 1:500,000 (319 points, 33 arcs) | 49.00 | _ | _ | 66 | 79.32 | 79.32 |
| 1:1,000,000 (143 points, 2 arcs) | 83.33 | _ | _ | 30 | 79.03 | 90.60 |
| 1:1,500,000 (143 points, 2 arcs) | 128.30 | _ | _ | 26 | 89.90 | 91.85 |

**Table 6.7: Digitally simplified maps from their 1:25,000 source databases (relative) after features have been selected from original (319 points, 33 arcs) according to the PCLS. (study area 4 -Transport Coverage).**

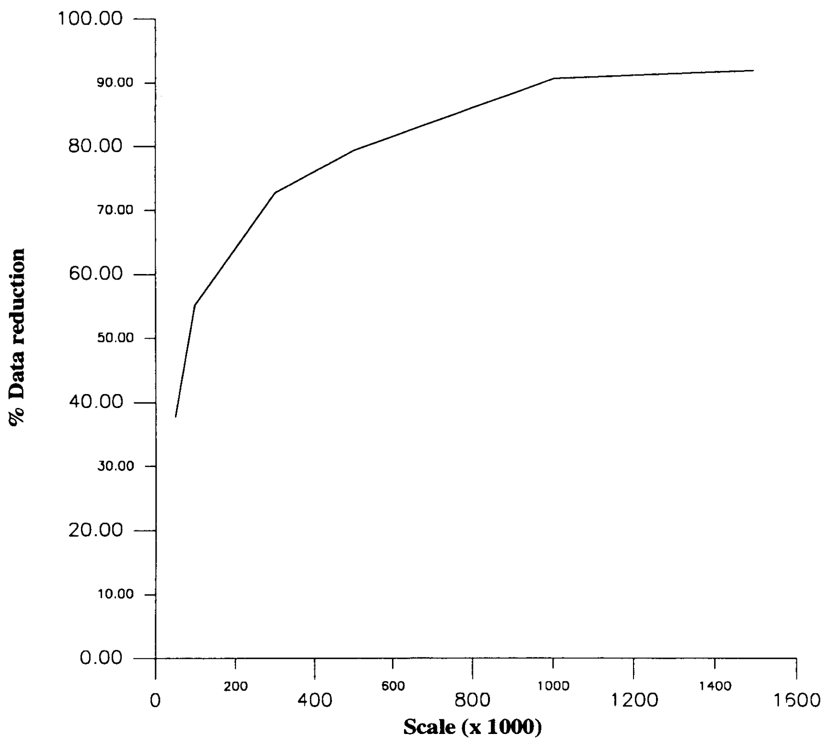



Figure 6.41: Scale against % Data reduction for digitally simplified maps (using the PCLS) from the original map (319 points, 33 arcs, study area 4 -Transport Coverage).

Figure 6.41 shows the effects of the processes of selection and simplification in terms of data reduction from the original database at all scales. As the figure reveals, the data reduction resulting from the simplification consistently increases with reduced scale, although it tends to be slower beyond the scale of 1:250,000 towards the last scale change. Although, there was a feature process performed at the last two scales, the data reduction resulted is largely due the simplification process.

Figure 6.42: Digitally simplified map at 1:50,000, 199 points, 37.62 per cent data reduction from 319 points original database. $T_t$= 4.000 m. Drawing line width: .15 mm, represents 7.5 m on the ground. (Study Area 4 -Transport Coverage).

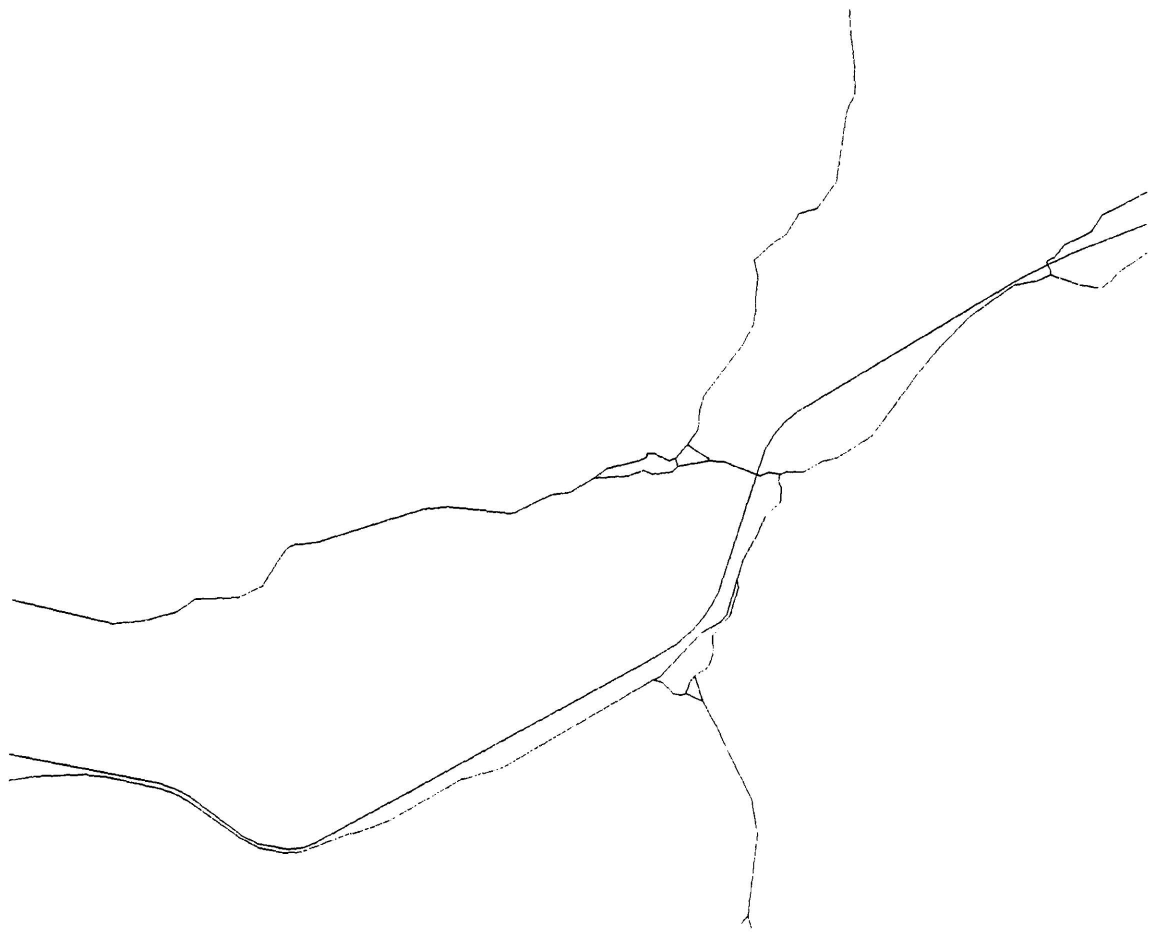


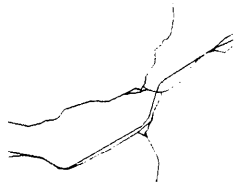
Figure 6.43: Two maps at 1:50,000. Original (in Red) without feature omission, 319 points, 33 arcs
Digitally simplified map (in Green), T 1= 4.000 m, 199 points, 37.62 per cent data reduction.
Drawing line width: .15 mm, represents 7.5 m on the ground.(Study Area 4 –Transport Coverage).

Simplified

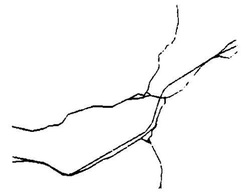
Simplified & Original

Scale: 1:100 000

Figure 6.44: Two maps at 1:100,000. Original (in Red) without feature omission, 319 points
Digitally simplified map (in Green), T 1= 9.000 m, 143 points, 55.18 per cent data reduction.
Width of drawing line= .15 mm, represents 15 m on the ground.(Study Area 4 -Transport Coverage).



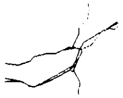
Simplified



Simplified & Original

Scale: 1:300,000

Figure 6.45: Two maps at 1:300,000. Original (in Red) without feature omission, 319 points
Digitally simplified map (in Green), T 1= 29.000 m, 87 points, 72.73 per cent data reduction.
Width of drawing line= .15 mm, represents 45 m on the ground. (Study Area 4 –Transport Coverage).



1:500 000

Digitally simplified map (in Green), T 1= 49.000 m, 66 points (79.32 per cent data reduction).
Width of drawing line= .15 mm, represents 75 m.

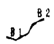
Digitally simplified map (in Green), and original map (in Red, 319 points, 33 arcs).

1:1000 000
Digitally simplified map (in Green), T 1= 83.334 m, 30 points, 79.03 per cent data reduction from reltive database (143 points), 90.60 per cent data reduction from original (721 points).
Width of drawing line= .15 mm, represents 150 m.

Digitally simplified map (in Green), and relative map (in Red, 143 points, 2 arcs).

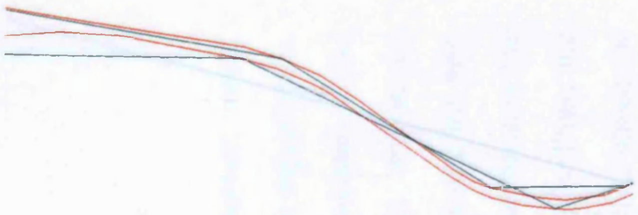
Distorted topology
Imperceptible: B

1:1 500 000
Digitally simplified map (in Green), T 1= 128.301 m, 26 points, 89.90 per cent data reduction from relative database (143 points), 91.85 per cent data reduction from original (721 points).
Width of drawing line= .15 mm, represents 225 m.

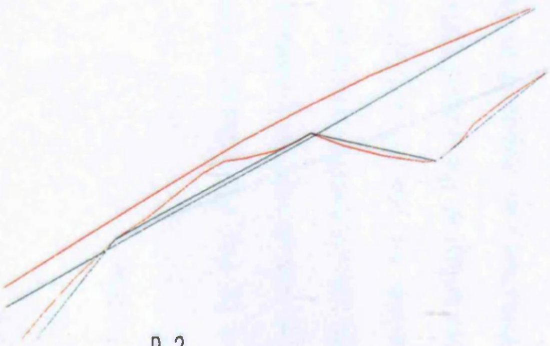


Digitally simplified map (in Green), and relative map (in Red, 143 points, 2 arcs).

Figure 6.46.1: Digitally simplified maps at 1:500,000, 1:1,000,000, and 1:1,500,000, according to the Simplification Model. (Study Area 4 -Transport Coverage).



