# *Error processes in the integration of digital cartographic data in geographic information systems.*

Krysia Rybaczuk

DO NOT FILM

**Error processes in the integration of digital cartographic data in geographic information systems.**

**Krysia Rybaczuk**

**Phd Thesis**

**1991**

**Volume 1**

I declare that the material in this thesis is entirely the result of my own work and has not been previously submitted to any other University.

Krysia Y. Rybaczuk.

# Error processes in the integration of digital cartographic data in geographic information systems.

## Abstract.

Errors within a Geographic Information System (GIS) arise from several factors. In the first instance receiving data from a variety of different sources results in a degree of incompatibility between such information. Secondly, the very processes used to acquire the information into the GIS may in fact degrade the quality of the data. If geometric overlay (the very raison d'etre of many GISs) is to be performed, such inconsistencies need to be carefully examined and dealt with. A variety of techniques exist for the user to eliminate such problems, but all of these tend to rely on the geometry of the information, rather than on its meaning or nature. This thesis explores the introduction of error into GISs and the consequences this has for any subsequent data analysis. Techniques for error removal at the overlay stage are also examined and improved solutions are offered. Furthermore, the thesis also looks at the role of the data model and the potential detrimental effects this can have, in forcing the data to be organised into a pre-defined structure.

## Acknowledgments

This thesis has taken its time in coming. Parts of it have been computed and written in several locations, and many of the people I owe debts of gratitude to, have long since moved on. I have still listed them, given that at some soiree I might whip the thesis out and prove that I appreciate their hard work, encouragement and assistance!

So here goes ... At **Durham University**, I would like to thank my supervisor, Michael Blakemore (and all his cronies) for putting up with me for five years. I can't say he was what one expected from an academic supervisor, but he has taught me more than I could ever hoped to have learned and taken me more places than I could ever hoped to have gone, and I hope I gave him a good run for his money! Secondly I would like to thank Lee Gould for his patience and resilience, for the help with obnoxious algorithms, for some of the meals, for all those lifts from the station and for the many happy hours on and off the dojo! Finally I would like to thank all those who took part in the digitising experiment; Lee, Albert, Mike, Norman and Claire.

At **Birkbeck College, London**, I would like to thank Prof. David Rhind for allowing me many weeks of computing time way back when, and more recently for taking the time to discuss aspects of GIS development. No longer at Birkbeck, but probably still lurking in the shadows I would like to acknowledge the help of Vince Andrews, Laurie Becker and Jim Wiggins.

Now we come to the **Geography Department, Liverpool University**. Firstly, I would like to thank Prof. Peter Lloyd, for providing me with the equipment I needed to continue this thesis, without that nothing would have been possible. Next I would like to thank Jim and Vicki Innes who looked after me and made sure I slept and ate during two very hard years of graft. I would also like to thank Vicki for typing some of the thesis when deadlines looked short. The technical staff at Liverpool have been ABSOLUTELY BRILLIANT. Of those I would especially like to thank are Ian Qualtrough for all those photographs and all the drinks. Claire Sullivan for various computing bits n' pieces. Derek France for the use of his software and for his patience. Bob Jude for watching me use Derek's software. Sandra Mather for her advice on the graphics and for help with figures 4.45 and 5.74 and last, but by no means least Pauline, Janet and Mary for their concern and support. Of the academic staff I would like to thank Peter James for lightening my tutorial load and for proof reading all those dreadful words he knew so little about. Thanks also to Dave Chester for taking my second years and to all those third years who took part in the generalisation project.

Academic institutions aside, I would like to thank my parents for their support and patience, Mark for helping me with food and panic attacks, Claire Hogan for typing and moral support and Marty J. for being the only one never to lose faith in me.

If pushed, I would dedicate this thesis to all those friends and acquaintances, who for good reason never completed ... you know who you are!

This thesis was supported by an NERC case award for three years. Academic visits to the States were supported by an Auto Carto London award via the RICS.

Krysia Rybaczuk, October 1991.

**Table of contents.**

Table of Figures held in Volume Two

**Chapter Five**

**Chapter Six**

List of Tables

**Chapter Six**



**Chapter Seven**

## Of Exactitude in Science

... In that Empire, the craft of Cartography attained such perfection that the Map of a Single province covered the space of an entire City, and the Map of the Empire itself an entire Province. In the course of Time, these Extensive maps were found somehow wanting, and so the College of Cartographers evolved a map of the Empire that was of the same Scale as the Empire and that coincided with it point for point. Less attentive to the study of Cartography, succeeding generations came to judge a map of such Magnitude cumbersome, and not without Irreverence, they abandoned it to the Rigours of sun and Rain. In the western Deserts, tattered Fragments of the Map are still to be found, Sheltering an occasional Beast or beggar; in the whole Nation, no other relic is left of the Discipline of Geography.

From Travels of Praiseworthy Men (1658)
by J.A. Suarez Miranda

## Chapter One - Introduction.

This thesis looks at the error processes associated with Geographic Information Systems (GISs). A GIS is a relatively recent computerised technology for storing, managing and manipulating spatial data, which has been promoted in the late 1980's and early 1990's as the answer to many spatial problems. Such systems form a suitable basis for the storage of information pertaining to positions on the earth's surface, and by virtue of the information storage techniques used, they can effectively be used to integrate and link geographical information. In addition to being a storage repository for geographical information, such systems can also perform basic spatial data analysis by looking for common patterns and searching on particular criteria. As a consequence GISs have become very popular with users of geographic data. Applications in the private and public sector have included forestry management, facilities management, emergency response and management, environmental monitoring, optimum routing, planning, urban utilities management, pollution monitoring, financial planning, health monitoring, land use change and inventories of land parcels. Major users have therefore tended to be planners, local councils, market analysis companies, environmental monitoring agencies and federal bodies.[1] In the academic community GISs have been applied to problems relating to landscape change, the monitoring of fragile ecosystems, the inventorying of soil and vegetation data and the analysis of epidemiological data.[2]

As with any new technology, praises are often loudly sung, long before shortcomings are fully appreciated. In the case of GISs, the focus of researchers, and more recently users, has been centred on the technology and its striking visual products, rather than on the operations themselves. In many ways this attitude is one of technological positivism, and it can be paralleled with the quantitative revolution that influenced the discipline of geography in the 1960's. As a consequence, relatively little attention has been paid to the way in which the

[1]Agencies responsible for the production of spatial data such as the Ordnance Survey in the UK and the US Geological Survey in the USA are also major users.

[2]Further examples of current applications are given in Mapping Awareness, Proceedings of Autocarto, Proceedings of GIS/LIS and the International Journal of GIS to name but a few sources.

techniques embodied within the system manipulate the information held by the GIS. Secondly, the information itself is rarely addressed. It is generally assumed by the majority of users, (who may have not been trained otherwise) that provided the information fits snugly into the GIS, it is suitable for use.

In many ways therefore, the quality of information within a GIS has, until the late 1980's been a neglected topic. Yet it is a vital one. If the products of GISs are to be used in a worthwhile and valid manner, they have to be reliable. The paradox is that the very nature of spatial data representation prohibits an accurate depiction of geographic space. The data will always embody a degree of error or uncertainty, due to cognition, acquisition, storage techniques and processing mechanisms. At the time of writing however, no GISs offer the user an ability to handle and deal with errors in their databases. If error is referred to at all, it is in the context of whether the data are fitting in with the system's model of the real world. This encourages users to view space in the same way as the system, which can place limitations on the way problems are addressed. In the absence of developments in error management within GISs, products and methodologies used within them have to be suspect.

> "The absence of facilities within GIS software for handling the effects of input data uncertainty and possible error propagation by GIS operations creates a question mark over the safe utilisation of many aspects of the technology"
>
> Green (1990, 106)

Furthermore, as the topic of inaccuracy and error propagation within GIS has been neglected for so long, users have tended to make no assumptions, or false assumptions about the accuracy of their material. This has resulted in a dangerous situation, in which the truth about data inaccuracy is still widely unperceived. As data quality issues move up the GIS research agenda it will become essential for system developers to pay more attention to building facilities that evaluate data integrity, scale effects and suitability of analysis.

> "If the burgeoning GIS industry is indeed being driven by false perceptions of data accuracy, then the truth will be devastating: even the simplest products will be suspect"
>
> Goodchild and Gopal (1989, xi)

> "we need explicit methods of tracking and reporting error in GIS software, and the algorithms and data structures which recognise uncertainty and inaccuracy directly: in short we need greater sensitivity to accuracy issues among GIS designers and developers"
>
> Goodchild and Gopal (1989, 1)

GISs are expensive products to purchase, with even small PC based products costing up to £10,000 when all the peripheries are included. Thus given the large investment already made by various public and private concerns in GIS, and the likely continued upward trend in sales, there is a greater need than ever to address the issue of error.

> "Considering the public and private investment in geographic information systems, additional research on the error of overlaid maps is required."
>
> Chrisman (1989,33)

User awareness of what problems exist, and how they might be overcome is equally important. GISs are becoming more frequently used to address a wider variety of problems, yet fewer people are being adequately educated and trained in the use and appreciation of the technology. Training is often provided at a cost by the vendors of the software, and usually deals with the mechanics of operating the system, rather than problems users might face with their data. As many users are not trained cartographers, they are unlikely to be aware of the nature and characteristics of cartographic representation, or of the errors embodied within such material. The approach of GIS vendors serves to nurture this state of affairs, by giving users the impression that they can mix and match data with such ease that the origin, error or scale of data are of little consequence.

> "With manual cartographic methods many of the problems are visible and the highly skilled operator makes the necessary adjustments and knows how far the information can be relied upon. With GIS the equivalent operations are transparent, the operators are no longer so knowledgable in or aware of all the limitations of the data, and the problems are more or less invisible. The apparent ease by which data from different scales and qualities of map document, with different levels of innate inaccuracy, can be mixed, integrated, and manipulated totally disguises the likely reality of the situation."
>
> Openshaw (1989, 263)

If error is not addressed promptly and comprehensively, the results for the GIS industry, the users and the reputation of the methodology could be serious. In terms of decision making,

> "uncertain data may ... be cherished and saved for use in subsequent operations, or else mapped, or input into a spatial analytical method or used with a complex model as part of a spatial decision support system. Yet all these products typically assume error free inputs. The results may well look nice. They may have to be used for decision making and planning despite possessing levels of uncertainty that are completely unknown and usually cannot even be guessed"
>
> Openshaw (1989, 264).

The consequence of such revelations, given the level of investment in the technology may provide a strong reaction against GISs and their associated methodologies.

> "the power of a GIS as a decision-making tool can be reduced if the accuracy of results can not be controlled ... user experience of such shortcomings seems likely to result in a backlash against the use of GIS-based methods and even against quantitatively-based planning tools in general."
>
> Joao et al (1990, 510).

At present the issues of error and accuracy are beginning to be addressed on several fronts. The National Center for Geographic Information and Analysis (NCGIA) in the USA have addressed the problem in a series of workshops notably in Goodchild and Gopal (1989). This thesis sets out to look at the problem of error in terms of data acquisition, and propagation. Early chapters review the adoption of GIS as a new methodology, its origins and technological expectations, followed by a discussion of information quality and the recent work carried out in that field. Chapters Four to Seven look at the problem of error empirically, with respect the ARC/INFO GIS.

Admittedly, the studies deal with one particular system, but the system in question is acknowledged by most as the world leader in the field of GIS. Furthermore, in the UK, the system is likely to be implemented in many Universities following the Combined Higher Education Software Team (CHEST) arrangement with the Environmental Systems Research Institute (ESRI) who market ARC/INFO. The deal allows universities to purchase the systems at a highly reduced cost, and is in itself an indication of the competition amongst vendors, and the importance vendors attach to promoting their product and the technology in higher education. It is likely that ARC/INFO will continue to dominate the market for some years to come, and that it will continue to set the standard in many aspects of the technologies and methodologies GISs use. Other GISs already advertise their product as being ARC/INFO compatible. Thus the work carried out in this thesis, although particular to one product, has wider implications and applications.

Chapter Four looks at the problems of secondary data acquisition and empirically measures errors that plague cartographic material entering a GIS as a consequence of the data entry process. Two areas are covered, those of precision and generalisation. In the first case, digitising precision is evaluated in terms of individual operators, and some pattern of error is searched for. In the second case, the way lines are acquired into a vectorised system is assessed. This involves precision in point selection, as users attempt to represent faithfully linear features, and in addition, there is an element of generalisation involved, as users are required to select the most suitable points to represent a line. This experiment sets out to demonstrate the way the same piece of data can produce differing results in the data set as a consequence of using a variety of digitising operators and material of differing scales.

One of the ways in which error can be monitored within a GIS, is to look at the variation in the cartographic features themselves. GIS features may be points, lines, areas or in three dimensional modelling, surfaces. Chapter Five sets out to assess the modification lines undergo as part of the data model building process. Linear change is investigated in terms of feature length, and also in terms of feature displacement. It might be argued that investigating the way that error affects the lengths and positions of lines within a GIS database is a relatively pedantic quest. Yet if one is dealing with lines, not as geometric features, but as geographic features then the difference can be crucial in terms of investigative analysis or as a basis for planning decisions. Two major types of analysis are the most prominent in GISs; those of networking and overlay.

Networking analysis looks at linear models in which flows and the transportation of commodities are investigated according to a series of constraints. If lines delimiting the network are inaccurate, flows will be wrongly calculated. Imagine a network in which the roads of a busy city formed the basis of the data set and that these were inaccurate; each to a varying degree. Imagine furthermore that a standard type of network analysis was carried out, in which the length of a particular journey was assessed. This would have combined the line lengths of all those roads that form part of the desired route, in addition to other data such as stop times and turning times. Inaccuracies in the cartographic lines would introduce a false impression of the true journey length. Alternatively, if hydrographic measurements were carried out, in which the length of a particular river section is used as a parameter in determining discharge; false assumptions as to stream length would be critical in terms of modelling possible scenarios. In fact if lines are to be used as representations of lengths of any features involved in network analysis, inaccurate placement or portrayal will jeopardise the validity of any GIS operation.

Overlay analysis can be directly compared to traditional cartographic overlay which involved several sheets of tracing paper, a light table and a skilled cartographer. Sheets are overlaid and common areas of interest are manually selected and redrawn. GIS overlay analysis performs essentially the same function, but mylar sheets are replaced with computer held graphical files, and the searching procedures are carried out by predefined algorithms. In many respects, overlay analysis is little more than an extension of traditional pattern analysis. Overlay analysis is also prone to the uncertainty of cartographic line depiction. Overlay deals primarily with areas, which in terms of vectorised GIS systems are made up of lines. Thus if the lines forming the bounding edges of polygons are inaccurate in length, the problem will be inherited by the areas. Standard types of areal analysis include quantity measurements (how much of an area in question is under cropland), comparative measures (what has been the area of change from cropland to light industry for a particular County over a particular temporal unit) and correlative methods based on topographic searching (how many schools have computing taught as part of the curriculum, where are they located, and how many of them are within areas in which the male unemployment rate is significantly greater than the National average). Essentially the last two operations rely on the same routines and algorithms, but the error component held within the data is manipulated in a slightly different way.

In the first case, as the data will pertain to the same variable, differences and discrepancies will be very apparent. It is then up to the user to decide whether these variations are due to a real change, or whether they are simply due to mismatches in the data. In the second case; since variables will all relate to different components of the data set, differences are not apparent, therefore data are more likely to be used and labelled as reliable. One way of reducing errors at the analysis stage is to adopt a strict procedure of data entry, ensuring that data are carefully selected and inputed. Despite attention and care however, errors will still arise, especially if the data are obtained from outside sources. A possible solution to this is to demand a complete history of the method and conditions by which the data were obtained. This route has received a lot of attention recently as the USA and the UK have published data quality and transfer standards.

After overlay analysis, assessing which areas are veritable differences between data sets, and which exist merely as part of the error process, is a difficult problem. Small spurious areas are often created as a consequence of data overlay, and traditional methods for dealing with this involve identifying them on the basis of size or shape and removing them. Chapter Six looks at the extent to which such spurious areas or polygons are present in overlaid data sets, and examines what processes are responsible for their initiation. The chapter also investigates the mechanism by which areas are 'removed', and questions whether such removal on the basis of mere geometry is justified.

Chapter Seven looks at possible alternatives to solving the identification and removal of areas of uncertainty. One of the solutions the chapter offers is to view the problem in a different light. Rather than to look for errors and eradicate them using specific technological tools, an alternative solution of error acknowledgement and minimalisation is proposed. This accepts that error is a feature of cartographic life and that in dealing with GIS databases, error has to be coped with in the best way possible. In this case, rather than look at spurious polygon identification and elimination, an alternative is to suggest that if areas are to be removed, by merging them with a neighbouring area, it is the most suitable neighbouring zone that should be selected for the merge.

Other issues the thesis attempts to cover is the way systems seem to be technologically deterministic, favouring 'tools' rather than 'information' in the quest for a solution. In the later chapters the thesis also covers some of the problems associated with temporality and versioning.

In terms of relevance to the future, studies of error and error propagation within GISs are vital if the systems are to continue to enjoy the technological status they currently possess. Rhind (1991) (table 1. 1) predicts that given current trends, the number of systems in use by the turn of the century will exceed half a million. Considering the amount of investment that will have been made by this stage, any increment in the knowledge of the error associated with spatial databases and GISs will prove vital, if these systems are to be viable.

Table 1. 1. Projected rise in GIS demand and sales worldwide.

| | 1990 | 1993 | 1996 | 2000 |
|---|---|---|---|---|
| Sales($m) | 1000 | 1750 | 3000 | 6200 |
| Number of systems ('000s) | 20 | 55 | 150 | 580 |

Rhind (1991)

An additional factor is the current move that society is making towards an information dominated society. For many future uses (as well as current ones), the geographic hook will become a fundamental one for storing and ordering the data an information society requires. This in turn will promote information managers to turn to GISs for storage and solutions. In such a climate any knowledge of potential inaccuracies will become vital to ensure that errors are not propagated within, or between systems. Alternatively, assuming that each task will have a supportable error level, a knowledge of the degree of error accumulation in any GIS analysis will help the user determine its acceptability.

It is predicted that geographic information will not only increase in quantity, but also in scope. One of the areas expected to increase in popularity are global databases (Goodchild 1988b). As global data become more readily available GISs will have a greater role to play in macro level applications. These will deal with information cutting across the bounds of data collection units (such as map sheets), administrative units and political units. Effective utilisation of GISs for macro level applications will rely on the universal acceptance of data transfer standards, but if data are to be easily amalgamated from different cultures, a knowledge of varying cultural perceptions will be an important area of concern.

As the adoption of GIS increases, it is likely that producers and distributers of geographic information will have a legal responsibility as well as a moral one to ensure that data are as error free as possible, or at least that they meet certain predetermined standards of acceptability. This raises questions such as;

- who will be responsible for erroneous decisions undertaken on the basis of data generated by GIS?
- what errors will be legally tolerated, and which will be deemed insupportable in the law?

For example, in the case of in-vehicle navigation systems, what happens if the user follows the advice of the data base, and involuntarily breaks the law? Whose responsibility is it? The user's or the supplier's? If producers are to be held responsible for the data they produce, then an understanding of the potential inaccuracies will be vital to data evaluation and dissemination. Interlinked with the issue of legal liability is the question of ethics. Can data producers really market data of uncertain quality? Furthermore can they market such products to areas in which they will be least suspect? As the third world and the eastern block begin to make an investment in GISs there is much current attention in GIS circles being turned towards technology transfer. The receivers of the technology may be more aware of the pitfalls associated with the technology; but alternatively they may rely entirely on vendors and consultants they deal with for advice and support. In either case, a knowledge of potential errors associated with GIS will be vital for planning and environmental monitoring in such areas.

**Chapter Two - Definitions and historical review.**

**2. 1. Introduction.**

The current use of GISs is a result of a short historical development of approximately thirty years. The development of systems has been based upon several factors, notably the development of the technological tools, the needs of key institutions and their subsequent research agendas, the role of individuals and finally the availability of information in digital form. The focus on system development has been one of use and analysis, and it was only in the 1980's that error and quality began to emerge as a major research theme. An understanding of the evolution and functionality of GISs is an essential component in determining how quality can be assessed. For example, error can be evaluated in terms of GIS technology, methodology or philosophy. Thus this chapter sets the historical scene in which these systems developed and provides some discussion of what these systems are capable of.

**2. 2. Definitions of a Geographic Information System (GIS).**

> "First let us get past what GIS is not. It is not just a computer program as its name implies; it is a technology. .... And it certainly is not simply map making by computer."
>
> Parker (1990, 1)

As GISs have gradually become utilised in policy, environmental and planning evaluations for a wide range of applications, the definition of a Geographic Information System has become more detailed and specific. Early on, users and researchers classified GIS as computer assisted systems able to

> "accept large volumes of spatial data derived from a variety of sources including remote sensors and effectively to store, retrieve, manipulate, analyse and display these data according to user defined specifications"
>
> Marble and Peuquet (1983, 923)[1]

---

[1]A detailed set of requirements for a GIS is outlined by Marble (1984). He defines GIS as incorporating all of the following criteria;

1. A data input subsystem which collects and/or processes spatial data derived from existing maps, remote sensors, etc.
2. A data storage and retrieval subsystem which organises the spatial data in a form which permits it to be quickly retrieved by the user for subsequent analysis, as well as permitting rapid and accurate updates and corrections to be made to the spatial

and for many users this ability for integration has become the accepted definition; Clarke (1986), Monmonier (1990b, 214), Bracken and Webster (1989, 137).

In many respects however, this was a definition made by the digital cartographic community, and by the developers of early GISs. As time has progressed however, the user community has increased and diversified so dramatically, that many of the more recent definitions have come from the users themselves. These have tended to revolve around the problems a GIS can be applied to, or the types of data a GIS can effectively deal with. For example Cowen (1990) states that;

> "a GIS is best defined as a decision support system involving the integration of spatially referenced data in a problem solving environment"
>
> Cowen (1990, 59)

whilst Scholten and Van der Vlugt (1990) state;

> "We see a geographic information system as essentially a tool for urban and regional research, policy analysis, policy simulation and planning."
>
> Scholten and Van der Vlugt (1990, 16)

For other users it is the ability a GIS has to handle spatial information that makes it distinctive from other computerised technologies. Furthermore, since so many current global problems have a spatial context, some see it as vital that GIS technology be employed to solve them.

> " The integration, storage and tracking of global environmental data must be done spatially if spatially meaningful solutions are to be forthcoming. Otherwise we are left with only a set of statistical data with no way to make meaningful decisions."
>
> Parker, (1990, 5)

As the functions and capabilities of GISs have become more clearly defined and accessible, some definitions have centred around the functionality of the systems and the 'tool box'

---

database.

3. A data manipulation and analysis sub-system which performs a variety of tasks such as changing the form of the data through user defined aggregation rules or producing estimates of parameters and constraints for various space-time optimisation or simulation models.

4. A data reporting system which is capable of displaying all or part of the original data base as well as manipulated data and the output from spatial models in tabular or map form. The creation of these map displays involves what is called digital or computer cartography.

incorporated within them. Of these 'tools', it is the ability to integrate information from several sources (DeMan 1984), combined with a facility to create new information from existing data using search and overlay techniques that perhaps best define a GIS.

> "This ability to both automatically synthesise existing layers of geographic data and to update a database of spatial entities is the key to a functional definition of a GIS."
>
> Cowen, (1990, 57)

A more comprehensive taxonomy is suggested by Rhind (1990a) in which he clarifies GIS functionality as the ability to respond to a series of questions.

| **Question** | **Task Type** |
|---|---|
| What is at ...? | Inventorying/Monitoring |
| How do I get from ... to ...? | Route Finding |
| Where is ... true [or not true]? | Inventorying/Monitoring |
| What has changed since ...? | Inventorying/Monitoring |
| What spatial pattern(s) exist(s)? | Spatial analysis |
| What if...? | Modelling |

Source: Rhind (1990a, 58)

For some users however, the differences between digital mapping packages and GISs is still blurred. This is not necessarily based on the functional ability of the two technologies, but on a superficial assessment of the final cartographic products. In some cases this leads users to utilise GIS techniques for tasks that could be easily accomplished with a digital mapping package. The background to such (mis)use may be a result of inadequate training, but in many cases organisational structures are responsible. For a department within any organisation to simply request funds for a mapping package equates the request with the desire to acquire a spreadsheet or simple graphics facility, however requesting funds for the purchase of a GIS is capable of triggering off all sorts of connotations. The words 'information system' imply wider impacts than do the words 'computer cartography'. Inherent implications include all the corporate investment tags of an increase in efficiency, added value, an increase in through-put and the ability to maintain an up to date system. Without a doubt GIS is equated with power on several levels, and as such it is regarded as more of a worthwhile investment than a mere digital mapping package. Yet the two are fundamentally different, and address different aspects of spatial data,

> "Although GIS technology arose in large measure from advances in computer aided cartography, fundamental conceptual differences separate these closely allied areas into two distinct subdisciplines with unique foci as well as shared interests...GIS concentrates on the analysis and integration of digital maps, whereas computer aided cartography more directly addresses the generation and visual effectiveness of map symbols. Although visualisation and map design can be important elements in a user interface, spatial analysis and data management are the central issues for GIS research and application."
>
> Monmonier(1990b, 215)

**2. 3. The historical development of GISs.**

The methodology underpinning most of the current 'topological' GISs arose due to a marriage of the techniques associated with computer aided cartography and database design.

> "Each of these functions has its own history. Some are extensive, for example, computer-assisted cartography, with well developed computing algorithms mainly concerned with putting traditional techniques into computerised form. Others such as the organisation of geographical data into efficient structures have a shorter history although they now represent an important branch of geographical history."
>
> Bracken and Webster (1989, 137)

However, much of the pioneering work carried out in the design and creation of GISs evolved from work carried out by those researching in the field of computer graphics and digital mapping, and as such, the two disciplines share a similar heritage. Thus any assessment of the history and evolution of GIS technology cannot be undertaken without a consideration of the historical development of computer assisted cartography.

> "the origins of Geographic Information Systems usually have been traced to early work in computer mapping."
>
> Cowen (1990, 52)

> "Computer aided cartography developed the theory and technique for capturing and editing digital cartographic data and GIS applies this methodology in establishing and maintaining geographic databases."
>
> Monmonier (1990b, 215)

**2. 3. 1. Motives responsible for the development of computer assisted cartography.**

The development, adoption and refinement of computer assisted cartography can be viewed through a variety of historical filters. For some users and agencies the developments were driven by the available technology, for example, the gradual improvement in computer hardware offered the carrot to those researchers in both academia and the public sector that were keen to utilise it. Although this was a clear impetus for some users it also prompted unnecessary dabblings in the technology by others, simply because association with 'computerised' technology was regarded as a great status enhancer for the user, the agency and the data[2]. In other cases, it was sheer user need to employ a more advanced technology that prompted users of geographic data to turn towards computerised processing and output.[3] For some users, needs revolutionised the way data was processed, but for others, especially agencies with very clear cut policies, the technology was often simply adapted to current procedures. In both the case of the Ordnance Survey (Rhind 1990) and the USGS (United States Geological Survey), there was a cautious evaluation of the technology given existing practices. For example in the USGS, the attitude was very much one of evaluating how the new technology could help them update their mapsheets, rather than investigating how the technology might help them revolutionise the way map data are stored and presented.

> "The approach adopted by the OS from 1971 onwards was to simulate manual production of its maps as closely as possible, an approach which facilitated map production but made it impossible to use the resulting digital data in an information system because they were unstructured."
>
> Coppock and Rhind (1991, in press)

In addition to the 'user needs' and 'technology driven' approaches, another factor responsible for development and in some cases adoption of the technology, was the decisive roles played by key individuals and institutions.

> "The social context of the individual working inside institutional structures has enormous impact on the research process. Innovation and invention have a complex history; both individual insight and historical milieu influence the results."
>
> Chrisman (1988, 291)

---

[2]It is this type of usage that prompted the numerous 'A census map of ...' Blakemore and Dewdney (1990). Another example of the 'slave of the technology' ethos that prevailed can be illustrated by the experience of a young Peter Lloyd (now professor of Geography at Liverpool University), when he was working as a lecturer at Queens University, Kingston, Ontario. He was sent in a specially hired car to collect the SYMAP punch cards from Harvard, and in effect became the chauffeur to this great technological invention; the computer mapping package.

[3]Coppock and Rhind (1991 in press) suggest that the creation of the ECU (Experimental Cartographic Unit) was prompted by David Bickmore's realisation that there was no other route for atlas publication.

A genealogy of initiators can therefore be defined.[4] Events on both sides of the Atlantic took on slightly different routes, with key institutions and personalities manipulating available manpower and technology. Often the early years were characterised by individual organisations (usually academic bodies or academic institutions) developing their own routines and software.The Department of the Environment in the UK for example, decided to develop its own in-house software (LINMAP) for specific tasks relating to the 1961 Census Data rather than using the available SYMAP (developmed by the Harvard Graphics Lab). Yet both packages perform similar tasks, although LINMAP was more geared to the needs of the Department. Chrisman (1988) also points out that there was a general reluctance for development agencies to share information or even to present papers at conferences for fear of new innovations being purloined and grants subsequently being lost to other competitors, thus papers in the early 1970's and late 1960's tended to be very general and often lacking in technical detail.

Rhind (1988b), in discussing the importance of the individual in the emergence of the discipline, cites the important role played by David Bickmore and the Experimental Cartography Unit (ECU) in advancing digital cartography in the UK.

> "Our chance to prosper was due largely to one individual. He obtained funds, brushed aside opposition, fended off detractors and insisted that we tackle the impossible. Life was never as easy or predictable working for David Bickmore. Without him and very few others, however, computer based cartography would certainly not be where it is today."
>
> Rhind (1988b,287)

Perhaps one of Bickmore's major achievements was in persuading the Ordnance Survey that the future of cartography lay not in the paper and pen based cartographic material they had been accustomed to producing, but rather in computer based methods. The Survey's response (apart from severing links with the ECU for about five years) was to devise and set up their own system based upon 'blind digitising', a decision based upon minimising costs and mimicking in every possible respect the traditional products they had been engaged in crafting for the past century. Again, however, this is an instance of isolated development as until the early 1980's this system treated cartographic lines as simply lines and made no

---

[4] Cooke (1989), devised a genealogical tree for GIS developments and dissemination in the US.

attempt to employ any means of topology or relationship between these lines whilst the ECU had adopted as a standard procedure the practice of multiple feature codes as far back as 1970.

Equally as important was the coming together of key individuals at a particular place in time. This acted as a catalyst for 'flash in the pan' insights. Rhind (1988b) cites an occasion when Ray Boyle had been demonstrating to David Bickmore an older type of digitiser, which was relatively cumbersome and unwieldy to operate. Bickmore replied that the design was unsuitable for cartographers as they were only interested in tracing lines as if with a pencil.

> "That night I realised I could give him just that. We took one of our standard servo-driven analogue tables and placed a piece of plywood over its top. We replaced the pen head below with two pairs of detector coils at right angles to each other and on top we wound a coil of wire on a pencil. Through this coil we sent an AC current and arranged the detector coils in a circuit which controlled the direction of the drive motors. Thus everywhere the pencil went, the power driven head followed."
>
> Ray Boyle, cited in Rhind (1988b,280)

The result of this operation was the first free cursor digitiser which became the pillar of cartographic data entry until the early 1970's. A similar encounter between Nick Chrisman and Dennis White in December 1975, resulted in the inauguration of the Harvard Graphics Lab Odyssey System. Chrisman recalls:

> "White and I were at the computing centre pushing the final large cities through our revised process. In that special second wind of late night computing on the back of a recycled listing we drew up a plan for a system of processes to treat cartographic data. We set out data structures functions and even names of each module. The system was named Odyssey."
>
> Chrisman (1988, 298)

As well as the drive from individuals, impetus came from the institutions functioning as grant receiving bodies and acting as an umbrella within which researchers could gather to design and develop. In the UK, such institutions included the Experimental Cartography Unit (ECU) and in the United States a crucial institution was the Harvard Graphics Lab. Here software routines such as SYMAP, GRID, POLYVRT and ODYSSEY were born.[5] Although SYMAP

[5]"From the creation of the first true GIS in Canada in the 1960's to the launch of the first sophisticated commercial systems in the early 1980's most of the activities were in University research laboratories." Rhind (1990a, 60)

was crude, in its heyday it could probably have been classed as the most widely used mapping package in the world. POLYVRT was in many ways responsible for our present day notions of topological relationships. Despite being a miserable marketing failure by ISSCO (a software firm in the business graphics field), ODYSSEY has been regarded by many as 'ARC/INFO without the manuals'. As one of its major programmers, Scot Morehouse was to move to ESRI in California when the Lab was dissolved in 1982, this observation is probably not far from the truth.

Advancements in hardware also influenced the evolution of computer-assisted cartography. In the early days hardware was both expensive and limiting in terms of data storage, the types of operations performed and the speed of execution. Today, however, sophisticated pieces of GIS software are comfortably catered for in terms of data storage and processing power by desk-top microcomputers. The essential hardware components in any GIS system can be defined as those pieces of machinery responsible for data entry, data storage, data processing, data display and hard copy output. Cartographic data entry generally takes place via a digitiser. These have not only become larger in terms of physical size and more accurate over the years, but the ergonomics of the table and tablet design have also improved making it less of a physically tiring process to digitise maps. In addition, digitising software has also allowed greater flexibility of data entry from the digitising tablet itself.

Changes in storage media and processing power have perhaps allowed the most dramatic acceleration in computer assisted mapping. The original versions of SYMAP for example, were operated using punch cards. Simple cartographic operations demanded several boxes of strictly ordered paper punch cards. Naturally any dislocated card would cause the operation to fail. Each punch card carried one record or one line of command in a program. A step forward from punch cards was witnessed in the form of paper tape. Although limiting and prone to tearing and warping, paper tape increased the density of data items held, improved input speeds and were easier to transfer between locations. Magnetic tape was the next breakthrough, enabling data to be encoded at densities varying from 1600 bpi (bits per inch) to 6250 bpi. An extension of magnetic tape technology was applied to magnetic disks:

here data are stored in sectors and tracks and data density is measured in bytes. Generally, disks hold between 360K and 3.5 megabytes (Mb). Hard disk capabilities on today's machines vary in storage capacity from 10Mb to several gigabytes depending on machine cost and type. More recent technological advances in this sphere include the advent of CD ROM's and optical disks, although few systems rely on such media.

Hand in hand with the innovations occurring in data storage were the rapid changes occurring in integrated circuit technology enabling speedier processing of data. Until the late 1950's the processing units were comprised of discrete transistors. These were unwieldy, power thirsty and awkward to program. In the early 1960's integrated circuits appeared, these were composed of several transistors on one chip, enabling operations to be performed at a faster rate. Equally, integrated circuits allowed the combination of about 10 components on a 5mm square surface, a procedure known as small scale integration (SSI). As the technology advanced, more transistors could be combined onto a chip resulting in medium and then large scale integration (LSI). LSI enabled the main functional units of the computer (arithmetic logic unit and program control) to be combined onto a single silicon chip, resulting in the first microprocessor. These then became the backbone of the microcomputer, and also allowed for processors to increase their word length from 4 to 8 to 16 to 32 bits. Consequently, data could be gathered and processed at faster speeds.

Such technology is often regarded by electronic engineers as passé, as sights are beginning to be focused on the role the transputer and parallel processing can play in accelerating the role of processing. This technology is based on the idea of a complex problem being broken down into compatible parts, and each part being dealt with at the same time by processors working in parallel, rather than being solved sequentially by an individual processor. At present such technology is still in its infancy therefore:

> "It is difficult to make firm statements about how GIS might avail itself of this new technology."
>
> Healey and Desa (1989,90)

Nevertheless, they go on to cite geometric parallelism as perhaps having the most to offer GIS applications.

> "Of the recognised approaches to algorithm construction geometric parallelism, possibly combined with limited algorithmic parallelism offers the

most promising route for initial work in parallel GIS processing."

Healey and Desa (1989,97)

Display media have also improved, early terminals were monochrome and had poor graphic resolutions. Modern workstations now extend to resolutions of 1600 x 1280 pixels[6] and even microcomputers with VGA cards are capable of 1024 x 768 pixels[7]. Output facilities have also undergone a generation leap, from the early days of line printer mapping where the density of the type face characters was utilised to represent tonal changes in choropleth mapping, to today's electrostatic colour plotters allowing for full colour glossy output.

All such changes in computer hardware have fostered an environment within which software could develop and reach new heights. In general adopted machines have become smaller and more powerful. In the early days of computer cartography however, most of the analysis, research and design was often carried out on mainframes. Due to the cost of such machines, they were beyond the means of most workers in the field, thus users often rented time on University mainframes.

> "Early on because of the cost of hardware, ESRI rented time on an IBM mainframe at a nearby University of California campus and digitising was subcontracted or done on rented digitising tables."
>
> Dangermond and Kent Smith (1988,301)

With the availability of microcomputers such as the PDP/9,10,11 smaller departments were able to become self sufficient and to support their own computing needs. Such processing power was still relatively expensive. Rhind (1988) recalls that an upgrade of the PDP9 from 16 to 32K main memory cost the ECU the equivalent of $40,000 at 1988 prices. In the early 1980's however hardware became available to all in the form of microcomputers. In the early stages of development these were relatively expensive and limited in terms of RAM. In the past decade however, they have become increasingly more sophisticated and economically feasible to purchase.

---

[6] The SPARC workstation has a resolution of 1152 x 900 pixels, but some workstations with super VGA can extend that.

[7] A recently advertised STRIDE graphics card availale for PC's allows for a resolution of 1530 x 1280, with a potential 2 million colours!

The increased accessibility to hardware able to cope with digital cartographic databases has resulted in a broadening of the user base and a subsequent increase in software demanding clientele. The last few years have therefore witnessed a rapid increase in packages able to run on PC's, as this type of product has a wider potential market than mainframe based systems. Users are also more likely to experiment with smaller packages as the risk involved is substantially reduced. Thus a 'think small, sell many ethic' has evolved.

### 2. 3. 2. The technological development.

Early adventures by users of geographically referenced data into the realms of computerised technology came about in the late 1950's and early 1960's, and consisted of simple algorithmic routines that dealt with particular isolated problems. The first internationally used package to appear was SYMAP, developed at the Harvard Computer Graphic Lab in 1966.

> "SYMAP is the patriarch of complete computer mapping programs particularly mapping programs designed for analysis of spatial data."
>
> Carter (1984,44)

SYMAP was built around a number of routines and dealt with two fundamental different types of cartographic data; points and areas. Points could be displayed as part of a contour surface or could be interpolated to produced trend surfaces. Area data could be displayed as choropleth maps and another product of the package was a matrix of values from point data output which could then be stored for use in other programs.

The main limitations of such a system were its data entry and output facilities. Few institutions possessed digitisers at that time, and so the system allowed for entry via x,y coordinates entered as punch cards. The other main limitation of the system were the methods used to output results. Data display was achieved by the use of lineprinters and the output was somewhat crude. None the less, SYMAP and its counterparts that also relayed the results of their labour to eagerly awaiting line printers set the ball rolling, for since then interest in computer cartography has never waned. The notion of maps, (no matter how aesthetically displeasing) being produced by a computer also lent them an air of scientific authenticity. Line printer maps proliferated, especially in localised productions introduced by the phrase 'A computerised Atlas of ...' (Blakemore 1985b).

Goodchild (1988a) also emphasises that the success of SYMAP lay in the fact that by using constant digital boundaries it proved that digital mapping could be cost effective, as most of the cost of map production lay in the input and editing of digital boundaries. SYMAP also had an important role to play in combatting traditionally subjective cartographic procedures with computerised techniques.

> "Its algorithm for spatial interpolation operationalised a standard but subjective cartographic process and remains one of the most useful of the large number of methods which have been devised for interpolating continuous surfaces from randomly distributed points."
>
> Goodchild (1988a,314)

The next generation of mapping packages emerged in the early 1970s. These were aimed at both mainframe and minicomputers such as DEC VAX. The major characteristic of such systems were, however, that they relied on vectorised data structures to capture, store and display cartographic data. Systems either stored data merely as lines with no formal relationship defined between lines, areas and points, or they were held in such a way that the relationship between all the digital lines held in a map could be identified and utilised. The former mechanism was common in both simple mapping packages and CAD systems, which were adequate for map production and display but which offered limited opportunities for data manipulation. The second type of system characterised the topological data structure that was to form the basis of most vectorised GISs. Topology was based on the interdependence of all geographic phenomena and the relative position of phenomena to each other in space.