B 1



B 2

Figure 6.46.2: Distorted topology resulted from the simplification produced for the 1:1,000,000 scale
The distortion is imperceptible (B) at the scale of 1:1,000,000, while it is enlarged at this figure to the scale of 1:20,000.
Original in Red and digitally simplified map in Green (see figure 6.46.1)

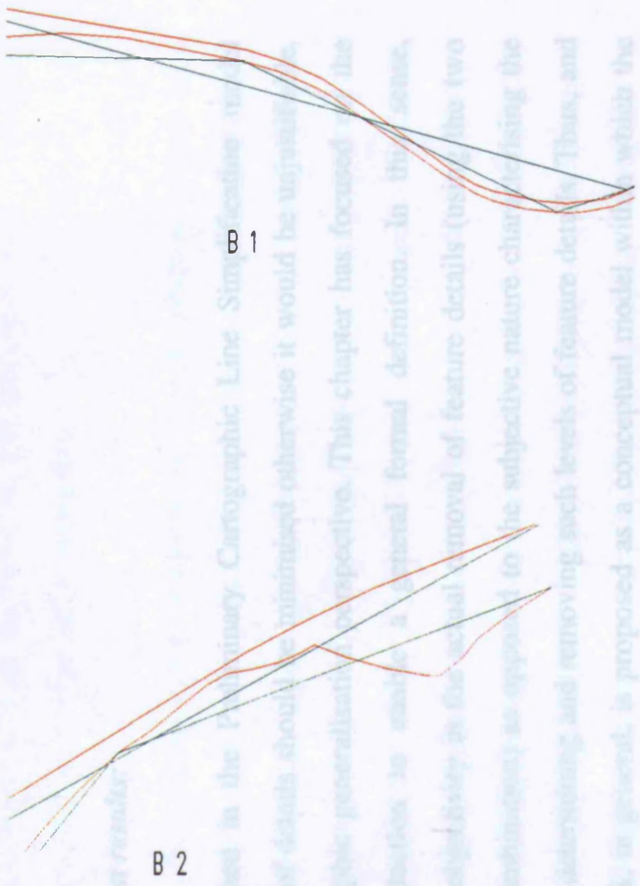



Figure 6.46.3: Distorted topology resulted from the simplification produced for the 1:1,500,000 scale
The distortion is imperceptible (B) at the scale of 1:1,500,000, while it is enlarged at this figure to the scale of 1:20,000.
Original in Red and digitally simplified map in Green (see figure 6.46.1)

## 6.5 Discussion:

### *6.5.1 Comparison of test results:*

As defined in the Preliminary Cartographic Line Simplification model (PCLS), the reduction of details should be minimised otherwise it would be unjustifiable, from a digital cartographic generalisation perspective. This chapter has focused on the magnitude of this reduction to enable a general formal definition. In this sense, formalisation refers to objectivity in the actual removal of feature details (using the two named algorithms in combination) as opposed to the subjective nature characterising the cartographer's work in determining and removing such levels of feature details. Thus, and as a solution, the PCLS, in general, is proposed as a conceptual model within which the formulation of such a *minimum* manipulation of line details is formally (or objectively) defined, enabling a fully automated process. Assessment of the proposed method should be performed against a backdrop of considerations including the stated requirements of the model. Some of these considerations should at least address the following questions: a) what if the model were to produce at the target scale largely or totally smoothed lines (i.e., line details are largely smoothed out), and/or; b) large or total removal of major forms or details of lines, and/or; c) large or total distortion of lines, or any other undesirable effect beyond the scope of simplification? The results produced by the model clearly showed that such effects did not occur in any test and at any scale. That is, the interaction between the Douglas-Poiker algorithm and the Cubic Spline algorithm could systematically and consistently reduce feature details according to the minimum manipulation criterion of line details proposed in the PCLS. On the other hand, the application of the process of feature selection for all mapping scales was performed subjectively, based on the existing generalised versions of the same line features, and in adherence to general cartographic guidelines. The amount of feature selection varied from one test to another and from one scale to another. This is consistent with a typical cartographic production. Thus, the implementation of the model allowed for reduction of small details that are usually regarded as redundant in terms of both cartographic communication (i.e., informative value), and digital processing (i.e., storage and manipulation).

The results showed that using the Douglas-Poiker algorithm as a hierarchical algorithm (i.e., different resolutions can be produced from a single database) is an effective method in scale-dependent digital mapping. This supports a current belief by many researchers that algorithms of this type "offer greater flexibility and efficiency in extracting reduced representations" (Weibel, 1995, p.262). However, this study stresses that this property is cartographically most desirable in the context of scale-dependent weeding or database simplification in the imperceptible realm, or according to the PCLS. In other words, the use of the Douglas-Poiker within the scale-dependent context is cartographically confined to the proposed simplification either in the perceptible or imperceptible realms. Furthermore, the proposed approaches in both this chapter and Chapter 5, present practical solutions as to the use of the Douglas-Poiker algorithm within the scale-dependent context. So far as Arc/Info is concerned, the proposed formulation of the integration between this algorithm and the Cubic Spline smoothing routine is, therefore, recommended for use especially during digital cartographic production. However, the distortion which results at some scales is undesirable and has to be interactively corrected where necessary before applying other generalisation processes. Furthermore, whilst the results of the last two scales in test 3 appear generally consistent with the model requirements, the results of the previous three scales fall short of an ideal reduction of data. That is, the databases could have undergone larger reductions whilst the effect would have been imperceptible at the target scales. Such a result highlights the problem of contriving general rules or formalising a flexible cartographic processes such as simplification, since some trade offs should be withstood for the sake of gains from the general rule. Nevertheless, the development and evaluation of the proposed formulation of the simplification role by the two algorithms used in this chapter, emphasise the fact that general rules are not valid within the simplification context. That is, various lines, map purposes and scales demand judicious considerations as to the application of either type of simplification, and even when the simplification is applied the formulation is designed to respond to basic principles in both the digital and analogue cartographic realms, and still has defined limits and merits. This observation is supported by a study presented by Monmonier and McMaster (1990) in which the authors studied the sequential effects of six digital algorithms in cartographic line generalisation. These processes are for: simplification, displacement, smoothing, enhancement, merging, and omission. They conclude that "... in a truly holistic approach to numeric generalisation, it will become increasingly necessary to view the generalisation operators as a complex interrelated package of techniques, not as isolated independent algorithms" (Monmonier and McMaster, 1990, pp. 93). For this reason, the interactive approach is most suited for this task, as it enables knowledge accumulation about both traditional and digital line simplification.

*6.5.2 Digital simplification versus manual simplification:*

Although, the process of the model simulates the manual work of a cartographer during a preliminary simplification of line details, the digital method surpasses the manual process in terms of consistency (i.e., uniform removal of line detailed) and efficiency (i.e., faster). For example, any simplified lines presented in the analyses, where the details were minimally reduced or processed, underwent a balanced reduction such that their general characters experience a uniform processing and similar levels of simplification. As noted in Chapters 2, and 4, manual simplification (in its complete role within generalisation) is essentially subjective, and hence inconsistent. If these reductions were to be performed by a cartographer twice or by different cartographers, different reductions would result each time. Joao (1994) shows examples where the inconsistency element in manual generalisation could produce negative impacts on a typical GIS-based operations mapping such as Overlay and Buffering. It is, therefore, crucial to aim for adopting digital processes where possible such as the form of simplification proposed, here.

The model allows for greater reduction of minor details when lines appear complex (e.g., see test 2, drainage coverage) while it tends to produce less reduction where the lines already contain less or no details (e.g., test 2, and 3, transport coverages). However, this study argues for the fact that, beyond this preliminary role of simplification, different generalisation processes and algorithm applications (using appropriate tolerance values) should follow in order to produce a consistent cartographic treatment of a line as well as other features. This conforms well to the principles of cartographic generalisation, in which all features are appropriately treated within the mapping context, in terms of achieving a consistent and balanced generalisation.

The significance of repeatability lies in the fact that users can always check results and parameters in the form of reports for two reasons. First, this should help in the process of acquiring knowledge about simplification, and hence formalising that knowledge. Second, it should serve as a reliable statement source describing the lineage of the cartographic product, which is increasingly gaining importance in digital cartography.

### *6.5.3 The digital simplification model in an operational digital generalisation scheme:*

As noted at the beginning of this chapter, previous schemes for digital generalisation proved hard to realise. Unlike some of the extreme and impractical cases that have been put forward in this regard (i.e., aiming for complete automation), this study argues for intermediate solutions whereby subjective and objective processes are mixed. In this section, a conceptual model is presented for a typical semi-automatic digital generalisation within which the proposed simplification model can be integrated. Figure 6.47 shows a set of digital generalisation processes that might follow the preliminary simplification process.

The first stage in the figure indicates that once the simplification has been performed on single line coverages, they and other features which have been selected for representation should be overlaid at either the same production scale or at larger scale. This should allow for resolving conflicts which resulted from the preliminary simplification, and to enable the application of other generalisation operations within the mapping context (i.e., scale, purpose, design requirements). Generalising a composite of overlays is made easier within digital mapping systems such as GISs. In the second stage of the model, the required generalisation processes are then considered and applied. In this stage, several iterations of the application of these processes might well be necessary according to the mapping context. The processes can be performed interactively where possible. Of specific interest here, is the possible application of other types of simplification algorithms. In this sense, simplification algorithms such as Brophy's (1973) that focus on the general form of line features are particularly recommended for application to small scale production. Brophy's algorithm requires more than one input parameter, allowing for more considerations about the lines being simplified. Many of these and other types of algorithms are supported by the MGE Map Generalizer programme, described in section 6.1.2.1. So far as the proposed generalisation model is concerned, this study supports the previous observation about the limitation of the Douglas-Poiker algorithm in producing satisfactory shapes at gross levels of simplification. This is a fact not so much related to the distortion that the algorithm tends to generate, but rather the fact that it tends to preserve minor details at the expense of the more important ones, i.e., major forms, which is contrary to the cartographic rule dictating that emphasis should be given to major forms of line features during small scale production (Thapa 1987, Whyatt, 1991, 1993, Visvalingam 1995). Like any applied generalisation algorithm or operation during generalisation, if another process of simplification is to be performed, checking for conflicts might well be necessary. Finally, the Douglas-Poiker algorithm might be used for removing any redundant data resulted from the simplification process. In this case, the third formula ($T_2$) in the simplification model (section 6.3.2.2) is recommended for use. The third step in the model is concerned with the actual process of symbolisation and lettering. In this process, all features are to be assigned the appropriate symbolisation in terms of graphic form (i.e., size, colour, and form), codes, and place names are attached to the features. This stage, obviously, starts during earlier stages, since the decision about their inclusion and their required space is firstly considered during the processes of feature selection, and classification. Furthermore, symbolisation requires further checking and adjusting.

The fourth stage is concerned with the process of the production itself. The process of colour proofing is performed first, where the generalised map in different colours is produced for checking on the mapping material (e.g., paper) at the required scale. If approved, the process of saving in stage 5 should be applied first, so that other printing processes including direct plotting, and colour separation for large scale production can be achieved through colour separation.

The final stage is related to the process of digitally saving this product and its associated *log* (lineage) file, in which all the processes performed are documented. The lineage file should serve as a significant contribution towards understanding generalisation operations, and in the case of the MGE program, such a file can be recalled and applied on similar types of generalisation tasks.

Based on the results and the proposed model of line generalisation, here, it can be seen that the simplification takes different forms in the digital realm. For example, at a particular scale and generalisation, the simplification can be in the perceptible as well as imperceptible realm. It can also be concerned with both the minor and major line forms. Given such a complexity, simplification is, therefore, achieved interactively and objectively where possible in a typical digital production. Such a structure and context should provide another effective scope for acquiring knowledge about the interaction of line simplification with other generalisation processes, as well as understanding other generalisation processes.

## Digital Line Generalisation Model

*For a particular scale production*

(1) **Overlaying different simplified (preliminary simplified coverages) and prepared coverages at equal or larger production scale for:**
- resolving conflicts, and
- mapping-context generalisation

(2) **Applying generalisation algorithms and processes:**
e.g., displacement, exaggeration, combination, etc.

Application of simplification processes:
*If necessary,* apply algorithms that are concerned with the general form of the line during small scale production. (e.g., Brophy's algorithm).

Checking for conflicts

*If necessary,* applying the Douglas-Poiker algorithm for weeding purposes as in stage three of the preliminary simplification model

(3) **Symbolisation and lettering**

(4) **Production:** colour proofing, saving and reproduction

(5) **Final saving and reporting**

Figure 6.47: A proposed Digital Cartographic Line Generalisation.

## 6.6 Summary

This chapter explored the scope of digital line simplification for producing cartographically acceptable products for cartographic production purposes. The chapter first indicated that previous studies do not adopt the definition and role of simplification as perceived within cartographic generalisation. Previous attempts in digital generalisation were first referred to, followed by presentation of a proposed model for cartographic line simplification within the digital generalisation concept. In this model, two digital simplification methods were employed; namely the Douglas-Poiker algorithm and Cubic Spline smoothing algorithm. This experiment was performed within the Arc/Info GIS, in which the two algorithms are implemented. The model was based on the design requirements of the two algorithms, the simplification requirements, and on the mechanism of digital map generalisation. The model was, therefore, set to reduce data and details on the line features before applying other generalisation processes. This reduction was proposed to be minimal so that line details in both the perceptible and imperceptible realm that are regarded as redundant at target production scales from cartographic and economic perspectives have to be removed. Since no objective definition of such reduction has been proposed, the formulation proposed in the model is regarded as a working definition as to how much details and data have to be reduced. The results showed that the integration of the two digital methods, generally proved acceptable with respect to these requirements (section 6.3.2). That is, for a particular scale line details and data were reduced without oversimplification which would be unacceptable. The model can be used directly within digital mapping systems and is, therefore, recommended for single or multiple preliminary simplifications during digital cartographic production. The results showed that performing a simplification, such as the one proposed, was not straightforward, however, highlighting the complex role of the process of simplification within the wider context of generalisation. It is, therefore, recommended that an interactive approach should be adopted for producing quality results and acquiring knowledge about line simplification. The results in this chapter re-enforce the premise of this study which states that both development and evaluation of digital techniques for cartographic generalisation should consider the whole context of this complex process. This implies that the implementation of the formulation proposed is limited to its design objectives, since processing line details within digital generalisation can not be objectively and entirely solved by the combination of the two simplification algorithms tested. It is, thus, suggested that further studies should focus on testing other types of simplification algorithms.

# CHAPTER SEVEN: CONCLUSION AND FURTHER RESEARCH

## 7.1 Conclusion

This study focused on the issue of cartographic line simplification, and to what extent it can be formalised in the digital realm. The process was examined within a cartographic context. That is, line simplification is seen as a sub-process within cartographic generalisation, and as such an understanding of its role had to be explored in order to embody it within automated systems. This study contends that the problem of formalising line simplification is not so much of the existing formulation of line simplification (i.e., simplification algorithms) as much as it is the lack of understanding of the simplification process itself in the traditional realm. A scheme was proposed as a methodology by which line simplification can be examined in both the traditional (manual) and digital paradigms. The scheme first involved an examination of the role of line simplification within cartographic generalisation. Second, it involved an evaluation of a typical widely used simplification algorithm (in this case, the Douglas-Poiker algorithm) according to its design specifications and objectives. Finally, the scheme included searching for cartographic quality in the algorithm's results within the context of the previous stages. Since this work deals with a visual perception-based cartographic process, the analysis is, therefore, principally perceptual. A list of the main findings and conclusions is given below.

### *7.1.1 Line simplification as a process within generalisation in the traditional realm:*

Chapter 2 presented a theoretical background about the distinction between the two processes of generalisation and simplification. The empirical study in Chapter 4, examined the role of line simplification within generalisation. Four study areas were selected and each was represented by different generalised maps. The tests were primarily focused on the process of line simplification in relation to the effects of map purpose, feature character, and mapping techniques. The tests lead to the following conclusions:

1) In spite of confusing factors and issues discussed Chapter 4, it was found that generally complex features required complex interaction and application of the generalisation processes, including simplification.

2) The analyses indicated that the interaction between the three factors looked at and other factors are difficult to ignore and isolate. It proved necessary to present general conclusions as to what might have contributed most to the simplification process in each test. In general terms, it was found that feature character is the most important factor influenced the process of simplification of line details, especially in study areas 1 and 2. Map purpose and mapping techniques were the focus of the analyses in study areas 3 and 4, respectively. The influence of these two factors can not be ignored in study areas 1 and 2. Similarly, the effect of feature character is implied in study areas 3 and 4. This implies that the line simplification was not influenced solely by single factors. Thus, it can be concluded that the relative influence of factors such as these discussed here on line simplification may be ranked. For example, given a generalised map, the influence of feature character on line simplification is the most important factor followed by the factors of map purpose and mapping techniques.

3) There were some important reasons for the difficulty in isolating the factors influencing line simplification. First, the subjective nature, and hence elusiveness of manual simplification. Second, isolation of every relative influence of each factor is difficult to achieve. For example, influence of factors of only map purpose and mapping technique may be isolated only if the same line features were simplified for different map purpose and by different mapping techniques. Apart from certain types of line features (e.g., international boundaries), it is, in fact, difficult to come across a line feature in several scale maps and for multiple map purposes, and processed by different mapping agencies. Third, line simplification is only a single context-dependent process. As such, a cartographer processes line details with reference to the mapping context, and this processing is, as pointed out above, subjective. Finally, a thorough

analysis of the effects of all possible factors that may influence line simplification is well beyond the scope of this thesis, due to its technical constraints imposed by the research work and topic.

4) It is essential to recognise that a given generalised line at a particular scale is only a composite product of all generalisation processes. In other words, representation of a line feature at a particular scale is not seen as a product of simplification alone, and hence any digital formulation of both simplification and generalisation must be evaluated according to this context.

### *7.1.2 Re-evaluation of the Douglas-Poiker algorithm:*

The effectiveness of the Douglas-Poiker algorithm for line simplification was examined in this study. The examination was based on the design objectives of the algorithm; that is data reduction (Chapter 5). In other words, the algorithm is designed to weed data that are deemed redundant at particular scales from economic and cartographic perspectives. Economically, data can require large storage space and lengthy processing time, whereas cartographically, they do not perceptually bear any information value. The evaluation was performed within the scale-dependent context on different sets of line features and at multiple scales. The findings of the evaluation are:

1) Effects of graphic reduction (or as termed here, *areal scale*) were significant according to the evaluation context; and hence it was necessary to consider the levels of data reduction suitable in representations at different scales. It was, therefore, proposed that for practical reasons the level of data reduction might reasonably be determined in relation to the function of linear scale for large and medium scale, while approximately a 90 % reduction would be at least sufficient at small scales. For example, if a line was reduced in scale by 50 percent, then its database should be reduced at least by the same percentage. However, it was proposed that this criterion should be applied to certain types of line features which are irregular lines and other

types of line curves (e.g., smooth curves representing, for example, railroads). In this case, straight lines or regular geometrical shapes (e.g., town plans and street guides) are excluded. This is for the simple reason, that those straight lines or regular-shaped features are usually represented only by critical points and further reduction of those points would necessarily lead to perceptible distorted shapes at target scales, which is an undesirable effect in terms of scale-dependent data reduction. Furthermore, this criterion should be applied to the number of points of feature databases but not to their length, and those points should be relatively large; for example, databases composed of hundreds or thousands of points. Based on this requirement of data reduction, a tolerance prediction formula (TPF) was proposed to produce a direct input parameter as a tolerance value to be used by the Douglas-Poiker algorithm to achieve this task for a given scale reduction (section 5.4.2.2). This is for two reasons. First, there is need for a formula that accommodates this context. Second, the Radical Law (RL) proved unsuitable, as it tends to yield redundant data (section 5.4.2.2), and it does not consider the evaluation context. Generally, the algorithm using the TPF, as opposed to a modified formula (MR1) based on the Radical Law, could successfully produce multiple data reductions from single databases consistent with the proposed requirements. Furthermore, the data reduction resulting from the algorithm reached up to 95 + %, yet no perceptible difference was detected between the original and simplified lines at the target scales. However, the evaluation revealed that the numerically desirable results do not necessarily coincide with the graphically desirable results throughout the analyses. This was due to factors such as the proportion of the number of arcs to the number of points (or, the degree of segmentation of the lines), rate of scale change, and complexity of the line details. For example, due to the effects of line segmentation, and line details (less details), the Douglas-Poiker algorithm produced data reduction of less than 50 % for the coverages of study area 3, and the transport coverage in study area 4 (see Table 5.28 and associated discussion). This raises an important observation regarding the difficulty of applying general rules in cartographic simplification and generalisation. Generally, the algorithm proved successful in selecting the points that are regarded as most critical for representation at target reduced scales. The algorithm is, therefore, shown to be suitable for weeding purposes within the scale-dependent context. The shape distortion that might have resulted in the imperceptible realm is regarded as irrelevant from the cartographic representation perspective.

2) The results showed that data reduction reaches a limit of between approximately 75 and 90 %, beyond which the effects of increased tolerance values were limited, irrespective of the line complexity. This is primarily due to the fact that the algorithm's behaviour is affected by the position of individual points but not by their numbers. Hence, data reduction is largely achieved by applying certain tolerance values, while beyond this large reduction a progressive increase in the tolerances tends to have little effect (sections, 5.4.2.2, 5.5.1.1, and 5.5.4.3). It was observed that this level of reduction approximately corresponds to a range of tolerance values from 20 to 30 m. Thus, and according to the evaluation context, the algorithm's performance on different types of line features was invariant. However, the results revealed that this limit is further intensified by the effects of factors such as the level of complexity of details and the number of arcs within the lines. For example, lines would undergo less data reduction if they were less complex (i.e., contain limited data) and consist of a relatively large number of arcs (i.e., largely segmented).