> "Because they are located in a particular manner on the surface of the earth, all geographic entities are related to one and other ... analysis of geographic information must take into account the basic connectedness and interdependence of all geographic entities."
>
> Chrisman (1977)

Very early systems did not accommodate topology into their data structures. and polygons were encoded by a system described as entity by entity digitising. This encoded every area or polygon individually and therefore often resulted in overlapping areas where the digitising operator had deviated slightly whilst entering two adjacent polygons.[8] On a large data base removing such errors would entail a vast amount of editing, therefore new methods were sought and the successor to the entity by entity method was one which made use of a point

[8] This is discussed in greater detail in Chapter Four.

dictionary. This encoded all the points that made up a series of polygons and then identified the polygons by calling up points that make up its boundary. Mapping packages making use of this system included the Harvard Graphics Laboratory's CALFORM program. Although sliver polygons do not appear in a point dichotomy, the subject of topology or neighbourhood relationships remained unaddressed.

Topology in cartography was first addressed by Cooke and Maxfield (1967) in Australia and then put into practice by the U.S. Bureau of Census. This was one of the first organisations to try to come to terms with the problems of topology by developing the Dual Independent Map Encoding (DIME) system to deal with errors arising in the Address Coding for the 1970 US Census.[9] This tabulated information for standard end nodes of roads and noted the census tracts and blocks that occurred to the left and to the right of the street in question. In effect attribute data was held for lines (streets), nodes (end and start of road) and polygons (block left and block right). Although neighbourhood relationships existed in such a system these were node dependent, rather than line dependent. Trying to assemble the outlines of a polygon or searching for all segments sharing a node requires a set of time consuming search procedures. It also poses a major drawback for cartographic areas made up of complex lines as it always assumes that the segments are straight and not crossed by any other line.

An improvement on this system was the POLYVRT data structure designed by Nick Chrisman (Peuker and Chrisman 1975) at the Harvard Graphics Laboratory (figure 2.1). This allowed for full topographical relationships to be incorporated into the cartographic system. The basis object of the POLYVRT is the 'chain'. Unlike the DIME segments (or streets) it allows the user to input as many points into the segment as are required for the representation of a complex line, but like the DIME segment it has a start and end node and ensures that no lines will be crossed. The points themselves are no longer in primary storage, but are referenced by the chains. Its major advantage is that topological checking is no longer dependent on points but on boundaries. Searches can also be speeded up as they are not uni-directional but can take place from the chain to the polygon or from the polygon to the chain.

---

[9]It has been suggested that the first 'cartographic' topological systems developed in the USA was the LOOMIS.

Thus neighbouring entities can be found through their 'bounding' or 'bounded' complements.

The topological data structure then formed the basis for the majority of the large computer mapping packages of the 1970's such as GIMMS and ODYSSEY. GIMMS was designed by Tom Waugh at the University of Edinburgh. Essentially GIMMS is based on two packages; a business graphics package that was enhanced by the Geocartographics Division at Statics Canada and the mapping package developed exclusively in Scotland. GIMMS has undergone many updates and it currently in its fifth update. The system allows for a variety of mapping styles including choropleth maps, dot maps and statistical graphs. Its ability to store data topologically means that nested hierarchies can be assembled for administrative units and census mapping. For example in the UK, at the lowest levels Wards may exist, followed by Districts, Counties and Regions. A line may therefore bound all three administrative units, but will only need to be represented once and then labelled accordingly. An example from the GIMMS 5-0 manual illustrates this point (figure 2.2). A portion of a map is drawn covering 2 counties (23 and 27). County 27 has districts 10, 12 and 15. District 12 has two wards 024 and 025 and district 10 has one ward numbered 107. County 23 includes district 19 and ward 045. Three segments are marked to denote ward boundaries only (A), ward and district boundaries only (B) ward, district and county boundary only (C). The segments may be labelled as follows:

Segment A W024W025
Segment B W025W107 D 12D10
Segment C W007W045 D10D19 C27 C23

Wards may be searched for and selected by calling up only those segments beginning with the letter W; a similar process may be carried out for districts and counties. The topographical structure of the data base also allows for the removal of lines that have the same polygon on either side, thus allowing various administrative polygons to be used from one datafile.

In the pre-microcomputer era, there were relatively few marketed mapping packages. Most of the digitally produced maps from that time were created using software routines and in-house agency specific algorithms. The major distributers of such systems included the Harvard Graphics Lab (CALFORM, SYMVU, POLYVRT, ASPEX, DOT.MAP and

ODYSSEY), the National Technical Information Service (NTIS), of the US Department of Commerce (CAM developed by the CIA), the US Bureau of Census, (DIME), the USGS (Digital Line Graphs and Digital Elevation Models) and the Geography Program Exchange at the Michigan State University and its European Subsidiary at the University of Karlsruhe, West Germany (Monmonier, M. 1982)[10]. Perhaps the most technically advanced system was the ODYSSEY package. In addition to performing similar options to GIMMS, ODYSSEY also allowed for three dimensional output, with the additional facility of a polygon overlay processor and the ability to search by entity on the basis of data attributes. In many respects ODYSSEY was really a raw GIS rather than a polished mapping package.

The main aims of the system are the managing, analysis and display of cartographic information. These functions are carried out through nine programs including map digitising, transformation, verification, overlay and display. Cartographic data base creation is managed through three programs. HOMER creates the cartographic file directly by digitising a paper source map. PENELOPE builds and verifies both geometry and topology of a set of unstructured digital output and crates a chain file from the digital line file. Finally CYCLONE reads and analyses a chain file for geometric and topological errors (both in digitised and converted DIME files). This component of the system reflects its origins as a vectorised mapping package, in that it has no facilities to embody satellite or scanned data.

Cartographic manipulation options include geometric transformations (projection, rotation and registration) in the PROTEUS program, whilst CYCLOPS can restructure data from chain to polygon format, generate topology files and convert standard chain files for entry into SYMAP. The Geographical Analysis package within ODYSSEY carries out the functions of coarsening, area and perimeter analysis, point in polygon and polygon overlay. which is regarded as one of the most important properties of the system.

> "Geographical overlay is the most advanced analytic capability of the system, ODYSSEY implements true polygon overlay of two or more CDBs (cartographic data base) of a region, creating a new CDB of the resulting boundary network with complete identification of the boundary intersections

---

[10]Brassel (1977), discusses the survey of 233 currently used (March 1976) digital mapping routines prepared for by the Commission on Geographical Data Sensing and Processing of the International Geographic Union (IGU).

and new polygons, and recording of the parent polygons. Congruent features are coalesced and insignificant areas eliminated using a user specified tolerance parameter".

Laboratory for Computer Graphics & Spatial Analysis (1985, 24)

This process is achieved by WHIRLPOOL, whilst CALYPSO the other program in the analysis section deals with the thematic attributes associated with coverages, and generates appropriate values of thematic variables for each polygon in a particular coverage, when values were originally recorded for different coverages. Map display can be either two or three dimensional and centres around the original cartographic mapping package of ODYSSEY. As with the original mapping system this allows interactive map creation, although it has been criticised for its lack of 'help' messages.

The applications of ODYSSEY have been varied but centre around thematic data projection in both static and dynamic frameworks. These include mapping demographic data by special geographic zones which do not conform to standard data reporting districts and overlaying maps in sequence to build and monitor intersections of multiple coverages for environmental site planning or residential assessment over time.

**2. 4. Early Geographic Information Systems and applications.**

As a consequence of (and sometimes parallel to), such developments in computer assisted cartography occurring in the technologically driven sector of the user community, agency needs were prompting research into methods of organising spatial data, in addition to allowing its display by digital cartographic means. However, in the early 1970's, such GISs were constrained by the technological limitations of the time. The result was that such GISs were very much concerned with uni-source information assimilation, rather than overlaying and data integration. What integration existed was only that which could be done by hand. Graphic output, if it existed at all, was in the form of line printer output with some users initiating the transfer to vector type data.

Tomlinson (1972, 1631) classified the GISs available into five families based upon the method employed in retrieving and storing data. The first incorporated those systems that recorded information about each small element of a map grid, and included systems such as

SYMAP, which is regarded today as a primitive graphic display system offering none of the qualities associated with GISs. Another member of this family was the Wildlife Resource Information System (WRIS) which was utilised for the Stanislaus National Forestry Inventory and included multistage sampling, a mapping system and a growth projection to enable the calculation of the allowable cut. This was carried out via the production of two maps, the first encompassing 'management components' such as areas suitable for logging, the variation of ownership and recreational sites. The second was concerned with 'treatment classes', for example poor stocking, mature timber and plantations. Although useful for the particular project in mind, the problems encountered using WRIS were similar to those faced today by many GIS users and designers; the quality of the map input and the cost of its digitisation.

Family two in Tomlinson's gang of five concerned systems with similar criteria as family one, but with the difference of using a larger grid for information assimilation and manipulation. An example of one such system was NARIS (Natural Resource Information System) developed at the Centre for Advanced Computing at the University of Illinois. This system sheds little light on the possibilities of data integration but rather placed heavy emphasis upon information storage. The data base contained natural resource information pertaining to 40-acre tracts of land comprising the information compiled for 8 counties in northeast Illinois. There are fifteen classes of information under the major headings of geology, land use, forestry, soil and water. Each of these in turn was made up of data elements, which are composed of the initial values. Storage of the information was via a three tiered hierarchical structure, whilst storage of tracts took place according to their geographic identification label, derived from legally established Rectangular Survey System.

Tomlinson's third group related to systems concerned with descriptive data entry enabling a degree of topology to be embodied within the map, the most obvious example of this is GBF/DIME (Geographic Base File/Dual Independent Map Encoding) which today would be regarded as more of an information structuring tool for US urban areas rather than a GIS in its own right. The fourth group centred on packages that had parcel coordinate manipulation. Finally the fifth group contained the systems using fine grids and producing some form of geographical model; a definition perhaps most akin to our present day ideas of GIS. Several 'marine systems' for obtaining oceanographic information can be cited as members of this group.

The Swedish Hydrographic System developed by Thurnberg in 1962 was perhaps the most primitive due to the limited computing facilities. This system merely coded oceanic depths by onshore digitising and utilised an early FACIT computer. The output was of matrix type on an ordinary line-printer, which was then passed on to draughtsmen to prepare contoured charts. HYPOS developed by Kerr, Evangelatos and Marshall in 1969 was another such system, which utilised a more advanced method of digitising than the Swedish system as positional information was input separately on punch cards into the 3100 control data computer, with the output arriving on an early Calcomp Drum plotter. Even more sophisticated was the HAPPS system developed by Douglas in 1971. The input to this system was though an edo-echo sounder using a compact and reliable C-tech digitiser. Output was processed on a PDP8 incorporating what was then regarded as a large 8k memory and was available on magnetic tape or via the on-line Calcomp drum plotter. More advanced was CGIS, which was first proposed in 1963 as a computer mapping system planned to facilitate the use of data gathered by the Canada Land Survey. This was the first full scale GIS to carry out the functions of reading measuring and manipulating spatial data within a computerised environment. In its early stages (early 1970's) this system recorded scanned data on a fine geographic grid and was capable of map overlay (although these were not computer generated but hand scribed); the new coverages created as the result of the overlay process were however capable of being stored as new mappable 'objects'. Output however was limited to a printed report, with the hope that as the system developed, graphic output would eventually be available. As with many early systems CGIS was refined as computer graphics capabilities increased, and eventually polygon data handling was incorporated.

The main areas of application for the results produced by the early GIS were firstly for mapping and surveying purposes, as illustrated by Boyle (1974) in his use of GIS in hydrography and oceanography to indicate three dimensional water bodies by equating darkness of colour with depth. Secondly as most of the early GISs discussed have shown, there was a sector concerned with land use information either for planning and conservation purposes, or simply as an information base for local civic agencies.

Displaying cartographic data was useful as far as it went, but it became apparent that once data was held in cartographic form, it could become the basis for the accumulation of information on that particular spatial area. Furthermore, if data could be integrated from a wide range of sources irrespective of scale or structure then the possibility of analysis and modelling arose. Coupled with this was the fact that during the late 1970's, the introduction of computerised information retrieval and processing resulted in an increased quantity and improved quality of data available for academic, commercial and civic use.

**2. 5. Modern day Geographic Information Systems.**

So far the discussion has revolved around very early usage of GISs and the advances that were taking place in the field of computer assisted cartography, both of which would form the methodological background for the GISs we see today. By the late 1970's, it became apparent that although standard mapping packages such as GIMMS and statistical analysis systems such as SAS, SPSSX and MINITAB could provide analytical tools or illustrative facilities, the configuration of such packages did not allow for the input of a variety of data types (such as raster or vector) from several sources, and consequently they were incapable of integrating such data. In addition, such packages were unable to handle large volumes of available geographically referenced data, resulting in what Marble (1979) and Beck (1977) term a lack of usage and manipulation of information by social scientists. To maximise the use and spatial analysis of multisource data it became apparent that a new type of technology was called for. The gap was filled by systems, characterised by some cartographic ability for spatial data display and interactive graphic queries, and usually fuelled by a relational data base modelling system, which would allow easy interlinking of features held as spatial data and features held as alphanumeric files. The core of such systems were their data structure, which allowed the relationships between the features to be maintained and utilised. Operations included geometric overlay of cartographic features, the ability to query the data set using spatial parameters and the ability to accept data from many sources.

The previous sections have illustrated the underpinnings for much of the technological research that was responsible for the foundation of the GIS technique, but its adoption by the wider user community as a useful tool for spatial analysis was not so much evolutionary as catalytic. In the early 1980's few GISs were in existence (table 2.1), and those that were tended to run on minicomputers and mainframe machines, which put them beyond the reach of many users, "Until recently... GIS were too expensive and cumbersome for most users." (DOE, 1987, 7). The adoption of the technology, has been very rapid, and in the most part this was due to the catalysts of cheaper hardware costs, improved user interfaces and the greater availability of digital geographic data. Parker (1990) goes as far as to state that without a drop in hardware costs, coupled with improved and easy to use user interfaces, GIS would never have hit the major user, and would have remained simply a 'lab. technology'.

Table 2.1. Details of established GISs.

| **System name** | **Year first installed** | **Approximate price** | **Number of installations** | **Number of users** |
|---|---|---|---|---|
| ARC/INFO (ESRI) | 1982 | NA | 4,000* | 10,000 |
| ERDAS | 1979 | $2,000+ | 1,200 | 5,000 |
| GFIS (IBM) | 1977 | NA | 200+ | 200+ |
| SPANS (Tydac Technologies) | 1985 | $6,000 | 1,000 | 1,200 |
| TIGRIS (Intergraph) | 1988 | $10,000 | NA | NA |
| GDS (McDonnell Douglas) | 1980 | $9,500 | 1,000 | 4,000 |
| System 9 (Computervision) | 1987 | $12,500-$20,000 | 50 | 200 |
| Strings (Geobased Systems) | 1979 | $9,000 | 300 | 200 |

* It is unclear whether the ARC/INFO figures include PC ARC/INFO

Source: GIS World Inc. 1990: The 1990 GIS source book. GIS World Inc: Fort Collins, Colorado.

In addition to the improving and ever decreasing price of hardware, another important factor in the enthusiastic uptake for GISs was the gradual proliferation of digital 'geographic' data. This included official statistics such as the Census of Population, and the Census of

Employment, natural resource data such as ecological inventories and thematic remotely sensed imagery and infrastructure data, such as cables pipelines and transportation networks.[11] Along with the availability of data in this form came a review of the way geographic problems could be perceived, treated and assessed.

> "The GIS phenomenon is a result of several factors coming together at this point in history. Perhaps the most obvious factor is computer technology ... Less obvious, but equally important, has been the incipient increase in need for a technology like GIS over the past quarter century. As a result of increasing environmental concern, national and local infrastructures and the complex of pressures resulting from increased human population, GIS is a technology needed now."
>
> Parker (1990, 1)

In addition to statistical data, access to digitised outlines has improved and surveys and official documentation have been computerised. In both cases availability has been supplemented by the developments in 'networking'.[12]

Classifying the GISs that are currently on the market, or in use today, is a difficult task. Factors of functionality, size, machine dependence, cost and ease of use are all worthwhile contenders for a typology. Dangermond (1983) offers yet a further alternative by classifying systems by the problem they address. Thus he would see systems as falling into one of the following groups;

- engineering systems; public facility management and utilities.
- property-based information systems to handle cadastral data relating to land parcels.
- generalised thematic and statistical mapping systems as used for natural resource management, census mapping and environmental planning.
- bibliographic systems which catalogue information about data sets and geographical documents.
- geographical base file systems relating to functional and administrative boundaries.
- image processing systems associated with LANDSAT and other remotely sensed data.

---

[11]The Chorley Report (DOE, 1987) also cites land and property data such as property titles maintained by the Land Registries and property transactions held by the Valuations Office.

[12]This enables information to be passed between remote geographical locations via computer linkages.

Yet some of the above applications may actually use the same generic GIS. These are developed by established software companies and as they are non specific in orientation, they have a high versatility component and may be implemented by a wide range of institutions to perform a large array of tasks. There are at present a multitude of GISs available on the market, with the more established ones being offered by ESRI, INTERGRAPH, SIEMENS, SCITEX, SYSSCAN, GEO SYSTEMS, MCDONALD DOUGLAS, ERDAS, GFIS, TYDAC-SPANS, SYNERCOM and LASERSCAN. Originally these systems were designed for mainframe or minicomputers only, but in the mid 1980's the demand from PC users was regarded as offering a potential revenue far greater than the few and far between institutions possessing mainframe or mini computers. Thus for example ESRIs PC ARC/INFO, TYDAC-SPANS and ERDAS are all PC based. In the late 1980's the trend has diverted towards workstation based systems. These offer the slightly more affluent client, greater processing power and storage facilities within an affordable unit.

It would be too time-consuming to review all the major systems available today and a good overview of the current market is offered in Shand and Moore (1989), Foster and Shand (1990), Monmonier (1990)[13] or the GIS World Sourcebook 1990. However, the following systems will illustrate the diversity of users and solutions sought.

## 2. 6. Examples of modern day systems.

The majority of these systems perform similar functions although some have emerged as more popular in certain fields than others. Without a doubt the leader of the field at the time of writing is ARC/INFO (table 2.1) produced by ESRI (Environmental Systems Research Institute).

> "This leadership is the result of firsts in the development of geoprocessing technology. The ARC/INFO software was the first GIS software to incorporate interactive graphics capabilities with a relational DBMS. It was also the first to utilise a topological and relational data model where not only map features are stored, but also their relationships to each other."
>
> GIS Sourcebook 1990 (1990, 96)

In some ways this system has become the industry's standard, with other GISs offering 'full

[13]Monmonier offers a critique of ARC/INFO, the Intergraph corporation, ERDAS, and Laser Scan.

ARC/INFO compatibility' as a selling point. Originally designed for large mainframe machines and minicomputers, ARC/INFO became available for PCs in the mid 1980's and workstations in the late 1980's, and has subsequently been installed in over 1000 sites. ARC/INFO is essentially made up of two constituent parts: ARC, which deals with the cartographic components and INFO[14] which is a data base management system in its own right, and allows the user to manipulate the attributes related to the cartographic entities. Although based on vector data, the system allows the user to import rasterised data from either satellite imagery or other data bases as well as to digitise information manually. Other functions executed by ARC/INFO include data analysis and management, data display and cartographic output. Green et al (1985) consider the most powerful set operators to be overlay and storage and the ability to build buffers. The overlay and storage option allows two or more coverages to be superimposed with their respective attributes, and the results to be stored as a new cartographic object. The buffer option enables a zone or buffer to be created around a predefined attribute to a predefined width specification. This can for example be utilised for the study of land use within a specified distance of a proposed road or aircraft flight path, and is therefore a useful planning tool. Although ARC/INFO users tend to be varied, a large number of clients are those dealing with land parcel data, environmental planning and monitoring and more recently those looking at networking problems.

SPANS, marketed by TYDAC Technologies is another PC based GIS. It differs from ARC/INFO in that it is raster based, and is therefore better able to handle overlays. One of the problems of storing data in raster form is that the storage requirements become extensive, and therefore often make such a system unsuitable for PC use, due to its limited storage capabilities. SPANS therefore utilises the quadtree data structure.[15] As a result of this

---

[14]ARC/INFO also operates with other data base management systems such as ORACLE, DBASE (3,4) and INGIS.

[15]This is based on the idea of representing areal blocks of information which hold the attribute as one data item, rather than assigning an item value to each of the pixels making up the area in question. The graphic is broken into two equal sectors and if either of the sectors contain entirely within their boundaries a uniform area, then that sector is not subdivided further. Alternatively, if a sector has no such contiguous area, it is halved, the halved again until it meets the condition.

facility, this system has proved popular with users who have a majority of their data in raster format. In addition to raster input, the system also caters for vectorised data and allows overlay at a variety of levels. Unlike ARC/INFO, SPANS is oriented more towards modelling rather than simple integration. Overlays can therefore be performed by assigning weights to each of the component coverages, or alternatively on the basis of a pre-set equation. Other features include the ability to generate surfaces from point data, using interpolation routines. Unlike ARC/INFO, the database is not as transparent and entities cannot be manipulated with the same degree of ease. Output is also generally of a poorer quality as the system relies upon shaded raster output rather than vectorised plots. Never the less, it is the upsurge of systems such as this that have prompted ESRI to enter into a formal agreement with ERDAS, thus enabling both companies to market a product with strengths in both vectorised data handling and rasterised image processing.

Intergraph's TIGRIS (Topologically Integrated Geographic Resource Information System), is another large turnkey GIS, launched in the late 1980's. It differs from its other turnkey counterparts in that it incorporates object orientated coding techniques in a unified topological data model. Object-oriented extensions to SQL are used to query the database, but the user interface still resembles that of a traditional relational database system, with users being able to select which of the major relational databases they wish to implement (Intergraph 1989, Maguire et al 1990).

At the other end of the market, and from the other side of the Atlantic are systems marketed by Laser Scan in Cambridge UK. Originally founded as a physics research unit investigating the applications of laser beams to particle paths, over the past fifteen years the company has branched out into the application of laser technology to GIS. The result has been a series of individually marketed components from semi automated digitiser to plotter. Over the last ten years these have been combined under a GIS umbrella, initially with LAMPS, and more recently METROPOLIS and HORIZON. HORIZON is one of the first GIS specifically designed to analyse, manage and monitor assets and resources, thus allowing it to predict and assess the effects of changes on the natural environment.[16] In addition to the standard display format of maps HORIZON offers the user the opportunity to view data as realistic three dimensional terrain models. Applications include work on environmental impact analysis, the assessment of the polar ice cap melt and issues of global warming.

In addition to GISs that exist to manipulate areal information, others focus on linear networks. Some may be geared to model shortest path algorithms, whilst others deal more with the management of facilities such as electricity, gas and water. IBM's GFIS product is an example of the latter.[17] This system allows geographic location to be represented in standard map form, with the additional bonus of graphic representations for facilities at those locations. Thus for example an electricity network can be stored in traditional map form, but particular installations such as pylons can have their own diagrammatic scheme in addition to tabular information describing their attributes, should the user require it. Although the system has been around for some time (table 2.1), it still lags behind some of its major competitors. Presumably, this is partly due to the fact that GFIS favours larger machines, making it more expensive in terms of hardware than a PC based product.

Naturally such large turnkey GISs have a kaleidoscope of price tags attached to them, depending on a menage both of the size of the host to be utilised and also the facilities the system can offer. Turnkey GISs may not however, totally fulfil user needs. The high initial cost of the more versatile systems may outweigh their benefits, whilst the cheaper systems may not offer all the facilities required. The compromise is often a hybrid, consisting of an affordable commercial system with specific additions, making it tailor made to user needs. For example the GIS responsible for managing the routine of Municipal Service Vehicles in Oyster Bay, Long Island, New York was achieved by utilising

> "a database developed using a commercially available Mapping Information System (MIMS) and applications software were developed specifically for this project."
>
> Fagan (1986, 21)

---

[16]Mapping Awareness 1990, 4, 7, pp3-6.

[17]Often referred to as an Automated Mapping/Facilities Management (AM/FM) system.

For those agencies unable to afford the incorporation of commercially marketed application software, the alternative is to design a system from scratch, specifically suitable to the application in mind. IEMIS, is one such system designed by the Federal Emergency Management Agency in Washington DC to assess its emergency preparedness. The system enables decision making to be aided by a suite of simulation models involving the diffusion of contaminants to air and water, evacuation dynamics and hydraulic flow. Combined with a resource database it is hoped that this system will help to maximise local government awareness and ability to deal with emergencies.

**2. 7. Data structures.**

A geographic information system is made up of several distinct but compatible components. These may be classified as mechanisms for 'data entry', 'data storage', data manipulation', 'data display' and 'hard copy output'. Although all these components are important in ensuring the smooth functioning of the system and the production of quality material at the end of any analysis, it is how the data are stored and what structures are superimposed onto them, that will often determine the way the data can be manipulated, searched or accessed.

There is therefore a level of technological determinism that prevails in the structuring of geographic information. Systems force the user to structure spatial reality within the limited confines of a rigid data model. In many respects this requires the user to abandon any preconceived model of reality, in favour of an imposed concept that fits in with accepted digital data storage techniques.

Peuquet (1984) describes a GIS data base as the final result of a series of abstractions. The first of these is a data model, which incorporates only those properties of reality thought to be applicable to the application in question. The data model is then represented as a data structure, which is usually expressed in the form of diagrams, lists or arrays. Finally the file structure becomes a representation of the data in storage hardware. The difficulties arise in trying to maintain the description of sets of entities and the multi-dimensional relationship between those sets of entities in a form suitable for incorporation into computer storage and memory.

As a guide to evaluating the way in which data bases are able to sustain the relationship between entities held, Peuquet (1984) suggested a five point test, comprised of completeness, robustness, versatility, efficiency and ease of operation. Naturally the relative importance of each factor will vary with the type of data being used and the operational requirements of the system, thus she points out users handling a large volume of data will look for strengths in the last three factors and probably override weaknesses in the first two. Generally however, the tendency has been a trade off between efficiency and globality. Due to its very raison d'etre no model can mimic reality to perfection, rather it must aim to be as comprehensive as possible within certain predefined limits of efficiency.

Assuming that accuracy can be equally maintained in either case, efficiency may be defined in either temporal or monetary terms depending on the aims and needs of the agency using the data base. If time is relatively plentiful compared to financial outlay, then Chrisman (1984) indicates that it is in the storage of co-ordinates where savings in running costs are most likely to be gained, as storage contributes to 90% of the cost of geo-processing. The possibilities for storage reduction rely first and foremost on compaction. The drawback however is increased processing time, as compacting routines often carry high overheads. Other compacting possibilities include using an alternative type of hardware technology that uses fewer bits per co-ordinate. Chrisman concludes that

> "choices involve complex tradeoffs between programming simplicity, data accuracy and storage compression."
>
> Chrisman (1984a, 269).

In present day systems, a variety of structures have evolved to tackle the problem of data storage. Two commonly used structures in GISs are the quadtree data structure and the relational data structure. Quadtree structures store information in a condensed form. Rasterised images can be subdivided so that pixels representing the same feature are not repetitively stored. The image can be divided into four components (hence *quad*tree), and any areas that are contiguous can be represented by one value. Areas that are not contiguous can be sub-divided into four again and the same question reapplied. The process can be continued until a suitable level of resolution has been reached. The areas and a suitable indicator showing the status of the area are held in a tree based structure.[18] The relational data base structure treats each block of data items as tabular files. Items and their attributes are interpreted as rows and columns and each file or block of data can be attached to or 'related' to any other. Although each optimises a particular aspect of data storage, it is the relational data base model that has evolved as the most popular by offering greatest efficiency through its versatility.[19]

This popularity is due to several factors such as access paths, operational use and conceptualisation, all of which are well illustrated by Sandberg (1981). In the first instance data are represented in tables allowing them to exist as columns and rows, where the rows represent records or segments and the columns, fields within records. Therefore whilst access paths in network or hierarchical data structures are pre-defined via parent or child segments, in relational data bases there are no pre-defined routes, as all access is accomplished by the matching of field values, hence many different paths potentially exist. Therefore the relational approach

> "has considerable potential for extensions and restructuring and provides a very high level interface to the data structures as compared with data models that use pre-defined paths."
>
> Sandberg (1981, 28)

The major disadvantage of undefined access paths is the potential increase in processing time and therefore in costs, since the user has no idea of which access paths are internally favoured over others, and therefore cannot decide whether optimum, cost effective paths are being followed. The pragmatics of cost aside, the relational data base does offer a symmetry of access to all features and entities and also enables an equality of fields to exist. This in turn allows for easy programming access and the ability for painless restructuring of items, unlike network or hierarchical implementations which often require different coding techniques, depending on the predefined paths being followed, and in which restructuring often affects existing programs and procedures.

---

[18] A good discussion of this provided in Burrough 1986, Mark 1986, Samet 1984 and Samet et. al. 1986.

[19] Undoubtably, the fact that ARC/INFO utilised such a system also popularised it.

Operationally, relational data bases function as entire tables, with a command affecting a whole set of records, and the result of each command being a new table. As a contrast networked and hierarchical data structures operate on one record, rather than a table of operations, and the result of the operation is a single record. Therefore when dealing with sets of entities operations become both locally and practically easier to perform. Apart from technical considerations, the major strong points of the relational data base revolve around the ease with which the user may implement it. Sandberg (1981) sees these as identifiable in five main areas. Firstly there is the ease of understanding and easy conceptualisation associated with relational data bases, as most people have an intuitive idea of what a table is. Secondly, there is the high level of data independence, as all fields are explicitly known and seen by the user. Thirdly, there is the power and ease of use due to the lack of procedurality. Fourthly, the theoretical foundation of using relative data allows for clearer implementation of high level data access, whilst the heavy emphasis currently being placed on retrieval theory, rather than the complex operations of updating, has also served to increase its appeal to the less mathematical user. Finally, the language of table operations and data definition is uniform throughout the company of users ranging from data base administrators to query users and programmers. This contrasts with the variety of languages often used within networking and hierarchical systems, making interaction within those systems more complex. In addition, the relational data base definition can allow one user's view to be expressed in terms of another users' views, thus reducing programming effort and breaking down the barrier between data definition and programming.

Another method of storing time and space simultaneously in geographical data bases lies in the application of object oriented programming to GIS. This looks at objects in their entirety and adapts an holistic view of the phenomena in question. Thus an object will have attached to it a series of operations that belong only to that particular object as opposed to traditional programming systems where data and procedures are separate entities and the operator becomes responsible for applying active procedures to passive data structures.

> "Object oriented programming is based on a paradigm of objects responding to messages, rather than one of operators performing actions on operands as is the case with procedural languages."
>
> Kjerne and Deuker (1986,149)

The state of a particular object is implemented by being given private persistent memory, and its behaviour is implemented as a series of procedures or operations that have access to that private memory. Therefore the computer can be divided up into a number of smaller computers, each of which can be given a role like that of an actor in a play. When signalled to, each object then performs its part. Other important features of object oriented programming are encapsulation, message passing and inheritance (Egeinhofer and Frank 1989).

Encapsulation allows for the data to be packaged with the procedures that access that data rather than being organised into procedures that share global data. Thus particular geographic phenomena may have certain conditions and tolerances built in, which ensures data are appropriately overlaid or dissolved into neighbouring regions. In effect this allows for a data oriented principle for program design. Messages are organised into public and private categories. Public messages are generally accessible, but private messages are only executed by the object itself. Thus objects serve to group operations with the data they will transform. Inheritance allows new objects to be added which are slightly more specialised than the one above it, a property which has obvious applications to recording change in geographic data. Furthermore, Piazza and Pessaro (1990) suggest that object oriented techniques will provide a better solution to the problems of contextuality and cognition than currently implemented techniques. For example, OOPs is highly suited to the development of a cognitive model which can take account of qualitative factors influencing a particular scenario, furthermore they can easily handle semantics and templating which traditional methods find very hard. The major object oriented GIS on the market today is the Smallworld System and this will be discussed further in later chapters.

**2. 8. Storing time and spatial change within a GIS.**

Current trends indicate that the problem many GISs face with respect to the accuracy of the data they hold is its currency. Not only is there concern with currency at a particular moment in time, but also with the way the temporal nature of all geographical data can be accommodated and maintained. Maps and the data they display are only snapshots of an ever changing world. Data pertaining to land use for example will vary greatly depending on what part of the year it is gathered. The burden of dealing with temporal phenomena is therefore rarely handled by geographers or cartographers in the map making process, instead, the responsibility is shifted towards the user. The same can be said of the major data base systems used in GIS as these are based upon the way geographical phenomena have traditionally been recorded.

Langran and Chrisman (1988) suggest that digital data bases could provide a means of not simply storing static and non-current information but also of allowing the user to store and manipulate sequent states of geographic information, enabling the effects of time to be suitably represented. The first step in such a process would be to ensure that data are superceded but never deleted. A possible method to achieving this goal might be to reconstruct topological data structures in terms of temporal topology rather than spatial topology. If contiguous temporal neighbours could be topologically connected, then a temporal topological data structure could be devised to link one to the next thus avoiding exhaustive searches through time and space to answer questions such as; 'what has changed?', 'what is the periodicity of change?' and 'what trends are evident?'. Langran and Chrisman (1988) review three possible methods of storing 'cartographic time' with a data base structure; time slice snapshots, basemap with overlays and a space-time composite.

Time-slice snapshots rely on maps as 'photographic snaps' of a particular phenomenon in time, the main problem here is that boundaries of temporal continuity are difficult to locate since subsequent snapshots store both the change and all the redundant data. A basemap with overlays are an improvement in time slice snapshots. This method relies upon a map of the initial data followed by overlays of successive changes, thereby allowing neighbours to be located on adjacent overlays. This allows for changes at time $t_I$ to $t_{II}$ to be analysed without the need to go back to $t_0$. In this type of system the temporal structure is evident as neighbours can be located by finding or searching for the mutation that separates them. Space Time Composites offer even greater flexibility as here each change causes the changed portion to break away from its parent object and thus become a discrete object with its own history, thus the representation decomposes over time into smaller and smaller components with histories distinct from their neighbours.

> "In other words the representation decomposes over time into the areas' greatest common spatiotemporal units; each unit's history is described by a variable length list attribute sets bracketed by effective data."
>
> Langran (1989a, 19)

Such a system would find simple search procedures such as 'what is the state of a region at time t?' and 'what happens to a region over given time period?' as the temporal access mechanism for a space-time composite operates primarily on an attribute database which in turn is cross referenced to the spatial representation. Langran (1989b) outlines the specific steps involved in accessing and manipulating a spatiotemporal data structure.

More recent research on the problem of managing time in spatial data sets has tended to look towards the incorporation of object-orientation and version management. Most GIS systems today utilise RDBMS which operate by allowing users to only make additions to the most recent spatial state, and which therefore find managing several versions simultaneously a difficult problem (Easterfield et al 1990). They advocate an hierarchical data model (incorporating object oriented techniques), in which several versions of the same database could be 'open' at the same time, thus allowing several users the option to interrogate it. Even so, the onus still seems to be on *management* techniques rather than on technological devices.

## 2. 9. GIS and the future.

> "Such a system is as significant to spatial analysis as the inventions of the microscope and telescope were to science, the computer to economics and the printing press to information dissemination. It is the biggest step forward in the handling of geographic information since the invention of the map."
>
> DOE (1987, 8)

Such was the optimism expressed by the Chorley report in the late 1980's when it seemed that GIS was to revolutionise the handling of geographic information. Unfortunately, although the technology has become more advanced, the tool boxes easier to use, and the output more impressive, GISs still lack many true spatial analysis functions. Openshaw (1990) states quite emphatically that GISs perform few, if any true spatial analysis.

> "There is not much emphasis on hypothesis testing. Instead, there is a focus on data description, data simplification and exploration in search of interesting results, without any a priori clues of what data might exist, where to find it or what variables might be most important."
>
> Openshaw (1990, 14)

He states that map operations commonly found in GISs such as buffering and overlay should not be viewed as spatially analytical, but merely descriptive. Furthermore, a consequence of such depravity is that in the fullness of time, whole areas of missing functionality will be discovered, but in the short term, users will adapt increasingly to the lack of appropriate technology. Thus for the 1990's vendors should seek to answer simple questions such as;

- is there any pattern to this database?
- where is the pattern strongest?
- what, if any, strong relationship exists?

In many respects therefore, GISs are used as systems capable of sorting and managing spatial data, more often than they are used for spatial analysis or data modelling, and the following statement made in 1987, can still be said to hold true.

> "Existing Geographic Information Systems and the technology underlying them are still limited in a number of ways. The majority of such systems are still inventory systems, where all that is needed is the spatial referencing of data such as land parcels or plant, to facilitate spatial retrieval or cartographic reproduction."
>
> DOE (1987, 50)

Aside from improved functionality and easier user access, another important issue that shadows GISs into the 1990's is the question of accuracy and data quality. How is information being obtained and used in GISs? What are the quality measures undertaken? How is the information treated? In the next chapter, the accuracy, error, quality and generalisation of geographic information as used in GISs will be discussed.

## Chapter Three - Error and GISs.

> "For the sake of simplicity, divergence between the map and the earth's surface is termed 'error'. This word may be a bit harsh, but it should not imply that the map is 'wrong' just that the map is limited."
>
> Chrisman (1982c, 159)

### 3. 1. Introduction.

Whilst Chapter Two reviewed GIS capabilities in the light of the systems and their functionality, this chapter seeks to look at the systems in terms of the information they accommodate, and the way in which the technological capabilities of the systems can be downgraded by the quality of information fed into them. The degree to which information can affect the outcome of any analysis will depend upon its accuracy. Accuracy can be defined as a measurement of exactness or correctness. In terms of spatial information, this can be interpreted as the degree of precision with which features on the earth's surface can be measured and represented in cartographic form. The result of such measurements can then be translated into an accuracy statement. Historically, this has taken the form of a statistical value, indicating the degree of agreement between the represented value and ground truth. In terms of cartographic material, obtaining such a value will rely on both sampling and comparative techniques, and as such, it may not reveal enough about the suitability of the material for entry into a GIS. Thus although a mathematical figure can define accuracy in terms of deviation from ground truth, it cannot define it in terms of whether the product ought to be used or not. For many GIS users therefore, it is not necessarily a measure of absolute accuracy that is important, but one of relative or representative accuracy. In fact, it is the quality of a product that will determine its use. Assenting that a map is of suitable quality does not necessarily mean that the map is accurate, merely that it is deemed appropriate for a particular application. This in turn depends upon the amount of 'error' a particular application can sustain. Error is generally regarded as a measure of inaccuracy, or deviation from the absolute truth. Thus whilst a measure of accuracy might define how correct a feature is, a measure of error might define how incorrect a feature is, and again, in terms of determining the appropriateness of use, this implies more than a mere statistical statement.

> "the ultimate arbiter of cartographic error is the real world, not a mathematical formulation. I define error as the deviation of our representation from the actual state of affairs."
>
> Chrisman (1989a, 522)

Any data held within a GIS are inherently inaccurate, as it is merely a conceptualisation of the reality it seeks to represent, but for users, the fact that it can adequately answer a problem, may make it accurate enough. Trying to represent the real world, or at least the world as the majority of people see it in a data model is a difficult task, which can only be answered by imperfect solutions. Philosophers might argue that there is in fact no such thing as the real world, as the real universe is unknown, thus all worlds are abstractions of an abstract reality, and the only measure of accuracy we have is how well a product can represent such an abstraction. Such muses are beyond the scope of this thesis, and are also somewhat irrelevant since users of GISs regard the world in which they have to perform the tasks of planning, inventory or management as being very real indeed.

The most common geographical abstraction that we use to communicate information about space is the map. A framework of how the map acquires its information and how this in turn is incorporated into a GIS is illustrated in figure 3.1. Initial information is translated into abstracted information via scientific and cognitive abstraction, which calls into play a series of filtering processes. One of the most important of these will be the cultural perspective in which the map is conceived, and the intended use a map is to be put to. For example, aboriginal maps, medieval T and O maps, political maps used for propaganda and organisational maps such as those used by the Ordnance Survey or the USGS, all have diverse audiences and were drawn from various perspectives with different goals in mind. Naturally such information will have to generalise reality, in order to suit the purpose in question, and undoubtably all cartographic products will have spurious additions as well as unintended and intentional generalisations attached to them. These are therefore the first potential sources for intentional modifications of the real world and also for unintended intrusions into cartographic products.