3) The results indicated that applying a single tolerance value for weeding purposes on a single coverage composed of different line features would be a reasonable choice, since applying suitable single tolerances for each segment (i.e., arc) of line features is unjustifiable according to the evaluation context, cartographic and economical considerations. Cartographically, it would lead to unbalanced data-reduced or simplified line segments or lines, whereas line features of different forms and complexity stored in a digital file are only a geometrical problem according to which the algorithm presents a consistent result if a single tolerance value was applied. On the other hand, the coverages, as can be seen from the analyses, are essentially compiled from segmented lines. Economically, processing every line segment in a coverage is time consuming and costly.

4) A general method was proposed to quantify the effects of graphic reduction on the resulting simplification by the algorithm. The method involves comparing the resulting

areal displacements (in map units) between the original and simplified lines with the actual graphic extent of the drawing simplified lines. All results showed that the total areal thickness (in map units) of the simplified lines at every target scale are larger than the total areal displacements (in map units) by a ratio of the difference to the areal thickness of the lines extended at some scales well above 99 %. This general quantification supported the graphical results in which the total areal displacements between the simplified and original lines were imperceptible at target scales.

5) The significance of this evaluation lies first in the applicability of the TPF as a practical means for predicting the tolerance value needed when using the Douglas-Poiker algorithm for a weeding process for a given scale. Second, the evaluation re-emphasised that the algorithm has proved again its ability as being a hierarchical method that is most valuable in the process of digital generalisation as generalisation processes are assumed to work on minimum databases. This is important, since it would be economical to process these data and store them, especially for small scale productions.

### *7.1.3 Digital simplification within digital generalisation:*

Given the above conclusions about line simplification in the traditional realm and the Douglas-Poiker algorithm, the study looked at the possibility of utilising the capability of the algorithm for a simplification role in digital cartographic production. Chapter 6 was dedicated for this task, and a number of conclusions arise:

1) Digital cartographic generalisation is different in practice from traditional cartographic generalisation. Whilst digital generalisation operators or processes have to be performed in a holistic context as that in traditional generalisation, they have to be applied repeatedly. This is due to the sequential mode of the digital process itself. This implies that the role of each operator differs from one application or stage to another. Thus, digital line simplification during digital generalisation should be carried out in different stages, since exhaustive processing of line details is not the task of simplification alone.

2) Given the requirements of digital generalisation, the role of simplification within generalisation, and the sequential mode of the digital methods used, the Douglas-Poiker algorithm integrated with the Cubic Spline smoothing algorithm was utilised to define a simplification role in the formal realm of digital cartographic production. A minimum processing of line details is proposed as an appropriate criteria for simplification by the integration of the two algorithms. Initially, this formulation of minimum processing was determined visually; that is qualitatively, so that line details at a particular reduced scale should be minimally reduced either in the perceptible or the imperceptible realm, depending on the complexity of the details and the map purpose and scale. Both types of simplification are concerned with removing the redundant details at the target scales; that is details that have no cartographic value and are economically costly. If the line feature appears smooth then the simplification required is to be in the form of *database simplification* (i.e., data reduction or weeding), whereas if the line appears complex then a process termed *shape/data simplification* is to be applied. According to the model, the role of simplification should vary according to the mapping context; i.e., scale, purpose, and type of line features. The model was referred to as the Preliminary Cartographic Line Simplification (PCLS) model, implying that further generalisation processes including simplification will follow this first step (section 6.3.2). An interactive approach within the Arc/Info mapping system was utilised for implementing the PCLS, so that the mapping context can be considered, as in a typical cartographic production. For this purpose, a program was written for allowing the processes of feature selection and simplification to be executed by the user; i.e., the cartographer.

3) Implementation of the simplification model on a selected set of line features proved successful in minimally reducing line details objectively at multiple scale productions from single databases, accompanied by a desirable minimal exaggeration of the minor details which were retained. The results were based on perceptual assessment, because no existing measurements in regard to this minimum reduction of line details exist. It was suggested that the proposed formulation can be regarded as an objective method for determining and measuring the level of minimum reduction proposed. However,

the process of shape/data simplification caused some shape distortion in places where line features were close to each other. The distortion took the form of crossed and/or connected lines. Unlike in the weeding process, this distortion whether in the perceptible or imperceptible realm, has to be corrected where necessary before other generalisation processes take place, as these processes are assumed to work on topologically-correct features.

4) Implementation of the model clearly suggested that the role of line simplification is complex in a digital production context. For example, the use of the Douglas-Poiker algorithm took two forms: one as an integrated use with another algorithm of different type of simplification (smoothing), and the other as an independent use, as for data reduction or weeding. Also, the weeding process took two forms: one as an integrated step in the process of shape/data simplification, where small tolerance values were applied, and the second as an independent use of the algorithm applying the TPF values for the process of database simplification.

5) The significance of this model relates to its practicality, so that at target production scales the model can be utilised within GIS-based generalisation programs such as the MGE Generalizer. For example, in such interactive programs, the user (cartographer) can interact with various generalisation operators, but a useful step would be to work in databases that have had their redundant data and/or details removed. As might be expected, the process of generalisation would be a lot easier and faster than working on large details and/or data. Furthermore, it can be recommended that the greater the complexity of line details the greater the benefit from the application of the PCLS. The significance of this model also lies in the fact that its framework highlighted the real merits as well as limits of the Douglas-Poiker algorithm in producing cartographically acceptable or even tolerable results in the digital cartographic production. For example, the model shows that the algorithm has a role in the actual process of line simplification which should be <u>minimal</u> and most importantly should be <u>guided.</u> The concept of minimum here relates to the minimum usability for particular scales, line features, and purposes. It is therefor suggested that the use of the algorithm beyond

this proposed application is not recommended from the cartographic point of view, and other algorithms will be the appropriate subject of further evaluations.

6) Given the conclusions from analyses in Chapters 4, 5 and 6, the Douglas-Poiker algorithm should not to be used for generalisation purposes without the intervention of the user (cartographer). Applying this algorithm alone on certain cartographic features without considering the complex requirements of generalisation is, as the thesis has highlighted, an inappropriate cartographic practice. The algorithm should, therefore, be judiciously utilised for simplification purposes, since manipulation of line details is a task shared by other generalisation processes. Furthermore, the application of the algorithm should be confined to minimal simplification, since uncontrolled simplification leads to production of caricatural shapes that are perceptible at target scales. Uncontrolled application of the algorithm simply ignores the wider context of simplification, let alone generalisation.

## 7.2 Further research

In concluding the research presented here, there is no belief that the complexity of line simplification can be solved by a single research project. Further research is therefore required in key areas of this subject highlighted by this study:

1) Line simplification as a manual process has not been thoroughly studied, and it has been studied only as a digital process. Further exploration of the traditional generalised products is, thus, essential to gain further knowledge about line simplification, utilising GIS capabilities. In this sense, specific analyses are recommended for different types of line features with varied complexities which are generalised for different map scales and purposes. Such analyses provide sound understanding of this process upon which the simplification algorithms can be adequately tested.

2) Other simplification algorithms of different types should be evaluated in the context suggested in this thesis, as this will further the understanding about the algorithms as well as about the simplification. Powerful interactive visualisation techniques are highly recommended for this purpose.

3) It is of great value that the integration between simplification and generalisation operations are cross-referenced. This implies that a meaningful understanding about simplification alone would not be completely realised unless improved knowledge about the other generalisation operators is acquired.

## REFERENCES


Abraham, I., (1988), *Automated Cartographic Line Generalisation and Scale-independent Databases*. Unpublished Master's dissertation, Polytechnic of Wales, C.N.A.A., U.K.

Arnheim, R., (1974), *Art and Visual Perception: A Psychology of the Creative Eye*. Los Angeles, CA: University of California Press.

Attneave, F., (1954), Some Informational Aspects of Visual Perception. *Psychological Review* 61: 183 - 193.

Barber, C., Cromley, R., and Andrle, R., (1995), Evaluating alternative line simplification strategies for multiple representations of cartographic lines. *Cartography and Geographic Information Systems*, Vol. 22, No. 4, pp. 276 - 290.

Batty, M., Fotheringham, S., and Longley, P., (1989), Urban growth and form: scaling, fractal geometry and diffusion limited aggregation. *Environment and Planning A*, Vol. 21, No. 11, pp. 1447 - 1472.

Beard, K., (1988), *Multiple Representations From a detailed Database: a Scheme for Automated Generalisation*. Unpublished PhD thesis, University of Wisconsin, Madison, Wisconsin.

Beard, K., (1991), Constraints on Rule Formation. In Buttenfield, B. P. and McMaster, R. B. (Eds.) *Map Generalisation: Making Rules for Knowledge Representation*, pp. 121 - 135, Longman: Harlow.

Bertin, J., (1983), *Semiology of Graphics: Diagrams, Networks, Maps*. Translator: W. Berg. Madison, W I: University of Wisconsin Press.

Blakemore, M., (1984), Generalisation and Error in Spatial Data Bases. *Proceedings AUTO-CARTO 6*, Fall Church, Virginia, pp. 313 - 322.

Board, C., (1967), Maps as Models. In: Chorley, R., and Haggett, P., (Eds.) *Models in Geography*, Methuen, London. pp. 671 - 725.

Boyle, A., (1970), The Quantised Line. *The Cartographic Journal*, Vol. 7, No. 2, pp. 91 - 94.

Brassel, K., (1985), Strategies and Data Models for Computer-Aided Generalisation. *International Yearbook of Cartography*, Vol. 25, pp. 11 - 29.

Brassel, K. and Weibel, R., (1988), A review and conceptual framework of automated map generalisation. *International Journal of Geographical Information Systems*, Vol. 2, No. 3, pp. 229 - 244.

Brophy, D., (1973), An automated methodology for linear generalisation in thematic cartography. *Proceedings, ACSM 33rd Annual Meeting*, pp. 300 - 314.

Buttenfield, B., (1985), Treatment of the Cartographic line. *Cartographica*, Vol. 22, No. 2, pp. 1 - 26.

Buttenfield, B., (1986), Digital definitions of scale-dependent line structure. *Proceedings, AUTO-CARTO 6*, London, England, Vol. 1, pp. 487-506.

Buttenfield, B., (1989), Scale dependence and self-similarity of cartographic lines. *Cartographica*, Vol. 26, No. 1, pp. 79 - 100.

Buttenfield, B., (1991), A rule for describing line feature geometry. In: Buttenfield, B. and McMaster, R., (Eds.) *Map Generalisation: Making Rules for Knowledge Representation*, Longman: Harlow, pp. 150 - 171.

Buttenfield, B. and McMaster, R., (Eds.) (1991), *Map Generalisation: Making Rules for Knowledge Representation*, Longman: Harlow

Carstensen, W., (1988), Fractal self-similarity in cartographic line generalisation. Paper presented, *Annual Meetings Association of American Geographers*, Phoenix, Arizona.

Chrisman, N., (1983), Epsilon filtering: a concept for automated scale change. *Proceedings of American Congress on Surveying and Mapping, 43rd Annual Meeting*, pp. 322 - 331.

Chrisman, N., (1989), Error in categorical maps: testing versus simulation. *Proceedings, AUTO-CARTO 9*, Baltimore, Maryland, April 2 - 7, pp. 521 - 529.

Christ, F., (1978), A program for the fully automated displacement of point and line features in cartographic generalisations. *Nachrichten aus dem karten- und Vermessungswesen* Series, Vol. 35, pp. 5 - 30.

Clarke, K., (1990), *Analytical and Computer Cartography*. New Jersey: Prentice Hall.

Cuenin, R., (1972), Cartographie Générale: Tome 1, *Notions générales et principles d'elaboration*, Editions Eyrolles, Paris.

Delucia, A. Robert, T., (1987), A comprehensive approach to automatic feature generalisation. *Proceedings, 13th International Cartographic Conference*, Morelia, Mexico, pp. 169 - 192.

Dettori, G. and Falcidieno, B., (1982), An algorithm for selecting main points on a line. *Computers and Geosciences*, Vol. 8, No. 1, pp. 3 - 10.

Deveau, T., (1985), Reducing the number of points in a plane curve representation. *Proceedings, AUTO-CARTO 7*, Washington, DC., pp. 152 - 160.

Douglas, D. and Peucker, T., (1973), Algorithms for the reduction of the number of points required to represent a digitised line or its character. *The Canadian Cartographer*, Vol. 10, No. 2, pp. 112 - 123.

Duda, R. and Hart, P., (1973), *Pattern Classification and Scene Analysis*. John Wiley: New York.

Dutton, G., (1981), Fractal enhancement of cartographic line detail. *The American Cartographer*, Vol. 8, No. 1, pp. 23 - 40.

Eckert, M., (1908), On the nature maps and map logic. Translated by W. Joerg, *Bulletin of the American Geographical Society*, Vol. 40, No. 6, pp. 344 - 351.

Environmental Systems Research Institute, *Understanding GIS: the ARC/INFO Method*, Redlands, CA: ESRI, 1991.

Fisher, P. and Mackaness, W., (1987), Are cartographic expert systems possible? *Proceedings, AUTO-CARO 8*, Baltimore, Maryland, pp. 530 - 534.

Forrest, A., (1985), Computational Geometry in Practice. Fundamental Algorithms for Computer Graphics, NATO ASI Series, Edited by R. A. Earnshaw, Vol. F17, pp. 707 - 724.

Freeman, H., (1978), Shape description via the use of critical points. *Pattern Recognition*, Vol. 10, pp. 159 - 166.

Geraint, L., Jones, C., and Furse, E., (1995), Holistic generalisation of large-scale cartographic data. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 106 - 119, Taylor & Francis: London.

Goodchild, M., (1980a), Fractals and the accuracy of geographic measures. *International Association for Mathematical Geology Journal*, Vol. 12, No. 2, pp. 85 - 98.

Goodchild, M., (1980b), The effects of generalisation in geographical data encoding. In: Freeman, H., and Pieroni, G., (Eds.), *Map Data Processing*. Academic Press: New York. pp. 191 - 205.

Goodchild, M., (1988), The issue of accuracy in global databases. In: Mounsey, H., (Eds.) *Building Databases for Global Science*. Taylor & Francis: Basingstoke, pp. 31 - 48.

Goodchild, M., (1991), Issues of quality and uncertainty. In: Muller, J., (Eds.) *Advances in Cartography*, Barking, Essex: Elsevier, pp. 113 - 139.

Goodchild, M., and Mark, D., (1987), The fractal nature of geographic phenomena. *Annals Association American Geographers*, Vol. 72, No. 2, pp. 265 - 278.

Gordan, N., McMahon, T., and Finlayson, B., (1992), *Stream hydrology: An introduction for ecologists*. John Wiley & Sons: London.

Grunreich, D., (1993), Generalisation in GIS environment. *Proceedings, 16th International Cartographic Conference*, Cologne, pp. 203 - 210.

Grunreich, D., (1995). Development of computer-assisted generalisation on the basis of cartographic model theory. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 47 - 55, Taylor & Francis: London.

Hakonson, L., (1978), The length of closed geomorphic lines. *Mathematical Geology*, Vol. 10, No. 2, pp. 141 - 167.

Huggins, S., (1991), *Programming in C*, transputer@prg.oxford.ac.uk mailing list, Software Development Group, INMOS Ltd, 1000 Aztec West, Almondsbury, Bristol.

Hole, F., and Campbell, J., (1985), *Soil Landscape Analysis*. Totowa, N.J.: Rowan and Allanheld.

*International Cartographic Association (ICA) 1973, Multilingual Dictionary of Technical Terms in Cartography* (Wiesbaden, F. R., Germany: Franz Steiner Verlag).

Jenks, G., (1979), Thoughts on line generalisation. *Proceedings of AUTO-CARTO IV*, PP. 209-220.

Jenks, G., (1981), Lines, Computers and Human Frailties. *Annals of the Association of American Geographers*, Vol. 71, No. 1, pp. 1 - 10.

Jenks, G., (1989), Geographic Logic in Line Generalisation. *Cartographica*, Vol. 26, No. 1, pp. 27 - 42.

Joao, E., (1994), *Causes and Consequences of Map Generalisation*. Unpublished Ph.D. Thesis, Birbeck College, University of London.

Johannsen, T., (1973), A program for editing and for some generalising operation for derivation of a small scale map from digitised data in 1:50,000 scale. *Nachrichten aus dem karten- und Vermessungswesen* Series, Vol. 2, No. 30, pp. 17 - 22.

Jones, C., and Abraham, I., (1986), Design considerations for scale-independent cartographic data base. *Proceedings $2^{nd}$ International Symposium on Spatial Data Handling*, Seattle, W., pp. 384 - 398.

Jones, C., and Abraham, I., (1987), Line generalisation in global cartographic database. *Cartographica*, Vol. 24, No. 3, pp. 32 - 45.

Keates, J., (1989), *Cartographic Design and Production.* (second edition). Longman: London.

Keates, J., (1996), *Understanding Maps.* (second edition). Addison Wesley Longman Ltd: London.

Krantz, S., (1989), Fractal geometry. *The Mathematical Intelligencer*, Vol. 11, No. 4, Springer-Verlag, New York.

Lang, T., (1969), Rules for the Robot Draughtsman. *The Geographical Magazine*, Vol. 42, No. 1, pp. 50 - 51.

Lee, D., (1993), From master data base to multiple cartographic representations. *Proceedings, 16th International Cartographic Conference*, Cologne, Germany, Vol. 2, pp. 1075 - 1085.

Lee, D., (1995), Experiment on formalising the generalisation process. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 219 - 234, Taylor & Francis: London.

Leopold, L., Wolman, M., and Miller, J., (1964), *Fluvial processes in geomorphology.* Freeman and Company: San Francisco and London.

Li, Z. and Openshaw, S., (1992), Algorithms for automated line generalisation based on a natural principle of objective generalisation. *International Journal of Geographical Information Systems*, Vol. 6, No. 6, pp. 373 - 389.

Li, Z. and Openshaw, S., (1993), A natural principle for the objective generalisation of digital maps. *Cartography and Geographic Information Systems*, Vol. 20, No. 1, pp. 19 - 29.

Lichtner, W., (1978), Locational Characteristics and the Sequence of Computer Assisted Processes of Cartographic Generalisation. *Nachrichten aus dem karten-und Vermessungwegen*, series, Vol. 35, pp. 65 - 75.

Lundquist, G., (1959), Generalisation: A preliminary survey of an important subject. *The Canadian Surveyor*, Vol. 14, No. 10, pp. 466-470.

Maguire, D., (1986), Generalisation, fractals, and spatial databases. *Bulletin of The Society of University Cartographers*, Vol. 20, No. 2, pp 96 - 99.

Mandelbrot, B., (1967), How long is the coast of Britain? *Science*, Vol. 156, pp. 636 - 638.

Mandelbrot, B., (1982), *The fractal geometry of nature*. W. H. Freeman: San Francisco.

Marino, J., (1979) Identification of Characteristics of Along Naturally Occurring Lines: An Empirical Study. *The Canadian Cartographer*, Vol. 16, No. 1, pp. 70 - 80.

McMaster, R., (1986), A Statistical Analysis of Mathematical Measures for Linear Simplification. *The American Cartographer*, Vol. 13, No. 2, pp. 103 - 116.

McMaster, R., (1987a), Automated Line Generalisation. *Cartographica*, Vol. 24, No. 2, pp. 74 - 111.

McMaster, R., (1987b), The Geometric Properties of Numerical Generalisation. *Geographical Analysis*, Vol. 19, No. 4, pp. 330 - 346.

McMaster, R., (1989), The Integration of Simplification and Smoothing Algorithms in Line Generalisation. *Cartographica*, Vol. 26, No. 1, pp. 101 - 121.

McMaster, R., (1995), Knowledge acquisition for cartographic generalisation: experimental methods. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 161 - 179, Taylor & Francis: London.

McMaster, R. and Shea, S., (1988), Cartographic Generalisation in a Digital Environment: A Framework for Implementation in a Geographic Information System. *proceedings, GIS/LIS'88*, San Antonio, Texas, pp. 240 - 249.

McMaster, R. and Shea, K., (1992), *Generalisation in Digital Cartography*. Association of American Geographers, Washington, DC.

Monkhouse, F., (1975), *Principles of physical geography*. Hodder and Stoughton: London.

Monmonier, M., (1983), Raster-mode area generalisation of land use and land cover maps. *Cartographica*, Vol. 20, No. 4. pp. 65 - 91.

Monmonier, M., (1986), Towards a practical model of cartographic generalisation. *Proceedings, AUTO-CARTO 6*, London, England, Vol. 2, pp. 257 - 266.

Monmonier, M. and McMaster, R., (1990), The sequential effects of geometric operators in cartographic line generalisation. *International Yearbook of Cartography*, pp. 93 - 108.

Morehouse, S., (1995), GIS-based map compilation and generalisation. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 21 - 30, Taylor & Francis: London.

Morrison, J., (1974), A theoretical framework for cartographic generalisation with emphasis on the process of symbolisation. *International Yearbook of Cartography*. Vol. 14, pp. 115 - 127.

Muller, J.C., (1986), Fractal dimension and inconsistencies in cartographic line representations. *The Cartographic Journal*, Vol. 23, pp. 123 - 130.

Muller, J.C., (1987a), Fractal and automated generalisation. *The Cartographic Journal*, Vol. 24, pp. 27 - 34.