Mapped information is then derived from the abstracted information via a series of cartographic techniques. Such techniques in traditional cartography are well known and rely on several forms of generalisations, such as those associated with measurement, scale, drafting and display. Unintentional inaccuracies at this stage revolve around factors such as the skill of the draftsperson, the width of the pen nib used, and the skill with which the information has been interpreted from the base information. For digitally held information, further generalisation techniques come into play. As data are incorporated into digital systems using new technologies, new methods are required to manage and represent this information. Data collection techniques can result in too much data for the available storage space within the system. Thus generalisation procedures are required in order to maintain effective data management policies. Furthermore, with the potential to store information in a scale free database, displaying information at a larger scale than which it was originally designed for, may render the information unaesthetically pleasing, thus requiring it to be cartographically enhanced. Alternatively, the data may be drawn at a scale smaller than that for which it was designed, thus requiring aggregation, amalgamation, merging or displacement.

Transforming the mapped information into the user's cognitive reality requires processes of dissemination and diffusion. How effectively the information can be communicated will depend upon the cartographic language that is being used, and on the user's ability to understand and communicate in the same symbolic language. Here there is the potential for use error, incorrect interpretation or perception of the information and poor management decisions all to reduce the accuracy of the user's image.

Ingrained in the cartographic process are a series of feedback loops, which prompt the continual refinement of the information creation process. These rely on the user having an opportunity to voice failings within the process, and in the GIS circuit, these have been plentiful and include issues such as the demand for standards, the demand for improved technology and interfaces. From those researching in the field a demand for more work on issues of generalisation and data quality have also fuelled these feedback loops. In many cases however, it has been the technology vendors that have determined the 'better' and more worthy routes for information and users to follow.

This framework of information transfer will be used in this chapter as a basis for the analysis of how error intrusions appear in the building of traditional cartographic products and from that, in modern day GIS databases. The early sections look at the problems of perception and cognitive generalisation, with subsequent sections looking at the problem involved in cartographic processes and in the process of GIS database creation and analysis. Final sections will deal with the issues of error minimalisation and review the current state of research into the issues of error removal within GISs.

> "Error can be viewed as the combination or result of a series of transformations which are imposed on reality, from the data capture to the display and map reading stages."
>
> Muller (1987a, 2)

**3. 2. Stage One: Scientific and cognitive abstraction.**

**3. 2. 1. Perception and generalisation.**

To make sense of the complex world in which we exist, some degree of simplification is necessary. In coping with spatial relationships, we tend to generalise space and our relative position within it. Such generalisations will vary to suit different purposes, and they exist as a by-product of our own personal perceptions. In its most basic form generalisation may be regarded as a simplification routine, within which there is a compromise

> "between what is geographically relevant, what is technically possible and what is aesthetically desirable."
>
> Maling (1989).

To generalise something implies a simplification or a reduction by combining several similar objects under broad descriptive brush strokes. In terms of intellectual concepts, this may adopt the form of simplified descriptions or metaphorical comparisons. Organisationally, categories and classifications may be employed. Where space is concerned, generalisation involves selecting and portraying the most distinctive spatial features, whilst allowing smaller, less dominant characteristics to melt into the background. Recreating world space as a graphic, depends on perception and the aim of the representation itself. Most individuals attach varying levels of importance to the spatial features they see around them, thus when describing their position with respect to their surrounding space, or whilst functioning within it, some features predominate, whilst others, though physically perceived, are not retained within the cognitive processes operating at the time.

In transferring our perceived view of reality to another person, or to a different medium, our perception forms the generalised envelope within which that spatial information is communicated. At times this envelope can be reduced in size to ease communication, but often the amount of generalisation that occurs as part of an individual's perception is unrealised and undetermined making it difficult to evaluate and assess.

The environment is seen, appreciated and acknowledged via our perceptive senses. One can either accept the notion that the perceived world is the real world, or conversely that the perceived world is merely a filtered view, peculiar to each individual and existing as a consequence of his or her interpretation of information received via sensory perception. It is the latter that is the generally accepted view in both cognitive science and experimental philosophy. Mark and Frank (1989) clarify concepts of space and spatial cognition within the context of small and large scale models of space. They assert that we tend to view small scale space in the following way:

> "our cognitive models of small scale space develop from direct perception of our every day world, dominated by a combination of visual inputs and the interaction of our bodies with the objects in that space."
>
> Mark and Frank (1989, 541).

We then fit together these small blocks to provide ourselves with models of larger areas. Cartography can be viewed as one of the major tools enabling such cohesion to take place.

> "The power of maps comes from the fact that they represent space with space. In fact maps represent large scale space in a small scale space on a piece of paper or a computer screen, allowing us to 'vicariously experience' the geometry of the large scale space in a familiar way, that is the way we experience small scale space."
>
> Mark and Frank (1989, 544).

### 3. 2. 2. Perception and Uncertainty.

Bedard (1987) looks at the intrusion of 'uncertainty' at various stages in the process of coding and encoding data in information systems. In the first instance, data has to be stored as a cognitive image in the mind of the observer. This image or model is then communicated to the central agency or information pool, where it is encoded into 'data' and a physical model of reality is created. Finally the receiver of the information must perform the inverse operation of the encoder and decode the data. This will, however, be carried out with the decoder's own reference and not that of the original observer. Therefore, the correct intended meaning will have to be selected out of a list of potential meanings. Such problems arise due to the impossibility of trying to define a continuum with discrete labels, and also due to the limitations of human beings as information processors. Dutton (1984) further dissects the process of encoding. He asserts that error or uncertainty is generated by several layers of variability in perception. Firstly, there is conceptualisation and the problems of identifying one feature from another due to the fuzziness of data. Secondly, description poses difficulties with recognising attributes and data values. Thirdly, there is the uncertainty associated with location and finally a meta-order of uncertainty, which is an accumulation of all of the three other groups. In addition most people adhere to a 'satisfying' concept, whereby all the possibilities for representation are not explored, rather the first solution to fit all the criteria is accepted. So by the very creation of the map itself subjective error is injected.

> "Our models of reality, including cartographic databases, are highly conditioned by our cultural and institutional consensus concerning 'what exists'. But that consensus varies across space, differs among groups, shifts over time and is coloured by our concerns".
>
> Dutton (1984, 179)

**3. 3. Stage Two: Techniques for creating a GIS database.**

**3. 3. 1. The problem of scale.**

In creating a model of reality, cartography veers away from a 1:1 representation; the irregularly shaped earth becomes a two dimensional plane; vast areas are reduced to portable images; local relief is substituted by a two dimensional description and all states and processes become static.

Thus any map, regardless of type, begins with a scalar reduction of the earth's features. The scale at which the information is to be represented within the map then determines how much, and at what detail information can be displayed. Not only will it curb the measured length of linear features, but it will also restrict the amount of attribute information, the symbols used and the number of valid classification classes.

> "The generalisation of cartographic features to support scale reduction must obviously change the way features look in order to fit them within the constraints of the graphic."
>
> Shea and McMaster (1989,61).

Naturally, the larger the map scale, the smaller the effects of scalar reduction, yet even at large scales maps will still be generalised, and thus it is untrue to assume that generalisation only happens at certain scales, whilst at others, data are empirically correct (Steward 1974).

In the creation of computerised digital cartographic databases, scale becomes an interesting issue as digital databases are allegedly supposed to be scale free. Yet some features have scale dependence, implying that they are designed to be represented at one scale only.

Buttenfield (1987) asserts that geomorphic processes have scale dependence, and that the graphic structure and character of a cartographic feature will change with the resolution of its display. This implies that the reduction of many features from reality to the map model,

> "is more than the mathematical relationship between the earth and the map; scale implies a specific decision about generalisation and aggregation."
>
> Chrisman (1982c, 161).

**3. 3. 2. Measurement of geographic features.**

Creating cartographic images calls for features to be measured empirically in the real world and then scaled down to fit into a reduced version of reality. No matter how accurate the measuring equipment is and how rigorous the measurement techniques are, inaccuracies will still persist and even in the best possible cases;

> "For many cartographic objects, however, the true value can never be determined either because no matter how much care is exercised in measurement, small random errors will be present, or because the ground truth is an abstruction rather than a tangible phenomenon."
>
> Muller (1987a, 1-2).

Cartographic entities may take the form of point features, linear features (be they naturally occurring lines or artificially defined borders), or alternatively they may be areal features. The measurement process can therefore be divided into the quantification of positional features and the quantification of attributes.

In terms of linear features, the geography of empirical lines can only be approximated and never precisely measured (Steinhaus 1954, Maling 1968). The problem, often referred to as the Steinhaus paradox, asserts that the length of any line depends upon the precision of measurement. The greater the precision, the greater the length of line. Perkal (1966ba, 1966b) also discussed this problem in terms of where the dividing line between two regions should go. He cites the example of measuring the line between the land and the sea. Even using a stick to measure the definitions between pebbles is not truly accurate as the areas could be more stringently investigated by observing grains of sand through a microscope.

> "A line separating two empirical regions is called an empirical line. There is no point debating its true nature, since its shape depends on the precision with which the observations are made."
>
> Perkal (1966b,2).

Therefore at some point a compromise will have to be made between accuracy and feasibility, and consequently a degree of generalisation must be accepted.

In terms of attribute or classificatory data, suitable generalised classification schema have to be developed and then the boundary of change between assigned classifications has to be defined.

> "Many soil scientists and geographers know from field experience that carefully drawn boundaries and contour lines on maps are elegant misrepresentations of changes that are often gradual, vague or fuzzy."
>
> Burrough (1986, 103)

Selecting suitable classifications is a well known geographical problem[1], and irrespective of what techniques are used, generalisation occurs. Classifying remote sensing imagery poses fewer positional problems than surveyed material, for measurement, but the size of the pixel resolution will determine the quality of data obtained, and the ease with which it can be accurately interrogated. When categorical data are being used, zones of change have to be represented by a straight line. Thus three dimensional features become reduced to two dimensional entities. Furthermore, if ground survey techniques have been employed, defining the boundary of change may rely on interpolation which will further generalise the boundary.

### 3. 3. 3. Source material.

In traditional cartography, the processes of map creation and their associated errors are well known and can be controlled. In building a GIS database, these processes are less well documented, and the error associated with each of them still requires further research. The abstracted information for incorporation into a GIS database will be available in several forms; these may be from primary information sources such as remotely sensed data, surveyed material, global positioning systems or cellular networks for easy survey updates.

Although such methods are now widely available, constraints abound, mainly in the form of cost, resolution and time. Remotely sensed data for example, is often unsuitable for large scale urban mapping, due to its poor resolution, or in cases where it might prove useful, the

---

[1] The problems of classification is still high on the research agenda for saptial analysis. See Openshaw 1990b. It includes statistical methods, different subjective classifications, (such as the Government's 'Standard Industrial Classification') and scale dependent processes.

cost of the data and the processor may be so prohibitive that the nationally surveyed topographic plan will be used in conjunction with local surveys.[2]

For the small time user, the majority of cartographic data entered into a GIS will not be from primary sources, but in the form of secondary data, derived from maps surveyed outside the users agency.[3] Topographic data are generally taken from the national mapping agency sheets at an appropriate scale, with other cartographic surfaces such as geology, agriculture, and utility networks coming from other local and national sources. The accuracy of such maps will depend on the efficiency of the equipment used to carry out the survey and also upon the skills and subjectivity of the surveyor.

In the UK most topographic maps used for digitising will be those surveyed and distributed by the Ordnance Survey. In many cases material displayed on Ordnance Survey maps may date as far back as the turn of the century, despite continuous revision;

> "Production is based on the principle that unchanged detail should not be redrawn, but should be carried forward from one edition to the next by photographic methods."
>
> Harley (1975, 12)

> "Ordnance Survey map makers have always had to work within the constraints of inherited triangulations, levellings, topographical field sheets and published maps. The fact for example, that much of the post 1945 large scales programme has taken the form of an overhaul of a pre 1939 survey has, despite experiments to achieve the best methods of adaptation, imposed absolute limits to the accuracy which can be achieved. The role of earlier survey operations in determining present day accuracy should always be borne in mind."
>
> Harley (1975, 161)

In the vast uncharted expanse of the United States and Australia original surveys carried out in the mid eighteenth century often formed the basis for state boundaries. The marking of one such boundary (Illinois) from the fourth principal meridian East to Lake Michigan, is described in a letter sent by John Messenger, Commissioner of the State of Illinois and

[2] Data capture methodologies and techniques are discussed in greater length in Chapters Four and Six.

[3] Although data may often be supplemented by internally surveyed material.

Lucius Lyon, Commissioner on the part of the United States to President Andrew Jackson. At the starting point a monument of stone about seven feet long, weighing about five tons was set into the ground at Mean High Watermark (MHW). At Lake Michigan the monument was less durable;

> "..an oak post a foot square and nine feet long was set firmly into the ground to the depth of five feet, in a small thicket of willows, about one chain from the water's edge. This post stands at the termination of the surveyed line, and from it a black oak tree, fifteen inches in diameter, bears south twenty degrees east, three hundred and sixty one links distant; a black oak twelve inches in diameter, bears north forty-five and a half degrees west, on hundred and fifty six links distant; a black oak twelve inches in diameter, bears north fourteen and a half degrees west, two hundred and thirty-one links distant; a black oak fifteen inches in diameter, bears north seventy five degrees west, on hundred and seventy two links distant; a black oak twelve inches in diameter, bears north one degree west, three hundred and eighty four links distant; and a black oak ten inches in diameter, bears south thirty degrees east, two hundred and twenty-eight links distant.
>
> The three first trees are marked with a notch in a blaze near the ground and facing the post, the three last are not blazed but have nails driven into them, out of sight about three feet from the ground on the side nearest the post."
>
> Reinhard (1984, 4)

This was a major landmark establishing a state boundary, it is easy to envisage the frailty of the markers in between.

Such inaccuracies came to light when cartographic material is used to make definitive legal statements, and linear features on a map are taken to be a perfect representation of a natural phenomenon. In Southampton in the state of New York during the mid 1970s a case known as the Dolphin Lane Case came to light whereby legal title to property was defined by the Mean High Water Mark. Various surveyors had used differing methods to determine the true course of the mean high water mark and subsequently arrived at a different positional value. The case went to court and the verdict was that the traditional method used by surveyors in Southampton was the 'correct one' and would become the law of the state. This relied not upon tidal observations, which were described by the New York surveyors as 'imprecise', 'impractical', and 'unsuitable', but;

> "instead they proclaimed the edge of the grass 'where the grass stops growing' is the traditional and only practical method of determining the boundary line in a salt marsh area. Since 1975 New York has a legal line of MHW which "one must go wading to see".
>
> Greulich (1980 MS-3-c-1)

The basis of this decision was that this had always traditionally been the way MHW was measured and the introduction of photogrammetric and remote sensing technology would result in a discrepancy between tenants whose land was surveyed by traditional methods and those who employed the more modern techniques.

> "The prime concern of the courts is the stability and predictability of title to real property. When a party conveys property by defining one of its boundaries by reference to a high water mark, that boundary cannot become an "unknown" or "variable" because the "means" for determining the HW mark have changed over the course of time or because a subsequent owner employs a different technique or procedure for its location."
>
> Glazner (1980)

Another famous example of cartographic discrepancy can be found in the length of the Yugoslavia-Greek border.

> "To the Yugoslavians their boundary with Greece is 262.1 kilometres long. To the Greeks the same boundary is 236.1 kilometres long. The difference is 25.5 kilometres."
>
> Nystuen (1967, 118)

This type of discrepancy reflects well the problem of measurement and scale. The more frequently a line is sampled the greater its length will be and the larger the scale onto which that line is transformed, the more accurate it becomes. Surveyors on both sides were probably employing different techniques, differing levels of precision and using different technology.

In addition to positional accuracy, classification accuracy can also waiver, in both topographic and categorical maps. In topographic maps attribute accuracy usually centres around labelling cartographic features such as making a building a church or defining a road as belonging to a particular class. For the lay person it is omissions or lapses in this qualitative information that condemns a map as being inaccurate.

"To the man in the street a map is accurate if it is generally correct in a qualitative sense as, for instance, if it shows an inn and the church and the path across the fields to them. Quantitatively, he is only concerned that they should be recognisably in their correct relative positions. He is likely to condemn the map as inaccurate if the "Inn" of the map proves in fact to be the "church" but less likely to do so if the path proves to be a few hundred yards shorter than shown on the map."

Crone (1953, 64)

Such inaccuracies become important when the map becomes employed as an instrument of truth. Sleigh (1976) cites several examples form the USGS volume map maintenance and accuracy standards including; 'your map shows a passable road up Skink Hollow, and I tried to follow it and broke an axle', and 'your map designates our local pond as Poison Lake. Everybody round here knows that the original Cajun settlers named it "Lac des Poissons" because there were so many fish in it. Your map with the "poison" label is ruining my bait business.'

In categorical maps, accuracy is far harder to pinpoint or assess than in their topographical counterparts. Here one faces the problem of defining a sharp precise boundary between two zones which actually drift into each other, as well as the difficulty of classifying a region which probably has variations within it. MacDougall (1975) believes that it is representations such as these which are responsible for the greatest error input in GIS.

"Error in source materials is highest in maps where boundaries occur in zones of transition, rather than along distinct edges between regions. Three cases where this commonly occurs are soil maps, slope maps and vegetation maps."

MacDougall (1975, 24)

Maps obtained from National Mapping Institutions such as the Ordnance Survey in the UK or the US Geological Survey will also suffer from what may best be termed as institutional error. That is, particular institutions have always relied on a particular method for data acquisition, organisation and map production. In the US the Geological Survey has relied on the scale quad-sheet, whilst in the UK data has also traditionally been viewed as isolated map sheets, although this is becoming less true for large scale maps, the bounds of which can be defined by the purchaser. However, the majority of outlets for Ordnance Survey material still retail traditional map sheets, and these are still used for most small GIS projects.

Problems arise however, when such data are used as input into what is effectively a seamless data base. That is a method of storage which joins all the map sheets of interest together to create a database that can then be drawn upon in whatever dimensions the user requires. Due to previous practices however, this often proves problematic since map sheets stored as individual entities never before had a requirement that they should perfectly match the lines and areas on its four bounding corners. In some cases the mismatch may be one where lines that are intended to meet run parallel to each other, in others lines may simply disappear at the bounding edge with no continuation on the next sheet. Resolving such problems usually entails the user either omitting detail and thus reducing the accuracy of the data base, or fudging an intermediate join between two non matching lines and thereby increasing the variability of accuracy within the data base. Furthermore, if a subsequent user then chooses to focus upon areas falling mostly within sheet boundaries an accurate representation of ground features will be seriously jeopardised.

It has been proposed therefore that mapping agencies no longer abide by the quad by quad mentality (given that such delimitations are relatively arbitrary anyway), but rather that any updates carried out on digital databases be administered entity by entity. For example, highways, then water features, then contours. Such notions, whilst being considered with a view to acceptance at the Ordnance Survey, have been regarded with some scepticism at the USGS, who still maintain that updates and production shall be continued on a quad by quad basis (Guptill 1989).

> "The main unit of data collection and revision will remain the map quadrangle."
>
> Guptill (1989, 438).

It would appear that the problem of adopting a scale free, seamless data base lies not in the technological implementation of such a scheme, but in pragmatic impediments. Even though the demand may be great from the user community the producer would have difficulty in justifying such developments, unless they were vital to their own internal use.

Not only are mapping agencies responsible for positional inaccuracies, but they are also prone to in-house conventions with respect to classification schemes. Thus maps obtained from one agency are quite likely to display a particular phenomenon using a different series of classification divisions than a map of the same area produced by a different agency.

Determining the degree of compatibility between two such data sources then becomes the responsibility of the data user, and is often based on intuition rather than factual background.

### 3. 3. 4. Digital map preparation.

Cognitive and institutional issues apart, traditional map creation involves physical input from what Chrisman (1982a) defines as the three p's; pen paper and person. As the traditional map form provides a substantial input into GIS databases, and since many GIS users abstract information from a map source by tracing onto a stable medium, a consideration of some of the drafting processes is worth undertaking.

Converting surveyed data into map space suitable for input into a GIS database usually requires some human input, often in the form of drawing or tracing. As we are all fallible most draftpersons and cartographers will suffer from eyesight variation, hand-shake and perceptual idiosyncrasies. The human eye, for example, is usually only accurate to 0.1mm. The pens used from such purposes also provide a further error input as in the best possible case the pen width would be 0.1mm, although a width of 0.2mm would be more common. In scaled terms this takes on the dimensions of an area rather than a line, one which can vary from its surveyed position by up to one pen width, before the operator has any hope of observing the discrepancy.

Finally there is the paper; the traditional map medium. Braund (1980) carried out an empirical study of the effects of temperature and humidity on a variety of drawing media. His conclusions stated that moving a medium from one environmental condition to another caused changes to take place in the lengths of lines, and this was most prominent when temperature and humidity were increased, where he found as much as a one percent change in length is possible. Secondly he noted that fluctuating temperature and humidity produces corresponding fluctuations in the media, whilst uniform conditions promoted minimal changes to the lengths of lines drawn on media. Finally on being returned to standardised conditions, many media returned to their original dimensions. It would appear therefore that under certain environmental conditions the reliability of information held upon paper media is variable and already limited representations of the earth's surface become even more deflected from the true nature of things.

### 3. 3. 5. Digital data capture.

> "Error and uncertainty have always been a feature of cartographic information so it is hardly surprising that these aspects are also present in digital versions of analogue maps."
>
> Openshaw (1989, 263).

Once data has been selected and presumed suitable for the application in mind, the transfer of source data with all its inherent 'source errors', to digital cartographic material can begin. Assimilating cartographic features into a computer environment involves either digitising or line scanning; depending on the available resources, and whether data are to be stored in vector or raster form. Digitising is essentially a sampling process in which the human eye, or set machine parameters select the most representative points on a line for storage.[4] Labels, or feature codes can then be added to denote what these points, lines and areas represent. Due to its very nature the process involves an element of generalisation and error;

> "Manual digitising resembles drafting so similar errors should be expected."
>
> Chrisman (1982c, 163)

Instead of using a pen to trace the length of a line, the operator will use a digitising tablet with a fine crosshair to pinpoint the characteristic points on a line. As with drafting, handshake, loss of concentration or eyesight deficiency may contribute to the injection of error. Unlike drafting however, the operator cannot see the discrepancy between the effectively 'traced' copy and the original map, as the map will be fixed to the digitising table and the digitised version will appear on a computerised screen. Pinpointing discrepancies at the time of digitising, therefore, becomes almost impossible. Such problems are often further compounded by the low rewards offered to those engaged in such tasks, as generally the jobs are underpaid and attract unskilled workers, whose only instruction is to press the keys on the console, and who therefore have little or no conception of the net result of their effort. As errors injected during this part of digital processing are entirely random in nature (unless there is a fault with a particular portion of the digitising table), removing them is a very time consuming and exhausting process, and one few organisations can afford in terms of time or money.

### 3. 3. 6. Making the data fit.

---

[4]This is discussed in detail in Chapter Four.

> "Many types of errors and uncertainties are a direct product of how we aggregate and represent natural phenomena in geographic information systems."
>
> Maffini et al. (1989, 66)

Once the data are 'captured' into the computerised environment, various management decisions are required to ensure that the data retain the accuracy they had on originally entering the system. Some features of computer storage are inescapable. For example, computer handling of data gives the digits held finite precision; a characteristic the features they portray do not necessarily deserve, given the less accurate methods of feature measurement and survey. However, others, such as the degree of post digitising generalisation adopted, and the level of geometric tolerancing sustained in cleaning up digitising errors are within the user's control.

The onset of digital methods attached a new demand to the process of cartographic generalisation, not only was generalisation required in the traditional cartographic sense, but it was also required as a solution to the very real problem of reducing the storage space taken up by redundant cartographic data. Easier means of transferability between systems and data structures, and the increase in the availability of 'raw' data also required a generalisation mechanism for the removal of redundant data. Furthermore many agencies were storing data on what could effectively be termed a scale free data base. Thus finished plotter products could theoretically be produced at any scale from the stored information. The possibility of a multitude of products from one stored set of information required various generalisation routines of both reduction and enhancement.

In manual cartography, generalisation was traditionally carried out by hand, and relied on the skill of the cartographer. In digital cartographic databases, as the number of points suitable for removal was often great, and since the computing power was available, manual methods of cartographic generalisation were superseded by automated algorithms and routines. McMaster (1989b), divides these into the broad categories of generalisation for data storage and generalisation for aesthetic improvement.

Aesthetics will be covered in section 3.4, but in terms of McMaster's first category, the major concern is to remove any errors introduced by the digitising process, and wherever possible to ease the load that redundant data places on available disk space. In the early days of digital cartography, computer storage space was often at a premium, as machines were smaller and had relatively feeble quantities of free storage space. Thus the advantages of reducing the quantity of stored points acquired through the digitising process were that storage space and plotter time would be reduced, and faster vector processing would be enabled. Holding data in digital form also meant that the data base was effectively scale free. Once the data had been digitised at a particular scale the product could then be generalised in order to make it adaptable at a variety of scales (Rhind 1973).

For vector systems, the initial techniques revolved around simplification and smoothing, later to be followed by enhancement, merging and displacement. In raster based systems, techniques dealt with smoothing and enhancement using high or low frequency filters. Primary needs involved reducing the number of points needed to represent a line, with secondary considerations being the need to compensate for scale reduction.

Problems associated with automated simplification routines is their potential to reduce the quality of the data, by removing points essential to the character of the line itself, thus whilst early simplification routines were Euclidean in nature, looking simply at the distance between points, later methods were more sophisticated, relying on complex geometric or angular controls to retain the 'character' of the line.[5]

For example, early algorithms were independent point routines and viewed the cartographic line as a set of equal points, any one of which was as important as the one preceding or succeeding it. This resulted in algorithms that removed every nth point and were therefore viewed as objective and in tune with the new wave of quantification and the scientific method that was sweeping through geography as a discipline in the mid 1960s.

When it became apparent that important points ware being lost Boyle (1970) suggested that

[5]Examples of such algorithms include the Douglas-Poiker algorithm, the Lang algorithm and the Reumann-Whitkam algorithm. McMaster (1989b) offers a good discussion of the major simplification and smoothing techniques.

weightings should be added so that important points did not disappear. Localised processing routines were next to emerge and these relied upon parametric descriptions that are automated by applying spline functions to lines (Buttenfield 1985). Finally there were more global routines that dealt with a line in its entirity and set out not to reduce information but to filter it. Such routines were based upon setting arbitrary tolerance limits within which points along the line could be eliminated. Perhaps the most famous of such algorithms is the Douglas Poiker algorithm of 1973, which selected coordinates by means of band width encoding, and as such the routine tried to match the capabilities of the manual cartographer.

> "The prime purpose of the routine ... 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."
>
> Douglas and Peuker (1973, 122)

Some attention has been paid to subjectively reducing points on a line by identifying and retaining the 'characteristic' points of a line. The objective is that lines emerge as being geographic rather than geometric. In this way it is hoped that the important point or message a line is trying to portray will be retained and not removed as redundant. Marino (1979) was perhaps the first to experiment with the identification of points on a line. She defined these in the following way.

> "Traditionally line generalisation has been a subjective practice. "Rules of thumb" existing as guidelines, stress, personal judgement, experience, skill and knowledge of geography. When generalising cartographers seem to stress certain points on a line an smooth portions in between. Those points which remain more or less fixed in position, resembling a precis of the written essay, capture the nature or character of the line, and in that sense, can be referred to as characteristic points."
>
> Marino (1979, 70)

Her experiments took the form of asking volunteers, both with and without cartographic training to select what they felt were the points that best defined a line. Two types of line were used; coastlines and rivers, and points were selected by the means of a predefined number of pins. Marino found that even when the pins were reduced in number, the same set of characteristic points were apparent. Furthermore there was essentially no difference between selections made by the cartographers and those made by the other group. Mark (1989) takes this idea further by suggesting that people with varying types of graphic training (such as a biological illustrator, draftperson, cartographer) may all be able to appreciate scaled drawings in each others' fields but are unlikely to attach the same kind of meaning to the representational lines held within them, and are therefore unlikely to be aware of what to retain in a selective generalisation process. He defines two subsets of generalisation; firstly visual generalisation, which works on the principles of computational vision and geographical generalisation which takes into account the knowledge an individual may have of geometric structures of geographic features or feature classifications.

> "Geographical generalisation aims to preserve the recognisability of geographic features of the real world and their interrelations."
>
> Mark (1989, 68).

Mark outlines three major objectives for line generalisation; the reduction of data volume, the preservation of visual appearances and recognisability and finally the preservation of geographic features and relations. Most objective algorithmic routines have attempted to achieve the first two goals, but few have addressed the final aim. Any attempt to fulfil such an aim would require the subjective knowledge of a geographer able to treat cartographic features an interrelated phenomenon, rather then isolated or abstracted lines. Mark (1989) points out that geographic phenomena often inherit components of their geometry from features of other kinds, for example contours are part of complex surfaces and shorelines are in effect eroded contour lines. Other man made features such as roads and railways also have connections with other phenomena, roads for example often follow rivers or shorelines, whilst also having constraints of curvature placed upon them, depending on their class or speed limit.

Buttenfield (1987) highlighted the problem of the interrelationship of features and processes in generalisation. She discovered that tolerances suitable for generalising one type of geomorphic feature may produce inappropriate results in other features due to the different processes responsible for their creation. Geomorphic processes were scale dependent and

> "automatic methods should therefore be sensitive to several levels of resolved detail for accurate definitions."
>
> Buttenfield (1987, 9)

The way in which geometric information is stored within most GIS data structures forces the user to make the data fit into a predefined framework. For the data structure to serve its goal, geometric rules defined by the topological data structure have to be observed; a situation which almost amounts to technological determinism. For example, in a vectorised, topological data structure, lines defining an area, *must* meet perfectly at their start and end points. Expecting that kind of accuracy from a digitising operator is unrealistic, and so post processing routines exist to ameliorate such situations. These will match the node points of *all* lines within a linear tolerance together, or alternatively remove all lines less than a linear tolerance on the assumption that they are no longer required by the user. In this way the data changes; locations move and the character of features alter.[6]

**3. 3. 7. Overlaying Errors.**

So far, the errors that have been discussed relate primarily to digital acquisition and storage. Naturally, using error prone data sets for analysis, in which data are combined to provide new information, results in an amplification of existing errors. Most modern GISs allow for the combination of data based on the elements of mathematical set theory. Layers of data can therefore be geometrically laid on top of one another in much the same way that a traditional cartographer would overlay a set of tracings on a light table and the following operations can be performed; intersection, union and identity. Intersection allows for everything that is common to both coverage A and coverage B to be retained. The union operation includes everything in A and everything in B, whilst the identity looks at everything in A plus everything in B that is also common to A. When such operations are employed to evaluate change where A and B refer to the state of features at different points in time, it becomes imperative that errors are not mistaken for feature change.

Generally however, the process of geometric overlay will result in small areas of uncertain origin or meaning, known as sliver polygons.[7] The majority of these will not arise due to veritable change between the two datasets, but as a result of the varying routes that the two or more data sets will have taken to reach the process of digital overlay. These in turn form a confusing mass of lines or areas, which have to be thinned out if any sense is to be made of the data. The method used to carry out such winnowing of redundant data is one based on a series of tolerance values. A tolerance is a value set by the user which determines the minimum allowable size of the feature in question. Lines may be thinned by setting a buffer tolerance around each line, and negating all other lines that fall entirely within that

---

[6]These are discussed in greater detail in Chapter Five.

[7]These are discussed in greater detail in Chapter Six.

predetermined buffer Chrisman (1982a). Sliver polygons resulting from overlay techniques are in many respects more problematic than lines or points as they enclose areal units, which in categorical coverages will respond to a classification value. Such areas are usually resolved on the basis of an areal tolerance, that is, any area that falls below a particular value is assumed to be an erroneous sliver polygon and is removed. Eliminating an area would, however, leave a black hole of sorts, therefore such areas are merged with neighbouring zones. Deciding which zone an area is to be coalesced with is usually based on purely geometric criteria, and so in most commercial systems sliver polygons are dissolved into their nearest largest neighbour, or into the neighbour with which they share the longest boundary. In effect the segment that divides the suspected sliver polygon from its nearest neighbour disappears. The basis of such logic is that larger areas can accommodate extra areas with less liability than smaller ones, or that long boundaries imply greatest similarity. Thus with this system certain areas get bigger on the basis of geometric criteria alone.[8]

Tolerancing procedures such as these are based upon pure geometry, and consequently are likely to sweep away some true cases along with the unwanted spaghetti data. This leaves the resultant map prone to gaps, or missing items of information. Worse still, in the case of areal data, polygons may be removed or reclassified in a way that belies the very components of change that are under investigation, and resulting in error prone maps.

### 3. 4. Transferring the message to the map user.

#### 3. 4. 1. Generalisation for effective communication.

> "In manually generalizing line data, cartographers have traditionally applied many decisions simultaneously. The most important features of the line are selected, simplified or exaggerated and perhaps smoothed and displaced from other features. In total it is a cognitive process and very difficult to understand.... the application of automated methods to generalization requires a logical breakdown of this holistic manual approach into discrete, precisely identifiable steps and ultimately into algorithms."
>
> McMaster (1987b, 331).

Once the GIS database has been established, the result of any analysis undertaken within it needs to be communicated. This may take the form of a written report, but more commonly

[8]This is discussed in greater detail in Chapter Seven.

it will be in the form of a graphic presentation. As with traditional cartography, features may be generalised to produced the desired aesthetic effect, but, with digital systems, a further issue arises, that of scale change. In what is effectively a scale-free environment, data can be graphically reproduced in a variety of scales. The consequence for the geographic features represented by the geometry of the lines is that they may be reduced in accuracy or aesthetic value. For example any of the following situations might arise; congestion, coalescence, conflict, inconsistency or imperceptibility. Each of these will require some form of generalisation to prevent clutter or aesthetically displeasing display. Common techniques are cited by Shea and McMaster (1989) and include smoothing, aggregation, amalgamation, merging, collapsing, displacement enhancement and exaggeration figure 3.2. There has been less work carried out on automated displacement and enhancement, than on smoothing and simplification, with research by Buttenfield (1989) and Monmonier (1989) being notable exceptions.

### 3. 4. 2. Displacement, merging and enhancement.

If there is a need to reduce digital information to a smaller scale format, it will be unlikely that features can retain either the same number of points in their depiction, or that they can simply be reduced without any loss of detail or any increase in clutter. The most common use of displacement techniques is when two linear features are planimetrically shifted to prevent coalescence (McMaster 1989a), with some of the most recent work in this field being carried out by Monmonier (1989). Sometimes, however, the converse is desirable, and routines exist to merge or omit features when they are depicted below a certain scale.

Enhancement has been justified on the basis that the visual characteristics a map reader expects to see are improved. Buttenfield (1985) believes this trend is also due to a shift away from determinism, in considering not only the line, but also the area surrounding it, thus acknowledging that generalisation is made up of perceptual as well as mathematical components.

> "For the map reader, generalisation should be considered as a transformation which preserves a delicate balance between measurable accuracy and recognisable character."
>
> Buttenfield (1985, 20).

Enhancement may simply be concerned with the symbology of the map, or alternatively with the spatial component (Shea and Mcmaster 1989). In the first case, as map scales are reduced, the original symbology may no longer be clear and accurate as in the original scale. As a consequence, the symbology will need to be altered to improve its readability. In the second case spatial enhancements may be performed via fractalization. This is based on the supposed self-similarity of geographic features, especially geomorphic features such as coastlines.

> "A feature is said to be self similar if any part of the feature appropriately enlarged is indistinguishable from the feature as a whole."
>
> Goodchild (1987, 268).

This implies that irregular self similar objects have constant fractal dimension and are unaffected by enlargement, thus giving them the property of being indistinguishable at all scales. Lines with properties of self similarity can readily be generated by a number of suitable recursive procedures. Goodchild (1987) cites those of Fourier, Fussesll and Carpenter in the early 1980s as being examples of such algorithms. This fractal simulation can add realistic yet spurious detail to angular generalised lines created in the digital process or determined by pixel size. Another potential application is in curvilinear data compression. The current dissenters at to whether fractal enhancement should be accepted or not rest their argument with the fact that simple metrics alone cannot sufficiently analyse the character of a line and furthermore it seems than many geomorphic lines are not self similar but scale specific (Buttenfield 1987).

### 3. 4. 3. Error and the map user.

So far, the error components discussed have referred directly to the paper map, and inaccuracies or errors within this medium have been placed at the feet of map producers and the technology they employ. Although cartographic error can always be minimised by the producer, it is the responsibility of the map user to assess the suitability of a particular map for certain uses and also to arrive at an acceptable error tolerance beyond which the information held is deemed to be unsuitable for the application in question.

> "If map makers are human, so too are map users. The qualities of integrity, judgement, critical acumen and the like are as much required in the interpretation of maps as in the preparation of them. Like carpenter's tools, maps should not be misused. More should not be expected of them than they can perform."
>
> Wright (1942, 543)

In this way, although cartographic error is absolute in nature, it becomes relative in value.

> "Error and accuracy are entirely relative issues, whose meaning and importance vary according to map usage and objectives."
>
> Muller (1987a, 2)

Such ideas on user error were put forward as far back as 1942 by J. K. Wright, he cites for example that;

> "A map exact enough for measuring the air-line distance from Boston to New York may be quite unsuitable for measuring the total length of the Maine coastline."
>
> Wright (1942, 531)

Ultimately therefore, the responsibility for determining how or whether a map should be used lies with the user, as the producer cannot bear the responsibility of incorrect conclusions being derived from their products since they have no control over how their product will be used once it leaves their hands.

> "The conscientious producer attempts to control misuse by maintaining scrupulous quality control during production, and hopes that once the map goes out for distribution it will be used in a reasonable manner."
>
> Beard (1989, 808)

**3. 4. 4. Use and generalisation.**

It is therefore the ultimate use of the map or spatial representation that will determine the selective elimination of its original detail, and so it becomes impossible to state one particular methodology as correct and of the greatest global applicability.

> "Because there are innumerable kinds of maps, there may be innumerable possibilities in the formulation of rules about the generalisation of information."
>
> Miller and Voskuil (1964).

Muller (1987) identifies three different map classes based on 'map purpose', and argues that each will support a different generalisation threshold. Firstly he cites reference maps, whose prime function is to store data. Maps in this category would include nationally surveyed map series, such as those carried out by the Ordnance Survey and US Geological Survey, as well as detailed distribution or categorical maps created as a result of large scale surveys. Secondly there are more simplified generalised representations, whose main function is to communicate specific geographic information. Included in this category would be the majority of thematic maps employed in geographic literature as well as more schematic cartographic representations such as cartograms and tactile maps. Finally there are working maps used to handle or analyse data.

In the case of reference maps, generalisation must be kept to a minimum with the determining factor being the scale at which the information is to be held. Maps that are intended for communicative purposes can withstand a higher degree of generalisation, in fact it may be the case that the removal of redundant detail portrays a message more concisely than a 'less accurate map'. Board (1967) suggests the more generalised a map the more useful it might be as does Blakemore: "The most generalised message seems to be the most powerful" (1985, 346).

Working maps that are consistently being used and updated are often as generalised as the user needs them to be, but as the producer is generally the consumer, generalisation needs can be targeted as and when they arise. GISs can hold maps in all three categories and often transfer data, between them. Thus unless data are specifically intended for only one use, generalisation should generally be kept to the highest limits.

### 3. 5. Feedback loops to improve spatial representation.

#### 3. 5. 1. Upgrading techniques of data capture.

Embodied in the model are a series of feedback loops which allow the GIS user to improve the collection and digital acquisition of material. The first of these would involve upgrading the techniques of data capture. This is essentially a technical issue, involving the technology associated with data acquisition. As such it is very much a concern of the vendors of data collection systems, and the suppliers of digital cartographic data. In recent years there has been a technical response to this requirement by an increase in the resolution of remotely sensed images. Although in terms of availability, problems still exist, as data are often expensive[9], difficult to procure, and problematic to process and classify. In terms of surveying, advances have clearly been made in terms of GPS (Global Positioning Systems) and Electronic Distance Measurement (EDM), implying that the more recently surveyed material should be of a greater accuracy than that used before.

[9]This has especially been the case since the privatisation of Landsat imagery.

Similarly, secondary methods of data capture such as digitising and scanning from secondary source material (already created paper maps) have witnessed an improvement in technical advances during the past thirty years, yet the data that these techniques are acquiring has remained the same. For example the most recent surveys by national mapping agencies may be over thirty years old.

Aside from directly acquiring digital cartographic information, either from primary or secondary sources, data may be obtained through a process of direct transfer. Receiving information in this way from other agencies provides the user with easily obtainable information of dubious quality. Thus testing data for quality and accuracy and ensuring a level of consistency in transferred digital cartographic data became a major concern for major mapping agencies in the early 1980s, and has led to the setting up in the United States of Federal bodies to investigate the production of national data standards.

> "Proliferation of digital topographic data ... necessitates establishment of national standards so that this terrain data could be exchanged for everyone's benefit."
>
> Zarzyeki (1984, 29).

Given the plethora of digital cartographic data producers and users (Hardy 1985) it became apparent that if data were to be efficiently exchanged between agencies, a national standard would be required to control both the quality of the data and also the means of its transfer.

Several bodies were set up in the US to evaluate and devise a national digital cartographic standard. The first of these was the National Committee for Digital Cartographic Data Standards (NCDCDS). This was set up in January 1982 under the auspices of the American Congress for Surveying and Mapping (ACSM) as a result of a formal request from the USGS. The second was the Federal Inter-agency Coordinating Committee on digital Cartography (FICCDC), created in April 1983 by the directive of the Office of Management and Budget to eliminate duplication and waste in the development of Federal digital cartographic data bases and also to serve as a focal point for the coordination of digital cartographic activities. The Standards Work (SWG) was formed later by the FICCDC specifically to develop digital cartographic data standards for use within Federal Government.

In February 1987 the draft proposal was issued by the NCDCDS and the FICCDC - SWG in a form that was expected to 'facilitate the use and exchange of digital cartographic data bases' (Moellering 1986, 1). This was sent to a joint mailing list of 1200 for comment and empirical testing of the standard took place throughout 1988. To resolve differences received from feedback and testing, the USGS convened the Digital Cartographic Data Standard Task Force (DCDSTF).

Cartographic data standards have of course prevailed though throughout the recent history of cartography, but as they dealt with paper maps, their concerns were more simplistic than those relating to digital maps. Generally such standards were dictated by the available technology and not user's desires; thus the producing agency was responsible for all the steps required to meet a quality threshold. Accuracy was defined using a statistical numerical value for a particular scale, and a map was deemed worthy if it fell within the acceptable 'error' limits and therefore conformed to the standard. An example of the components of such standards can be found in the comments made by Thompson (1960) when reviewing the current national standard as it was in 1960. Firstly he points out that accuracy should be couched in terms of standard error. Secondly, that vertical tolerance should be expressed as a formula in which the allowable standard error is a function of the contour interval and the ground slope. Finally that the horizontal tolerance should be changed to 1/50 of an inch or 1/40 of an inch for all scales.

The NCDCDS has moved away from such static, traditionalist ideas of accuracy to look more towards the quality of the data rather than to a prescribed accuracy value. Defining data quality has been the remit of Working Group II. This change of emphasis has partly been due to the wider range of applications and techniques spatial data are now employed in, and partly to the difference in the very nature of the data. Digital data has uses that will far extend the producer's initial estimates, and although this widens the user community and allows for a greater potential use of the data, it also opens the door to a greater potential misuse of the data, as information acquired second or third hand may be inappropriately handled.

> "The root of data abuse is; not in the quality of the data, but in the awareness and understanding of the quality of the data.... The evaluation and judgement

of fitness for use must be the responsibility of the user not the producer. To carry out this responsibility the user must be presented with much more information to permit an informed decision."

Chrisman (1986)

Furthermore, the fact that digital data bases are effectively scale free and can be stretched and squeezed to suit the users needs, renders redundant the notion of a deviation from the true ground value being representable in terms of measurable units on the ground.

"This 'digital' map has no scale in the traditional sense; therefore the criteria of accuracy and content associated with map scale must be replaced by new concepts which are applicable to the digital map."

Zarzycki (1984, 30).

Such considerations led the NCDCDS in the USA (and subsequently the Ordnance Survey in the UK) to adopt a quality control based on standards of information rather than on cartographic entities. The main concept employed is that of 'fitness for use'. The producer will not be obliged to achieve any particular threshold of quality, but will be required to inform the user of all the processes used in the creation of a particular data set.

"The philosophy behind the proposal ;is that the user of data should know as much about the history of it as the producer."

Sowton and Haywood (1986).

This fitness of use concept has been applied through a method entitled 'truth in labelling' Chrisman (1986).

"This means that the producer of the data has the responsibility of verifying the quality levels in a particular data base and providing that information to the prospective user."

Moellering (1986).

'Truth in labelling' can therefore be seen as a permissive standard (Chrisman 1986) as it allows the producer to distribute whatever product is obtained without having to meet a fixed threshold of performance. Describing such 'quality' parameters is, however, notoriously difficult and so some form of standard is needed for terms, classification and quality (Sowton and Haywood 1986). Working Group II of the NCDCDS defined five categories of quality information that the data producer is obliged to provide in addition to the data itself. These are, lineage, positional accuracy, attribute accuracy, logical consistency and completeness.

"The goal is a comprehensive scheme that can serve all forms of data, so the standards will prescribe specific numerical thresholds for; any particular product."

Chrisman (1985, 113).

This quality report will be issued as a free text paper document (or encoded on computer compatible media) and should be obtainable separately from the data so that the user can assess its fitness for use prior to purchase. The digital data itself may contain the quality report in whole or in part, but as a minimum it must contain a reference to the report and how it might be obtained (DCDSTF, 1988).

Perhaps the most vital component of the quality report is lineage of the data. This can best be described as a report containing the historical record of data production. It includes a description of the source material, how the data was derived from it and all the transformations used to obtain the final product. Any information obtained by merging data sets must be defined in enough detail to allow for the identification of each source element used. Control points, transformation and coordinates and registration points should all be documented. In addition as 'each section of the report will contain reference to temporal information and currency' DCDSTF (1988, 132), a period of validity should be attached to such a report.

Positional accuracy can be satisfied using several criteria. As a basis the existing standards were used and extended for the needs of digital cartography. Firstly, provided that all the procedures used in production have been carefully calibrated, then a deductive estimate of positional accuracy can be aimed at, based on the knowledge of errors in each step of data gathering and processing. Secondly, accuracy levels can be provided on the basis of repeated internal testing by the producing agency. Other methods include greater comparison to the source material when digitising and repeated testing of the data against an independent source of higher accuracy.

"The preferred test for positional accuracy is a comparison to an independent source of high accuracy."

DCDSTF (1988, 133).

This type of test must however be carried out using the rules prescribed in the proposed Spatial Accuracy Standard for large scale topographic maps. If one of the maps is of a different date to the other, then dates must be given so that temporal effects can be taken into account. Results of such tests must be rigorously documented in both ground units and the number and location of test points. NCDCDS are keen to point out that such tests can only be applied to well defined points.

Attribute accuracy is somewhat harder to quantify and assess, as this is one of the more subjective components of digital cartographic data capture, and is dependent to a large extent upon decisions taken by the individual data coder.

> "Unfortunately, coding of data is often the most subjective phase of the digitising process and requires human interpretation of the source graphic and coding scheme."
>
> Beck and Olson (1986).

Of prime importance in the quality report should be the date of the test and the rates of change in the phenomena classified. In addition any of the following tests may be used to define attribute accuracy; firstly as with positional accuracy, a deductive estimate can be given. In its lowest form this may be a guess in qualitative measures such as poor, good, fair, although obviously a quantitative measure is of greater use. Secondly, tests based on independent samples can be used, provided the sampling procedure and location of units are described. A misclassification matrix can then be produced as counts of sample units cross-tabulated by categories of sample and of test material. Finally comparative tests may be based on polygon overlay.

Logical consistency describes the fidelity of relationships encoded in the data structure of the data base (DCDSTF, 1988). Tests and results of such relationships must be detailed. General tests for graphic data include answering questions such as-

- do lines intersect only where intended?
- are any lines entered twice?
- are all areas completely described?
- are any polygons too small or are any lines too close together?

(DCDSTF, 1988)

Naturally the results of such queries must be included as well as the type of software used to perform the tests, and more importantly if any graphic failed the test it should be corrected, or if not at least identified. Specific topological tests may also be performed to assess whether data are topologically 'clean' or not. Data can be certified as clean if all chains intersect at nodes, if cycles of chains and nodes are consistent around polygons and if inner rings embed consistently in enclosing polygons.

Completeness should include information about selection criteria, the definitions used and other mapping rules. Tests to establish the exhaustiveness of a particular set of features should be carried out, and geocodes present should appear on the master list.

Another area of data quality deemed to be of importance is that of currency. Any data gathering process reflected in any cartographic representation is merely a general snapshot of reality as it exists in a particular moment in time. Certain features will 'age' considerably faster than others and given such circumstances some indication of the expected duration of the data as a piece of quality information should be included in all stages of a quality report.

Standards were not only being evaluated and tested for their ability to define quality specifications, in addition a data standard implies an aura of uniformity and consistency in the data being handled and in the methods by which it is transferred. Thus other tasks of the NCDCDS and FICCDC has been centred around defining a set of primitive and simple cartographic objects (NCDCDS Working Group I) and a unified set of cartographic features (NCDCDS, Working Group III). Once these are suitably catalogued, an effective set of data exchange methods can be considered.

As digital cartographic users are no longer limited to surveyors and map makers, with users coming from such a wide variety of backgrounds, it becomes imperative that terms used to define and manipulate cartographic entities are standardised and an acceptable glossary of terms is adopted. The NCDCDS definitions of cartographic features are intended to produce a comprehensive and compatible set of features for users. Examples of such definitions include; a feature, defined as an entity of interest that is not further subdivided; an attribute, classified as a characteristic of a feature; an attribute value, regarded as a quantity or quality assigned to the attribute and feature/attribute classes which contain specified groups of features or attributes. The structure within which the features are held is non-hierarchical, with the total number of features definition being reduced by simplification to between 150 and 200. Details pertaining to each feature are then provided by the attributes. It is hoped that such a structure will be more amenable to computer processing and machine data base systems.

Evolving a set of digital cartographic exchange forms and modules is more problematic than defining a glossary of terms. The NCDCDS and FICCDC-SWG went about this task with four aims in mind (Moellering, 1986). Firstly that a mechanism for the interchange of digital information had to accommodate non communicating parties using dissimilar computer systems. Secondly, to ensure that the digital objects and relationships which represent real world cartographic entities will be translated into the set of objects, relationships and information specified by the standard for the purpose of interchange such that their meaning will be preserved and can be understood by the receiver of such data. Finally to ensure the following characteristics hold true;

- the ability to transfer raster, vector, grid, attributes and other ancillary information
- media independency
- the ability to process any description of data types or formats into the users native system
- data an media formats are based on existing FIPS, ANSI or ISO standards.

The first phase of the NCDCDS proposed standard culminated in 1985 with an interim standard, which has recently undergone tests in various sectors of the digital cartographic community and were approved as a national standard in 1989.

The United States have not been alone in developing digital cartographic data standards. In the United Kingdom similar moves are being made to achieve some standardisation of cartographic material, following recommendation 25 of the House of Lords Select Committee on Remote Sensing and Digital Mapping (House of Lords 1983), initially under the auspices of the Ordnance Survey (Sowton and Haywood 1986). The philosophy behind such transfer standards is similar to that advocated by the NCDCDS, in that the transfer format must be able to handle all of the data generated by any cartographic system; techniques used for the transfer format must be capable of being adapted easily for any computer on the market and finally that all costs of implementation and use be kept to a minimum. As with the NCDCS, the main areas of research are feature classification, data quality and transfer formats.

Other countries developing standards include Finland, South Africa and Australia. Generally these all have similar aims, although administrative and institutional difficulties are often peculiar to each nation. Vahala (1986) for example illustrates the difficulties of exchanging data collected in Land Information Systems between separate administrative authorities in Finland, and outlines the efforts underway to create a set of classification standards that will allow a common reference system based on location to be built and hopefully to overcome difficulties in data transfer logistics. Cooper (1989) sketches the main components involved in the South African standard. As this is one of the later standards to be developed the designers have benefited from experience obtained by early innovators such as the USA and Australia. In this respect the authors believe it to be the first standard to cater for all forms of geo-referenced information (rather than limiting itself to either cartographic or hydrographic information as other standards have). The only drawback in such a flexible standard is that users wishing to transfer simple data sets may suffer from a lack of simple guidelines and in such cases good worked examples would go some way to ease this problem.

It is inevitable that the continuous production of data in large quantities such as those generated by remote sensing platforms, and the users' desire to address a wider range of applications without having to engage in data capture thereby duplicating the data acquisition exercise will necessitate and even encourage the adoption of such standards. The main impediment to the speedy adoption of such standards is a monetary one. Many smaller users may be unable or even unwilling to meet the high initial cost of adhering to such a system, as they may feel their needs do not justify data of high quality. Such viewpoints may be justified where the user operates as a discrete unit in a closed system, not receiving or exchanging information. Even so, some internal system of quality control and standardisation of production should be applied. Chrisman (1986), however demonstrated that the proposed standard did not necessarily imply an increase in the cost of production or maintenance.

> "Over the process of designing the standard the Working Group has had to consider cost of implementation along with technical needs. The Bell South test provides some evidence that the standard will not create large dislocations. Because it was retrospective the test shows that Donohue and Aero Metric Engineering were able to assemble the information for the quality report from existing archives."
>
> Chrisman (1986, 55)

As such standards become more widespread it has also been suggested, that the user community should not have to pay for the burden of implementing standards, especially if they lack programming skills. Rather the responsibility should shift towards the software manufacturers, who should incorporate such features within their products.

The refinement and adoption of data standards is progressing fast, both nationally and internationally, and once such standards become widely accepted and adhered to by mapping agencies in various countries problems of data transfer should start to diminish.[10]

### 3. 5. 2. Upgrading GIS processing techniques.

GIS processing techniques have a fundamental problem in that they operate on data as if it were scale free, and yet the majority of data input into a GIS is scale determined. Thus any upgrade of the way data are acquired, edited, stored or processed in a GIS has to be enacted within the constraints of an imperfect source material. At present the situation within the GIS community is one in which naive approaches to error predominate (Openshaw 1989). These are due to a variety of reasons.

Firstly, there is no real consensus that error and accuracy are problems which actually matter, and until users fully comprehend the consequences of poor data quality, little can be achieved. In many cases, current practice is a continuation of past attitudes. These have tended to take the form of looking to the producer of cartographic material for an accuracy statement, or simply relying on the software to flag topological inconsistencies in the data. Both of these options attach little importance to the quality of the information itself. Secondly, even if the user was aware of the problems associated with the use of material from several sources, there are no widely accepted methods for measuring the uncertainty of GIS products, or of tracking error through a coverage. Thus users are at a loss as to how potential problems can be flagged, and even if users devise their own means of doing this, there are no established methods for dealing with error in GIS coverages. Thirdly, due to the lack of empirical research, the extent of the problem is unknown, and there is little knowledge of the effects spatial operations may have on data. This in turn fosters a lack of awareness that a problem actually exists, and so a vicious circle of ignorance and inertia

---

[10] Papers such as Mason (1990) give an indication of current developments.

builds up.

> "The typical end-user of GIS output will probably care or know little about the cartographic uncertainty and error characteristics of the map data being used, whilst the GIS itself is perfectly happy to ignore varying accuracy and reliability of the digital map data being processed."
>
> Green (1990, 106)

Perhaps the first step towards improving GIS error response would be an undertaking on the part of GIS vendors to enhance user awareness of the possible risks associated with ceratin GIS techniques.

> "In reality a GIS may encourage poor analysis by separating the data collection, compilation and analysis functions, and failing to make the user aware of the possible dangers of indiscriminate use of such functions as scale change, reclassification and overlay."
>
> Goodchild (1989, 107)

At present, aside from having any mechanism for dealing with the problems of error propagation, GISs may actually encourage users to carry out operations that the data itself does not warrant. There is a growing adage amongst some users that the technology of GISs lulls the user into exploring more and more functions, which perform visually impressive tasks, but which allow us to 'produce rubbish faster and with more elegance than ever before'.

> "Geographic Information systems encourage the user to do things which are clearly not justified by the nature of the data involved."
>
> Goodchild and Gopal (1989, xiii)

A second initiative would be the introduction of facilities into the GIS 'tool box' that are capable of dealing with error and uncertainty. At present, no system offers the user a mechanism to estimate the error associated with imperfect data, neither are there opportunities to trace the propagation of error as a consequence of GIS techniques.

> "a problem arises because GIS packages fail to offer any means of keeping track of the effects of input data uncertainty and error propagation throughout sequential operations."
>
> Green (1990, 107)

> "none of the currently operational GI systems have functions to specify the errors associated with GIS data and to calculate the propagation of errors through GIS operations."
>
> Heuvelink et al (1990, 453)

> "No current software system tempers its results with an error estimate."
>
> Chrisman (1989b, 22)

> "No current GIS warns the user when a map digitised at 1:24,000 is overlaid with one digitised at 1:1,000,000, and the result is plotted at 1:24,000, and no current GIS carries the scale of the source document as an attribute of the dataset."
>
> Goodchild and Gopal (1989, xiii)

Therefore what the user requires is some ability to evaluate the information that is to be utilised. In part this will be addressed by the gradual adoption of digital cartographic data standards, but in addition the user needs to be able to make decisions such as how the functions incorporated into the system might affect the reliability of the data.

> "While the goal of eliminating errors from output products may at present prove to be infeasible, decision-makers should at a minimum be provided with a means of assessing the accuracy of the information upon which their decisions are based."
>
> Veregin (1989, 3)

This can only be achieved by empirical studies which measure error at each stage of GIS analysis. The execution of such research will allow a model of digital cartographic error to evolve, which in turn would offer clues as to how error can be avoided and dealt with.

> "Once errors are understood they form the basis for probabilistic statements. The analytical strengths of many sciences are based on stochastic models, not deterministic structures with zero error."
>
> Chrisman (1989, 22)

One of the major areas that needs to be addressed is that of overlay analysis. This is perhaps the most widely used GIS operation, and also one in which error can rapidly accumulate.

> "The routine compositing of map layers without regard to accuracy will then give rise to composite products of unknown integrity."
>
> Newcomer and Szajin (1984,62)

In general the accuracy of the composite map used in overlay analysis will generally be less than the accuracy of the least accurate map that is utilised in the analysis (Newcomer and Szajin, 1984). Furthermore, the greater the number of layers that are utilised, the greater the number of possible unknowns and inaccuracies.

"Conventional wisdom would say that as you add more data to the solution of a problem, the likelihood of getting an accurate solution increases. However, if each additional data layer degrades the quality of the combined data set, and hence the accuracy of the solution, then adding data sets may be counterproductive. It may be the case that there is an optimal collection of data sets for a given problem; adding or subtracting data lowers the probability of a correct solution."

Guptill (1989, 97)

Several writers in Goodchild and Gopal (1989) put forward suitable research agendas. Veregin (1989) for example classifies studies in error analysis as falling into a hierarchy of error needs;

1. Strategies for error reduction
2. Strategies for error management
3. Error propagation modelling
4. Error detection and measurement
5. Error source identification.

Maffini et al (1989) suggest that three main issues should present themselves in any study of error in GIS databases and operations;

- illustrate the size and range of some types of errors that are generated in the digitising process
- explore ways of dealing with these errors in derivative GIS products
- make some observation and suggestion for research that may help to comprehensively cope with errors

Openshaw (1989) synthesises this further by stating that in fact there are only two issues on the agenda. Firstly the development of adequate means to represent and model uncertainty, and secondly, the development of GIS related methods and techniques that will take error into account during their operations with spatial data.

Several models of digital cartographic error have come to light in the last decade. Yet these have tended to be more deductive than inductive in nature, and have therefore relied upon stringent testing rather than abstract design. Earlier models referring to the cartographic line per se were more abstract and relied on arithmetical geometry. One such model was that devised by Perkal (1965, 1966), which looked at the idea of an epsilon distance around a cartographic line as a basis for its generalisation and measurement.

The epsilon distance or 'e' is the radius of a small circle, the edges of which when rolled along an empirical line such as a coastline, generates a new line approximating the empirical line, but having a curvature at any point not greater than that defined by the 'e' radius of the circle. The 'e' neighbourhood of the line is measured by taking into account the area between the path of the circle on either side of the boundary thus producing a definition of the 'e' length. All lengths however, depend on stating the 'e' radius, and any discussion of the length of such lines without doing so has little meaning. Although Perkal's ideas acted as a springboard and led to many models of cartographic lines and their properties such as those proposed by Steinhaus and Mandelbrot (1977), there was until recently a sparsity of work on modelling digital cartographic error. Chrisman (1982a, 1982b, 1985, 1989) was perhaps the pioneer in this field. He recognised that the quality of the data held within any digital cartographic data base was dependent upon the processes and methods used to collect, assemble and integrate the data.

> "Each error effect relevant for a particular map can be treated as a random variable, perturbing the true line to obtain the observed line. A crucial part of this analysis hinges on combining these separate error effects."
>
> Chrisman (1982b, 163).

Chrisman (1982a, 1982b) devised a model based on Perkal's idea of epsilon distance. This assumed an error band around a line, within which the true line would lie. The error or epsilon band was defined by taking into account the error-bound processes that would have been employed in constructing the map.

> "Each of the separate sources of boundary error should be combined to form a comprehensive model of the overall variability or reliability of the map. The epsilon model takes the expected deviations of lines and gives a geometric form for the location of error bands. By measuring these bands the expected variability of area measurement can be determined."
>
> Chrisman (1982a, 115).

The model was used to assess the accuracy of USGS maps by measuring the area of the epsilon bands, created by taking into account the USGS map production process. He found that his figures come within 85% of the classification accuracy goals stated by the USGS, but the error model only accounted for those errors arising due to boundary errors and took no account of errors induced into coverages as a result of misclassification. Since this study, the epsilon model has proved popular with several researchers looking into the problems of error accumulation in GIS databases, notably Haemers (1990), Green (1990), Dunn et al

(1990) and Chrisman himself readdressed it in 1989[11].

Other statistical models of error such as using a Monte Carlo simulation technique have been addressed by Green (1990) and Openshaw et al (1990), with other quantitative techniques being explored by Veregin 1989 and other contributors to Godchild and Gopal (1989). However, models of error are still few and far between. Furthermore Chrisman (1989) warns of the danger involved in statistical techniques if the foundation for them has not been clearly laid.

> "other sciences engage in stochastic modelling from a firm foundation in measuring their phenomena of interest.....by contrast, the mapping sciences have not developed a comprehensive taxonomy of what errors occur and what processes control the amount of particular kinds of error.....it is premature to develop stochastic models for a field without a clear understanding of the fundamentals."
>
> Chrisman (1989a, 522)

Rather than the adoption of error models as an estimation of the worth of a piece of GIS material, Chrisman asserts that empirical tests were of greater import, and that efforts should be concentrated in that area.

> "While other research emphasises mathematical models to simulate error, a practical test provides a more useful judge of cartographic data quality. A comprehensive test, overlaying two categorical maps intended to be the same, can provide an estimate of separate components of error including positional and attribute accuracy along with scale effects."
>
> Chrisman (1989a, 521).

### 3. 5. 3. Upgrading effectiveness of use.

> "Two GIS capabilities that excite enthusiasm among potential users are the ability to change map scales and the ability to overlay maps at random. Both capabilities are indeed exceedingly useful ... Both capabilities may also mislead decision makers who are unaware of the imprecision inherent in all cartography and who are untutored in the ways errors compound when map scales are changed or when maps are merged."
>
> Abler (1987, 305)

Part of the process of improving GIS processing techniques relies on a change in user attitudes, and this has been discussed in the previous section. A newly developing area of study has been into use error (Beard, 1989). This places the onus for maintaining accuracy

[11] Chrisman (1989b).

on the shoulders of the user, as it becomes the user's prerogative to assess the suitability and quality of a particular piece of data for analysis. This is important as the consequences of misuse can be quite serious, for example, "a single case of misuse can cancel all investments in source and process error reduction" (Beard 1989, 810). Limiting such errors could be achieved by more extensive data quality documentation, improvements in updates to maintain accuracy and structuring data to avoid illegal or illogical operations (Beard 1987). The adoption of digital cartographic standards with its emphasis on data quality and lineage should also go some way to improving the current state of affairs.

An obvious route to further improvement lies in education. Here the NCGIA (National Center for Geographic Information and Analysis) in the US, the RRLs (Regional Research Laboratories) in the UK and the AGI (Association for Geographic Information) also in the UK should be instrumental in raising user awareness, and hopefully in supplying some of the solutions. Part of the educational process should involve an admission that error is inevitable, and therefore impossible to entirely eradicate using present technologies.

> "Error in digital geographic data is inevitable. Errors cannot be eliminated by simply improving the performance of instruments and/or operators."
>
> Maffini et al (1989, 66)

> "The statistical concept of error is hard to grasp by those who come from the long tradition of manual cartography. As the term is used by traditional cartographers, error is a bad thing, and the profession has the intention to wipe out all error."
>
> Chrisman (1989b, 22)

### 3. 6. Conclusion.

Confronting error as an issue within GISs will depend initially on the type of data used, and the application to which it is to be put. For example, users of small scale cartographic information are less likely to be concerned with detailed inaccuracies than users of large scale sensitive cartographic material. Table 3.1 illustrates this point.

Table 3. 1. Typology of error with respect to data requirements

| | Application area | Example of data use | Perceived importance of spatial/temporal error |
|---|---|---|---|
| Military | Terrain models | Precision bombing | High |
| Utilities | Network models | Pipeline repairs | High |
| Navigation | Traditional navigational charts/networked models and routing in real time | Air/sea/traffic routing and journey management | High |
| Local authority planning/environmental assessment | Geo-referenced large scale maps | Project assessment | Relatively high; especially in 'sensitive' areas |
| Resource monitoring | Distribution mapping and temporal recording | Monitoring change in a particular resource | High-Medium; depending on the value of the resource in terms of retail/conservati on premium |
| Resource management | Distribution mapping and cataloguing | Inventorying | Medium-Low Error is less important with a broadly distributed resource |
| Market Research | Georeferenced material based on geodemographic divisions of the population | Geographic location of 'Target areas' in terms of maps or address lists | Low, given that data is so highly aggregated. Results are evaluated in terms of increased profit rather than spatial accuracy |
| Administrative Mapping | Small/Medium/Large scale maps | Qualitative cartographic representations of particular phenomenon | Low |

In recent years, several methodologies have been used to address error in GISs. If error intrusion is viewed as an illness worthy of eradication, then a number of possible treatments exist. Perhaps the most naive way of addressing the intrusion of error has been by regarding it as a problem of geometry alone. For example most commercial GISs deal with sliver polygons by removal on the basis of area or area to perimeter ratios[12]. This looks at the symptoms rather than the illness. It is not the lines necessarily that are at fault, but the underlying philosophical foundations on which they are built. and any amelioration routines based on such ideas can never really address the root of the problem. The problem for GIS is not merely assessing geometric and numerical differences, for all they represent is a surrogate of an inter-reacting series of institutional, policy, sampling and other issues. What really needs to be addressed is the problem of how information is evaluated. The main drawback here is that education is a management skill, and management skills are not easily automated in a database environment.

Another methodology would be to look at the origin of the lines and what their original meaning was. This can be equated with the lineage paradigm embodied in the NCDCS recommendations, and is akin to looking at the history of the illness. The medical record is produced and a diagnosis performed. A final approach might be to look at an epidemiology of error. An explanation of the error could be given, and measures could be developed to limit its extent and its effect. The remainder of the thesis is dedicated to such an approach. Chapter Four deals with error processes and digitising, Chapter Five looks at the limitations that the system's data structure places on the data, Chapter Six looks at problems associated with overlay and Chapter Seven looks at some of the solutions for sliver polygon removal.

[12]This will be dealt with more fully in Chapter Seven.

## Chapter Four - Errors associated with digitising.

### 4. 1. Introduction.

The error in GIS begins with user decisions as to what resources are available. These decisions will determine the material to be utilised, how the information will be linked and integrated, and what type of system the project will be carried out upon. This includes the area of GIS benchmarking, selection and performance testing. User-based decisions will then have a series of technological implications with various error components attached to them.

The first technological opportunity for error in a GIS arises with the need to input spatial material to the system. Various techniques are available for map information, the most popular being scanning or table digitising. Each of these techniques has its merits and problems. For users dealing with small projects, the high cost of scanning technology will usually outweigh the high labour input required by table digitising techniques. This chapter assesses the types of error that can be encountered in table digitising, by looking at the two components involved in the digitising process. These can be defined as the 'mechanical' or 'unavoidable' error and the 'selective' or 'generalised' error. The first was analysed using a point coverage, in which operators were asked to identify the points, and their deviations from the true location of the point were then assessed. The second was analysed by asking a set of operators to digitise a line at various scales a number of times. The variations between operators and between scales was then noted and discussed.

For both these experiments, the working coverage was accepted as the absolute product that required entry into the system. That is, no consideration was given to the primary data capture techniques used to originally acquire the data. A discussion of the methodological impact of primary data capture techniques will be made in Chapter Six.

### 4. 2. Data Capture.

Acquiring locational data for use within a GIS entails representing locational data in digital form; that is the translation of points, lines and areas into numbers easily stored within a computer. Primary data capture involves surveying techniques, which may be carried out manually, digitally or remotely, relying on the use of aerial photography or scanners.[1] For most GIS users this primary stage of data acquisition is rarely encountered, and features of interest are often obtained from maps or processed images. This is often referred to as secondary or derived data capture. Methods used to achieve such ends have been primarily driven by the current 'state-of-the-art' in technology, with user needs determining which of the available methods is utilised. Essentially two routes to data capture exist. Firstly, there is the rasterised notation of 'picture cells' that make up the map coverage, involving the process of scanning. Secondly, there is the option of a series (or vector) accumulation of points with additional information as to how the points should be linked (topology) and classified (feature coded). This process is vector digitising.

### 4. 2. 1. Scanners as a means of data entry.

Scanners move across the extent of the map and react to the tone or colour of cartographic material, operating on the principle of measuring the amount of light transmitted through, or reflected from, a small part of the image. The scanning operation may be performed in two ways. Firstly the spot of light may be moved from picture element to picture element and the entire amount reflected or transmitted is measured, such as occurs in a TV camera. Secondly the entire image can be illuminated and any light reflected or transmitted from one picture cell at a time is measured, this is the method commonly used for digital map data acquisition. The main components of the scanning operation are shown in table 4.1.

---

[1]Methods such as EDM, GPS, field surveys, photogrammetry and remotely sensed data are discussed further in Chapter Six along with the problems of utilising multi-source cartographic data.

Table 4.1. Components of the scanning operation.

| Equipment | Function |
|---|---|
| Uniform and accurate controlled field of vision | Illuminates data |
| TV camera with high resolution lens | Emits electrical signal as light acts upon it |
| Analogue to digital converter | Quantises electrical signal to digital form |
| Scan process | A. Flat bed (scanner mounted on rails and is moved backwards and forwards over document surface).<br>B. Drum (scanner moves in one direction only; map is mounted on rotating drum). |

Accuracy of scanned *material* will initially depend on the source material,[2] however the accuracy of the scanning *process* will depend on the techniques used to transfer the data from source product to digital product (table 4.2). Given that few scanners can scan directly from original cartographic products,[3] relevant data generally need to be transcribed from original map sheets. The quality of these drafts will in turn play a part in governing the accuracy of the final output, for example any arbitrary smudges or dirt marks appearing on the scribed map may be represented as undesirable real features, alternatively closely packed individual lines that occur in narrow portions of space may be represented as one thick feature as sometimes occurs in contour maps where many closely spaced lines represent a steep slope.

A second factor that will determine the accuracy of the scanned image is the timing and spacing of the sampling techniques (table 4.2). The smaller the step size (sampling frame) and/or, the higher the sampling rate, then the more accurate the representation of the map will be. However, the price paid for such accuracy will be the large amount of storage space required to store such data. Finally, there are the specifications of the equipment itself (table 4.2). Essentially these are beyond user control, and will probably be determined by the

---

[2]This is discussed in greater length in Chapter Six.

[3]Although Scitex claim their Scitex response scanner is able to achieve this.

amount of available budget the user has to work with. For the most part, the more expensive a product, the higher its quality and resolution.

Table 4.2. Factors determining the accuracy of the scanning process

| | |
|---|---|
| •Source data | •The accuracy of the source data |
| | •Cartographic quality of the source map |
| •Sampling techniques | •Spacing of the sampling frame |
| | •The spatial sampling rate |
| •Equipment resolutions | •Sensitivity of the detector |
| | •Signal to noise ratio of the detector |

As with digitised maps, scanned output cannot guarantee perfection, and mistakes will inevitably arise, usually due to the fact that the scanner does not have the inbuilt intelligence to discriminate between intentional cartographic lines and darker areas which appear on the map, but are not desired on the final image. One alternative to this problem is a semi-automatic method which allows the user an option of 'directing' the scanner over misleading zones. This approach is utilised in the Laserscan Vtrak[4] scanning system. Here lines are 'scanned' individually by a laser beam, which makes the process fast, efficient and undemanding in terms of post processing requirements, whilst allowing operator intervention when the laser reaches an intersection and is uncertain which path to continue tracking. Contour labels for example can therefore be by-passed by the operator guiding the cursor over them, rather than them being scanned, which is often the case with automatic scanners.

**4. 2. 2. The process of operator controlled digitising.**

Perhaps the least technically advanced form of collecting cartographic information is via operator controlled digitising. This involves the use of a digitising table, a cursor with crosshair and an operator. The map is placed on a digitising table (which can vary in size from A4 to A0) and secured firmly with tape. The operator then denotes the geographical features to be captured by positioning the crosshair over the entity in question and pressing a sequence of buttons on the cursor. The digitising table acts as a piece of electronic graph

[4]Formerly Fastrak and then Lasertrak.

paper and notes the X,Y co-ordinate of each point chosen. Points are simply denoted by an X,Y value, (figure 4.1) whilst lines are seen as being defined by points which the operator thinks best describes the line. These points are then connected by straight segments (figure 4.1). Due to the element of generalisation involved, curved lines will naturally lose some of their sinuosity. Areas can be constructed by defining lines as boundaries rather than features in their own right (figure 4.1).

A modification of the selective procedure of manual digitising is a semi-automated alternative[5] which picks up the points for storage at intervals either defined by units of time or units of distance. If the intervals are defined by time then points are collected in units of time, for example every one tenth of a second; if the intervals are defined by distance, then points along the arc of a line are recorded every nth of a millimetre or inch. The amount of decision making and subjective choices demanded of the operator are relatively few as the major task is to ensure that the cross hair is tracing the line correctly with a minimal amount of mismatch. The degree of post processing such mechanisms demand, is however, relatively high as both techniques usually collect too many points for simple features, thus requiring some form of algorithmic line reduction to be performed. Linear stream mode for example collects too many points on long straight stretches of lines, whilst too few are apparent on sinuous portions (figure 4.2). Temporal stream mode on the other hand is more satisfactory on curves as operators usually slow down on such lines because their complexity requires greater concentration. As with the linear case too many points are captured on long straight segments (figure 4.2). Furthermore the accuracy of the line collected is very much dependent upon the speed of the operator, and is therefore unlikely to be uniform for an entire coverage, as digitising speed is a variable commodity.

### 4. 2. 3. Costs and benefits of the various digitising methods.

Assessing the costs and benefits of the various digitising processes involves isolating the components that are of most concern to the user. Of paramount importance to commercial users will be limiting the cost of incorporating material whilst maintaining acceptable levels

[5] Often called stream mode digitising, or line following techniques.

of accuracy. Other factors of importance are shown in figure 4.3 with their corresponding responses from automatic scanners, line followers and manual digitising techniques.

Scanners therefore offer the means for rapid map input to GIS and digital mapping packages with a relatively low level of labour intensity compared with manual digitising (Peuquet and Boyle 1984). However, the time responsiveness of the system is low due to the high degree of source material preparation that is required and the level of post-processing that needs to be performed. Yet the speed of scanning systems, especially the more advanced systems such as Lasertrak still makes them appealing to major users of digital geographic information, as they are the only institutions able to justify the high financial outlay such systems demand. Furthermore they are more likely to be able to cope with the high demand on storage space these systems make.[6]

Line following techniques have a lower cost than scanning apparatus, and are similar in price to selective manual digitising systems. Trade offs are made however in terms of amount of time necessary for material preparation, the speed of the process and subsequent labour demands, and the degree of intelligence that the system embodies. The demands placed upon computer storage by line followers prior to post processing are also relatively high, whilst the time required to eliminate 'excess data' or augment 'sparse' sections can prove costly.

Selective manual digitising is in many ways the antithesis of the more technical, software dependent automated techniques. The cost of purchasing the hardware can still be relatively high,[7] but the software demands, the preparation time and the volume requirements are relatively low. The price paid for this however, is a process that requires high labour

---

[6]Scanned data uses up large amounts of storage space, which can prove cumbersome and expensive. The most common approach to overcoming this problem is to utilise data compaction techniques or to convert the data to vectors, as they take up considerably less storage space than rasters. Even if it is economically feasible to hold topographic map data in raster format, plotting such data using pen plotting routines will still require conversion.

[7]A0 digitising tables can cost as much as £5,000, in addition to the software necessary for data capture.

intensity[8] and which is slow in comparison to the other techniques discussed.[9] Digitising is neither a fast nor a cheap process (Lai, 1988). Yet the major benefits to be gained from the system are the high degree of intelligence it assimilates into the data capture, and the high time responsiveness with which the system operates. However the situation is continually changing, and the old 'high cost of labour' argument against manual digitising is gradually being overcome by using cheap third world labour to perform these tasks.[10]

**4. 2. 4. Meeting user needs.**

User needs and software demands will determine which of the techniques is preferable for particular organisations and applications. Firstly there is the way the data is structured within the GIS, as different data structures place different demands on data models. Secondly there is the question of quantity, large organisations with a high turnover of digital cartographic material will often require automated speedy methods of acquisition. Finally, and most importantly, there is the question of cost as agencies may have their decision influenced by the monetary funds that are available to them.

Therefore, large organisations such as mapping agencies or the military will tend to favour high cost automated methods as they will be able to justify the financial outlay and easily cope with the pitfalls of such systems. For the majority of smaller users however, there is a reliance upon manual digitising as the major method of digital cartographic data entry, as this

[8]Lai (1988) estimates that the capture of one layer of information from a 1:24,000 USGS sheet averages five hours. Breaking coverages down into their feature components revealed that point coverages could be digitised and processed in one third of the time required to process polygon coverages (2hrs vs 6.5hrs per quad sheet). In cash terms, he estimated conversion costing $116, on average, with point data costing $38 per quad sheet and polygon data costing $153 per quad sheet. Thus assuming an eight hour working day, he calculated the daily outlay payable to a digitising operator to be $151-$187.

[9]Cimon, N et al (1990) discuss a regression method for predicting the time required to digitise any given map by using the number of polygons and the total perimeter of polygons (estimated by the area of these polygons) as input parameters.

[10]Indian, Chinese and Malaysian labour have been used by North American companies.

is cheaper and faster, requiring little additional work or cost. Therefore it is this method of digitising that this study focuses attention upon, and the problems which are highlighted are essentially those of the small time users, of which there are now an increasing number. For example, in the UK the majority of local authorities have access to digital data, as do most planning and market research companies.

As figure 4.3 illustrated, the error processes encountered in the differing data capture routines are also variable. The scanners and line following techniques are largely predictable in their generation of errors, however the cost of the added intelligence of a manual data assimilator, who allegedly has some knowledge of the points, lines and areas being dealt with and the relationships that exist between them, is that the errors will vary with the whims, abilities and individual decisions of the operator.

Despite the attempts at GIS 'humanisation' and the general tendency towards user-friendly components within many GIS processes, the task of digitising has remained and essentially technical one, with few attempts being made to make it user driven, although voice recognition in the 1970s was a notable exception. As a consequence, users who are forced to utilise table digitising technology are bound by a series of repetitive technical operations. Although the 'human' involvement in selecting which points to retain is an important intellectual component of the process, it often becomes obscured by the technical feats within which it has to be performed. Furthermore, the demands made by the process on the individual often serves to weaken the human input-users to become physically tired. The repercussion of these combined factors leads to the introduction of errors.