Muller, J.C., (1987b), The concept of error in cartography. *Cartographica*, Vol. 24, No. 2, pp. 1 - 15.

Muller, J.C., (1991b), Generalisation of spatial data bases. In: Maguire, D. and Goodchild, M. and Rhind, D., (Eds.), *Geographical Information Systems: Principles and Applications*, Vol. 1, pp. 457 - 475, Longman: Harlow.

Muller, J.C., Weibel, R., Lagrange, J., and Salge, F., (Eds.), (1995a), *GIS and Generalisation, Methodology and Practice*. Taylor & Francis: London.

Muller, J., Weibel, R., Lagrange, J., and Slage, F., (1995b), Generalisation: state of the art and issues. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 3 - 18, Taylor & Francis: London.

Nickerson, B., (1988), Automated cartographic generalisation for linear features. *Cartographica*, Vol. 25, No. 3, pp. 15 - 66.

Nickerson, B. and Freeman, H., (1986), Development of a rule-based system for automatic map generalisation. *Proceedings, 2nd International Symposium on Spatial Data Handling*, Seattle, pp. 537 - 556.

Opheim, H., (1982), Fast data reduction of a digitised curve. *Geo-Processing*, Vol. 2, 33 - 40.

Pannekoek, A., (1962), Generalisation of coast lines and contours. *International Yearbook of Cartography*, Vol. 2, pp. 55 - 74.

Penck, A., (1894), *Morphologie der Erdoberflache*. Stuttgart.

Perkal, J., (1966), An attempt at objective generalisation. Translated by W. Jackowski. In: Nystuen, J., (Eds.), *Discussion Paper 10, Michigan Inter-University Community of Mathematical Geographers*, Ann Arbor, Michigan.

Powitz, B. and Schmidt, C., (1992), 'CHANGE', Internal Technical Report, Institute for Cartography, University of Hannover, pp. 16.

Raisz, E, (1962), *Principles of Cartography*. John Wiley: New York.

Ramer, Urs., (1972), An iterative procedure for polygonal approximation of plane curves. *Computer Graphics and Image Processing*, Vol. 1, pp. 244 - 256.

Ratajski, L., (1967), Phénomènes des points de généralisation. *International Yearbook of Cartography*, Vol. 7, pp. 143 - 151.

Reumann, K. and Witkam, A., (1974), Optimising curve segmentation in computer graphics. *Proceedings, International Computing Symposium*, North-Holland Publishing Company, Amsterdam, pp. 467 - 472.

Rhind, D., (1973), Generalisation and realism within automated cartography. *The Canadian Cartographer*, Vol. 10, No. 1, pp. 51 - 62.

Roberge, J., (1985), A data reduction algorithm for planner curves. *Computer Vision, Graphics and Image Processing*, Vol. 29, pp. 168 - 195.

Robinson, A., (1953), *Elements of Cartography*. 1st Edition, John Wiley: New York.

Robinson, A., and Sale, R., (1969), Elements of Cartography. 3rd Edition, John Wiley: New York.

Robinson, A., Morrison, J., Muehrcke, P., Kimerling, A., and Guptill, S., (1995), *Elements of Cartography*. 6th Edition, John Wiley & Sons: New York.

Robinson, G. and Zaltash, A., (1989), Application of expert systems to topographic map generalisation. *Proceedings, AGI 89*, 'GIS - A Corporate Resource', pp. A.3.1 - A.3.6.

Shea, K., (1991), Design considerations for an artificially intelligent system. In: Buttenfield, B. P. and McMaster, R. B. (Eds.) *Map Generalisation: Making Rules for Knowledge Representation*, pp. 3 - 20, Longman: Harlow.

Smith, N., (1980), The automated generalisation of small scale to cartographic maps, with particular reference to the Ordnance Survey 1:50,000 scale series. Unpublished M A thesis, Durham University.

Speiss, E., (1995), The need for generalisation in a GIS environment. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 31 - 46, Taylor & Francis: London.

Srnka, E., (1970), The analytical solution of regular generalisation in cartography. *International Yearbook of Cartography*, No. 10, pp. 48 - 60.

Steward, H., (1974), *Cartographic Generalisation, Some Concepts and Explanation. The Canadian Cartographer*, Monograph No. 11.

Swiss Society of Cartography (1987), *Cartographic Generalisation - Topographic Maps*. Cartographic Publication Series, No. 2, 2nd edition. Zurich: Swiss Society of Cartography.

Thapa, K., (1987), Critical Points Detection: The First Step to Automatic Line Generalisation. Unpublished Dissertation. Department of Geodetic Science and Surveying. The Ohio State University, Columbus, Ohio.

Thapa, K, (1988), Automatic line generalisation using zero-crossings. *Photogrammetric Engineering and Remote Sensing*, Vol. 54, No. 4, pp. 511 - 517.

Tobler, W., (1964), *An experiment in the computer generalisation of maps.* Technical Report No. 1, ONR Task, ASAD Contract 459953, Washington, D C.

Tobler, W., (1966), Numerical Map Generalisation. In: *Michigan Inter-University Community of Mathematical Geographers, Discussion Paper 8*, (Eds.), John Nystuen. Ann Arbor: University of Michigan.

Topfer, F. and Pillewizer W., (with notes by D. H. Maling) (1966), The principles of selection, a means of cartographic generalisation. *The Cartographic Journal*, Vol. 3, No. 1, pp. 10 - 16.

van Horn, E., (1985), Generalising cartographic data bases. *Proceedings, AUTO-CARTO 7, Baltimore, Maryland, pp. 532 - 540.*

van Oosterom, P., (1989), Reactive data structure for Geographic Information Systems. Proceedings, AUTO-CARTO 9, pp. 665 - 674.

van Oosterom, P., (1991), The reactive-tree: a storage structure for a seamless, scaless geographic database. Proceedings, AUTO-CARTO 10, pp. 393 - 407.

van Oosterom, P., (1993), *Reactive Data Structures for Geographic Information Systems.* Oxford University Press: Oxford.

Visvalingam, M. and Whyatt, J., (1991), The Douglas-Peucker algorithm for line simplification: re-evaluation through visualisation. *Computer Graphics Forum*, Vol. 9, No. 3, pp. 213 - 228.

Visvalingam, M. and Williamson, P., (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, pp. 264 - 275.

Watson, W., (1970), A study of the generalisation of small scale map series. *International Yearbook of Cartography*, Vol. 16, pp. 65 - 95.

Waugh, T. and McCalden, J., (1983), *GIMMS Reference Manual (4.5)*, Gimms Ltd, 30 Keir Street, Edinburgh.

Weibel, R., (1986), Automated cartographic generalisation. In: Sieber, R. and Brassel, K., (Eds.), *A Selected Bibliography on Spatial Data Handling: Data Structures, Generalisation, and Three-Dimensional Mapping*, Vol. 6, pp. 20 - 35. Geo-Processing Series, Department of Geography, University of Zurich.

Weibel, R., (1991), Amplified intelligence and rule based systems. In: Buttenfield, B. P. and McMaster, R. B. (Eds.) *Map Generalisation: Making Rules for Knowledge Representation*, pp. 172 - 186, Longman: Harlow.

Weibel, R., (1995), Three essential building blocks for automated generalisation. In: Muller *et al.*, (Eds.) *GIS and Generalisation: Methodology and Practice*, pp. 56 - 69, Taylor & Francis: London.

Weibel, R., (1996), A topology of constraints to line simplification. *Proceedings, 7th International Symposium of Spatial Data Hnadling*, pp. 533 - 546.

Williams, C., (1978), An efficient algorithm for the piecewise linear approximation of planer curves. *Computer Graphics and Image Processing*, Vol. 8, pp. 286 - 293.

White, E., (1985), Assessment of line generalisation algorithms using characteristic points. *The American Cartographer*, Vol. 12, No. 1, pp. 17 - 28.

Whyatt, J., (1991), *Visualisation and Re-evaluation of Line Simplification Algorithms*. Unpublished Ph.D. Thesis, University of Hull, U. K.

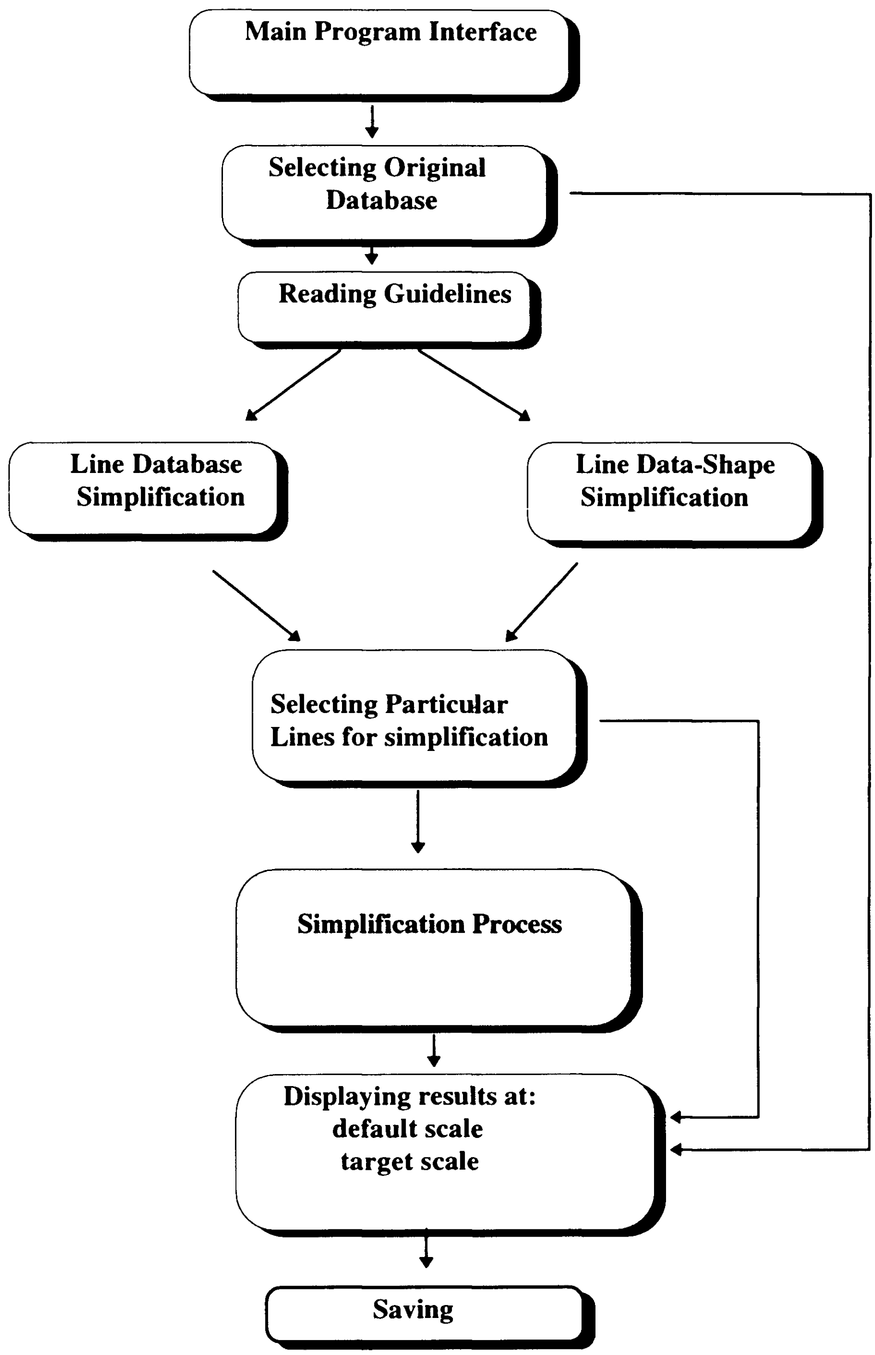
Whyatt, J. and Wade, P., (1988), The Douglas-Peucker line simplification algorithm. *Society of University Cartographers Bulletin*, Vol. 22, No. 1, pp. 17 - 25.