**4. 3. Error and the digitising process.**

The opportunities for error intrusion into the digitising process are severalfold (figure 4.4). In the first instance there is the operator who is always prone to 'human frailties' (Jenks 1981) and these cause him to introduce a variety of error components into the digitising process. These may consist of personal failings, such as poor concentration, poor emotional or physical state, and bad eyesight. Even an operator with good vision cannot see to an accuracy of greater than 0.001 of an inch. Furthermore, unlike drafting or tracing, where the operator can spot the deviation of lines, in digitising the operator does not have the physical means of comparing the two entities of paper map and digital image; therefore assessing areas of mismatch becomes a matter of guesswork.

Secondly, there is the issue of generalisation and subjectivity. In deciding which points best represent a curved line, the operator has to exercise a degree of personal cognition. Representative points will vary from operator to operator depending upon factors as diversified as individual aesthetics, the number of hours spent digitising and the level of cartographic expertise. The number of points chosen to represent the line will depend on the level of generalisation the operator is employing. Naturally the operator could select an infinite number of points, but this would render the whole process ineffective so some degree of generalisation has to be allowed. The fewer the number of points chosen the cruder the resultant features are, but at the same time the faster the digitising task is performed.

Thirdly there are 'mistakes'; genuine errors which result from the monotony of the task or due to tiredness. Unfortunately all such errors or deviations from the truth are not constant throughout the extent of the map, especially since due to the number of possible pitfalls involved, errors can easily be cumulative or can even negate each other.

In addition to the operator there are problems associated with the equipment itself, which may in turn affect the operator's performance. For example, the ergonomics of the cursor may make it easy to press the wrong button or the combination of keys required by the software driving the digitising table may be cumbersome to use and may cause the operator to lose concentration easily or to forget the extent of the line already covered. The digitising table may be placed in such a position that reaching the outermost corners may put the operator in a physical posture where his or her eye is not directly over the crosshair and therefore points digitised will be subject to a slight lateral error. In rare cases there may actually be a fault with the digitising equipment, for example there may be 'dead' patches on the digitising table where no or little data is picked up, or the cursor may be faulty and may only be registering arbitrary points.

Finally there are problems with the data itself. The media from which the digitising is carried out may also exacerbate operator error. If maps are placed on a digitising table in their original colour printed form, then the operator is faced with an unnecessary plethora of detail both in the form of lines and colour. In such a situation, not only will the operator be faced with all the decisions peculiar to the digitising process, but he or she will also have the extra mental task of weeding out the significant pieces of information. Furthermore such information may not appear in the form desired on the digital version. Roads for example are often denoted by relatively thick lines on Ordnance Survey maps, therefore selecting points to represent such lines with a cursor that is accurate to 0.001 of an inch will cause the operator extra labour as he or she will have to decide whether to follow a centreline, which would involve total guesswork, or to follow either the left or right bounding line (figure 4.5). This will often result in intersecting lines falling short, or overshooting the desired meeting point. Original maps will also often have folds in them, these in turn will cause a lateral deflection, since there will be a gap between the apex of the fold and digitising table which will cause that area of the map to move. The matching of map edges 'on the fly' also results in extra uncalled work for the operator and may cause incorrect, uninformed decisions to be made. The problem of paper stretching and warping is also an issue to contend with; paper maps can quite easily crinkle under certain environmental conditions, so a break in the digitising process may in effect result in a slightly different data source being used on each occasion.

Detecting errors that result due to presence or absence of features such as polygon labels, lines, points, even node mismatches, is to a certain degree straightforward as the resultant digital image can be directly compared and verified with the original. A rigorous digitising procedure would ensure that digitised coverages were checked line by line and polygon by polygon, thus eliminating or at least reducing such errors to an acceptable level. Locating errors that rely on the operator's representation of the line is more difficult, as some of these errors are a result of a physical visual mistake, whereas others are associated more with the way an operator perceives and replicates a geographic feature in space.

Perceptual errors can be defined as relating to the generalisation procedures an operator chooses to enforce in translating a linear feature on the mylar tracing into a digital vectorised image. In some cases the generalisation may be intentional and in others, it may be involuntary. Such deviations will be looked at in detail in the second part of this chapter, alternatively some of these deviations may be relatively constant in nature, such as those that might arise due to poor eyesight, and if this were the case, then the results could perhaps be modelled and rectified according to an error 'signature'.

For the most part however, operator inaccuracies arise due to a combination of factors acting in conjunction with each other. The error response from such an idiosyncratic potpourri of factors, makes such errors very difficult to model and predict, due to the randomness of their nature, and their variability from person to person. To assess the extent of such errors an experiment was devised to look at the way a variety of operators performed in digitising the same coverage. The software used for the following experiments was PC ARC/INFO. Further details of this software are given in appendix 4.1.[11]

**4. 4. Assessing the precision of operator digitising.**

**4. 4. 1. Coverage creation and data used.**

To eliminate the possibility of subjective line digitising, the created coverage for this experiment was a 30x30 cm grid from which the operator had to digitise the point intersections. The points were 5mm apart and were denoted by the intersections of the 5mm divisions on standard green graph paper. Under normal digitising conditions, operators would rarely digitise directly from the paper source, as this overwhelms the operator with detail and allows for the instability of paper as a storage medium. In this case however, transferring the grid intersection points onto tracing paper would have incorporated an extra error factor in the experiment as drafting is as prone to displacement as selecting points using a crosshair. Tic registration points were selected prior to digitising and these were entered to an accuracy of three decimal places from the keyboard into a dummy file . The dummy file was then used as the input source for the registration points. Once again this eliminated the possibility of further error introduction from the operator to conform to a predefined accuracy, rather than allowing the operator to define these points from the digitising cursor.

**4. 4. 2. Operator selection.**

[11]The accuracy of the digitising tablet used for this and subsequent experiments was +/- 0.005 inches, and the resolution of the tablet was 0.001 inches.

Six operators were used. The selection of operators was not random but relied on availability and willingness to volunteer for an arduous experiment. The six people selected varied in age, height, eyesight as well as in their degree of computing and digitising expertise. Two of the five were geographers with considerable skill in GIS, digitising and computer mapping, whilst a further two were not geographers, had no skill in GIS or digitising and had only experienced computing through word processing. The final two subjects were not geographers, but had considerable computing experience and had used CAD (Computer Assisted Design) systems extensively. Two of the subjects had eyesight problems that were sufficiently severe to warrant the use of spectacles for detailed work.

### 4. 4. 3. Methodology

The computer coverage and registration points were set up by the author, thus limiting the operator to the digitiser only. The operator was then given the piece of graph paper and with the 30x30 grid clearly marked and was asked to place the graph paper on the most convenient position on the digitising table. It was pointed out to the operator that the table could be moved vertically up and down and also that the angle of the table could be adjusted to suit the operator's individual needs. Should the operator have wished to reposition the table, he or she was quite welcome to do so, either before the experiment started, or at any point throughout the experiment itself. The six tic registration points were marked on the graph paper, and the operator was told to identify each one using the digitising cursor. The residual mean square error was noted down for each case, as this would determine the relative accuracy of the points to come. The RMSE values for all five operators are shown in table 4.3.

Table 4. 3. RMSE values for each of the digitising operators.

| Operator | RMSE value (inches) |
|---|---|
| GLG | 0.004 |
| ALS | 0.002 |
| CEH | 0.002 |
| NJP | 0.003 |
| MJB | 0.004 |
| KYR | 0.004 |

The operator was told to begin digitising at the bottom left corner of the predefined grid square and to proceed along thirty points until the last point on the first row had been entered. The next row could then commence from the far end of the square, or the operator could return to the beginning and continue as before, but one row higher. As the operator proceeded vertically up the coverage, certain points may have been difficult to reach, therefore a high stool was provided, offering the operator the opportunity to maintain his or her arm level at a constant height with a minimal amount of effort. No time limit was imposed on the each operator, but they were told to aim for a digitising time of more or less an hour. The method of holding the cursor and also of positioning it on the point to be selected was left to the individual whim of each operator. Some chose to hold the cursor so that the cross hairs were at 45 degree angles to the vertical and some preferred to hold the cursor so that the cross hairs were at 90 degrees to the vertical. The operators were told to concentrate on the digitising table and not to stray towards the VDU screen displaying the gradual accumulation of points.

The procedure for digitising each point was for the operator to position the cursor over the point in the most accurate fashion he or she could muster, and then to press the key on the digitising cursor that recorded the points (in this case the number 1 key). Next, the operator was to take the cursor away from his or her immediate field of vision and reposition it over the next point, thus hopefully ensuring that each point selected was mutually exclusive from the one that preceded it. If, alternatively, the operator was to follow the horizontal line on which the points lay with the cursor, then each point might to some extent, be dependent on its predecessor. If the first point was accurate then the subsequent points might mimic that accuracy on the x-axis, if however, the starting point was inaccurate this could then influence the rest of the row.

The experiment was always carried out at approximately the same time in the afternoon, and whenever possible on days that did not require an additional light source to natural sunlight. None of the operators were told how to delete the last point digitised, or how to rectify any obvious mistakes. To ensure that every point was digitised the author sat in front of the screen and observed the accumulation of points. If a point had been missed, the operator was told immediately and the point would be redigitised. After twenty one rows (620 points) had been digitised the data set was saved and then the process resumed. This ensured that the work carried out up to that time was safe from the unlikely event of a system crash and it also allowed the operator a short break to stretch and change his or her focal length.

### 4. 4. 4. Results

The aims of this experiment were severalfold. Firstly it was hoped to illustrate the high degree of variation obtained in digital output as a result of operator inaccuracy. This was looked at both in terms of the variability of one operator over space and time, and also with respect to several operators performing the same task under the same conditions. Secondly these variations were assessed to determine whether such differences are random or whether they have a spatial, or temporal pattern that could perhaps be modelled and used as a mechanism for error prevention or reduction.

The variation of each denoted point on the grid from the true point on the grid could have been calculated by comparing the results with a perfect grid created within ARC/INFO. However, this would have entailed the use of the ARC/INFO overlay program which may have influenced the results. Thus the error for each point was compared to its rounded value within a separate Pascal computer program (see appendix 4.2).

The variations in the sum of error encountered by individual operators for the whole experiment is illustrated in figure 4.6. Amounts varied from 10.936 inches (NJP) to 25.385 inches (ALS); a substantial difference. An accumulated difference from the graph paper (in this case 'reality'), of such dimensions is disturbing, as in many ways this situation is an optimistic scenario, in which the operator was only asked to perform the mechanical task of point notation, and no demands were placed on the need to select suitable points. Although it would be impossible to state that the situation would only get worse when point selection was involved, it would seem that there would be no grounds to assume that it would be any better. The composition of the sum of errors deserves further investigation. For example, how extreme are the greatest deviations? These are shown in figure 4.7. A similar pattern of operator fallibility can be observed as in the previous case. Extremes vary from 0.081 inches to 0.038 inches, with all operators achieving a minimal error deviation of zero. This is important in terms of the potential for feature misrepresentation, as the accuracy of a map can be evaluated not only in terms of the total amount of error encountered within a particular map sheet, but in terms of the features that the map represents, grave mismatches in a particular feature cannot be ironed out, or smoothed away by a low error component in another.

An impression of the tendency of the operator to deviate by a particular amount can be obtained by looking at the modal and mean value of error. These are shown in figure 4.8 for all six operators. Here there is a relative degree of uniformity with most operators tending to a mean and modal value of between 0.01 inches and 0.02 inches. The operator with the highest sum of error and the highest deviations exceeds that, but a more interesting case is that of operator MJB, whose modal value indicates a generally very low error deviation, but the higher mean implies odd deviations being responsible for most of the accumulated error.

Operator error can be analysed further by examining the distribution histograms. These are shown for all operators in figures 4.9 to 4.14. Distributions vary, with some operators (KYR, MJB, NJP) tending towards a skewed distribution, with primarily smaller error values being incorporated. Whilst the other operators tend towards a more normal distribution. Yet if the values of the midpoints are examined, the peak of the distribution is usually on a par with the heavily skewed portion of other operators. The case of GLG for example, demonstrates that the operator achieved a few points that were extremely accurate, which operator KYR (with a skewed distribution) was unable to accomplish. Without going into the detailed variations of each individual operator, it is possible to say that distributions of error do vary between operators, and that in a true digitising situation, the consequence would be varying map representations.

Why are some operators more accurate in digitising some points than others? Possible factors contributing to diminishing accuracy could be the increasing amount of time that the operator has spent working at the digitiser, thus implying a positive relationship between an increase in time and inaccuracy. Other factors that could contribute to the variability of the operator include, poor eyesight, poor concentration and also the position of the point on the digitising table itself (some points are more physically accessible than others). It is also feasible that the increasing time spent in front of the digitiser could exacerbate such elements.

To assess any such possible relationship, the correlation coefficient between the order in which points were digitised and the degree of displacement attached to each of the points was carried out,[12] (table 4.4). The highest correlations can be seen in the performance of operator ALS and operator CEH, although these are still relatively low, this might imply that there is a cyclical nature to the accuracy of the points. To investigate in detail the variations with respect to time, the errors of each operator were plotted against the order of digitising (figures 4.15 to 4.20). It would appear from these, that great variations exist from point to point. Whilst one point might approach extreme accuracy, its predecessor and successor may both be highly displaced.

[12]Unfortunately, it was not possible to do a real time series, as the digitising package could not easily be configured to record the time at which each of the points was entered.

Table 4. 4. Correlations of error and increasing time spent digitisiing.

| Operator | Correlation coefficient |
|---|---|
| MJB | -0.100 |
| NJP | 0.065 |
| GLG | 0.127 |
| KYR | 0.136 |
| CEH | 0.299 |
| ALS | 0.504 |

On closer inspection however, it would appear the operators seem to work in bursts of accuracy. Most of the error graphs (4.15 to 4.20) show a cyclical pattern, with operators tending to have periods of precision, followed by periods of inconsistency. This is perhaps most dramatically evident in the case of ALS, where error peaks occur after periods of relative accuracy at approximately 100 point intervals. It would almost seem that operators become lulled by the monotony of the task, and only return to giving it 100% concentration after they make an error that they can visibly recognise. Another feature that seems to come to the forefront, is that although there is only a low positive correlation between the increase in time spent at a digitiser and the increase in accuracy, in some cases (ALS, GLG and CEH) there does seem to be an initially healthy performance which diminishes after approximately 150-200 points. This could perhaps be due to the gradual deterioration of the initial high standards one sets oneself when approaching a new type of task as the tediousness of the task becomes apparent.

A further break-down of the degree of error incurred as the number of points imputed increased can be obtained by looking at the error distribution on the x and y coordinate individually (figure 4.21 to figure 4.32). It is interesting to note that many operators have difficulty with one particular axis, whilst being relatively accurate with the other. Examining the errors of each operator in this fashion, reveals whether failure on one particular axis is a contributory factor to poor positioning. Regarding total error alone often blurs such issues, as random variation in one axis, may obscure a trend in another axis, or alternatively high errors in one axis may be cushioned by low error in the other and vice versa. In addition, instances where the operator always errs by a certain factor may be revealed. For example if changes in the accuracy of the position of the point are mirrored in both the x and y coordinates, but with one coordinate having a small deviation from the true value and the second a large deviation from the true value, there may be a case for assuming that the operator's error is relatively constant. This is most likely to be true of those people who suffer from eyesight deficiencies, especially astigmatism, and could possibly be alleviated by transforming the inaccurate coordinates by a suitable parameter.

GLG, KYR and most dramatically ALS all showed tendencies to be more inaccurate on one axis than on another. GLG was more consistently inaccurate on the Y coordinate than on the X coordinate (figure 4.26 and figure 4.32). The distribution of his errors on the X coordinate fluctuated quite substantially, yet on the Y coordinate they were steadily inaccurate by 0.01", except for a period at the start of the experiment, the middle of the experiment and the end of the experiment. The psychology of this phenomenon could be that the operator started off with intentions of accuracy and concentrated hard when he was at the beginning of his task, and in the same way as he approached the end of the task he might also have put in a little more effort.[13]

The distribution of the errors belonging to KYR showed a similar pattern, but on a more grandiose scale (figure 4.25 and figure 4.31). The errors in the X coordinate (that is the vertical) were far greater and more randomly distributed that those in the Y coordinate (the horizontal line). Even more spectacular was the difference between the axes displayed by ALS (figure 4.22 and figure 4.28). Although the variations in the X coordinate were relatively small, only twice exceeding 0.03", the deviations on the Y coordinate were consistently high, following an initial period of concentration. All the values after this period (barring one) were underestimations of the position of the coordinate, and although these were high they rarely varied much from one to the next. Their values seemed to increase and then

[13]Making such psychological assumptions of the operator is however very dubious, as they are varied and random, and are not predictable without further analysis of the individual. For example, in the case of GLG, an improvement in the end of the experiment may be synonymous with renewed vigour after a small break the operator had chosen to take once he realised that he was half way through his task.

decrease in well defined cycles. These cycles are evident in the positions of the X coordinate of the point, but on a much subtler level. Such deviations can perhaps be put down to the poor eyesight the operator claimed to have in his left eye; a condition that might perhaps make him undesirable as a digitising operator.

The plots of the deviations on the X and Y coordinates for the operators CEH and NJP both showed little deviation between the two coordinates. In the case of CEH (figure 4.23 and figure 4.29), both the X and the Y coordinates show a very small deviation from the absolute position of the point; only rarely does the error exceed 0.02". There is also a tendency in both coordinates, to move away from the true value as the number of points increases, although this is more prevalent in the X rather than the Y coordinate. There is also a very small variation between the errors themselves, as indicated by the low value for the range of the data.

NJP showed a similar tendency towards a small range in error values. He was also fairly consistent in his digitising as his low standard deviation indicates. Unlike the other cases however, the position of his points actually tended towards the true value as time increased (figure 4.24 and figure 4.30), rather than moving away as one would assume given the high degree of concentration required for this task. Finally, these cases can be compared with that of MJB, (figure 4.21 and figure 4.27) and in which the errors seem to have no pattern at all; displaying what would appear to be a completely random pattern.

The majority of the operators had error levels that varied randomly and did not err in a predictable fashion. That is to say there was no way of predicting the accuracy of the next value. Errors such as these would be impossible to model or ameliorate. In some cases however, there may have been a case for arguing that due to a particular inability to accurately pinpoint a particular axis, all the values of the operator could be transformed by a particular function, which would leave the resultant points more accurately positioned than if they were left in their original raw state. There might also be a case for using such a method to determine the accuracy of the physical ability of the operator. Conclusions could then be drawn as to the probable error margins likely to be incurred by each individual. Such indices could however only be used with a degree of caution as the whim of the operator may change from second to second.

In addition to the 30x30 grid square one operator (GLG) was asked to digitise a 10x10cm grid twice, once with as much time as the task required and once under timed conditions. The results of the 10x10 grid square are interesting, in that they would tend to indicate that freedom with respect to the amount of time an operator can spend digitising is not necessarily accompanied by a significant increase in the operator's accuracy (figure 4.33 to figure 4.36). This is perhaps due to the fact that the operator has an accuracy threshold that cannot be exceeded irrespective of the amount of time he or she spends attempting perfection.

Another interesting fact is that the differences in the error levels between the two coverages were quite significant. Whereas in the 10x10 grid GLG did not have any points that were greater than 0.006", in his 30x30 coverage this rose to 0.042. This would imply that the larger coverage was perhaps more daunting than the smaller one. Also the points were twice the distance apart on the smaller coverage, which may have aided their identification.

The spatial distribution of the errors were also investigated to see whether there was any possible correlation between the position of the points and their degree of error. Of special interest was whether some areas of the digitising table were especially difficult to reach, thereby influencing the accuracy of points that occur in that region. To show the extent of the errors as they appeared on the digitising table each of the operator's error values were plotted on a three dimensional graph, with the horizontal and vertical axes of the digitising table depicted as the x,y dimensions of the graph and the errors elevated onto the z plane (figure 4.37 to figure 4.42).

There was a general tendency for the areas at the top of the digitising table to be marginally more inaccurate than those at the lower portion. NJP and CEH (figure 4.37 and figure 4.38), both showed slightly higher error values in the top right hand corner of the table, although these points were also the last to be digitised and so may also be suffering from the operator's depreciated concentration.

KYR and MJB (figure 4.39 and figure 4.40) on the other hand, had peaks in the top left hand corner, although in both these cases there were so many random peaks, that these cannot be regarded as significant. GLG (figure 4.41) seemed to show no relationship between the error values and their position in space. ALS (figure 4.42) had markedly higher error values in the top right hand quadrant of the digitising table, although he tended to perform well in cycles. Certainly his lowest 80 values were the most accurate, but this could be a result of the threshold of his visual concentration, rather than the ergonomics of the digitising table itself.

**4. 4. 5. Conclusions.**

The aim of this experiment was to assess whether operator accuracy is influenced by any specific parameters. Parameters investigated were those of time, and spatial position. It would appear, that despite a superficial trend for accuracy to diminish with the number of points digitised, this is not a significant pattern, and is obviously delimited by thresholds for each operator. Spatially, there seemed to be little evidence to link in a consistent error with the position of the point on the digitising table, although some areas may have been difficult for some individuals. In some cases there was a degree of periodicity, with operators having several bad rows followed by an accurate one. It would seem therefore, that in general, operator errors vary randomly with respect to time and space.

If the performance of the individual operators is compared, there is also very little evidence to suggest any correlation between their achievements. Once the experiment was completed, some operators had similar results, for example CEH and NJP as well as MJB and KYR, but there is very little indication to suggest that these are linked in some way. A greater number of operators with defined characteristics, might enlighten the situation, but unfortunately such a luxury is beyond the realms of this study. One can only assume therefore, that in the mainstay, random factors proliferate.

Errors incurred through the physical difficulty of locating points on a two dimensional plane, as occurs in digitising, are therefore difficult to account for, model and predict. Accepting them must be regarded as a hazard of cartographic data collection, although care and attention to detail will naturally help minimise such errors. This is especially vital when maps of small scale are being used as cartographic input, as discrepancies of 0.1" can be grossly magnified when translated into units on the ground.

Admittedly very few people partook in the experiment, and to obtain enough data to make assertive assumptions about the 'average' person's digitising capabilities would require a very exhaustive study of operators[14]. Furthermore, all the results obtained within the experiment have to be viewed in the light of the constraints placed on the individuals by the software and hardware used. For example, the accuracy of the digitising table was 0.005 inches, which implies that some of the error encountered is negligible. In addition, although the coverage was not transformed for comparative means, initial tic points were registered and this may have affected the final results. Yet this study has attempted to show that variations between operators do exist, and such variations when extended into GISs can be responsible for further problems of error eradication.

Unfortunately, the majority of agencies that distribute digital data often contract the 'dirty' job of digitising to outside contractors, who in turn employ the only people willing to indulge in such a task on a long term basis; unskilled, young and often female labour. As these operators are unlikely to ever see the finished product of their efforts, there is usually very little incentive to maintain the highest standard possible. Usually the merely 'acceptable' will suffice.

Most agencies in which a large amount of in-house digitising takes place, have tests for their digitising operators. However smaller and medium sized GIS users will often simply assign operators on the basis of availability, and in such situations accuracy levels might benefit from a study of this type, which is relatively cheap and easy to perform.

**4. 5. Assessing line definition in operator digitising.**

A further error component that affects the accuracy of the digitising process is one of line definition. This deals more with the perception and cognition of the line than problems in physically locating a line. As the operator decides which points to select using the digitising

---

[14]See Traylor 1979.

cursor and table, he or she is in effect generalising the infinite points on a curve into a few representative ones. This process is further diversified by problems of scale. As all maps are a scaled down version of reality, the amount of available 'space' for the cartographer will determine the level of detail that a map can depict; if the ratio between the map and the real world is relatively small (ie a large scale map is being used), then more detail can be incorporated, however if the same features have to be shown on half the space, then the amount of detail depicted will depreciate. Therefore maps of the same area drawn at small scales will have undergone a greater degree of generalisation at the cartographer's hands than their large scale relatives.

The methods of detail omission are varied and occur in both positional and classificatory schema. (Shea and McMaster, 1989)[15]. In terms of absolute position, the depiction of linear features such as roads or rivers as cartographic lines may have to be smoothed or simplified to allow them to present the desired impression on the small scale map. Detail of point features may be aggregated, whilst groups of similar features may be amalgamated or merged. In terms of classified features, what on the larger scale map might have represented a positionally accurate depiction of a feature, on a smaller scale map, this may simply be replaced by a symbol. Also areal features may often be reduced or 'collapsed' into linear ones. Such disparities in the data represented in different scale maps may be more apparent in some cases than in others. The implications of gathering data for a GIS from a variety of scales into what is effectively a scale free data base will have some severe implications if the data is to be accurately integrated or overlaid. Extra 'wiggles' or curves in the data sets originating from the large scale maps could be viewed as areas of interest, instead of being rightly dismissed as a scale based error component.

Another likely consequence of digitising from different scale maps is that the lengths of lines will actually vary. On small scale maps there is far less space to include all the inflections of a linear feature on the ground. Although a large scale map will still never embody the line in its totality, more of the detailed curves of the line will be accounted for and therefore the feature will be longer than representations on a small scale map. In applications where the

[15]See Chapter Three for a more detailed discussion.

length of features is important for analysis, such disparities could cause a certain degree of anxiety.

In the previous digitising experiment, the errors examined were only those relating to the physical positioning of the cursor itself. All the points requiring digitisation were already selected and all the operator had to do was to identify them. When digitising a line however, the circumstances for the operator are slightly different, not only does he or she have to position the cursor over the exact tract of tracing paper covered by the traced line, but there is also the subjective process of selecting the lines themselves and then picking out the most representative points on that line. If the coverage has been diligently traced, then the first task is accomplished away from the digitising table and the operator's only concern is with selecting the points themselves.

In digitising a line therefore the operator juxtaposes the pure problems of placement he or she may have (as highlighted in the previous experiment), with his or her own personal interpretation of the line and what it is trying to portray. Assuming each individual has a different perception of the line itself and a randomised pattern of error generation, then each operator's rendition of lines of differing scales will vary from those created by his or her counterparts. If such products are overlaid at a later stage of GIS analysis, lines representing the same features will not coincide. As the majority of integrated material used in GISs are obtained from varying scales using different digitising operators, investigating the disparities in lines digitised by a variety of individuals may provide some insight as to the degree of mismatch that can be expected.

Digitising experiments have been performed before, notably by Traylor (1979). However, whereas he evaluated a methodology for assessing digitising error, this study seeks to point out the relative importance of the digitising procedure in error accumulation within spatial data sets, as errors created at this stage may require the use of error removal techniques later on.

The following experiment focuses on the variation in the lengths of features with respect to changes in scale, operator and each individual operator's subjective interpretation of the line.

Understandably, length is only one of a possible number of surrogates for error, given the way it is affected by the changing shape of the line itself. A more comprehensive measure might be achieved by an analysis of the way the character of the line has actually changed, however, this would prove problematic to measure empirically.

It is a well known fact that scale and line length have a high correlation. In traditional cartography, scale affects the length of any line represented. In a digital environment, both digitising operator and scale will interact to affect the resultant length of a digitised line. This experiment attempts to analyse the extent of these interactions.

**4. 5. 1. Data source and coverage creation.**

The line chosen for this experiment was the River Dove in Derbyshire (UK), stretching from north of Mayfield to Newton Sloney and extending in length from OS grid coordinate 164480 to OS grid coordinate 260280 (Landranger sheet 128). The river was traced from three different scales onto plastic film using a 0.35mm pen. The scales chosen were 1:25,000 - Ordnance Survey (appendix 4.3), 1:50,000 - Ordnance Survey (appendix 4.4) and 1:100,000 - Bartholomew (appendix 4.5). As the Bartholomew map did not employ the Ordnance Survey Grid Coordinate System, the length of the river had to be interpolated into Grid Coordinates from the Latitude and Longitude values provided on both maps.

**4. 5. 2. Operator selection.**

Ten operators (final year geography undergraduates) were selected to take part in the experiment all with the same expertise in digitising and GIS techniques. All the students were familiar with the theory of generalisation and realised that in digitising they were not aiming for an infinite number of points, but rather they were trying to achieve an optimum representation. The experiment was carried out in groups of two or sometimes three operators (due to equipment limitations), with one person digitising and another operating the keyboard. The digitising process would then be rotated within the group. Each operator was told to adhere to the same procedure and they were instructed to aim for accuracy, but not to be laborious.

**4. 5. 3. Methodology.**

The operator would first securely fasten the 1:100,000 scale traced map onto the digitising table and then adjust the height and angle of the table to suit his or her preference. A template coverage consisting of tic registration points and boundary points in OS grid coordinates was then prepared within the software system (ARC/INFO), by keyboard entry. This ensured that there was an accurate control for the tic points against which the operator error could be measured. Six tic registration points were used for the 1:25,000 and the 1:50,000 coverages and only four for the 1:100,00 coverage, as the line was relatively small. The same template was used for all the operators and for all the scales, thus providing a common base line from which to draw comparisons between individuals and scales.

For each coverage digitised the operator was told to identify and digitise the tic registration on the coverage and to note the RMSE. Even if the error was relatively high, the operator was told to continue. Once into the 'addition of arcs' module, the operator was instructed to avert his or her eyes from the VDU and to concentrate wholly on the line and its salient points, working from top to bottom. After the addition of the last vertex in the line the operator would inform the keyboard operator who would note down the number of points constituting a line. The line would then be saved and the coverage exited. Once each member of the group had digitised the 1:100,000 map three times, the process was repeated with the 1:50,000 map and the 1:25,000 map.

This produced nine coverages per operator and a total of ninety versions of the same line. Each of the lines then had to have some topology built into them, so that their length could be assessed and their AAT files queried. The results of the experiment are displayed in table 4.5.

**4. 5. 4. Results.**

The aims of this experiment were to show that the lengths of the digitised line and their pattern in space, is different with respect to operator and scale, and that in some cases this can be quite severe. The implications for such errors is that in data overlay within a GIS, the lines themselves are not merely lines in vector storage, but rather, they represent features which may need to be pinpointed accurately if they are to be used for planning and developmental analysis.

**Table 4.5 Results of repeated representations of linear features for ten operators**

| OPERATOR | RMSE | POINTS | LENGTH | SCALE | ATTEMPT |
|---|---|---|---|---|---|
| HAV | 0.009 | 308 | 33.9764 | 100000 | 1 |
| HAV | 0.012 | 303 | 33.4688 | 100000 | 2 |
| HAV | 0.013 | 283 | 33.5138 | 100000 | 3 |
| HAV | 0.020 | 381 | 37.5230 | 50000 | 1 |
| HAV | 0.021 | 381 | 37.3545 | 50000 | 2 |
| HAV | 0.019 | 383 | 37.2438 | 50000 | 3 |
| HAV | 0.008 | 454 | 37.9666 | 25000 | 1 |
| HAV | 0.009 | 459 | 38.0516 | 25000 | 2 |
| HAV | 0.009 | 519 | 38.0295 | 25000 | 3 |
| EG | 0.013 | 360 | 34.0569 | 100000 | 1 |
| EG | 0.011 | 351 | 31.5519 | 100000 | 2 |
| EG | 0.011 | 393 | 32.4074 | 100000 | 3 |
| EG | 0.018 | 410 | 35.5493 | 50000 | 1 |
| EG | 0.019 | 387 | 35.6496 | 50000 | 2 |
| EG | 0.018 | 415 | 35.8725 | 50000 | 3 |
| EG | 0.007 | 494 | 37.8445 | 25000 | 1 |
| EG | 0.007 | 544 | 37.7867 | 25000 | 2 |
| EG | 0.007 | 396 | 37.5230 | 25000 | 3 |
| GBF | 0.013 | 374 | 32.8347 | 100000 | 1 |
| GBF | 0.012 | 306 | 32.8927 | 100000 | 2 |
| GBF | 0.012 | 318 | 32.6570 | 100000 | 3 |
| GBF | 0.019 | 444 | 36.8265 | 50000 | 1 |
| GBF | 0.018 | 504 | 36.7137 | 50000 | 2 |
| GBF | 0.020 | 371 | 36.4839 | 50000 | 3 |
| GBF | 0.008 | 562 | 37.4335 | 25000 | 1 |
| GBF | 0.011 | 604 | 37.2631 | 25000 | 2 |
| GBF | 0.009 | 582 | 37.4112 | 25000 | 3 |
| GMP | 0.014 | 494 | 33.7051 | 100000 | 1 |
| GMP | 0.014 | 402 | 33.6805 | 100000 | 2 |
| GMP | 0.012 | 293 | 33.5057 | 100000 | 3 |
| GMP | 0.020 | 411 | 37.4315 | 50000 | 1 |
| GMP | 0.019 | 382 | 37.2764 | 50000 | 2 |
| GMP | 0.020 | 387 | 37.3531 | 50000 | 3 |
| GMP | 0.011 | 528 | 37.9200 | 25000 | 1 |
| GMP | 0.011 | 570 | 37.8427 | 25000 | 2 |
| GMP | 0.011 | 520 | 37.7154 | 25000 | 3 |
| IT | 0.011 | 358 | 34.3656 | 100000 | 1 |
| IT | 0.012 | 284 | 33.8397 | 100000 | 2 |
| IT | 0.013 | 280 | 34.1817 | 100000 | 3 |
| IT | 0.020 | 367 | 37.6051 | 50000 | 1 |
| IT | 0.020 | 343 | 37.5971 | 50000 | 2 |
| IT | 0.019 | 293 | 37.4430 | 50000 | 3 |
| IT | 0.009 | 384 | 37.8129 | 25000 | 1 |
| IT | 0.009 | 443 | 37.7889 | 25000 | 2 |
| IT | 0.008 | 479 | 37.8809 | 25000 | 3 |
| KE | 0.014 | 241 | 32.8708 | 100000 | 2 |
| KE | 0.011 | 233 | 32.9926 | 100000 | 3 |
| KE | 0.021 | 294 | 36.7619 | 50000 | 1 |
| KE | 0.019 | 297 | 36.6239 | 50000 | 2 |
| KE | 0.021 | 315 | 36.8278 | 50000 | 3 |
| KE | 0.008 | 591 | 37.3478 | 25000 | 1 |
| KE | 0.010 | 504 | 37.2551 | 25000 | 2 |
| KE | 0.011 | 399 | 37.3557 | 25000 | 3 |
| SJE | 0.011 | 245 | 33.6207 | 100000 | 1 |
| SJE | 0.013 | 530 | 34.1035 | 100000 | 2 |
| SJE | 0.013 | 353 | 33.7527 | 100000 | 3 |
| SJE | 0.019 | 603 | 37.8711 | 50000 | 1 |
| SJE | 0.019 | 549 | 37.8029 | 50000 | 2 |
| SJE | 0.020 | 530 | 37.6331 | 50000 | 3 |
| SJE | 0.008 | 652 | 38.0001 | 25000 | 1 |
| SJE | 0.012 | 879 | 38.1573 | 25000 | 2 |

table 4.5 continued

| Operator | RMSE | Points | Length | Scale | Attempt |
|---|---|---|---|---|---|
| SJE | 0.009 | 925 | 38.2011 | 25000 | 3 |
| PNA | 0.012 | 519 | 33.7525 | 100000 | 1 |
| PNA | 0.012 | 348 | 33.9001 | 100000 | 2 |
| PNA | 0.010 | 334 | 34.0989 | 100000 | 3 |
| PNA | 0.020 | 415 | 37.6802 | 50000 | 1 |
| PNA | 0.019 | 513 | 37.8430 | 50000 | 2 |
| PNA | 0.020 | 530 | 37.6058 | 50000 | 3 |
| PNA | 0.009 | 804 | 35.8605 | 25000 | 1 |
| PNA | 0.012 | 717 | 37.9808 | 25000 | 2 |
| PNA | 0.005 | 762 | 40.0032 | 25000 | 3 |
| KOB | 0.014 | 232 | 33.3167 | 100000 | 1 |
| KOB | 0.011 | 238 | 33.5010 | 100000 | 2 |
| KOB | 0.011 | 252 | 34.1068 | 100000 | 3 |
| KOB | 0.018 | 403 | 37.4529 | 50000 | 1 |
| KOB | 0.020 | 364 | 37.3961 | 50000 | 2 |
| KOB | 0.019 | 458 | 37.5728 | 50000 | 3 |
| KOB | 0.008 | 825 | 37.9531 | 25000 | 1 |
| KOB | 0.009 | 743 | 38.0803 | 25000 | 2 |
| KOB | 0.008 | 626 | 38.0969 | 25000 | 3 |
| NNM | 0.011 | 340 | 33.4396 | 100000 | 1 |
| NNM | 0.011 | 333 | 32.7846 | 100000 | 2 |
| NNM | 0.012 | 304 | 34.1074 | 100000 | 3 |
| NNM | 0.018 | 510 | 37.5243 | 50000 | 1 |
| NNM | 0.016 | 538 | 37.1103 | 50000 | 2 |
| NNM | 0.018 | 564 | 37.3138 | 50000 | 3 |
| NNM | 0.008 | 790 | 37.7614 | 25000 | 1 |
| NNM | 0.009 | 617 | 37.5041 | 25000 | 2 |
| NNM | 0.009 | 746 | 37.6434 | 25000 | 3 |
| KE | 0.012 | 319 | 32.9101 | 100000 | 1 |

As expected there was a great deal of variation in the line length and the number of points used to depict lines. The range of lengths recorded varied from 31.552km to 40.003km, thus embodying a disparity of 8.451km. The mean length was 36.137km with a relatively low standard deviation of 2.014km. If the histogram of the distribution of the data values (figure 4.43) is looked at in depth however, it becomes apparent that there is a split distribution. There is a distinct break between the values of the line that are less than 35km long and those that are above that length. If figure 4.45 is observed, it can be seen that all of the values below this value are from the 1:100,000 scale map, and in fact that there is no mixing or fuzziness at the defining the boundaries between the 1:100,000 scale lines and the 1:50,000 scale lines.

This phenomenon may be due to two factors. Firstly that both the 1:25,000 and 1:50,000 scale maps are from the Ordnance Survey, whilst the 1:100,000 scale map is from the Bartholomew map publishing company. The two cartographic agencies may have different methods of data collection and presentation, which in turn may affect the compatibility of maps of the same area. Secondly as small scale maps are known to be generalised to a greater degree at drafting, it is no surprise to encounter a strong relationship between small scale maps and shorter line lengths.

The general trend was for the line lengths to increase as the scale of the map increased. This is as was expected, given that the greater detail in which lines are depicted inclines the operator to select more characteristic points than he or she would on a smaller scale map. The average length of the line digitised from the 1:100,000 scale map was 33.4632km, whilst from the 1:50,000 and the 1:25,000 scale maps they were 37.1648km and 37.7842km respectively. Although the trend was for the lines digitised from the 1:25,000 scale map to be the largest, there were still some lines from the 1:50,000 scale map that superseded some of those from the 1:25,000 (table 4.5). The number of points used to define these lines also increased with increasing scale and the average number used for each scale are shown in table 4.6.

Table 4. 6 Average number of points used to represent lines from the different scales.

| Overall | Scale of line | Average Number of points |
|---|---|---|
| | 1:25,000 | 604 |
| | 1:50,000 | 425 |
| | 1:100,000 | 331 |
| Individual operators | Name of operator | Average number of points |
| | HAV | 386 |
| | EG | 417 |
| | GBF | 452 |
| | GMP | 443 |
| | IT | 359 |
| | KE | 355 |
| | KOB | 460 |
| | NNM | 527 |
| | PNA | 549 |
| | SJE | 585 |

Overall the smaller number of points per line occurred on the smaller scales; although there is a degree of mixing. More apparent however, is that nearly all the lines falling in the last thirty (in terms of the number of points that they contain), are taken from the 1:25,000 scales map. It would seem therefore, that the scale of maps does affect the number of points that can be extracted; generally the more points that are used to describe a particular line then the longer the line becomes.[16]

This 'difference' between the scales can further be highlighted by looking at the distributions of the length of the line with respect to each of the scales that were used (figure 4.45). As can be seen, although there is some intermixing between the 1:25,000 and the 1:50,000

[16]If the number of points becomes extremely large, then it could be argued that the length of the line is simply increasing by the zig-zagging of points along a line's extent. However, this is unlikely for two reasons. Firstly, the operators used in this experiment were experienced enough not to digitise in that manner, and secondly, the pen-nibs used for the tracing of the lines were sufficiently fine enough to make such a scenario difficult to achieve.

lengths, implying the values come from the same population, the lengths for the 1:100,000 scale representation are isolated and carry the implication that they do not adhere to the same parent population.

There is therefore a relationship between the number of points used to define a line, its length and the scale at which it was drawn, but this is by no means perfect. It would seem that an increase in the number of points does not necessarily denote an improved accuracy, as the increased number of points could also apply to portions of the line that were relatively straight, but within which some operators were still recording many points (figure 4.46). For example operator SJE recorded a 1:100,000 line with 528 points (the 66th largest collection of points out of 90 samples), yet the line itself was only the 27th largest line out of the ninety samples. This may partially be accounted for by collecting too many points in straight line sections of the line, but it would also appear to indicate that certain scales have a ceiling on them, and that no matter how many points are actually collected, the line cannot increase in length. This can be highlighted by SJE collecting (purely by chance) the same number of points for the line on the 1:50,000 scale map. This line shows a length of 37.8029km as opposed to the 34.1035km recorded on the line digitised from the 1:100,000 scale map. Scale certainly seems to be a major determining factor in the length and subsequent shapes of lines, but the operator himself is also responsible. Great variations exist between individuals on the same scale and also between the same individual on different scales.

In the same way, the relationship between the scale at which a particular line is digitised and the length and detail of that line is by no means perfect. Variations contrary to the rule exist, and they exist sporadically as a result of operator variability and fallibility. This can be illustrated by firstly looking at the variations that exist within each operator's spectrum of lines on one scale. Secondly, the results obtained by several operators digitising the same line at the same scale can be contrasted. Thirdly, the performance of one operator through a variety of scales can be illustrated. Finally, the most realistic case of information acquisition in a GIS can be analyzed; that of many operators digitising information from a variety of different scales that will be integrated into a uniform data base within a GIS.

Table 4. 7. Average length of line digitised by individual operators

| Overall | Scale of line | Average length of line | Line length range |
|---|---|---|---|
| | 1:25,000 | 37.7824km | 40.0032km to 35.8605km |
| | 1:50,000 | 37.1648km | 37.8711km to 35.5493km |
| | 1:100,000 | 33.4632km | 34.3656km to 31.5519km |
| Individual operators | Name of operator | | Average line length |
| | HAV | | 36.3476km |
| | EG | | 35.3602km |
| | GBF | | 35.6129km |
| | GMP | | 36.2700km |
| | IT | | 36.5017km |
| | KE | | 35.6606km |
| | KOB | | 36.3863km |
| | NNM | | 36.1321km |
| | PNA | | 36.5250km |
| | SJE | | 36.5714km |

The average line length of the lines digitised by the individual operators spanned a very small range of between 35.6129km (GFB) and 36.5714km (SJE) (table 4.7). It is however not the average line that is used in geographical data integration in GIS. Most users do not have methods for extracting the 'average' line from a series of attempts at a digitising task. In addition, given that digitising is such an laborious task, lines making up an individual cover are only digitised once, and if the operator happens to have been somewhat more inaccurate than usual, then the situation is very much a 'fait accompli' and errors simply become accumulated in the GIS data base.

The highest range within any one scale for a single operator occurred, surprisingly enough with the line digitised at the 1:25,000 scale, by the operator PNA. He ranged from a low of 35.8605km to a high of 40.0032km, a disparity of 4.1427km. However, this disparity is not visually apparent in a non-scaled representation of the line (see figure 4.47). Presumably, if the line were drawn to scale, or enlarged this difference would become more discernible. Alternatively, operator HAV digitised the same line with resulting lengths of 38.0516km and 37.9666km, a discrepancy of only 0.085km. Similar situations exist in the other scales; for example at the 1:100,000 scale operator EG achieved a range of 31.5519km-34.0569km with a resulting variance of 2.505km, whilst GBF managed a range of 32.8927km-32.6570km; a divergence of only 0.2357km. Similarly, at 1:50,000 NNM had a deviation of 0.414km with line lengths ranging from 37.5243km to 37.1103km, whilst at the same scale KE achieved values ranging from only 36.8278km to 36.6239km. It was however, at this scale that the smallest variation in the entire group sample occurred.[17]

Such measures give an indication of the variability of the operator within a specific scale. The degree of variability the operator exercises is however also subject to change. If, for example, the variance of operator PNA within the 1:25,000 scale is compared with his performance in the 1:100,000 scale, (table 4.5) it can be seen that his variation in digitising the 1:100,000 scale is far less than his efforts on the 1:25,000 line, a situation that one would not normally expect to arise, but which inevitably occurs due to operator subjectivity.

The variations between several operators dealing with the same scale were relatively similar irrespective of the scale concerned. The maximum discrepancy at the 1:100,000 scale was between operator EG (31.5519km) and IT (34.3656); a difference of 2.8137km. At 1:50,000 the greatest variation existed between EG and SJE with lengths of 35.5493km and 37.6331km respectively, giving a variation of 2.0838km. The greatest margin of error at the 1:25,000 scale occurred within the realms of the same operator; PNA as discussed in the last section. If however his smallest length is seen as an outlier and the next smallest length is taken into account (that produced by KE) the lengths have a disparity of 2.7481km.

The digitised lines of most operators (eight) had ranges of between 4km and 5km (table 4.5), whilst two had ranges of just over 6.2km. It would appear therefore that some operators have a much wider variance in their approach to digitising the same line many times at different

[17]Perhaps this implies that the length of the 1:50,000 scale line was more manageable, in that operators felt more daunted by a longer and potentially more difficult task.

scales. In practice, this implies that even the same operator cannot be relied upon to reproduce the same level of accuracy every time he or she digitises. There may however be some gain in determining the degrees of variation people are likely to produce. Armed with such information, users would have some indication of the reliability of certain operators.

The case most prone to integration difficulties in a GIS, is when information arrives digitised by a series of different operators or agencies, having already been extracted from several different source scales. In this experiment, this can be visualised by looking at the maximum difference in lengths and character of line between two operators using different scales. The disparity in terms of line length is 8.4513km, which if compared to the mode value of the total line lengths (37.523km) is quite a severe margin of ambiguity and comprises 23% of the mode value.

The scales of the different maps will allow the data to be grouped into ascending lengths; the smaller the scale then the smaller the line; the larger the scales then the longer the line. There is however, a degree of overlap depending on the operator, and also on the attempt; generally people are more careful on the first attempt although longer lines at the greater scales tend to appear at the second or third attempt. This is probably because each time the line is looked at anew more detail can be discerned.

**4. 6. Points for discussion.**

The results of both the experiments suggest that errors in digitising arise, but that their occurrence is random and unpredictable. The errors that arise as a consequence of digitising can be subdivided into those that materialise as a result of 'mechanical' failure, and those which arise due to poor 'intelligent' decision making (which points best represent a line).

Errors in mechanical point assimilation have a variety of sources, but tend to rely on mechanical human failings, such as poor eyesight, tiredness, and an inability to concentrate, in combination with equipment faults and poor ergonomic design. In addition to these, errors in 'intelligent' feature assimilation are due to generalisation factors, which in turn depend upon scale, the number of points used to signify a line and the subjective decision as to where those points should be placed. The greater the number of points used to represent a line, then generally the longer the line will become, and this may be viewed as an increase in accuracy. However the number of points an operator will use is largely unpredictable. For both mechanical and intelligent errors, 'error signatures' could be developed for operators with consistent problems, such as poor eyesight. However the operational benefit to be gained for such an exercise is questionable.

The consequences of the introduction of errors at this stage is severalfold. Firstly areas are represented in a geometrically incorrect fashion and therefore a degree of positional inaccuracy ensues. In short the features on the map are not as they appear on the ground. Secondly, in producing inaccuracies, problems are created for the user when an attempt is made to build the data into a desirable data model. Inaccuracies in terms of the 'intelligence' or the subjective path of the line are of no consequence to the data model which the software will require for the creation of a working dataset. What is of importance, is that nodes at which lines meet should match perfectly. It is therefore at the end points of lines that the software is forced to focus the attention of its error elimination routines. Thus when all the endpoints of lines are satisfactorily joined, the data are pronounced 'clean', without any regard for the shape of the features themselves. Feature appraisal is therefore left very much up to the individual user.

It is only when data are used for the overlay of time series data that the error components of digitising come to the forefront, as in this type of analysis, the discrepancies in feature delimitation becomes easily visible. If on the other hand similar datasets are not being overlaid, the discrepancies will never become known and inaccurate, or unsuitable data may be used for GIS analysis. This chapter has attempted to portray the base from which most data are accumulated into the GISs of small users, and as such it forms the basis of all the errors and inaccuracies that will be explored in chapters Five and Six.

> "Haphazard digitising can result in digital data with a short life, or require a great deal of remedial effort to be converted into a useable form. Careful planning and an understanding of how the digital data are to be used will not only ensure the timely delivery of useful databases but will help to eliminate unnecessary waste of personnel and computer resources."

Lai (1988, 343).

## Chapter Five - Errors associated with building topology.

### 5. 1. Introduction.

Reality as we perceive it, often relies on some means of simplification if it is to be effectively communicated. In Chapter Three, the breakdown of cognitive processes, combined with the use of an accepted cartographic language was used to illustrate the common use of the map as a model of reality. In adopting such a model as the base for any input into a GIS, the process in many respects becomes reversed, as the aim is once again the construction of reality from an imperfect collection of sources. The first stage of creating a computerised version of reality begins with the acqutisition of information, but as Chapter Four has illustrated, problems arise due to the scale effects and the inherent generalisation present within the maps themselves, as well as due to the subjective generalisation of the digitising operator and the idiosyncrasies of the digitising process. As Muller (1990) effectively points out, the attainment of a scale free database cannot be envisioned in the immediate future.

Following the aquisition of information, is the need to construct the computerised model to hold that information in a way that best represents reality as dictated by the user defined needs of comprehensiveness and policy. For example, National Mapping Agencies responsible for a national mapping strategy, will implement an information model that relies heavily on comprehensiveness and the implemenation of strict policy procedures, such as the DLG-E executed by the USGS. Agencies with an organisational priority, such as those dealing in forestry management, will opt for a model that has an emphasis on the structural aspects of its data. Other agencies, such as those responding to commercial or user needs will in turn select different criteria for the composition of their models.

The purpose of the data model should seek to address several needs, notably topographic, topological and attribute. Topographic needs concern the accuracy of the material that is being used. For example there is a need to ensure that contour lines do not cross or that rivers do not flow uphill. Such needs concern the comprehensiveness of the information and have a bearing on the potential legal liability of the organisation utilising the data. Topological needs relate to the adjacency of the features being represented, and in some cases to the hierarchy of the relationships that are maintained within the model. If for example the database consists of line data, to be used in a network format, then the lines will all have to be joined so that they are topologically contiguous, enabling flows to take place along them (figure 5.1). Alternatively, if points and lines are to be built up into polygons for the purpose of areal analysis, areas must be topologically closed, with no gaps between lines at the endpoints of nodes. Furthermore, there must be no lines that enter the polygon, or 'dangle' into it (figure 5.2). Attributes are often added after the entry of the geometric features, and they are attached onto them via a network of specific relationships, rather than being incorporated as a fundamental part of the feature itself. Finally there is another set of needs which may be viewed as an extension of the attribute needs, these may be termed 'information needs', and are in many respects external to the features themselves. These are represented by pieces of geographic information that hang onto the spatial hooks such as unemployment data or transport flows.

In many repects however, the data model is constrained by the technology it seeks to serve. As the relational database model relies on an atomic rather than an object oriented approach, the links between features are presented implicitly, rather than explicitly. A factory, for example is a two dimensional area, which in turn will be comprised of several component parts, such as a processing unit, administration, a car park and so on. Using a relational data base, as the scale at which the information is represented decreases, the location of the factory will tend to an x,y coordinate, and the elements that make up the factory will be lost, whereas with an object oriented model, features and their corresponding elements are retained as an entity. Although some of the more contemporary GISs are tending to the adoption of an object oriented approach[1], the majority of systems still retain the relational database facility, irrespective of whether this is the most efficient and suitable way to model data, but rather because it is acknowleged as a defacto standard. The adoption of the relational database concept arose with Cooke's paper in 1967, which discussed the early DIME philosophy and prompted its adoption as a pragmatic interim solution to problems of the time, and which then became the international standard.

---

[1]Notably the Smallworld system, and allegedly the TIGRIS system marketed by INTERGRAPH.

In Chapter Four, the problems associated with the manual digitising process were looked at in detail, especially aspects associated with operator inaccuracy and generalisation. This chapter seeks to look at these 'errors', why they need to be eliminated, the options available for their removal, and what such routines do to the data.

In the digitising operation, the inherent information that is present in the data is destroyed, as the geometrical features are abstracted from the reality they represent. Once they are assimilated into a computerised environment, these relationships have to be restored if the data model is to function. Rebuilding the relationships makes certain demands upon the digitised information and this in turn will depend on the nature of the model itself. As this study will concern itself with polygon data, 'errors' are viewed to be lines which intrude into polygons,[2] or lines that define polygon boundaries which do not perfectly meet.[3] These are not necessarily errors of consequence in terms of positional accuracy, but in terms of the topological model, they are intrusive and undesirable.

**5. 2. Methods for dealing with digitising error.**

The methods for eliminating such errors in GIS packages and most digital mapping packages are twofold; either such errors are avoided in the first place, or some method of selective tolerancing is carried out after the digitising process has taken place. The first method involves the elimination of such errors at the digitising table. This can be done by assigning a high snap tolerance or by assigning polygons priority numbers.

Setting a high snap tolerance will allow the endpoint of the last line digitised to be snapped onto the nearest node, or alternatively it will allow the start point of the new line to be snapped onto the endpoint of the nearest line within the tolerance set. Selecting the nearest node is done by defining an areal tolerance, within which a search for a node to snap onto is made (figure 5.3). If several nodes fall within the tolerance, the decision as to which of the nodes might be the correct one to snap onto is arbitrary. Often the node chosen is the

---

[2]Often termed overshoots, as the operator, in all likeliness has 'overshot'the intended stopping point for the line.

[3]Often termed undershoots, as the operator has probably underestimated the required matching points for the lines.

first one that the search comes across; the order of searching usually being determined by the order in which the nodes were placed into the system, rather than on any indicator of their relative accuracy. Naturally, the higher the snap tolerance becomes, then the greater the chances of the system attaching the digitised line onto the wrong node. To overcome the possibility of node error completely, snap tolerances have to be set to relatively high levels. This proves problematic as it reduces the potential of a high level of detail for the entire coverage, so that the possibility of errors in a few cases can be eliminated. Therefore if accuracy rather than speed is the aim, such a process of data assimilation is undesirable.

Alternatively, (if polygon data is to be used), polygon priority codes could be assigned at the time of data entry. This reduces the need for topological digitising (points to lines to areas) and relies instead on the user digitising the polygons as whole entities, rather than viewing them as being made up of individual lines bonded by a series of relationships. The problem with this approach is that polygon boundaries have to be digitised twice, frequently resulting in the creation of sliver zones where the operator has failed to define the same line in exactly the same way (figure 5.4). Circumventing this undesirable effect can be achieved by assigning priorities to polygons. The higher the priority number, the 'more accurate' the outline of the polygon. This method allows the line of the polygon with the higher priority to override the line of its neighbour, should a degree of overlap occur. Digitising procedures making use of such an option may follow several routes. With the first, the operator digitises all the polygons as accurately as possible and then assigns priorities visually. Although seemingly maintaining accuracy levels at a premium, such a method may in fact lead to tiny loopholes between polygons which then results in unassigned areas. A more reliable method for guaranteeing error free regions is to digitise one polygon first and then to build upon than, increasing the priorities sequentially with each polygon. Once a polygon is digitised, a subsequent polygon that shares the same boundary will have that boundary falsely extended within the first polygon's boundaries (figure 5.5). This procedure is often utilised by systems that do not rely on the data being in a topological format.

As most vector based GISs use the topological data structure as the basis for data storage and data entry, strategies for 'cleaning' data and ensuring topology is correct are essential before further analysis can take place. Assuming that high snap tolerances are unacceptable, (since they promote a crudity in the data greater than the inaccuracies that the operator may in fact have instilled), other methods of error removal have to be sought. These routines generally fall into two categories; those concerned with undershoots and those used to resolve overshoots.

**5. 3. ARC/INFO as a case study of the technological tools available.**

ARC/INFO is one of a number of leading GISs, and as such demonstrates the standard type of technology utilised in the 1980s and 1990s for resolving these types of topological problems[4] . Furthermore, ARC/INFO has been a clear world leader in the PC market since ESRI released the system in the late 1980s. The adoption of the system as part of the ESRC Regional Research Laboratory (RRL) Initiative in 1988 has only served to strengthen that position. In UK academic circles the consequence of the CHEST agreement allowing academic licences of the software to be purchased at greatly reduced prices, has yet to be realised, but it is likely that this too will provide an added boost to the systems' impact.

In ARC/INFO these two categories of error can be dealt with by the use of two sets of algorithms that modify the coverage in a blanket manner, or alternatively through an interactive edit facility. As an average 1:25,000 Ordnance Survey digitised sheet will have as many as 300 to 400 node errors[5], interactive editing often proves to be impractical and therefore blanket algorithms are relied upon to alleviate the majority of the error, with only the odd mistake being allayed using interactive edits.

Two mechanisms are commonly used for coverage modification in ARC/INFO, both of which rely on the use of a linear tolerance value. Features falling short of the prescribed tolerance value will be altered. The first of the mechanisms will match nodes together and therefore removes undershoots, whilst the second eliminates dangling or overshooting arcs and also enables a weeding operation to be enacted on the lines in the coverage.

---

[4]Other vector based systems widely available on the market at present include INTERGRAPH, PRIME System 9, GENASYS and GIFIS.

[5]Figures obtained during test runs using Ordnance Survey data.

The two types of node error are approached in separate but similar ways. Once a coverage has been digitised, the numbers of errors residing within it need to be ascertained. Node errors may be defined as dangling nodes or pseudo nodes. *Dangle* nodes imply that the arc onto which the *dangle* node is attached hangs isolated in space. *Dangle* nodes may therefore define both undershoots or overshoots (figure 5.6). Since ARC/INFO only recognises a node as a graphic object that marks the intersection of three or more lines, any points that are defined as nodes which occur only at the junction of two lines are flagged. These may indicate the place at which a missing arc should intersect.[6] Unlike dangling nodes however, pseudo nodes do not prevent the creation of polygons, since the lines on which they occur can form topologically closed areas.

**5. 3. 1. The node matching operation.**

Removing the *dangle* node errors may either involve the bringing together of the end points of two lines, or actually removing small spurious lines that are unintended. Within ARC/INFO, it is advised that gaps between nodes should be removed prior to overshoots as the operations for dealing with overshoots can also be used to generalise lines at the same time. Eliminating gaps between two nodes uses an action called *mnode*. An areal distance is selected, to represent the extent which end nodes are allowed to stray from each other and any gaps between two endpoints that fall within this distance are assumed to be unintentional and to have come about due to digitising inaccuracy. Such gaps are therefore eradicated by stretching the two endpoints so that they converge at a common point. However, the nodes do not move in isolation; wherever they travel, the lines and vertices attached onto them are also obliged to move in that direction. It works in much the same way as the snap tolerance within the digitising module, but rather than setting an error limit within which the digitising process must operate, it can be used to respond to individual situations. Thus the *match node* can be set to be significantly smaller than the largest gap, in the knowledge that the majority of the errors will be removed and those that remain can be eliminated by a manual editing procedure.

- **Vertex movement and resultant errors.**

---

[6]Section 5; Understanding GIS (User Guide provided by ESRI)

The moving of a vertex (the last section of a line) to ameliorate errors at the endpoint of the line can often result in the start of the vertex being moved to compensate the extended distance that the vertex has to migrate in order to close a gap (figure 5.7). In situations where the majority of the lines are straight features such as in a coverage that contains information on field boundaries or drainage ditches (as is the case with the OS maps for much of the Fenland regions), the feature is represented by a single vertex, therefore when the end of the vertex has to be moved to accommodate an undershoot in digitising, the start of the vertex pulls away from its node; sometimes taking the line with it and sometimes detaching itself completely (figure 5.8).

- **Effects of line collapse.**

Nodes which are within the snap tolerance will also be snapped together, even if they are endpoints of the same line (figure 5.9). Thus lines which contain only two points, a start and an endpoint and which fall within the match tolerance, will be collapsed onto each other, and disappear altogether. This results in the obliteration of small intentional lines or areas.

- **The random nature of node matches.**

As the match tolerance acts on the first node it comes across, the results of the process can never be predicted. Three near accurate endpoints can be snapped onto one erroneously positioned node simply because the erroneous node has a higher sequential number than its co-nodes. This results in software contortions of what was an essentially accurate depiction. One supposed way of reducing the harmful side effects of this algorithm is to overlay the coverage to be cleaned with one that just consists of the point locations of the node end and start points. The algorithm can then be asked to use these points as the accurate meeting points of the lines in the coverage rather than simply relying on sequential node numbering in the coverage. If the points are entered into the system by virtue of their true x y location in space, then some of the danger associated with node movements can be allayed. However since it is unlikely that many users will have the time to devote to such an endeavour, the points themselves would probably have to be digitised. As with the original digitised material the operator will be as prone to the misinterpretation of locations as when digitising the line coverage.

• **Effects of arbitrary tolerance size.**

Debating what size the tolerance ought to be can also prove problematic. On the one hand, the tolerance width must be able to cope with the majority of the errors, but on the other hand, it must also be sensitive to the needs of the coverage and to the individual features that the lines represent. Tolerance gaps that are too conservative will still leave large proportions of errors unaltered, whilst gaps that are too generous may result in gross feature distortion and the loss of small lines altogether.

Such node movements can cause unintentional distortions to manifest themselves in the coverage, often influencing areas away from the node points. Yet to eliminate such node errors individually, (therefore allowing some measure of control as to the most acceptable movement of lines and vertices), would require inordinate amounts of time, which is a commodity few users have to spare. Therefore some degree of compromise must be reached, in which the majority of the errors are eliminated by one quick enactment of the algorithm, whilst preserving as much of the original coverage topology and character as possible.

• **Cartographic errors as features on the ground.**

If such topological alterations are then related to the real features represented by these processed cartographic lines, considerable distortion to the way that real world elements are portrayed can occur. The positional accuracy of real lines and consequently areas will therefore be brought into doubt, but in an unpredictable way. If the degree of alteration that the process inflicted upon coverages could be predicted then certain guidelines could be adhered to. Unfortunately, this is not the case as this operation causes the topology of coverages to change in a random and unforeseeable manner.

**5. 3. 2. The Clean operation.**

The second algorithm deals with lines that unintentionally dangle[7] or overshoot their intended stopping point. Firstly these have to be identified and secondly they have to be eliminated. The algorithm takes both these tasks on board. Initially, it searches all lines sequentially, if

[7]ARC/INFO defines dangling arcs as those which have the same polygon number on either side of it.

any lines cross without a node identifier being present at the intersection, it then splits these lines and places a node at the junction point. More lines than originally existed are therefore created and the lengths of those initially present are altered. The second part of the procedure is to then compare all lines to the *dangle* tolerance, anything falling below that tolerance is removed. This procedure, although less drastic than that carried out by the match tolerance will still alter the geometry of features and remove the evidence that lines were supposed to intersect at a particular place by eliminating the *dangle*s that might have offered a tenuous link to the proper endpoints of the lines.

- **Consequences of arbitrary values for the *clean* operation.**

The consequences of this algorithm are far less severe than those which arise due to the employment of the node snapping algorithm. If the tolerance is too small, then *dangle* errors will persist and delay the creation of homogenous polygons. If the tolerance is too large then lines which are intended to dangle might disappear (figure 5.11).

- **Topological changes following the use of *clean*.**

The topological consequence of this activity is to change the number of lines that are actually stored, since the intersecting of arcs brings about the creation of a number of small arcs that might still fall above the tolerance limit. In addition, the lengths of arcs change. These effects in turn influence the area held within a certain polygon and use the perimeter of the polygon. In terms of linear features; extra routes will emerge that do not exist in reality, and which in turn may need to be edited to avoid error-prone assumptions to be made at the stage of data analysis.

### 5. 3. 3. The *weed* or *fuzzy* tolerance.

In addition to the techniques that deal with the endpoints of lines, one further algorithm exists for the modification of digitised data. Sometimes when digitising too much detail is assimilated. The operator might for example have digitised a particular line several times, or alternatively too many points might be used to represent a relatively simple feature. A weeding procedure will easily erase such excesses by eliminating points that actually make up the character of the line (figure 5.12).This is carried out by a linear tolerance defined as the minimum distance between two vertices, if this distance is too small then superfluous information might be retained, causing the unnecessary waste of computer storage and processing time during later analysis. Alternatively, a tolerance value that is set artificially high will cause arcs to alter their paths and become severely generalised so that their curves are spiky and less pronounced. Not only are 'dangling' lines that fall within the *weed* tolerance removed, but so are individual vertices that make up the lines themselves. Therefore this function is able to perform operations associated with both of the two preceding algorithms, in that if a gap exists between two arc endpoints, and that gap is less than the weed tolerance, they will be snapped together. If however, a 'dangling arc' falls below the tolerance value, it too will be removed.[8] A more common problem associated with this routine however is the disappearance of lines altogether. If small connecting lines that exist in a narrow gap between two greater lines fall below the *weed* tolerance, then the two endpoints to which they are attached may be snapped together and lines may simply disappear.

These algorithms are rarely used in isolation, and are often used in a variety of combinations to achieve the elimination of an assortment of errors that may prevail within the coverage, however there is little guidance from suppliers to aid the user in tolerance selection. In general, it seems to be an iterative process of trial and error. If the length selected is too great, then not only will the errors disappear, but so might bona fide gaps between lines or intentionally small arcs such as a small T-road. Alternatively, making the tolerances too small results in many remaining errors which then have to be removed individually by manual methods. Ideally, tolerance levels should be set so that they eliminate the majority of errors, leaving the remainder to be dealt with individually. Naturally this takes more time, but unlike the former approach does not bring all the errors down to the lowest common denominator. Consequently the accuracy of the entire coverage is not reduced by the existence of what might possibly be one inaccurate outlier.

### 5. 3. 4. Manual editing of node errors.

Correcting node errors is by no means limited to using blanket tolerancing. If desired, each

[8]This does not just apply to straight line segments; that is lines that only have one vertex, but also to lines that are made up of several vertices, as the vertices will gradually be coalesced until only one remains.

node error may be dealt with individually, however time constraints are often a crucial factor in determining the level of accuracy data is able to attain, and eliminating four hundred errors by hand would devour a large proportion of any operator's time, therefore sacrificing a few points to the jaws of inaccuracy in order to save time on the majority is often the chosen route.

Yet even the process of manual editing is subject to a degree of generalisation and subjectivity, as each individual is likely to implement a different route to the elimination of a topological error. Undershooting arcs may for example, be treated in several ways; one may choose to add an extra arc to make up the distance, add an extra vertex, move the last arc or vertex, or any of these actions combined with other options such as splitting or removing arcs or vertices.

Other digitising errors such as digitising a line twice or missing out arcs may be dealt with interactively, using a manual editing procedure. The editing facility comprises an entire module of PC ARC/INFO and can therefore be expected to offer a wide range of options to ameliorate any single graphical problem. For example lines that don't quite match, even after match noding and *dangle* tolerancing, may be brought together, by extending the line with an extra vertex. Alternatively, the line might be deleted and the whole entity redrawn. Another option would be to allow an extra arc to be added to the end node of the line. Finally the last vertex of the arc might be moved so that it extends to meet the node it is supposed to intersect with. With such a wealth of geometric manipulations available to the user, it is small wonder that even the same map when digitised and edited by the same person will not appear to be entirely identical.

Several factors need to be taken into account when these tools are applied. In the first instance there is the misuse of terms such as 'error' and 'clean'. In the ARC/INFO package, the denotation of node mismatches as 'errors', may also focus the attention of the user in false directions. Node errors become perceived as satanic intrusions into the coverage space, and with evangelistic fervour elimination is preached. Such elimination often occurs at a cost to the coverage topology; a fact the user may be entirely unaware of. Also of importance is the false assumption that routines which remove node error somehow make the coverage more 'accurate'. This is not really the case; such routines merely make the coverage suitable for analysis by that particular system, as it is the system that defines node mismatches as 'errors'. To some extent therefore, this chapter seeks to challenge that assumption, and to point out that although mistakes associated with endpoints may well be obliterated, they disappear at a positional, geometric and topological cost.

A second factor is the disparity that exists between the human visualisation of information and the system's model of representing reality. A computer generated line may appear as a line when it is plotted, but the internal information that gives that line its value and in effect its 'geography', may view a line in many differing ways. For example, a line may be used to represent an artificial zone, such as an administrative unit, or the line may reflect a physical feature that exists on the ground. Topologically, a line may be used to form part of an areal boundary, or it may exist as a line or feature in its own right. Furthermore, a continuous line on paper or on the ground may not be stored in a similar fashion in a computerised system. Thus roads for example, are viewed in the mind's eye as continuous features, and there is no thought of seeing them as separate entities demarcated by their intersections with other streets or thoroughfares. If, for example a motorway is to be represented by a linear feature, one would never contemplate viewing its portrayal by several discrete lines with their start and end points at the location of slip roads. GIS storage however, is quite likely to do this as every intersection has to be accounted for. Thus seemingly singular features on the ground may entail depiction by several segments within the system.

Therefore what the user comes up against, is the difficulty of reconciling the visualisation of the line as it exists in the real world with the storage of the line in a computer system. When the line is perceived, it is visualised as a real entity, complete with all its geographic value, rather than as a graphic entity. Thus lines are digitised as detached features in space, and although they might cross over (and therefore have some relationship with) other graphic features, they have not been given the distinction of having a true intersection with that feature. With its entry into a GIS the user may assume that the added value associated with that line is inherent, but unfortunately, such information has to be built and acquired.

In many cases however, it can be argued that the final practical visualisation of the line lies with the digitising operator. For example, since operators often devote more time to endpoint definition these may actually be more accurate than those points which define the mid-points of a line. If, however the operator is told to follow the ESRI instruction to overshoot [9], which anticipates the user pinpointing the intersection of the two lines rather than the true intersection of the node endpoints, the line length of lines will be deliberately extended to allow the error correction processes a greater chance of success. Consequently, the linear features in the coverage will be a contorted version of the map.

**5. 4. Assessment of change rather than diminishing accuracy.**

The following experiments attempt to evaluate the impacts of the tolerancing process. However, no statements about the varying accuracies of the coverage will be made, as there is no means with which to statistically compare the results of these procedures with the real representation of the ground surface. As discussed in Chapter Three, a 'real', error free and exact cartographic representation does not exist, but is merely an ideal that is strived towards[10]. Admittedly, it might be argued that the processed coverages could be overlaid with the original map from which they were traced and subsequently digitised. This would however only serve the purpose of visual interpretation, and no quantities of difference in line placement or length could be derived.

Thus the experiments in this chapter seek to make statements about change, rather than accuracy. It is impossible to say how much more accurate one coverage is than another, but it is possible to evaluate the differences between them. It is such differences adopted at this stage of data preparation that are responsible for problems of mismatch at the later stage of data integration.

---

[9] ESRI unofficially tell their users that they should aim to overshoot at node endpoints if possible, as this will reduce the need to use the potentially dangerous MNODE options and helps the CLEAN algorithm to be more effective. In the past year, this has gone into print in the Understanding GIS manual, Section 4-11.

[10] As discussed in Chapter 3, this is a view which the National Map Accuracy Standards are also tending towards, as they have tended to steer away from statistically defined measures and towards user based evaluations.

## 5. 5. A visual evaluation of the effects of tolerance application in ARC/INFO.

The following experiment sets out to investigate what the possible topological effects on a coverage might be when a variety of tolerance values are applied to it. A further aim is to find a situation in which the damage to topographic features is minimised, whilst the removal of errors is maximised. Both the algorithms of node snapping and the removal of dangling lines are considered, as well as to a lesser extent the *weed* option. Initially however, rather than applying these to a set of topological relations taken from a true cartographic coverage, a test coverage was designed that embodied a large selection of situations in which linear features susceptible to delicate changes in tolerance values were included. These features were also representative of geographical patterns that often emerged in maps, such as closely indented coastlines, lakes or shorelines with narrow inlets, roads or rivers that run parallel, small offshoots from linear features and narrow zones are just some of the characteristic entities that are depicted (see figure 5.13).

### 5. 5. 1. Methodology.

The rule for tolerance application in the ARC/INFO software is that the *mnode* operation should be performed first followed by *dangle* and *weed*ing tolerances. This is because *mnode* is viewed as part of the digitising corrective procedure and the *clean* operation which deals with the *dangle* and *weed* tolerances is seen as part of the topology creation process.[11] A suitable procedure to follow, is an iterative one that involves setting tolerances, observing the results and then repeating the processes with slightly increased values.[12]

The coverage itself consisted of 31 node points and 20 areas and was drawn with a 0.25mm nibbed pen onto plastic film. It was then digitised using an initial *weed* tolerance of 0.05cm and a snap tolerance of zero. This ensured that very little of the error was eradicated at the data acquisition stage, where feature alteration would effectively be invisible to the user.

Once the data had been digitised the coverage was examined to see what the maximum

[11]Capter Four, PC ARC/INFO Starter Kit, Users Guide (1988).

[12] An initial tolerance can be obtained by selecting the distance corresponding to the largest node error or the smallest retainable arc.

difference between unmatched nodes was. The largest mismatches were selected and their lengths measured. The gaps encountered are shown in table 5.1.

Table 5.1. Gaps between nodes in the test coverage

| Distance | Type |
|---|---|
| 1.53mm | Unintentional |
| 0.73mm | Unintentional |
| 2.79mm | Unintentional |
| 1.28mm | Intentional |

The largest unintentional gap[13] is that of 2.79mm, however this is greater than the intentional gap of 1.28mm. A dilemma now presents itself. Should the user elect;

- to use a gap tolerance of greater than 2.79mm, thus eliminating all the gaps?
- or should only tolerances less than 1.28mm be employed, enabling the integrity of the coverage to be maintained?

The dangling arcs also provided a similar impasse, in that the largest undesirable *dangle* was greater than small arcs which needed to be retained (table 5.2).

[13]That is to say a gap which exists between lines in the coverage that should not be present.

Table 5.2. Dangles within the test coverage.

| Size (mm) | Action |
|---|---|
| 3.075 | remove |
| 2.240 | remove |
| 1.130 | keep |
| 0.350 | keep |
| 0.420 | keep |
| 0.360 | keep |
| 0.560 | keep |
| 0.110 | keep |

It was decided to test a variety of tolerances with the *mnode* option first. Selected tolerances are shown in table 5.3 in conjunction with the number of node errors that they leave behind.

Table 5.3. Node tolerances used and the subsequent number of errors and nodes retained.

| Tolerance (mm) | Number of nodes retained | Number of errors |
|---|---|---|
| 5 | 33 | 3 |
| 3 | 37 | 3 |
| 2 | 39 | 4 |
| 1 | 39 | 4 |
| 0.5 | 51 | 17 |
| 0.2 | 67 | 35 |
| 0.1 | 67 | 35 |

In the 2mm and 1mm tests exactly the same results were obtained. The same applied to the 0.1mm and the 0.2mm tests (table 5.3). Of this group therefore, it was decided to only utilise the 2mm and 0.2mm tolerances for further testing. As the distribution of number of errors removed by these initial tests were quite skewed, further tolerances of 0.8mm, 0.4mm and 0.3mm were tested in an effort to obtain a more evenly distributed sample of the effects of *match node* tolerances. It was decided to drop the 0.02mm tolerance as this did not substantially reduce the number of node errors from the raw digitised file. The final selected values for dealing with node gaps were are shown in table 5.4 along with the values of the *dangle* and *weed* tolerances which were selected using a similar procedure.

Table 5.4. Tolerance values selected for use.

| |
|---|
| Match node tolerance values selected.<br><br>• 0.3mm<br>• 0.4mm<br>• 0.5mm<br>• 0.8mm<br>• 2mm<br>• 3mm<br>• 5mm. |
| Dangle tolerance values selected.<br><br>• 0.2mm<br>• 0.8mm<br>• 3mm<br>• 6mm |
| Weed tolerance values.<br><br>• 0.3mm<br>• 0.5mm<br>• 0.8mm<br>• 3mm<br>• 5mm |

Initially, the *match node* tolerances were applied to the test coverage in isolation, to see what the effects of progressively increased tolerance distances would be. Then the coverage was submitted to further tolerancing using *dangle* and *weed*ing algorithms in addition to the initial *match node* tolerances. Coverages can be divided into four groups based on the type of tolerance used;

- those that were only submitted to *match node* tolerancing,
- those that were *match node* toleranced and then submitted to a *dangle* tolerance;
- those that were *match node* toleranced, *dangle* toleranced and *weed* toleranced

- those that were *weed*ed and *dangle* toleranced.

The outcome of the tolerances were divided into those coverages which had few node errors remaining and those which still retained a large proportion of their original errors. Analyzing the results of these 'tolerance cocktails', is crucial to the integrity of the coverage. If the criteria of node error elimination is the only one to be used, then distortions manifested within the geometry of the coverage will pass by unheeded. This study therefore aims to look at the resultant coverage modification following error cleaning using several methods. As it is the change in the number of node errors and node points that is usually of paramount interest to the user, coverages will be initally assessed on this basis.

For coverages with few errors, the reduction of 38 node errors to zero was almost achieved by two tolerance combinations;

- that of a *weed* tolerance of 5mm plus a *dangle* tolerance of 6mm
- and also by a match tolerance of 3mm followed by a *weed* tolerance of 0.5mm and a *dangle* tolerance of 3mm.

Both managed to reduce the number of node errors to 2. Although the number of node endpoints on the two coverages do not differ greatly (table 5.5), it is the difference in the number of segments that offers a clue as to the potential coverage destruction that might have taken place.

The results of the processed test coverages are shown in figure 5.14 and figure 5.15. The brutality of the *weed* tolerance can be seen in full evidence in figure 5.14. If this is compared with the original coverage (figure 5.13), the discrepancy becomes even more apparent. The resultant image is only a mere skeleton of its former self, with a great deal of detail being lost. The coastline no longer exists in any recognisable form. The sinuosity of the lines has been completely destroyed, and in its place new polygons have been formed. On the right hand side of the coast, a large *dangl*ing arc has appeared, which is the remainder of the inlet that has collapsed onto itself. Another line that has had its detail removed is the very jagged line in the southwest of the coverage. All lines that existed in close proximity to each other, such as the river feature in the north west of the coverage, the narrow section of land between two polygons in the centre of the coverage, the thin spit like feature in the extreme northeast and the rectangular piece of land in the north east of the coverage have been collapsed onto themselves and these areas are now represented with lines.

Small gaps between areas have been dealt with either by being brought together, so that entire polygons have been formed instead of narrowly open areas, or by being dragged apart so that they no longer represent the feature they were intended to. Examples of the former case can be seen in the south east of the coverage, where the inlet has had its endpoints brought together and also in the centre of the coverage, where a narrow gap between two relatively parallel lines has been lost altogether. Intentionally small areas, such as the two tiny polygons in the centre of the coverage and small lines, such as that attached to the western extreme of the coastline have naturally disappeared completely. Furthermore, the initially straight coastline at the north of the coverage has become indented where some of the arcs that were attached to it have pulled this line downwards.

Although the coverage has changed in character, all the dangling arcs that existed have been eliminated; so in effect, the aim of the tolerancing techniques has been achieved, but at the cost of positional accuracy. The two *dangles* that still exist in the coverage, are not those that were originally present as a result of the initial digitising errors. Rather, they are a by-product of the churning up of features by the large *weed* and *dangle* tolerances employed.

The use of such a high *weed* tolerance can therefore be very destructive on particular features. However the majority of the map features have still remained in approximately the same location, and the larger features are relatively unscathed. Therefore if the test coverage were only made up a small fraction of a much larger coverage, comprised mainly of much larger features, such omissions and discrepancies might go unnoticed, or may be regarded as relatively unimportant.

The second coverage to produce a result of only two *dangles* (figure 5.15), has been far less detrimental to the coverage features than its counterparts. The coastline has remained intact, with only a moderate amount of generalisation having taken place on its features; notably in the finer detail of the protrusion on the eastern spit. Other narrow features, small areas and inlets have also emerged unscathed. Only three small arcs have been displaced or lost; the small *dangl*ing arc on the western coastline, a slight displacement on a node in the centre of the coverage and a slight collapse in the spit in the extreme north east of the coverage. The two *dangle*s that still remain are those which the *match* tolerance has been unable to address and which could be removed using an interactive editing procedure.