Wright, J., (1942), Map makers are human. *Geographical Review*, No. 32, pp. 527 - 544.

Zoraster, S., Davis, D., and Hugus, M., (1984), *Manual and Automated Line Generalisation and Feature Displacement*. US Army Engineering and Topographic Laboratory Report, ETL-0359, Belvoir, Virginia.

# APPENDIX A

## A.1 A flow diagram of the operation sequences of the program for the Preliminary Cartographic Line Simplification model:



## A.2 Preliminary Line Simplification program

```
/* ***************************************************************************/
/* Program for: Line Simplification                                          */
/* Date: September 1995.                                                     */
/* Author: Ali M. Al-Ghamdi                                                  */
/* Description: This an AML code which performs an Interactive Simplification */
/*              of line features, within the Arc/Info GIS. It runs different graphical */
/*              interfaces allowing the user to choose the line features and the */
/*              simplification required.                                     */
/* ***************************************************************************/


/* Running the main menu

  &terminal 9999

/*--- Check which program is running

   &if [show graphic] &then
     display 9999
   &else
     ap
     display 9999


/*-- SET THE PATHS TO THE AML AND MENU FILES

   &amlpath amls

   &menupath menus

/* background

  shadeset color

  shadesymbol 5
  patch 0 0 10 7

  move 1 4
  textsymbol 35
  textsize 0.3

  textcolor 2
  text 'INTERACTIVE SIMPLIFICATION ROUTINE (ISR)'

  &menu continue.menu &sidebar &position &below &display &stripe 'CONTINUE'

  &menu main.menu &form &size 550 500 &position &left &display ~
       &stripe 'INTERACTIVE SIMPLIFICATION ROUTINE'

  QUIT
```

```
1 pull down menu for continue
'-- OK --' &return

&RETURN

7 MAIN.MENU
 INTERACTIVE SIMPLIFICATION

 %today

 INPUT COVERAGE %1

 PARAMETERS :          %help

 SOURCE SCALE  %2

 DERIVED SCALE  %3



 %first        %second

 %third         %save

 %cancel



%today  display datevar 32

%1 input .incov 32 help 'Enter coverage to simplify' ~
  required cover * -all -other 'Select a coverage'
%2 input .sscale 7 help 'Enter source scale' ~
  required integer
%3 input .dscale 7 help 'Enter derived scale' ~
  required integer
%first    button '  DATA REDUCTION    ' &menu firstm.menu
%second   button '  SIMPLIFICATION    ' &menu secondm.menu
%third    button 'Display Original Coverage' &menu disp1.menu
%save     button '       SAVE        ' &menu  saves.menu
%cancel   button cancel '  QUIT   ' &return
%help     button return ' READ FIRST   ' &popup help.txt
%formopt  nextfield same
%forminit &s datevar := [date -cal]

/* Database simplification */

7 firstm.menu
 LINE DATABASE SIMPLIFICATION

 %today


 %first        %second


 %third         %forth
```

```
%fifth        %sixth

%cancel

%today  display datevar 32
%first    Button ' WHOLE COVERAGE  ' &run redfull.aml
%second    button ' SELECTED LINES  ' &run redpart.aml
%third   button '  Display To Scale  '  &menu disp.menu
%forth   button 'Default Display Scale' &run  result.aml
%fifth   button 'Display Original Coverage' &menu disp1.menu
%sixth   button 'Display Original & Simplified' &menu disp2.menu
%cancel   button cancel '  GO BACK   ' &return
%formopt  nextfield same
%forminit &s datevar := [date -cal]


/***********************************************
/*
/* If the whole coverage is to be simplified this
/* AML file will copy the original coverage into
/*a new coverage name. And then runs the relevant
/* AML files and MENUS
/*



   &if [exist result1 -cover] &then
     arc kill result1


/*  COPY THE ORIGINAL COVER TO NEW NAME

   arc copy %.incov% result1


/*  RUN THE RELAVENT FILES AND MENUS

   &run dcalculate.aml

   &type ' derived scale is ' %.dscale%


 &RETURN


/*****************************************************
/*
/*        This AML file is used to make calculations on
/*        the Original coverage.  These calculations
/*        includes finding the number of ARCS, ARC
/*        SEGMENTS, AND POINTS.  These values are
/*        assigned as global variables
/*



   &describe result1 .
   &s .Ar1 = %DSC$ARCS%
```

```
&s .Ns1 = %DSC$SEGMENTS%


&type ' the result is ' %.Ar1% %.Ns1%

&s .S     = %.Ns1% + 1

&s .T2  =  %.dscale% / %.sscale%

&if %.sscale% >= 25000 &then

   &s .T2 = %.T2% * 2.0

 &else
  &do

     &s .T2 = %.T2%

 &end

&s T  = ( LN %.S% ) / 2.3

&s .T1  = ( %T% * %.T2% ) - 1


&type 'THE VALUE FOR T IS ' %.T1%
&type 'THE LOG OF NUMBER OF POINTS IS ' %T%

 &if [exist result2 -cover] &then
  arc kill result2

arc generalize  result1 result2 %.T1%

arc build result2 line

&if [exist result4 -cover] &then
  arc kill result4

arc copy result2 result4

&RETURN

/*****************************************************
/*
/*        If only selected lines from a particular coverage
/*        are to be generalised, the this AML file will
/*        allow the user to select those lines writeselect
/*        to a file.  And then form a new coverage from
/*        those lines.  It also runs the relevant AML
/*        files and MENUS
/*



  &if [exist result1 -cover] &then
   arc kill result1

/* DISPLAY ORIGINAL COVERAGE ON SCREEN
```

```
    reset

    mapext %.incov%
    linecolor 1
    arcs %.incov%

/*  ALLOW USER TO MAKE SELECTION

    &thread &create ' MESSAGE ONE '  &menu message.menu &position &UR

    reselect %.incov% line many *
    writeselect selected %.incov% arc

    &thread &delete ' MESSAGE ONE '

7  message.menu
  TO SELECT A LINE

  --- POSITION THE MOUSE POINTER OVER IT ---
  --- CLICK LEFT MOUSE                     ---
  --- SELECT AS MANY LINES AS YOU WANT  ----

  --------- ONCE FINISHED PRESS 9 ----------


/* CREATE result1 COVERAGE FROM THE SELECTED LINES

   arc reselect %.incov% result1 arc selected


/* BUILD THE NEW COVERAGE

  arc build result1 line


/*  RUN THE RELAVENT FILES AND MENUS

   &run dcalculate.aml

&RETURN



7 disp.menu
 DISPLAYING TO SCALE



 DERIVED SCALE %1

 %display              %cancel

%1 input .scale1 10 help 'Enter Scale' ~
  required char
%display    button       '  DISPLAY    '  &run scale.aml
%cancel  button cancel '  GO BACK  ' &RETURN
```

```
/* scale.aml
/*--- Check which program is running

   &if [show graphic] &then
    display 9999
   &else
    ap
    display 9999


   mapext result4
   mapscale automatic
   linecolor gold
   arcs result4


&RETURN

/*
7 disp1.menu
 DISPLAYING TO SCALE



 DERIVED SCALE %1

 %display          %cancel

% 1 input .scale1 10 help 'Enter Scale' ~
  required char
%display   button     '  DISPLAY   ' &run scale1.aml
%cancel  button cancel '  GO BACK  ' &RETURN


/*--- Check which program is running

   &if [show graphic] &then
    display 9999
   &else
    ap
    display 9999

   mapext %.incov%
   pageunits cm
   pagesize 30 20
   maplimits 0 0 30 20
   mapunits meters
   mapscale automatic
   mapscale %.scale1%


   linecolor 1
   arcs %.incov%

/* Overlay of Original & Simplified
7 disp.menu
```

DISPLAYING TO SCALE

DERIVED SCALE %1

%display %cancel

```
%1 input .scale1 10 help 'Enter Scale' ~
  required char
%display   button      '  DISPLAY    '  &r overlay.aml
%cancel  button cancel '  GO BACK  ' &RETURN

/* Overlay.aml



/*--- Check which program is runing
&if [show graphic] &then
     display 9999
   &else

     ap
     display 9999


   mapext result4
   pageunit cm
   pagesize 30 20
   maplimits 0 0 30 20
   mapunits meters
   mapscale %.scale1%
   linecolor 1
   arcs result1
   linecolor gold
   arcs result4

&RETURN


&RETURN
/* Line data-shape simplification */
7 secondm.menu
```