There are eight coverages that exhibit only three *dangle* node errors (table 5.5) and these can be divided into three distinct groups. Firstly there are those coverages which have high *weed* tolerances, no *match node* tolerances and which irrespective of the different *dangle* tolerances attached to each still only have 51 nodes and 71 arcs. The character of these coverages is the same as that first discussed, with the exception of one extra self imposed *dangle* (figure 5.16). This is due to the fact that the *dangle* tolerance in these two cases is smaller than in the original example, (less than 6mm) and is therefore unable to obliterate large *dangl*ing lines.

The next group consists of those coverages which have only been submitted to the *match node* tolerance. These coverages have surprisingly few errors (either in terms of *dangle* or pseudo nodes), but it must be remembered that they are not as yet topologically complete, since all crossing lines have not yet been intersected and no PAT file exists for these coverages. There are two such coverages which only yielded a total of three *dangle* errors. The first (figure 5.17) was submitted to a *match node* tolerance of 3mm and has survived relatively unimpaired. The only deviations form the original occurring in two small zones, both near the centre of the coverage where some form of node collapse has taken place. One of the *dangle*s still remaining is an intentional one in the south west of the coverage, which has been obliterated in all the previous trials. Its presence here and in the following case is due to the absence of a high *dangle* node tolerance.

A slightly increased *match node* of 5mm can be seen in figure 5.18. Here more drastic node collapse has taken place. Collapse has occurred in several places in the top right hand corner of the coverage, and as a result of one of these a new *dangle* node has been created. Furthermore, the river feature at the top of the map has been collapsed onto itself, and in an effort to remove the *dangle* in the south west of the coverage, the algorithm has managed to cause the line to collapse onto the nearest node. In addition a new *dangle* node has been created, in the east of the coverage.

The final set of coverages with only three *dangle* node errors are all characterised by a high *dangle* tolerance of 6mm. Prior to that they were all treated using different procedures. Both the coverages that were match toleranced first, retained all the characteristics of those operations, (figure 5.19 and figure 5.20) with the added removal of one of the lines in the east of the coverage that fell within the *dangle* tolerance, resulting in a newly contrived *dangle*. The last coverage which managed to achieve a result of only three *dangles* was not match toleranced, but had a *weed* tolerance of 0.8mm, which caused the beginnings of a generalised line to emerge from what was once the extremely jagged line in the south west of the coverage and also in the southern coastline. In addition, the intentional gaps created during digitising on the southern extremity of the river feature have not been snapped onto each other, but rather they have been dragged further away from each other.

At the opposite end of the spectrum are those still riddled with errors as a consequence of insufficient tolerancing values. Figure 5.21 illustrates the case where in which twenty *dangle* errors still remain. The gradual demise of the coverages can be seen by looking at figures 5.22 to figure 5.25. Progressively more severe tolerances have been applied to the original coverage. The result is a gradual loss of character and an increasing tendancy towards inaccuracy.

**5. 5. 2. Effects of the individual operations.**

By analysing the geometrical change in those coverages that have resulted in few errors, some preliminary repercussions of various error tolerancing routines can be inferred. It would seem that high *weed* tolerancing results in dramatic segment collapse, capable of reducing the number of segments in a coverage to an eighth of its original dimensions. High *match node* tolerancing results in segment and sometimes arc collapse, whilst high *dangle* tolerancing removes lines that are accurately placed on the assumption that they are in error because they do not exceed the tolerance limit. More detailed analysis can be deduced by looking at each of the routines in turn and also by assessing the statistics associated with each coverage.

The results of four different *match node* tolerances applied in isolation to the coverage are shown in figures 5.26 to 5.29. The smallest match tolerance of 0.3mm (figure 5.29), can be

Table 5.5 Statistics for all test coverages

| COVERAGE | MNODE | FUZZYTOL | DANGLETOL | NODENO | ARCNO | SEGNO | DANGLENO |
|---|---|---|---|---|---|---|---|
| TOLF5D6 | 0.0 | 5.0 | 6.0 | 51 | 70 | 248 | 2 |
| TOL3D3 | 3.0 | 0.5 | 3.0 | 57 | 69 | 1752 | 2 |
| TOLF5D3 | 0.0 | 5.0 | 3.0 | 51 | 71 | 249 | 3 |
| TOLF5D8 | 0.0 | 5.0 | 0.8 | 51 | 71 | 249 | 3 |
| TOLF5D2 | 0.0 | 5.0 | 0.2 | 51 | 71 | 249 | 3 |
| TOL3D6 | 3.0 | 0.5 | 6.0 | 57 | 68 | 1743 | 3 |
| TOL3 | 3.0 | 0.0 | 0.0 | 37 | 56 | 1823 | 3 |
| TOL5 | 5.0 | 0.0 | 0.0 | 51 | 55 | 1822 | 3 |
| TOL5D6 | 5.0 | 0.5 | 6.0 | 59 | 72 | 1745 | 3 |
| TOLF-8D6 | 0.0 | 0.8 | 6.0 | 61 | 70 | 1499 | 3 |
| TOL5D3 | 5.0 | 0.5 | 3.0 | 59 | 73 | 1751 | 4 |
| TOL2 | 2.0 | 0.0 | 0.0 | 39 | 56 | 1823 | 4 |
| TOLF3D6 | 0.0 | 3.0 | 6.0 | 75 | 99 | 448 | 5 |
| TOLF-8D3 | 0.0 | 0.8 | 3.0 | 61 | 72 | 1504 | 5 |
| TOL-8D3 | 0.8 | 0.5 | 3.0 | 55 | 62 | 1753 | 5 |
| TOL-8D6 | 0.8 | 0.5 | 6.0 | 55 | 60 | 1737 | 5 |
| TOL3D8 | 3.0 | 0.5 | 0.8 | 58 | 73 | 1760 | 6 |
| TOLF3D2 | 0.0 | 3.0 | 0.2 | 75 | 101 | 450 | 7 |
| TOLF3D8 | 0.0 | 3.0 | 0.8 | 75 | 101 | 450 | 7 |
| TOL5D8 | 5.0 | 0.5 | 0.8 | 60 | 76 | 1756 | 7 |
| TOLF3D3 | 0.0 | 3.0 | 3.0 | 75 | 101 | 450 | 7 |
| TOL3D2 | 3.0 | 0.5 | 0.2 | 58 | 75 | 1762 | 8 |
| TOL-808 | 0.8 | 0.5 | 0.8 | 56 | 66 | 1761 | 9 |
| TOL-8 | 0.8 | 0.0 | 0.0 | 46 | 56 | 1823 | 10 |
| TOLF-8D8 | 0.0 | 0.8 | 0.8 | 61 | 77 | 1512 | 10 |
| TOL5D2 | 5.0 | 0.5 | 0.2 | 60 | 79 | 1759 | 10 |
| TOLF-8D2 | 0.0 | 0.8 | 0.2 | 61 | 77 | 1512 | 10 |
| TOL-8D2 | 0.8 | 0.5 | 0.2 | 56 | 69 | 1764 | 12 |
| TOL-5D6 | 0.5 | 0.5 | 6.0 | 59 | 59 | 1726 | 12 |
| TOLTED3 | 0.0 | 0.5 | 3.0 | 61 | 62 | 1747 | 13 |
| TOL-5D3 | 0.5 | 0.5 | 3.0 | 59 | 62 | 1748 | 13 |
| TOL-3D3 | 0.3 | 0.5 | 3.0 | 61 | 62 | 1747 | 13 |
| TOLTED6 | 0.0 | 0.5 | 6.0 | 61 | 59 | 1725 | 14 |
| TOL-3D6 | 0.3 | 0.5 | 6.0 | 61 | 59 | 1725 | 14 |
| TOL-5D8 | 0.5 | 0.5 | 0.8 | 60 | 67 | 1757 | 17 |
| TOLTED-8 | 0.0 | 0.5 | 0.8 | 62 | 66 | 1755 | 17 |
| TOL-5 | 0.5 | 0.0 | 0.0 | 51 | 57 | 1824 | 17 |
| TOL-3D8 | 0.3 | 0.5 | 0.8 | 62 | 67 | 1756 | 19 |
| TOL-5D2 | 0.5 | 0.5 | 0.2 | 60 | 69 | 1759 | 19 |
| TOL-3D2 | 0.3 | 0.5 | 0.2 | 62 | 70 | 1759 | 20 |
| TOLTED-2 | 0.0 | 0.5 | 0.2 | 62 | 69 | 1758 | 21 |
| TOLF-3D3 | 0.0 | 0.3 | 3.0 | 65 | 58 | 1866 | 21 |
| TOLF-3D6 | 0.0 | 0.3 | 6.0 | 65 | 55 | 1841 | 23 |
| TOLF-3D8 | 0.0 | 0.3 | 0.8 | 65 | 61 | 1874 | 23 |
| TOLF-3D2 | 0.0 | 0.3 | 0.2 | 65 | 66 | 1880 | 24 |
| TOLF-3 | 0.0 | 0.3 | 0.0 | 65 | 66 | 1880 | 24 |
| TOL-4 | 0.4 | 0.0 | 0.0 | 57 | 57 | 1824 | 24 |
| TOL-3 | 0.3 | 0.0 | 0.0 | 62 | 57 | 1824 | 29 |
| toltest | 0.0 | 0.0 | 0.0 | 69 | 57 | 1824 | 38 |

regarded as relatively ineffective in removing digitising errors, as 29 errors still remain; a reduction of only nine from the original coverage (table 5.5). Removing such errors interactively would be too time consuming and therefore a reduction of the *dangles* could only be achieved with excessively high *dangle* or *weed* tolerances. Assuming a desire to maintain a respectable segment number, high *weed* tolerances would be undesirable and therefore only a variety of *dangle* node tolerances have been applied to the coverage (figure 5.30 to figure 5.32).

As can be seen, the more conservative *dangle* tolerance of 0.8mm (figure 5.32) is unable to effectively reduce the number of *dangles* that exist as a result of nodes having gaps between them, and these remain at 19. Even a relatively high *dangle* tolerance of 3mm is still unable to resolve the majority of the errors, as they still remain at 13. An increase of the tolerance to 6mm results in a *dangle* increase to 14 (figure 5.30).

An increase of the *match node* tolerance to 0.5mm (figure 5.28) results in some improvement, with node *dangles* having been reduced to 17, whilst 0.8mm has caused the number of errors to fall to 10 (figure 5.27). A *match node* tolerance of 2mm has reduced errors to 4, although some node collapse is apparent here (figure 5.26). Adding a *dangle* tolerance onto the coverage produced by the *match node* tolerance of 0.8mm, would probably remove some of the true *dangles* that remain, which cannot be addressed by the *match node* tolerance without the expense of loss of detail through segment collapse.

The coverage created by a *match node* tolerance of 0.8mm was then submitted to a variety of *dangle* node tolerances to see how these might influence the error distribution and geometry of the coverage. The *dangle* tolerances applied were those of 0.2mm, 0.8mm, 3mm and 6mm, with a constant *weed* tolerance of 0.5mm. The *dangle* tolerance of 0.2mm, actually resulted in two more *dangle* nodes than were in the coverage prior to its application (figure 5.33). The *dangle* tolerance of 0.8mm managed to effect a reduction of one dangle (figure 5.34), whilst the two larger *dangle* tolerances have both enabled a reduction to five dangles, by removing arcs that overshoot in the lower coastline area of the coverage (figure 5.35 and figure 5.36). The 6mm tolerance, has however caused some segment collapse in an effort to remove what are comparatively large arcs. If the 6mm and the 3mm coverages are compared; differences become apparent. For example the small island in the centre of the coverage has been removed, and a line in the eastern part of the coverage has also been deleted. Thus at this level, the damaging potential of the *dangle* tolerance becomes conspicuous.

As with the *match node* tolerance, the higher the *dangle* tolerance becomes; the greater its potential for disaster. Increasing the *dangle* tolerance any further in this example, would only aggravate the situation, since the remaining errors could quite easily be dealt with interactively, as they are few in number and are a consequence of gaps, rather than overshoots.

If table 5.5 is analysed, several themes relating to the *weed* tolerance emerge. Firstly that high *weed* tolerances result in a dramatic drop in the number of segments attached to each coverage. Tolerances of 3mm and 5mm have reduced the number of segments in the coverage from 1824 to 450 and 249 respectively. A tolerance of 0.8mm keeps the segments at the 1500 mark, whilst a tolerance of 0.5mm is less of a determining factor in the number of remaining segments, as coverages submitted to this level of *weed* tolerance are more susceptible to the influences of the other two tolerances in force, and the segment numbers for a *weed* tolerance of 0.5mm vary from 1725 to 1764. No *weed* tolerance leaves coverages with segment values very close to the original, only dropping slightly where high *mnode* tolerances have eliminated some of the segments in the process of snapping lines together.

Curiously enough *weed* tolerances of 0.3mm, which is less than the digitising *weed* tolerance, actually causes an increase in the number of segments that exist in the coverage. Although this is geometrically imperceptible, segment numbers are being increased by the insertion of segment breaks at a shorter distance than in the original file and as a consequence more *dangle* nodes are beginning to appear (figure 5.37). Given the drastic contraction in feature detail that occurs through setting high *weed* and *dangle* tolerances in an effort to reduce the number of node errors, it would perhaps be advisable to consider the number of segments as an important secondary criteria in addition to the number of *dangle* tolerances remaining.

To illustrate the potential hazards of the *weed* tolerance, the test coverage was submitted to a progressively increased *weed* tolerance. So as not to obscure the effects of the *weed* tolerance, the *dangle* tolerance was kept constant and the *match node* tolerance was not employed. *Weed* tolerances of 0.3mm, 0.5mm, 0.8mm and 3mm were applied. Two such tests were carried out, the first with a constant *dangle* tolerance of 0.2mm and the second with a constant *dangle* tolerance of 3mm.

By repeating the experiment using a relatively small, and then a relatively large *dangle* tolerance some idea of the relative import of each of the tolerances can be ascertained. For example, if the *weed* tolerance is set too high, does the dimension of the *dangle* tolerance actually make any difference? Or is there a point at which the *dangle* tolerance starts to become just as destructive?

The set *weed* tolerance of 0.3mm failed to have any significant impact on either the coverage geometry or in the reduction of the number of node errors in conjunction with both the *dangle* tolerance of 0.2mm and 3mm (figure 5.38 and figure 5.39). In both cases the coverages had more segments than the original coverage (table 5.5), and the number of errors were still high; 21 for the *dangle* tolerance of 0.2mm and 24 for the *dangle* tolerance of 3mm. In this case, it would appear that irrespective of the *dangle* tolerance set, the resulting modification to the coverage is minor. Yet one would assume that a large *dangle* tolerance such as 3mm would remove a large number of the errors, if the errors were overshoots. Naturally gaps would be unaddressed by the *dangle* tolerance.

The *weed* tolerance of 0.5mm when applied with the same pair of *dangle* tolerances, begins to show some degree of error removal, as node errors now stand at 17 and 21 respectively. The difference in the number of errors is due to the 'cleaning' up of small arcs by the *dangle* tolerance following the adjustments to the coverage of the *weed* tolerance. The disparity is especially notable in the lower section of the coverage (figure 5.40 and figure 5.41). At a *weed* tolerance of 0.8mm, the gap between the two coverages remains moderately constant at five node errors difference, with the reduction in the total number of errors being due almost entirely to the *weed* tolerance (figure 5.42 and figure 5.43). It is only in the cases of a few overshoots that seem to appear in every coverage, that the higher *dangle* tolerance is proving to be of any significance.

In the final case, when a *weed* tolerance of 3mm is used, both the coverages display exactly the same geometry (figure 5.44 and figure 5.45) being made up of only 450 segments. In addition both have seven node errors. In this case, the aftermath of the high *weed* tolerance has so severely altered the coverage, that even the relatively large *dangle* node tolerance of 3mm was incapable of removing the dangles, which are now no longer due to the digitising errors, but are purely a result of machine transformation of the coverage. Even the application of a *dangle* tolerance of 6mm (figure 5.46) cannot eliminate the node errors induced by the inapplicability of the high *weed* tolerance.

### 5. 5. 3. Discussion.

It is the *weed* tolerance that determines the number of segments that are retained within the coverage, and this tolerance can be used to eliminate unnecessary detail from a coverage, thus reducing the amount of storage space dedicated to coverages. At the same time, however, if used to excess, this tolerance reduces the character of a coverage, and can fundamentally alter its geometry. As it joins two segments that are within a set tolerance of each other, it can also eliminate gaps by causing segments to collapse onto each other. Paradoxically, however, this process can also have the opposite effect, as sometimes segments can be eliminated which does not allow their remaining counterparts to be collapsed onto themselves and consequently the same process can create gaps. This is evident in figure 5.46, where the *dangle* nodes at the bottom of the river feature have been removed by collapsing and yet a new dangle has been created in the extreme west of the coverage as a segment has been removed.

*Dangle* node tolerances generally act in conjunction with other tolerances. In isolation, the *dangle* tolerance is only suitable for eliminating *dangle* nodes that fall within its bounds. Gaps that have developed as a result of operator undershoot remain unaddressed by large *dangle* tolerances. Figure 5.30 illustrates this point, as here a large *dangle* tolerance of 6mm has been employed, and yet 14 node errors still remain. This coverage also highlights the drawback of implementing high *dangle* tolerances, as when small arcs fall within the tolerance value they are eliminated. This situation can be observed in the absence of several small arcs in the eastern portion of the coverage. This scenario in turn, results in further errors as arcs become unattached from their neighbours and more 'dangles' are created.

However, when used in association with a *match node* tolerance, or *weed* tolerance or both, this operation is useful in identifying and eliminating unintended overshoots.

The *match node* tolerance snaps together nodes that are within a certain width apart. In this respect it is useful for eradicating inadvertent gaps that have arisen through inaccurate digitising. The *match node* tolerance can be regarded as the antithesis of the *dangle* tolerance, as each deals with the errors unaddressed by the other. If the *dangle* and *weed* tolerances are to be used effectively, a suitable *match node* tolerance is usually a perquisite, as this then allows the other two operations to simply deal with errors within their realm, rather than being artificially augmented to cope with arcs that do not coincide as they should. However, in many respects, high match tolerances are as dangerous as over exaggerated *weed* tolerances, since coverage manipulation can very easily occur without the user being aware of it. Even *match node* tolerances that are slightly too large for the coverage can culminate in collapses that leave the coverage littered with small triangles which then require elimination as areal units.

Achieving an ideal combination of tolerance values for a coverage is often a result of an extended trial and error procedure. The aim is to make the coverage more accurate, by removing the 'errors' that exist in the digitised version. What actually happens though is that the procedures are often taken to their extremes in an attempt to obliterate all errors. Consequently, coverages end up becoming more inaccurate in their entirety, rather than just being inaccurate at the arc endpoints. As a minimum insurance policy against adopting an inappropriate tolerance value for any of the operations, a modicum of testing should be undertaken. This should involve the analysis of the geometrical changes that have been incurred by the coverage, as well as the number of errors that have been removed. As this chapter has highlighted, many combinations of coverage geometry can produce a similar statistical description and also different measures of each tolerance value can produce very similar coverages.

The extent to which the accuracy of the original coverage was maintained, would ultimately rest with the user's ability to assess the nature of the data, the scale at which the data were digitised and the potential use of the material. Data that are known to have been dubiously collected, or for which the scale is very coarse is likely to have a high error component entrenched within it, and meticulous care at this stage, may thus be viewed as a fruitless effort. Furthermore, if the eventual use of the data is to be directed towards large scale investigations, in which the accuracy of the conclusions can accommodate high error margins, fastidious attention to detail may be a wasted effort. However the contrary is true for small scale planning investigations, where even if data is digitised from a 1:10,000 scale map tolerances of 3mm will mean movements of 30 metres on the ground.

**5. 6. An evaluation of the application of tolerances to a non-controlled coverage in ARC/INFO.**

The previous experiment looked at the detrimental effects of the tolerance values on a test coverage that was artificially created to monitor the visually changing character of the lines. In the following experiment, the effects of the tolerances are also investigated, but the emphasis is on looking at the changes numerically, rather than visually. In addition, it is the mechanics of the process that are investigated, rather than simply their effects. The extent to which the location of the lines change, and how their length alters, are two fundamental questions which will be addressed.

The aims of the following set of experiments were to investigate the effects of a variety of tolerance combinations on a real-world coverage, in which the results would be relatively unforeseeable. In the previous experiment, the results were largely predictable, as particular features of high potential errors were introduced, and their subsequent response to the different tolerance combinations could be easily monitored. This enabled the topology to be consistently verified. In the real world, however, although certain sinuous features or small areas might be pinpointed as plausible problem zones, the majority of the coverage may undergo very little positional or topological checking, short of a cursory glance, or comparison with the original (although this might pose problems if the output is to be retained at a different scale).[14] In some situations however, even visual comparative techniques may be inappropriate, as the problems may lie not so much in the visual

[14]See Chrisman 1987b.

appearances of the coverage, but in the topological changes that have taken place within the coverage to accommodate the movement.

An Ordnance Survey map of the Fenlands (sheets TF12 SE and TF12 SW) at the 1:10,560 scale was digitised and a subset of this was selected for use in the following experiments (figure 5.47). Selecting an area smaller than a full OS sheet enabled faster processing and a smaller burden on storage space. The Fenlands was regarded as suitable for this project because the features depicted within the area consisted of mainly straight lines representing drainage ditches and field boundaries. As straight lines consist of a connection between two points, then any movement at the endpoints will affect the geometric position of the feature itself more than if a large number of sinuous lines were used. In this way this type of coverage clearly exposes the shortcomings of the routines.

The coverage was digitised using an A0 Summagraphics digitiser and the ARC/INFO digitising package with accuracy being the major digitising goal. The coverage was then subjected to a variety of tolerance techniques using PC ARC/INFO. The objectives of the experiment was to remove as many of the node errors as possible whilst maximising the coverage topology. Therefore of primary interest was the responding changes in the lengths of lines that made up the coverage, to the various tolerances used. To facilitate the observation of such modifications subsequent to the application of various error removal techniques, each of the lines within the coverage was individually labelled, allowing them to be 'tracked'.

Coverage units were maintained for map preparation in preference to real world units as these were regarded as more cognitively accessible. This was due to the fact that although the maps might have been traced and digitised from different scale maps, the error margins are often the same in terms of coverage units. Thus despite changes in the lengths of lines that arise due the level of generalisation associated with maps of different scales, the distances of overshoots or undershoots are likely to be twice as large on the ground for a

1:25,000 scale map as for a 1:50,000 scale map.[15] Once obtained, the values associated with coverage distortion can be translated into real world units for a variety of scales as an indication of how such processes move features on the ground.[16]

Initially, the user has only a limited idea as to how many node errors may be removed by the application of a particular tolerance. An initial test was carried out to assess the values (in inches) that were likely to prove effective in removing node errors.[17] In this case, as the results of a variety of tolerances was required, tolerances were applied on the basis of arithmetic progression in the range of 0.001in to 02in. Once these were applied then a group was selected on the basis of the number of node errors that the tolerance removed. Tolerances that provided the widest spread of node errors removed were finally selected.

Match tolerances were the first to be applied, and these ranged from 0.001 inches to 0.09 inches. The number of errors each of these processes removed is shown in table 5.6 and are illustrated in figure 5.48.

---

[15] The generalisation procedures associated with changes in scale will however reduce the amount of error incurred when larger scales are used, and increase it when smaller scales are employed.

[16] In a real digitising project, tolerances that are to be used should always be matched up with their real world values, so that an assessment can be made as to whether the distances they represent are indeed acceptable.

[17] This can be done by selecting the distance command in ARCEDIT and measuring the distances needed to overcome a selection of node errors. Alternatively it can be set to be below the smallest line or gap the user wishes to retain.

Table 5.6 Number of dangles removed from the real world coverage by the first set of match node tolerances.

| Name | Match tolerance (inches) | Number of dangles |
|---|---|---|
| FTEST | NA | 250 |
| T009 | 0.009 | 209 |
| T01 | 0.01 | 198 |
| T02 | 0.02 | 68 |
| T03 | 0.03 | 18 |
| T04 | 0.04 | 11 |
| T05 | 0.05 | 10 |
| T06 | 0.06 | 11 |
| T07 | 0.07 | 11 |
| T08 | 0.08 | 10 |
| T09 | 0.09 | 9 |

As the match tolerance is increased there is a drop in the number of node errors retained, although this is by no means proportional. There are two break points that occur in this distribution. Firstly there is one between 0.01in and 0.02in. Here the drop in node errors is quite dramatic, decreasing from 198 to 68. Secondly, between 0.02in and 0.03in, there is a substantial drop in the number of node errors sustained, although this is less than that encountered in the previous case. The remaining tolerance values are unable to affect the values in such a striking manner. As a consequence of this preliminary selection of *match node* tolerances, four more were experimented with to see if an improved representation of the variation of error distribution with respect to changing match tolerances could be distinguished. This was a trial and error operation and the results are shown in table 5.7.

Table 5.7 Number of dangles removed by second set of match tolerances.

| Name | Match tolerance (inches) | Number of dangles |
|---|---|---|
| T026 | 0.026 | 33 |
| T001 | 0.001 | 250 |
| T005 | 0.005 | 231 |
| T015 | 0.015 | 120 |

A subset of these tolerances were then selected as being suitable to proceed with. The criteria for selection was based on obtaining as wide a spread as possible in the distribution of the reduction of node errors. The subset of *match node* tolerances selected to be used for further analysis were those of

- 0.026in, 0.03in, 0.04in 0.09in and 0.009in.

The next tolerance to be applied to each of the *match node*d coverages was the *dangle* tolerance. This was also approached on a trial and error basis, as the only gauge to build the selected *dangle* lengths on, were a random sample of the lengths of the dangles themselves. Dangle tolerances renaged from 0.001in to 0.2in.

## 5. 7. Analysing the change in coverage features.

As coverages undergo software post processing, they experience changes in the form of node error removal. This can be viewed as a positive effect of the 'cleaning' process (and for many users, the only consequence of the procedure).[18] Naturally, the larger the number of errors removed, the better, or so the software implies. Unfortunately, there is also a negative repercussion of this 'error removal'. As the software strives to eliminate node errors and squeeze features into an acceptable form, the entire coverage changes geometrically and topologically in a manner that is unknown and unpredictable, and therefore beyond the user's control. Coverage change arising as a result of tolerances used can therefore be evaluated in terms of both its positive and negative aspects. Positive change can be equated with the removal of errors, and negative change can be associated

[18] ARC/INFO training, and the training on many other systems rarely approach this issues. Furthermore some users receive little or no training or backup, due to the expense of such courses or links.

Table 5.8 Tolerances used on the test coverages and the number of errors remaining.

| COVERAGE | NODEERROR | MNODE | DANGLE | WEED |
|---|---|---|---|---|
| T09C2W1 | 2 | 0.090 | 0.200 | 0.100 |
| T03C03W1 | 2 | 0.030 | 0.030 | 0.100 |
| T009C2W1 | 3 | 0.009 | 0.200 | 0.100 |
| T09C2 | 7 | 0.090 | 0.200 | 0.002 |
| T009C2 | 7 | 0.009 | 0.200 | 0.002 |
| T04C2 | 8 | 0.040 | 0.200 | 0.002 |
| T09 | 9 | 0.090 | 0.000 | 0.002 |
| T04C09 | 9 | 0.040 | 0.090 | 0.002 |
| T05 | 10 | 0.050 | 0.000 | 0.002 |
| T08 | 10 | 0.080 | 0.000 | 0.002 |
| T04C05 | 10 | 0.040 | 0.050 | 0.002 |
| T024C03 | 10 | 0.024 | 0.030 | 0.002 |
| T06 | 11 | 0.060 | 0.000 | 0.002 |
| T07 | 11 | 0.070 | 0.000 | 0.002 |
| T04 | 11 | 0.040 | 0.000 | 0.002 |
| T03C2 | 12 | 0.030 | 0.200 | 0.002 |
| T09C03 | 12 | 0.090 | 0.030 | 0.002 |
| T03C09 | 15 | 0.030 | 0.090 | 0.002 |
| T03C03 | 16 | 0.030 | 0.030 | 0.002 |
| T03C05 | 16 | 0.030 | 0.050 | 0.002 |
| T04C02 | 16 | 0.040 | 0.020 | 0.002 |
| T04C01 | 16 | 0.040 | 0.010 | 0.002 |
| T04C009 | 17 | 0.040 | 0.009 | 0.002 |
| T03C02 | 18 | 0.030 | 0.020 | 0.002 |
| T03 | 18 | 0.030 | 0.000 | 0.002 |
| T03C01 | 22 | 0.030 | 0.010 | 0.002 |
| T03009 | 22 | 0.030 | 0.009 | 0.002 |
| T026C2 | 23 | 0.026 | 0.200 | 0.002 |
| T026C09 | 27 | 0.026 | 0.090 | 0.002 |
| T026C05 | 28 | 0.026 | 0.050 | 0.002 |
| T026C03 | 28 | 0.026 | 0.030 | 0.002 |
| T026C02 | 31 | 0.026 | 0.020 | 0.002 |
| T026 | 33 | 0.026 | 0.000 | 0.002 |
| T026C009 | 37 | 0.026 | 0.009 | 0.002 |
| T026C01 | 37 | 0.026 | 0.010 | 0.002 |
| T02 | 68 | 0.020 | 0.000 | 0.002 |
| T015 | 120 | 0.015 | 0.000 | 0.002 |
| T009C03 | 146 | 0.009 | 0.030 | 0.002 |
| T01 | 198 | 0.010 | 0.000 | 0.002 |
| T009 | 209 | 0.009 | 0.000 | 0.002 |
| T005 | 231 | 0.005 | 0.000 | 0.002 |
| T001 | 250 | 0.001 | 0.000 | 0.002 |
| FENS1AAT | 250 | 0.000 | 0.000 | 0.002 |

with undesirable feature movement. Thus to assess the effects of the tolerancing routines, an evaluation of the number, and position of lines in the coverage, with respect to the number of dangles removed needs to be undertaken.

**5. 7. 1. Positive change - node error removal.**

As expected, the increase in tolerance values led to a corresponding decrease in the number of *dangle* nodes and also to a decrease in the number of lines associated with each coverage. Figure 5.49 illustrates the *match node* tolerances used and the ensuing drop in the number of errors associated with each coverage. After a point however, the number of errors that can be removed by a match tolerance alone begins to level off. Not all errors arise as a result of undershoots; many are due to overshooting the intended end point. Such errors, although they can be removed by an increase in the *match node* tolerance (at the cost of extensive coverage depreciation), are best dealt with by a *dangle* tolerance. In most cases, this additional tolerance has further reduced the coverage error associated with particular *match node* tolerances (table 5.8).

In some cases however, there has been an increase in the number of *dangle* nodes, notably in coverages that have been subjected to a *dangle* tolerance of 0.01in, 0.02in and 0.009in. The reason for this seeming inverted state of affairs, is that once the dangle removal routine has performed the initial function of intersecting all the overlapping lines, many of the newly constituted small lines are actually too large to be removed by such small *dangle* tolerances (figure 5.50). As a consequence, the lines remain as dangles, thus giving the impression that there are more errors in the coverage than prior to the *dangle* removal operation.

Table (5.8) and figure 5.51 show the number of errors associated with each of the coverages ranked from the lowest to the highest. Four major groups can be groups can be discerned. Cases 1-3 which consist of those coverages which have undergone *weed* tolerancing of one form or another. The second group consisting of cases 4-17, is made up of:

- high match node tolerance
- high match node tolerance with medium *dangle* tolerance
- high match node tolerance with high *dangle* tolerance
- Most of the medium match node tolerances (irrespective of *dangle* tolerance)
- High *dangle* tolerance when combined with certain match node tolerances.

The third group consisting of cases 18-35 had been subjected to the following:

- Medium match tolerances with a variety of *dangle* tolerances
- Medium match tolerances,

whilst the last group (cases 36-42) contained those coverages with a low match tolerance. In many respects this is an expected state of affairs as it is the destructive *weed* tolerance which is removing the majority of the 'errors'. This is followed by high *match* and (or) *dangle* tolerances, with the low *match* tolerances having little or no effect, even when combined with a medium *dangle* tolerance. Thus one could conclude that error removal is most effectively performed by the *weed* tolerance routine, followed by the *match* and *dangle* routines. For the *dangle* tolerance to be effective in error reduction it needs to be combined with a reasonable *match node* tolerance, otherwise it has to be undesirably high. So much for the desired effect of these routines, what of their consequences in terms of topological modification?

### 5. 7. 2. Negative change - topological modification.

Section 5.5 looked at geometric change visually, and highlighted the types of features that were most prone to geometric change and the resulting feature shapes following processing. This section seeks to look more at the empirical value of the lines and in addition does not make any apriori assumptions about the susceptibility of some features to change. Furthermore, since the test coverage consists primarily of straight lines, changes affecting the location of the lines themselves will be initiated from the endpoints. In the first instance the distribution of the length of lines in each of the coverages following processing can be analysed and compared to the original digitised coverage prior to 'cleaning'.

- **Variations in the number of lines per coverage.**

As could be predicted, the total number of lines held within the coverages varied with respect to the size and types of tolerance used on the coverage. Figure 5.52 and table 5.9 illustrate the variations in total line length.

**Table 5.9 Tolerances used on the test coverages and the number of lines remaining**

| COVERAGE | TOTLINES | NODEERROR | MNODE | DANGLE | WEED |
|---|---|---|---|---|---|
| FENS1AAT | 827 | 250 | 0.000 | 0.000 | 0.002 |
| T001 | 827 | 250 | 0.001 | 0.000 | 0.002 |
| T005 | 782 | 231 | 0.005 | 0.000 | 0.002 |
| T009 | 756 | 209 | 0.009 | 0.000 | 0.002 |
| T01 | 748 | 198 | 0.010 | 0.000 | 0.002 |
| T015 | 726 | 120 | 0.015 | 0.000 | 0.002 |
| T02 | 712 | 68 | 0.020 | 0.000 | 0.002 |
| T05 | 697 | 10 | 0.050 | 0.000 | 0.002 |
| T06 | 697 | 11 | 0.060 | 0.000 | 0.002 |
| T07 | 696 | 11 | 0.070 | 0.000 | 0.002 |
| T08 | 693 | 10 | 0.080 | 0.000 | 0.002 |
| T09 | 690 | 9 | 0.090 | 0.000 | 0.002 |
| T09C2W1 | 654 | 2 | 0.090 | 0.200 | 0.100 |
| T03C03W1 | 656 | 2 | 0.030 | 0.030 | 0.100 |
| T009C2W1 | 658 | 3 | 0.009 | 0.200 | 0.100 |
| T026C2 | 763 | 23 | 0.026 | 0.200 | 0.002 |
| T09C2 | 764 | 7 | 0.090 | 0.200 | 0.002 |
| T026C09 | 769 | 27 | 0.026 | 0.090 | 0.002 |
| T03C2 | 769 | 12 | 0.030 | 0.200 | 0.002 |
| T009C2 | 770 | 7 | 0.009 | 0.200 | 0.002 |
| T04C2 | 770 | 8 | 0.040 | 0.200 | 0.002 |
| T09C03 | 770 | 12 | 0.090 | 0.030 | 0.002 |
| T026C03 | 771 | 28 | 0.026 | 0.030 | 0.002 |
| T026C05 | 771 | 28 | 0.026 | 0.050 | 0.002 |
| T026C02 | 774 | 31 | 0.026 | 0.020 | 0.002 |
| T03C09 | 774 | 15 | 0.030 | 0.090 | 0.002 |
| T04C09 | 774 | 9 | 0.040 | 0.090 | 0.002 |
| T024C03 | 776 | 10 | 0.024 | 0.030 | 0.002 |
| T03C03 | 776 | 16 | 0.030 | 0.030 | 0.002 |
| T03C05 | 776 | 16 | 0.030 | 0.050 | 0.002 |
| T04C05 | 776 | 10 | 0.040 | 0.050 | 0.002 |
| T03C02 | 778 | 18 | 0.030 | 0.020 | 0.002 |
| T04C02 | 778 | 16 | 0.040 | 0.020 | 0.002 |
| T026C01 | 780 | 37 | 0.026 | 0.010 | 0.002 |
| T026C009 | 780 | 37 | 0.026 | 0.009 | 0.002 |
| T03009 | 782 | 22 | 0.030 | 0.009 | 0.002 |
| T03C01 | 782 | 22 | 0.030 | 0.010 | 0.002 |
| T04C01 | 782 | 16 | 0.040 | 0.010 | 0.002 |
| T04C009 | 783 | 17 | 0.040 | 0.009 | 0.002 |
| T009C03 | 791 | 146 | 0.009 | 0.030 | 0.002 |
| T04 | 821 | 11 | 0.040 | 0.000 | 0.002 |
| T026 | 827 | 33 | 0.026 | 0.000 | 0.002 |
| T03 | 827 | 18 | 0.030 | 0.000 | 0.002 |

The tolerances that resulted in the largest loss of lines were those that had *weed* tolerances applied to them. For example a *weed* tolerance of 0.01in was used on three coverages:

- one which had a small *match node* (0.009in) and a large *dangle* tolerance (0.2in);
- another with a moderate *match node* tolerance (0.03in) and a moderate *dangle* tolerance of (0.03in)
- and finally one with both a high match tolerance (0.09in) and a high *dangle* tolerance (0.2in).

all of which had a loss of approximately 170 lines as compared to the original coverage. This totalled approximately 20% of the coverage features (Figure 5.53).

Following these three coverages in the ranking, are the coverages that have been *match node*d with a high to medium tolerance, and then those *match node*d with a low tolerance. These lose approximately 15% of their features by retaining more or less 700 lines. The remainder of the coverages, (with the exception of the T001 coverage which was too small to extend any influence at all), retained a total number of lines which ranged from 763 to 791, (approximately 90% of the total number of lines) with the high match and or *dangle* tolerances losing more than the low match, low *dangle* combinations.

What is of interest however is how the lines retained in the coverages correspond to the number of errors their implementation has removed. The correlation between the two variables is shown in figure 5.54. In terms of match tolerance alone, the total number of errors retained, with respect to the number of lines removed increases proportionally. Thus as match tolerance increases, the number of errors decrease, as do the number of lines retained. Of the other tolerances used, the *weed* proves to be the most drastic, with the removal of error and lines being great. The combination tolerances however bunch together to form a small cluster. The error reduction is generally good, with a greater number of lines retained than in the match tolerances alone. The lower the *dangle* tolerances applied to these coverages, the greater the number of lines retained. Of the cluster, the medium match tolerance/low *dangle* tolerance seem to retain a greater number of lines for a similar amount of errors reduction achieved by the other tolerances. Following this tolerance group are:

- medium match tolerance/medium *dangle*

- medium match tolerance/high *dangle*

Thus it can be inferred that the *weed* tolerances, though effective in error removal, can drastically alter the coverage topology. Match tolerances seem to be very effective, but it must be noted that the error results for the match tolerance values apply to the original coverage only, and if the toleranced coverages were to be queried following the insertion of topology, the error total would be higher. Coverages which incorporate *dangle* tolerances are automatically subjected to the building of the data model; a process causing the creation of further errors and the definition of more lines.

- **Changes in line distribution.**

As well as investigating the changing total number of lines that are held within each of the coverages that have been toleranced, the distributions of the line lengths in each of the coverages were also analysed. The ARC/INFO database was interrogated for the number of lines in discrete categories. These were:

- less than 0.009in,
- less than 0.01in, but greater than (or equal to) 0.009in
- less than 0.02in, but greater than (or equal to) 0.01in
- less than 0.03in, but greater than (or equal to) 0.02in
- less than 0.05in, but greater than (or equal to) 0.03in
- less than 0.09in, but greater than (or equal to) 0.05in
- less than 0.2in, but greater than (or equal to) 0.09in
- greater than (or equal to) 0.2in.