INTERACTIVE LINE DATA-SHAPE SIMPLIFICATION

%today

%first %second

%third %forth

%fifth %sixth

%cancel

%today display datevar 32

```
%first    Button ' WHOLE COVERAGE  ' &run fullcov.aml
%second    button ' SELECTED LINES  ' &run partcov.aml
%third    button '  Display To Scale  '  &menu disp.menu
%forth    button 'Default Display Scale' &run  result.aml
%fifth    button 'Display Original Coverage' &menu disp1.menu
%sixth    button 'Display Original & Simplified' &menu disp3.menu

%cancel   button cancel '  GO BACK   ' &return
%formopt  nextfield same
%forminit &s datevar := [date -cal]


   &if [exist result1 -cover] &then
     arc kill result1


/*  COPY THE ORIGINAL COVER TO NEW NAME

   arc copy %.incov% result1


/*  RUN THE RELAVENT FILES AND MENUS

   &run oldres.aml

   &run newres.aml


   &describe result1
   &s .Ar1 = %DSC$ARCS%
   &s .Ns1 = %DSC$SEGMENTS%

&type ' the result is ' %.Ar1% %.Ns1%

   &s .S      = %.Ns1% + 1

   &s .T2  =  %.dscale% / %.sscale%

   &s T  = ( LN %.S% ) / 2.3

   &if %.sscale% <= 25000 &then

   &s .T1 = ( %T% * %.T2% ) - 1

   &else
    &do

     &s .T2 = %.T2% * 2.0

   &s .T1 = ( %T% * %.T2% ) - 1
    &end

   &type 'THE LOG OF NUMBER OF POINTS IS ' %T%
   &type 'THE VALUE OF T1 IS ' %.T1%

&RETURN
```

```
&if [exist result2 -cover] &then
  arc kill result2

arc generalize result1 result2 %.T1%

arc build result2 line


stat result2.aat info
sum length
&watch data1.dat
end
&watch &off

&s file2 [open data1.dat status2 -read]
&if %status2% = 0 &then
&s line [read %file2% error1]
&s line [read %file2% error]
&s .l [extract 3 [unquote %line% ] ]

&s dommy [close %file2%]



&describe result2

&s .Nsd = %DSC$SEGMENTS%

/*&type ' the result is ' %.Nsd%



&s A   = %.Nsd% / %.Ns1%
&s B   = %.l% / %.Nsd%

&s C   = %B% * %A%

&s D   = ( LN %C% ) / 2.3

&s .VD  = %C% * %D%

&type 'the value is'  %D%
&type 'THE VALUE OF VD IS'  %.VD%

&if [exist result3 -cover] &then
 arc kill result3


arc arcedit
mapext result2
ec result2 arc
select all
grain %.VD%
spline
save result3
quit
```

```
&if [exist result4 -cover] &then
  arc kill result4

&if %.T2% <= 10.00 &then

arc generalize  result3 result4 %.T2%

&else
  &do

    &s .T2 = %.T2% * 0.5
&type 'the value is' %.T2%
    arc generalize  result3 result4  %.T2%
  &end

arc clean result4


&RETURN


&if [exist result1 -cover] &then
  arc kill result1

/* DISPLAY ORIGINAL COVERAGE ON SCREEN

reset

mapext %.incov%
linecolor 1
arcs %.incov%

/* ALLOW USER TO MAKE SELECTION

&thread &create ' MESSAGE ONE '  &menu message.menu &position &UR

reselect %.incov% line many *
writeselect selected %.incov% arc

&thread &delete ' MESSAGE ONE '

/* CREATE result1 COVERAGE FROM THE SELECTED LINES

arc reselect %.incov% result1 arc selected


/* BUILD THE NEW COVERAGE

arc build result1 line


/* RUN THE RELAVENT FILES AND MENUS

&run oldres.aml

&run newres.aml

/* Save Menu
```

7 saves.menu
SAVING THE RESULT COVERAGE

OUT COVERAGE %1

%save %cancel

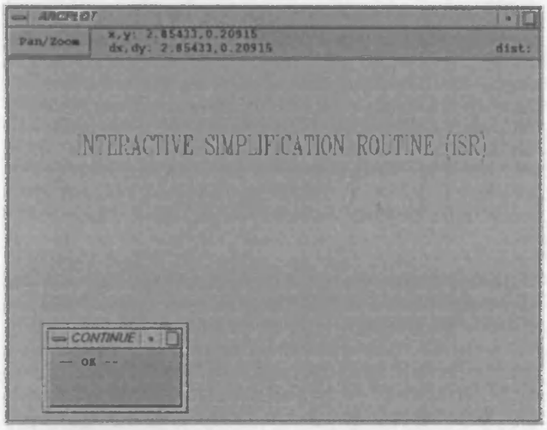
```
%1 input .outcov 32 help 'Enter cover Name' ~
  required char
%save   button     '  SAVE   ' &run saving.aml
%cancel  button cancel '  GO BACK  ' &RETURN
```

file is used to save the final

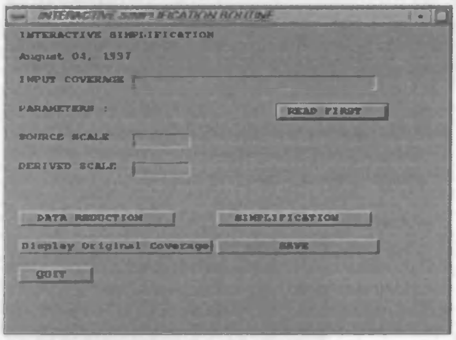
```
/* saving.aml

   arc copy result4 %.outcov%
   arc kill result4
   arc kill result3
   arc kill result2
   &RETURN
•
/*   end   */
```

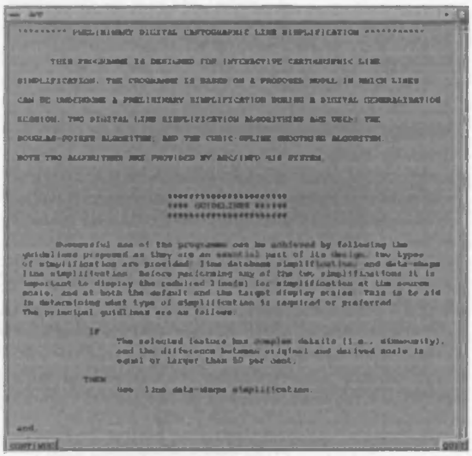
## A.3 Outputs of the User Graphical Interfaces generated by the program:



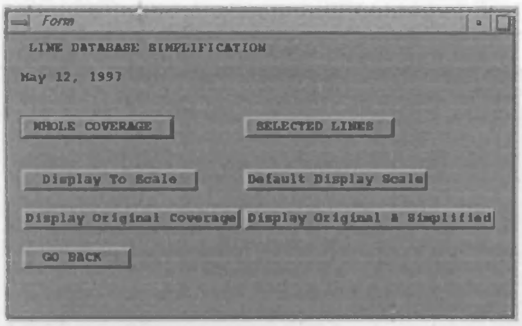

← Opening Screen Interface of the Interactive Line Simplification program.



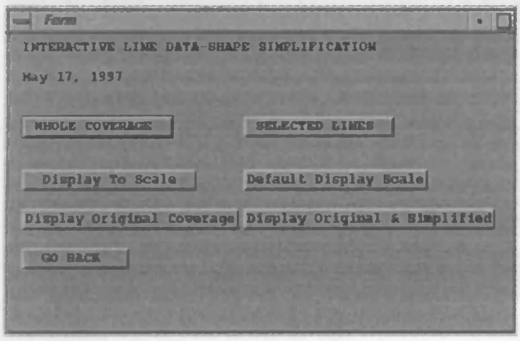

← Main Driving Menu of the program.



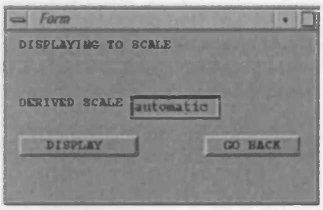
← Information and Guidelines of the purpose and use of the program activated by the READ FIRST button.



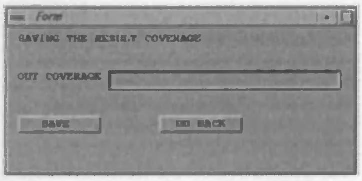

Driving Menu for Line Database Simplification.



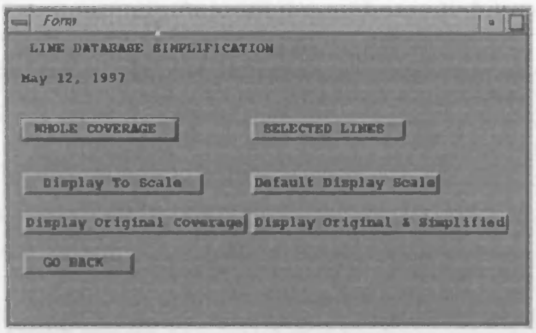

Driving Menu for Line Data-Shape Simplification.



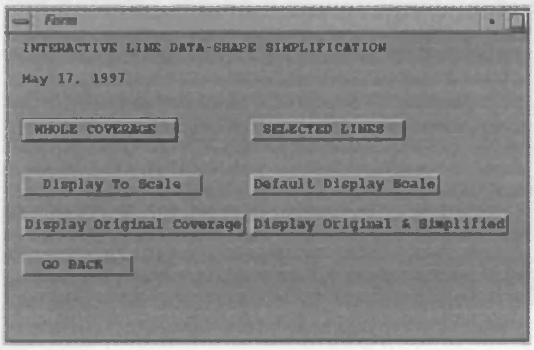

Menu for determining the Display Scale



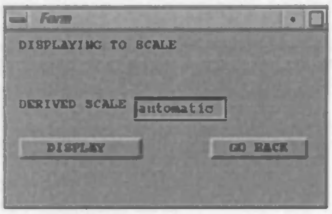

Menu for Saving the Simplification



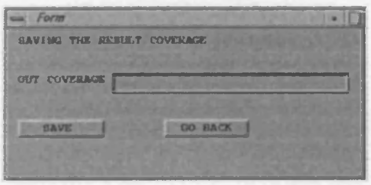

Driving Menu for Line Database Simplification.



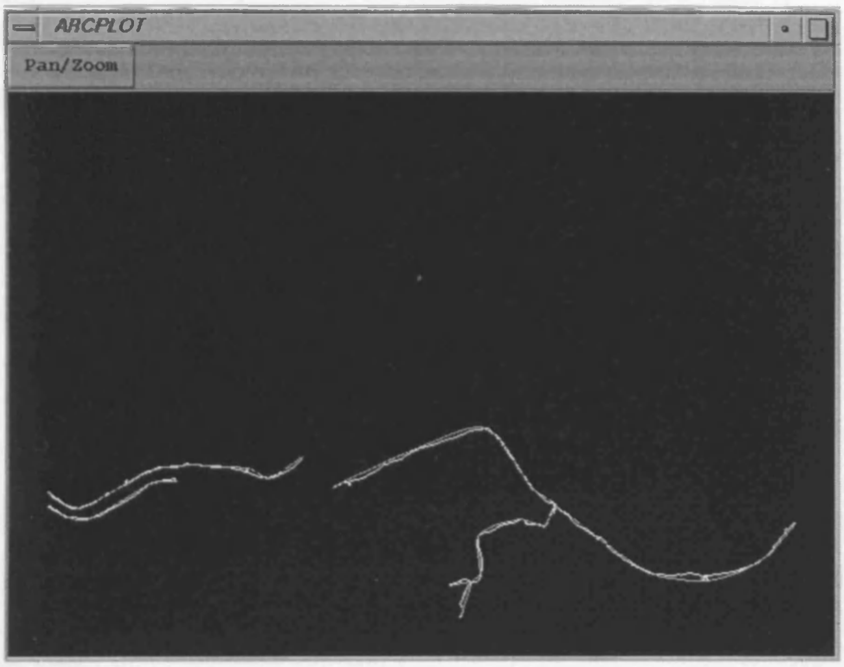

Driving Menu for Line Data-Shape Simplification.




Menu for determining the Display Scale




Menu for Saving the Simplification




← Display screen

------------------------------ End ------------------------------