These were chosen as they responded to the tolerance thresholds that had been used. The results are displayed in table 5.10. Figure 5.55 illustrates the distibution of lines in these categories with the exception of the last one, as this would have blurred the detail of the smaller lines. The coverages submitted to a *weed* tolerance loose all their small components, with the other coverages varying depending on whether the category is greater or less than their tolerance thresholds. The categories are shown in figures 5.56 to 5.63 with respect to the number of errors each of the tolerances represented have removed.In general there is a loose positive relationship between the number of errors removed and the number of lines retained. In terms of the relationship between the number of lines retained in each category, and the type and size of tolerance used; combinations can be ranked in the following order on the basis of most lines and most errors to least lines and least errors:

| No. errors | Total No. of lines | gt 0.2in | le0.2in gt0.09in | le0.09in gt0.05in | le0.05in gt0.03in | le0.03in gt0.02in | le0.02in gt0.01in | gt0.01in le0.009in | le0.009in | Cov. name |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 656 | 602 | 54 | 0 | 0 | 0 | 0 | 0 | 0 | T03C03W1 |
| 2 | 654 | 602 | 52 | 0 | 0 | 0 | 0 | 0 | 0 | T09C2W1 |
| 3 | 658 | 600 | 58 | 0 | 0 | 0 | 0 | 0 | 0 | T009C2W1 |
| 7 | 764 | 618 | 44 | 5 | 9 | 13 | 26 | 4 | 45 | T09C2 |
| 7 | 770 | 620 | 32 | 10 | 5 | 15 | 48 | 5 | 35 | T009C2 |
| 8 | 770 | 617 | 42 | 12 | 9 | 14 | 27 | 4 | 45 | T04C2 |
| 9 | 774 | 617 | 46 | 12 | 9 | 14 | 27 | 4 | 45 | T04C09 |
| 9 | 690 | 621 | 52 | 0 | 2 | 4 | 8 | 0 | 3 | T09 |
| 10 | 776 | 617 | 46 | 14 | 9 | 14 | 27 | 4 | 45 | T04C03 |
| 10 | 776 | 617 | 46 | 14 | 9 | 14 | 27 | 4 | 45 | T04C05 |
| 10 | 697 | 620 | 47 | 15 | 2 | 3 | 8 | 0 | 2 | T05 |
| 10 | 693 | 621 | 52 | 3 | 2 | 4 | 8 | 0 | 3 | T08 |
| 11 | 697 | 620 | 47 | 16 | 2 | 3 | 7 | 0 | 2 | T04 |
| 11 | 697 | 620 | 47 | 14 | 2 | 3 | 8 | 0 | 3 | T06 |
| 11 | 696 | 620 | 50 | 10 | 2 | 3 | 8 | 0 | 3 | T07 |
| 12 | 769 | 617 | 41 | 12 | 11 | 12 | 28 | 3 | 45 | T03C2 |
| 12 | 770 | 618 | 46 | 9 | 9 | 13 | 26 | 4 | 45 | T09C03 |
| 15 | 774 | 617 | 46 | 12 | 11 | 12 | 28 | 3 | 45 | T03C09 |
| 16 | 776 | 617 | 46 | 14 | 11 | 12 | 28 | 3 | 45 | T03C03 |
| 16 | 776 | 617 | 46 | 14 | 11 | 12 | 28 | 3 | 45 | T03C05 |
| 16 | 782 | 617 | 46 | 14 | 9 | 16 | 31 | 4 | 45 | T04C01 |
| 16 | 778 | 617 | 46 | 14 | 9 | 16 | 27 | 4 | 45 | T04C02 |
| 17 | 783 | 617 | 46 | 14 | 9 | 16 | 31 | 5 | 45 | T04C009 |
| 18 | 778 | 617 | 46 | 14 | 11 | 14 | 28 | 3 | 45 | T03C02 |
| 18 | 700 | 620 | 47 | 16 | 5 | 3 | 7 | 0 | 2 | T03 |
| 22 | 782 | 617 | 46 | 14 | 11 | 14 | 32 | 3 | 45 | T03009 |
| 22 | 782 | 617 | 46 | 14 | 11 | 14 | 32 | 3 | 45 | T03C01 |
| 23 | 763 | 617 | 40 | 12 | 10 | 16 | 27 | 3 | 38 | T026C2 |
| 27 | 769 | 617 | 46 | 12 | 10 | 16 | 27 | 3 | 38 | T026C09 |
| 28 | 771 | 617 | 46 | 14 | 10 | 16 | 27 | 3 | 38 | T026C03 |
| 28 | 771 | 617 | 46 | 14 | 10 | 16 | 27 | 3 | 38 | T026C05 |
| 31 | 774 | 617 | 46 | 14 | 10 | 19 | 27 | 3 | 38 | T026C02 |
| 33 | 705 | 620 | 47 | 16 | 5 | 8 | 7 | 0 | 2 | T026 |
| 37 | 780 | 617 | 46 | 14 | 10 | 19 | 33 | 3 | 38 | T026C01 |
| 37 | 780 | 617 | 46 | 14 | 10 | 19 | 33 | 3 | 38 | T026C009 |
| 68 | 712 | 620 | 46 | 17 | 5 | 15 | 7 | 0 | 2 | T02 |
| 120 | 726 | 622 | 44 | 17 | 3 | 15 | 22 | 0 | 3 | T015 |
| 146 | 791 | 620 | 45 | 17 | 6 | 15 | 48 | 5 | 35 | T009C03 |
| 198 | 748 | 622 | 44 | 17 | 3 | 15 | 45 | 0 | 2 | T01 |
| 209 | 756 | 622 | 44 | 17 | 3 | 15 | 45 | 8 | 2 | T009 |
| 231 | 782 | 622 | 44 | 17 | 3 | 14 | 47 | 8 | 27 | T005 |
| 250 | 827 | 622 | 44 | 17 | 3 | 12 | 49 | 7 | 73 | T001 |
| 250 | 827 | 622 | 44 | 17 | 3 | 12 | 49 | 7 | 73 | FENS1AAT |

Table 5.10 Lines of varying lengths present within the test coverages

- medium match tolerance/low *dangle* tolerance
- medium match tolerance/medium *dangle* tolerance
- medium match tolerance/high *dangle* tolerance
- high *match node* tolerance/medium *dangle* tolerance
- high *match node* tolerance/medium *dangle* tolerance.

Therefore it would appear that the number of lines varies with the number of errors removed and using this type of basic analysis could allow the user to define a threshold value on the basis of lines of a certain width that the user wishes to preserve. An error removal threshold could be set as the minimum requirement for a blanket routine. The corresponding loss of lines which ensue as a consequence of this desired figure after the application of a variety of tolerance routines could then be assessed, and a suitable selection made

The problem faced by this method of analysis is that the user is unable to discern between the unwanted lines that are being removed and the desired lines which are also eliminated. Furthermore, there is the problem of line creation, which is carried out as a function of the data model preparation routine. For further analysis some tracking mechanism is necessary, to ensure that the same lines are always being compared.

**5. 8. The need for more detailed analysis.**

So far, particular line lengths in the test coverages have been analysed in an attempt to assess whether the distribution of a variety of line lengths in the test coverages varied from those in the original, and in addition some attempt has been made to hypothesise the extent and causes of the variation. However, the information obtained so far cannot provide statistical details, nor can any lineage of the lines themselves be discerned. An important question is how each line varies from its original forefather. In an attempt to try and address these questions, a subset of the coverages used in the previous selection were selected for further analysis.

These coverages were selected on the following basis. Firstly the two extremes in terms of match tolerance were selected; 0.009in and 0.09in. Using these as a base, coverages were selected that had a high *dangle* tolerance of 0.2in, and a relatively low *dangle* tolerance of 0.03in. Furthermore, to further investigate the effects of the *weed* tolerance, coverages with *weed* tolerances of 0.1in for both the extremes were selected. In addition a mediocre match tolerance of 0.026in with a low *dangle* tolerance of 0.03in and a high *dangle* tolerance of 0.2in were also tested. It was hoped that this spread would represent both the best and worst possible case in terms of coverage manipulation as well as a potential compromise.

On the basis of what has been investigated already, it was hypothesised that the *weed* tolerances would perhaps dominate the movement of the topology, and that therefore the coverages with this tolerance applied to them would perhaps undergo the most radical transformation. Following this, the coverages with the higher match tolerances would perhaps be the next in line for coverage modification. Using these assumptions a hierarchy was derived as shown in table 5.11.

Table 5.11. Expected hierarchy of coverages

| | | Fuzzy tolerance (in) | Matched tolerance (in) | Dangle tolerance (in) |
|---|---|---|---|---|
| Least affected | T009C03 | 0.002 | 0.009 | 0.03 |
| | T009C2 | 0.002 | 0.009 | 0.2 |
| | T026C03 | 0.002 | 0.260 | 0.03 |
| | T026C2 | 0.002 | 0.260 | 0.2 |
| | T09C03 | 0.002 | 0.090 | 0.03 |
| | T09C2 | 0.002 | 0.090 | 0.2 |
| | T009C2W1 | 0.100 | 0.009 | 0.2 |
| Most affected | T09C2W1 | 0.100 | 0.090 | 0.2 |

On a more pragmatic note, it was also hoped to achieve a better understanding of which tolerances might provide a compromise between the elimination of as many digitising errors as were feasible, within the constraint of maintaining as much of the original coverage topology as possible. In this respect some sort of threshold value was desired, beyond which tolerance routines are no longer effective in terms of the negative changes being wrought upon the coverage.

The data pertaining to the lengths of the lines of the coverages in question and their individual identifiers were downloaded from the ARC/INFO package and the PC environment, and were mounted on a mainframe machine. The data was rewritten into a suitable format using Pascal programs (appendix 5.1), and a preliminary set of descriptive statistical procedures were applied to the data using both SPSSX and MINITAB.

The summary statistics of each of the coverages are presented with their histograms in figures 5.64 to 5.71. The smaller range and sum associated with the two coverages that have been subjected to *weed* tolerancing, are a preliminary indication of the difference between these and the other coverages as a consequence of the reduction of the smaller line lengths. This is further reinforced by the minimum values of line length for these coverages which equal 1.01 and 1.04 inches. The rest of the coverages display similar values, although the sum of the total line length is consistently less in the coverages with the higher *dangle* tolerance. Maximum values are all uniform with the exception of the T009c2w1 coverage, which is marginally higher. All of the toleranced coverages do however, display a smaller maximum value than the original coverage, a fact that can perhaps be accounted for by the intersecting and consequential shortening of crossing lines following digitising.

The mean, mode and median values for these coverages also indicate a high degree of dissimilarity between them and the other coverages, as well as between them and the original. For example, the mean value of the lines is almost 0.1in greater than the original and up to 0.63in greater than those of the other coverages. Of the other coverages, those with means that most closely approximate the original were coverages with the smaller *match* and *dangle* tolerances. The same is apparent in the median values, but in the mode value, the two coverages with the small match tolerance of 0.009in have a higher value than the rest of the non *weed* toleranced coverages, which can perhaps be attributed to this tolerance value leaving more of the larger lines than its associates. The coverages with *weed* tolerances have a smaller standard deviation than the other coverages, which reflects the fact that they are more compact and have fewer outliers than the others (figure 5.70 and 5.71).

Each of the coverages should display a similar distribution as they are derived from the same initial coverage and the same data. If the data distribution is different, this implies that the dataset has mutated quite significantly.

The use of statistical methods for testing the significance of the difference between the distribution of the original coverage and those that have been subjected to the tolerance operations, would indicate the degree to which the two populations corresponded. However, a result that was not statistically significant would not necessarily imply that the differences in the coverage were significant for a particular user at a particular time, or alternatively for a particular feature at a particular place.

In order to perform classical tests of significance, the distribution of the data needs to approach a normal or Guassian curve. Unfortunately, the distributions of the original coverage and some of the test coverages (notably the ones with low *match node* tolerances), do not conform to such a distribution, and are unable to be transformed onto it due to the bimodal nature of the data. However, the coverages with a tolerance greater than or equal to 0.026in were amenable to a square root transformation, and so the test was performed on the t026c2 coverage and the t09c2w1 coverage (figures 5.72 and 5.73). As both of the coverages were derived from the same distribution, one would assume that there should be no significant difference between them.

Therefore the null hypothesis (Ho) is that there is no significant difference between the coverage t026c2 and t09c2w1. The alternative hypothesis (Hi), therefore becomes that there is a significant difference between the datasets. This is two directional and does not assume that one will be greater than the other. The rejection level or significance level has been set to 0.05 (5%). If the t-test[19] is then applied to the two coverages, the result is

[19] The t-test being defined as the mean of dataset x minus the mean of dataset y, all divided by the standard error of the mean of data set x minus the standard error of the mean of dataset y.

$$\frac{\bar{x}-\bar{y}}{SE\bar{x}-SE\bar{y}}=t$$

$$\frac{0.741-0.807}{0.016}=4.125=t$$

This gives a probability, or p value of less than 0.01. Therefore at the 5% significance level, the null hypothesis, Ho, will have to be rejected, implying that the coverages are statistically different. This can be further highlighted by looking at the distributions of the two coverages for the 99% confidence interval (figure 5.74).

In the case of these two coverages it is fair to assume that there must indeed be grave geometric changes committed, for such a statistically different sample. A similar difference can be observed if the distribution (also for 99% confidence) are examined for the t09c2 and t09c2w1 coverages (figure 5.74). Whereas if the t09c2 and t026c2 coverages are compared, their t-test results are seen to be almost identical.

It is unfortunate that parametric tests could not be performed on all the coverages with respect to the original, but from this small sample, differences obviously exist between the t09c2w1 coverage and two of the larger toleranced coverages, t09c2 and t026c2. From this it might be assumed that, since t009c2w1 is so similar in its distribution to t09c2w1, an analogous situation might exist for this coverage. Also since t09c2 and t026c2 are the larger of the non-*weed*ed tolerances, it could be assumed that if a significant difference existed between t09c2w1 and these two coverages, then the same might hold true for the other coverages.

There are non-parametric test that can be performed on data that does not conform to the normal distribution. In general however, these rely on the ranking of the data. One such test, the Wilcoxon test was attempted, but problems arose with the nature of the data set, as the robustness of the test decreases with the number of equal ranked values. As the original coverages had so many values that were equal to 0.002, the test was not deemed to be reliable for further analysis.

**5. 8. 1. Line by line comparison techniques.**

The statistical information outlined above shows that there are differing statistical characteristics associated with each coverage, but still gives few clues as to the changing nature of the lines themselves. In order to carry out a direct comparison of the lines generated as a result of the tolerances with those in the original coverage, an attempt was made to directly compare each of the individual lines in the generated coverage, with the lines that they were derived from in the original coverage[20].

- **Visualisation vs topology.**

Two changes take place when map data are stored in a GIS environment. Firstly, the lines themselves may be moved and might be altered positionally in terms of shape and length. Secondly, the relationships and topology necessary for any GIS analysis, require that the data be redefined. In this way the visualised line is no longer equal to the stored line.

To look at positional change visualised positions of the line need to be compared, rather than comparing visualised positions with the topologically created ones. If the latter is performed, then the only thing measured is the ability of the system to intersect and relabel lines. Of real interest however is the amount of positional accuracy that disappears when this procedure is enacted. If lines are entirely eliminated, then there is an obvious change in length and the position of surrounding lines will change to compensate. If however the lines are split, then can the length of the line be said to have changed, when in effect the line actually remains as part of the coverage? What has changed is the computerised information held on that line.

If the line is to be used in linear analysis, then linear attributes (such as length), which are attached to that line will have changed. User attributes (such as roads, rivers or rail links) attached to the line will however remain and be recognised as such. If however, the lines are to form the boundaries of areas, then any excess lines which have been intersected, may reduce the size of the polygon it is intending to represent. In figure 5.75 for example, lines a and b are supposed to intersect at point z, instead, they cross before that point and thereby reduce the size of area I.

---

[20]This was done using a series of Pascal programs, which are illustrated in appendix 5.1 and 5.2.

- **The dangers of false software induced intersections.**

To overcome such problems of visualisation lines are intersected as part of the *clean* process, thus ensuring that all features are correctly encoded. This also presents a problem, as features no longer retain their 'visualised' relationship. Instead they incur a new set of topological labels as defined by the software. These may result in the lengths of lines becoming shorter or longer, thereby altering the length of the represented feature.

In addition to necessary intersections however, unintentional ones may also occur as a result of poor digitising. For example, consider a line that is to meet another in the form of a T-junction (figure 5.76), if the line that forms the horizontal part of the T is digitised slightly lower than its true position, the bisecting arc will be intersected and retained as two separate entities. If the small portion then falls below the *dangle* tolerance set, it will be eliminated entirely. If it falls below the tolerance, the intersecting component of the program will still diminish its total length as the smaller portion of the line will be given a completely new identifier and will retain no relationship to the attributes of the original line. This introduces problems of searching on the basis of lengths of lines and also makes the task of attribute coding more exacting.

- **The problems of line and polygon collapse.**

A more serious condition occurs when, as a consequence of the amelioration routines, lines are intersected and moved in such a way that causes them to collapse and to retain some, but not all of their original segments. Neither is this situation exclusive to line features. Small spurious polygons may also be created as a result of segment collapse, which may in turn lead to inaccurate results, should some form of area analysis be performed on the coverage. In both cases, features below a certain magnitude can be removed. Unfortunately, not only does this result in loss of intended features that also fall below this magnitude, but further coverage distortion may also arise.

- **The problems of relabelling and false indications of line lengths.**

When the user digitises the data, the length and shape of the line are perceived to be as close to the cartographic product from which it is being taken as possible. Naturally, however, errors arise along the path of that line. This not only affects the shape of the line, but can result in several possible endpoints occurring. Before the intervention of any error reduction processes all of these 'errors' or 'mistakes' are user- created and not software-induced. If however, any of the lines cross each other, the error reduction procedures in ARC/INFO will insert a node at that point and store the lines as four entities as opposed to two. Therefore as far as the software is concerned, the lengths that the user intended for the two lines will have been reduced, not necessarily due to the fact that the lines no longer exist in the coverage, but that they exist under a different label. With this fact in mind, assessments of the changing lengths of lines and their changing x,y locations cannot simply be made on the differing lengths of lines corresponding to the same label.

- **The need to isolate line intersecting and line removal techniques.**

As the process that deals with overshoot removal is also responsible for line relabelling, evaluating the effects of the line removal section requires that the two elements of the process be isolated. Therefore this study seeks to divorce the two components and to concentrate on the changing lengths of lines as a consequence of geometric movement within the coverage, be it node snapping or line dropping.

- **The problem of matching lines.**

This was not however the simple task that it first appeared. As has already been witnessed in previous analyses, the number of lines that exist in the toleranced coverage are different to the number originally held. This is due in part to the fact that *dangle arcs* have been removed, but it is also a function of line intersection as part of the *clean* process.[21] So a line that was once a single entity may be split into two or more lines. Thus in comparing the lengths of the lines in the processed coverages with their predecessors in the original coverage, a common indicator was created to ensure that the correct lines were being matched. Originally it was hoped to use the internal line identifier, but after initial tests it became apparent that these are not unique line indicators which remain tagged onto the line, rather, they are created as and when they are needed. New lines born of original lines therefore become very difficult to trace as new identifiers may even refer to lines that were previously deleted.

---

[21]This is a topological creation, rather than positional change.

To by-pass this problem, an additional identifier was added to each of the lines in the coverage. This was exempt from machine initiated change as whilst the *clean* process creates new internal identifiers, these will remain intact (figure 5.77). For comparison, however this proves to be problematic, as, if one wishes to compare the original with the generated coverage, which line does one take as being the same as that in the original? Furthermore, if one were to count only the part of the line remaining after portions of it had been clipped, only topological change and not positional movement would be accounted for. One would simply be monitoring the *clean* process itself and producing differences between the generated coverages and the original as a consequence of line identification, rather than differences due to variations in line length. Therefore, it was decided that if geometric change in the line length was to be assessed, then multiple identifiers should be added together to recreate the line as it existed in the original coverage. In this way, the only change monitored would be that occurring due to inadvertent coverage movement. Duplicate lines were not the only problem that was encountered in seeking to compare coverages. Some lines that existed in the original coverage were simply not present in the generated coverages. Identifiers not found in the generated coverages were therefore replaced by zeros (appendix 5.1).

- **Problems of defining positional difference.**

The presence of zeros highlighted further problems in analysing the differences between the coverages. If the total difference is taken each time, then no distinction is made between an increase in line length and a decrease in line length; instead every deviation from the original contributes to the accumulating discrepancy between the original and the toleranced coverage. This would be acceptable if one could be certain that once the difference was ascertained, it would not be compensated for under a different identifier. Unfortunately trials showed this not to be the case as the following example illustrates.

For example if a line under the identifier 17 is equal to 0.004in on the original coverage and on the test coverage it has a total length of 0.678in, (made up of both the 16 **and** 17 in the original coverage, figure 5.78); a comparison between the original id 17 and the test coverage id 17 will yield a result of 0.64in. This would imply that the test coverage had an increase in total line length of 0.674in. However only one line of length 0.674in exists in the test coverage and there has been no increase. This is affirmed when the original id 16 is compared with id 16 in the test coverage. As id 16 no longer exists in the test coverage a deficit of 0.674 arises. Summing both these values will reveal no change in coverage line lengths.

Under these circumstances if the difference in both cases were to be accumulated, then only database changes would be undergoing assessment. The following experiment tries to steer clear of such problems, and to concentrate merely on the geometrical change in the line length between the original coverage (irrespective of whether this was a true representation of reality) and the processed coverages created as a by-product of the implementation of the data model. Therefore it was decided to note whether differences were positive or negative as when the total difference is calculated lines which still existed would cancel each other out (appendix 5.1).

**5. 8. 2. Measuring the extent of geometric difference in the coverages.**

Data were downloaded from the PC environment and ordered on the basis of the unique identifier. Missing values were given a length value of zero, and each individual line was matched up with the line of the same identifier in the master file, thus giving 827 records in each case. Assessments of change were based on relative difference from the original, using a Pascal program (appendix 5.1). Lines which were longer than the original were given a negative value and lines which were shorter than the original were given a positive value. Thus lines which exist in both the test and master coverage, but under different indicators (due to topological redefinition), will not contribute to the cumulative difference between the two coverages. Conversely, they will counterbalance each other and result in a more normally distributed set of cases. Naturally, there may well be a situation whereby many line lengths are falling extremely short of the length of the lines in the original coverage due to error, and are being balanced by deviations in the opposite directions, but it is unlikely, especially if nearly perfect symmetrical distribution were to be maintained.

Missing lines indicate that some coverage features have been lost as a result of the preprocessing techniques. Some of the loss will be intentional, representing small *dangle* arcs that have overshot their intentional stopping point, other losses will be due to the relabelling processes, whilst some change will be unintentional. Examining the comparative loss between the test coverages gives an indication of the number of lines lost with respect to the number of node errors removed.

Table 5.12 illustrates the way in which the number of lines that are retained in the test coverages vary from the original digitised coverage.

Table 5.12 The number of lines and errors removed and the number of multiple lines, with respect to the size and type of tolerance used.

| | Lines Removed | Dangles Removed | No. of Multiple lines |
|---|---|---|---|
| T026C02 | 53 | 219 | 113 |
| T026C2 | 64 | 227 | 107 |
| T09C2 | 63 | 243 | 117 |
| T09C03 | 57 | 238 | 119 |
| T09C2W1 | 173 | 247 | 50 |
| T009C03 | 36 | 104 | 128 |
| T009C2 | 57 | 116 | 121 |
| T009C2W1 | 169 | 247 | 49 |
| T09 | 137 | 241 | |
| T026 | 122 | 217 | |
| T009 | 71 | 41 | |

As previously discussed, the coverages which show the most dramatic loss of lines are those submitted to a large *weed* tolerance. In both cases, over a fifth of the coverage features are no longer existent in their previous form, whilst the number of multiple lines in both coverages is relatively small in comparison to the other test coverages. This is probably due to the fact that the coverages with the more drastic tolerances are having a lot of the smaller lines created by the intersection routine removed and implies that there has been a notable change in the features held within the two coverages.

Such drastic measures have however, enabled the removal of 247 of the 250 *dangle* errors present in the original coverage.[22] This is commendable, but similar reductions can be made, without such a seemingly large distortion to the coverage. Consider for example the coverages that were not submitted to a *match node* tolerance of 0.009in. In these cases errors removed ranged between 219 and 243, whilst the number of lines no longer present in the coverage would appear to be smaller, and the geometrical change therefore less. The extent of change can only be sketched at this stage, and will be illustrated further in the following sections.

Of the other coverages, the numbers of lines removed varied between 64 and 36, with the greatest drop being associated with the larger *dangle* tolerances, although the magnitude of the match tolerance seemed to set the initial parameters. If table 5.12 is considered, it can be seen that although the t09 and t026 coverages remove a relatively similar number of lines, the smaller t009 tolerance is not quite so fierce in its reduction of coverage features.

Another feature which table 5.12 highlights is the drop in the number of lines removed once a *dangle* tolerance is added onto the match tolerance. In all cases the drop in the number of lines removed is quite significant. This drop is due to the intersection program incorporated in the *dangle* tolerance removal routine. As expected, the coverages with a match tolerance of 0.009in, have the larger number of multiple lines since too few have been snapped together during the node matching routine. It is however interesting to note that the coverages with a match tolerance of 0.026in have marginally fewer lines than those with a match tolerance of 0.09in. This can only be attributed to the distribution of the lines within the coverage being more susceptible to one coverage than another; a completely unpredictable factor.

As could have been expected, it is the coverages with the large *weed* tolerances that exhibit the greatest differences, both from the original coverage and also from the other test coverages. All coverages with the exception of these two display very small mean values,

[22]Removing *dangle* errors, or node errors does not necessarily imply that lines have been lost; merely that errors in the topology have been rectified. Sometimes errors are addressed through line movement (in the case of undershoots) or alternatively via line removal (in the case of overshoots).

in the range 0 to 0.006in. Differences of this magnitude in terms of line length can hardly be regarded as significant, statistically or otherwise. Whereas the mean of 0.059 and 0.06 inches demonstrated by the two *weeded* coverages, indicates a relatively large decrease in the lengths of lines. This is further substantiated by the maximum values of these two coverages (indicating a drop in line length), which peak at 2.815 and 3.007 inches as compared to the figure for the other coverages at 2.468 inches. Minimum values also paint a similar picture, with these two coverages showing a smaller minimum value than the others (table 5.13). Given that the maximum length of difference is 2.468in, the minimum length of difference should try to emulate this as closely as possible, as these line differences are in all probability due to topological relabelling, rather than true line change. In the coverages *match noded* with the small tolerance of 0.009 the minimum value is the same as the maximum value, other coverages show a slight decrease of 0.004in. These results are further emphasised by the larger range and standard deviations associated with the *weeded* coverage.

Table 5.13 Summary of statistics for the line comparison coverages

| Coverage name | Mean | Standard deviation | Maximum value | Minimum value | Range |
|---|---|---|---|---|---|
| T026C02 | 0.002 | 0.494 | 2.468 | -2.468 | 4.932 |
| T026C2 | 0.004 | 0.494 | 2.468 | -2.468 | 4.932 |
| T09C2 | 0.001 | 0.496 | 2.468 | -2.468 | 4.932 |
| T09C03 | 0 | 0.496 | 2.468 | -2.468 | 4.932 |
| T09C2W1 | 0.060 | 0.529 | 2.815 | -2.334 | 5.149 |
| T009C03 | 0.003 | 0.495 | 2.468 | -2.468 | 4.936 |
| T009C2 | 0.006 | 0.494 | 2.468 | -2.468 | 4.936 |
| T009C2W1 | 0.059 | 0.532 | 3.007 | -2.442 | 5.449 |

This produced some surprises, in that it was the coverages match toleranced with the highest value that had the lowest sum of difference (table 5.14), conversely, those *match noded* with the lowest value had the highest sum of differences.

5. 14. The sum of the differences for the line comparison coverages in inches and their representation on the ground in metres by scales of 1:50,000, 1:25,000 and 1:10,560.[23]

| | Inches in coverage units | Metres on the ground as represented by 1:10,560 map | Metres on the ground as represented by 1:25,000 map | Metres on the ground as represented by 1:50,000 map |
|---|---|---|---|---|
| T026C02 | 1.975 | 529.74 | 1254.12 | 2508.25 |
| T026C2 | 3.163 | 848.39 | 2008.50 | 4017.01 |
| T09C2 | 1.022 | 274.12 | 648.97 | 1297.94 |
| T09C03 | 0.371 | 99.51 | 235.58 | 471.16 |
| T09C2W1 | 49.366 | 13241.15 | 31347.41 | 62694.82 |
| T009C03 | 2.229 | 597.87 | 1415.41 | 2830.82 |
| T009C2 | 4.796 | 1286.40 | 3045.46 | 6090.92 |
| T009C2W1 | 48.802 | 13089.87 | 30989.27 | 61978.54 |

These figures demonstrate the total difference in the extent covered by the linear features. In all cases there is a positive difference, indicating that less ground is actually traversed in the processed maps than in the original coverage. In the case of the two coverages that have been submitted to a *weed* tolerance, the sum of the difference is large and equivalent to over 13,000 metres on the 1:10,560 map and over 6,000 metres on the 1:50,000 map. If the values in inches are compared to the total sum of lines in the original coverage (514.547 inches), then it can be seen that almost a tenth of the coverage has been lost. Such losses could only be acceptable for the crudest of uses, as much valuable detail will have been eradicated.

The other coverages show a much smaller loss, and although for the most part these variations may seem relatively small, when regarded in terms of inches, if they are converted to ground values for the scales of 1:10,560, 1:25,000 and 1:50,000, the distances they represent could be of significance to certain users. For example 4.796in in the case of t009c2 could mean a loss of 3,045.46 metres on a 1:25,000 scale map. Admittedly one might argue

[23] Conversion used: 1 inch = 25.4mm = 0.0254m

that it is unlikely for many users to be using such small scale maps for detailed analysis. Yet many users do implement such maps either through ignorance or through use error or due to the unavailability of any other source. A further indication of the dissimilarity of the generated coverages might be provided by looking at the correlations of the differences. These are shown in table 5.15.

Table 5.15 Correlations of the differences in line length for the line comparison coverages.

| | T026 C02 | T026 C2 | T09 C03 | T09 C2 | T09 C2W1 | T009C 03 | T009C 2 |
|---|---|---|---|---|---|---|---|
| T026C2 | 1.000 | | | | | | |
| T09C03 | 0.986 | 0.986 | | | | | |
| T09C2 | 0.986 | 0.986 | 1.000 | | | | |
| T09C2W1 | 0.878 | 0.878 | 0.874 | 0.874 | | | |
| T009C03 | 0.989 | 0.988 | 0.997 | 0.997 | 0.873 | | |
| T009C2 | 0.989 | 0.989 | 0.997 | 0.997 | 0.873 | 0.999 | |
| T009C2W1 | 0.881 | 0.881 | 0.873 | 0.873 | 0.980 | 0.873 | 0.873 |

As anticipated, the two coverages with the *weed* tolerance show a high correlation between each other (0.970), but much smaller correlations with each of the other coverages. The lowest correlation was found between the t09c2w1 coverage and the two coverages *match noded* with a value of 0.026, (a correlation of 0.794) thereby implying that these are the least similar in terms of the differences between them and the original. These correlations are further highlighted if the percentage of similarity is calculated. These are shown in table 5.16, and show that for the most part the other coverages bear no more than a 78% similarity to either of the two *weede*d coverages.

Table 5.16 Percentage of similarity between the line comparison coverages (on the basis of line length).

| | T09C2W1 (%) | T009C2W1 (%) |
|---|---|---|
| T026C02 | 77 | 78 |
| T026C2 | 77 | 78 |
| T09C2 | 76 | 76 |
| T09C03 | 76 | 76 |
| T09C2W1 | 100 | 96 |
| T009C03 | 76 | 76 |
| T009C2 | 76 | 76 |
| T009C2W1 | 96 | 100 |

**5. 8. 3. Positional movement; looking at node shifts.**

So far, the analysis has concentrated on the lengths of the lines in the coverages, yet these may in some cases remain the same, but be transformed onto different coordinates within the coverage, which in turn represent different points on the earth's surface (figure 5.79). Since the test coverage is made up of primarily straight lines, any movement in the position of the endpoints will cause the attached lines to alter their geometry.

To try and assess the coordinate movement, and possible line displacement, the xy coordinates defining both the generated and master coverages were downloaded from the PC onto the mainframe. The files were then restructured so that only the start and end coordinates of each line were retained (appendix 5.2). As with the line length data, a unique identifier had been tagged onto each of the lines, and these then formed the basis for ordering the start and end nodes of each of the lines. Again as with the line lengths, lines bearing the same identifier were seen to be part of the same line. When these lines were observed, it became apparent that most of the 'multiple' lines consisted of a large line, followed by any number of extremely short ones. Furthermore the lines match each other perfectly nose to tail, in that the x y coordinates of the end of the first line formed the start coordinates of the next line. Using this feature of the data, a program was designed to search for multiple occurrences of labels and then to retain only the extreme x y coordinates for that label (appendix 5.2).

Trials with the program showed that there were some cases in which multiple lines bearing the same identifier did not follow onto each other, but consisted of a situation in which line collapse had occurred (figure 5.80). These cases were dealt with separately by hand. The endpoint awarded to the line was that which showed the least difference. The number of such cases are shown in appendix 5.3.

Once the coordinate data had been prepared in this manner, the start and end coordinates of each line as designated by its unique identifier, were then matched with those points in the original coverage. Differences were found for the start and end coordinate movement and also for the total positional movement occurring within the line itself.

The data illustrating the total displacement occurring in the coverages was run through the SPSSX package and the following basic statistics derived (figure 5.81 and figure 5.82). All of the coverages show a difference distribution that is quite heavily skewed. In the case of the coverages *match node*d with a tolerance of 0.026, a double peak occurs, causing the statistic for the skew to be tempered slightly. The first peak is a mid point of zero, and the second peak occurs at 0.013 to 0.014, indicating that most lines suffered little or no displacement, but of those that did, almost a tenth of the lines (74 in t026c02 and 73 in t026c2), were displaced by between 0.012 and 0.014 inches. If the peak is extended laterally, the number of lines almost double. The t09c03 and t09c2 coverages both display and attempt at a double peak, but it is not as developed as that in the former cases (only 26 values) and it is concerned with higher displacements (0.08 inches). The coverages *match node*d with the 0.009 inch tolerance both showed a heavily skewed distribution of differences, with the majority of differences occurring in the 0-0.01 inch range, indicating very little true displacement. As expected, the coverages displaying the greatest degree of displacement were those with the exaggerated *weed* tolerance. These were also heavily skewed, but the average differences were 0.055 inches, as opposed to the 0.002-0.013 inch averages displayed on the other coverages. The range of the difference magnitudes was also greater, equalling 2.025in. The 0.09in coverages also had a higher range than the other coverages, but this was still relatively small (table 5.17).

Table 5.17 Summary of statistics for the coverages comparing positional movement.

| Coverage name | Mean | Maximum value | Standard deviation |
|---|---|---|---|
| T026C02 | 0.006 | 0.039 | 0.008 |
| T026C2 | 0.006 | 0.039 | 0.008 |
| T09C2 | 0.0013 | 0.320 | 0.025 |
| T09C03 | 0.013 | 0.320 | 0.024 |
| T09C2W1 | 0.055 | 2.205 | 0.171 |
| T009C03 | 0.002 | 0.098 | 0.005 |
| T009C2 | 0.002 | 0.98 | 0.006 |
| T009C2W1 | 0.0057 | 2.205 | 0.171 |

If the sum of the total differences between the test coverages and the original is examined, some idea of the total geometric movement can be obtained. Total displacements are shown in table 5.18.

Table 5.18 Total displacements for coverages in inches and their representation on the ground in metres by scales of 1:50,000, 1:25,000 and 1:10,560.

| Coverage name | Total displacement (inches) | Metres on the ground as represented by 1:10,560 map | Metres on the ground as represented by 1:25,000 map | Metres on the ground as represented by 1:50,000 map |
|---|---|---|---|---|
| T026C02 | 4.219 | 1131.64 | 2679.06 | 5358.13 |
| T026C2 | 4.184 | 1122.25 | 2656.84 | 5313.68 |
| T09C2 | 8.752 | 2347.50 | 5557.52 | 11115.04 |
| T09C03 | 8.712 | 2336.77 | 5532.12 | 11064.24 |
| T09C2W1 | 33.768 | 9057.39 | 21442.68 | 42885.36 |
| T009C03 | 1.325 | 355.40 | 841.38 | 1682.75 |
| T009C2 | 1.342 | 359.96 | 852.17 | 1704.34 |
| T009C2W1 | 35.074 | 9407.69 | 22271.99 | 44543.98 |

As can be seen, the displacement for the *weeded* coverages hover around the 35 inch mark, as compared to the t009c03 and t009c2 coverages, in which the total geometric displacement of end nodes is only just over an inch. Translated into ground values, the movement on a 1:10,560 map would be 9,407.69 metres and 359.96 metres respectively. As with the line length experiment, the difference achieved in the *weeded* coverages makes them too crude for detailed use. In addition, however, the coverages *match noded* with 0.09 inches also display worrying displacement values, which brings their integrity into question too.
The numerical and statistical figures attached to these coverages provide useful background information as to the magnitudes of change that occur when such post processing routines are enacted. All too often however, it is the type of loss and its actual location that might prove significant in further analysis. To illustrate the types of line movement that have occurred, subsets of the original coverage have been plotted (with the unique identifiers) for the original coverage and the following test coverages; t009c03, t009c2 and t009c2w1. These are shown in figures 5.83 to 5.86 together with the node errors attached to each of the coverages.

The original coverage is riddled both with node errors and small spurious lines adjacent to line endpoints. These are reduced in the T009c03 coverage, and most of the node errors that remain are not *dangle* errors, but pseudo node problems. For example, in the centre of the coverage, line 364 has disappeared, and with it the *dangle* node error that once existed. In the same way, line 378 towards the left of the coverage has disappeared taking with it two *dangle* errors and replacing them with a pseudo node. The t009c2 coverage has accenuated this process, removing line 373 in the left of the coverage and also line 408 in the centre. It is however, in the third coverage that widespread geometric change becomes apparent even without considering the individual identifiers. The drainage channels no longer take the form of parallel lines, and considerable collapse has occurred around the island in the left of the coverage. Line movement and length changes have occurred along the whole expanse of the horizontal channel, for example, lines 359 and 356 have merged, as have lines 450 and 454 on the vertical channel. The island on the left of the coverage no longer exists as such, but has been incorporated into the main body of the coverage. The lines that originally formed the island and the lines situated around it have completely altered their positions, for example lines 382, 430, 374 and especially line 398 have been intermixed and condensed with their neighbours.

### 5. 9. Points for discussion.

- **The location of the errors.**

In many respects, it is not so much the magnitude of the errors that the tolerancing algorithms create, which determines the detrimental effect on coverage integrity, as the position in which they arise.

- **The user faced with the black box scenario.**

Node 'error' removal in GIS packages is essential to allow completely closed areal units for overlay analysis, but for the most part the entire operation is carried out behind closed doors alienating the user from participating in the actual processes that are taking place. The only decision the user can make is the size of the tolerance to be used. In many cases this is rarely evaluated with respect to the coverage depreciation that takes place. Furthermore, in ARC/INFO the use of an option entitled 'clean' plays on user preconceptions and word association. Data is by no means cleaner or 'better' in some way than it was on entering the system, if anything the data will have lost some of its original character. Part of the problem facing the user is that in addition to not being aware of the way in which the processes are manipulating the data, there is no means by which the user can trace the origin of the lines or track their movements within the system algorithms.

Users of such systems often merely accept what the processes do, provided that the 'error' is removed. The mechanisms of error removal in GIS deserve to be looked at in greater detail as the removal processes themselves can geometrically change a coverage in a way that the operator could not foresee.

- **Evaluating coverages for analysis.**

Looking at the number of node errors that have been removed within the coverage in isolation is not a sufficient basis for declaring a coverage to be devoid of error. Potential problem areas should be singled out and even if statistical analysis exceeds the funds the user has available for the project, at least some detailed feature checking should take place. This should involve both the checking of shape and topological change. Any anomalies that are found can then be rectified manually. If statistical analysis can be carried out, then coverages displaying large statistical anomalies should be looked at, and potential problem areas analysed and reviewed. Alternatively coverages with problem features could be flagged. Thus sinuous coastlines and small jagged features could be highlighted for rigorous checking, as could lines that denoted important features.

Another factor that has been demonstrated, is that there are many routes to initiating the removal of similar numbers of node errors, each of which will have a slightly different influence on both the positioning of map features and their shape. Some of the line removal will undoubtedly have been carried out intentionally and in a manner which the user desires; so not all of this 'difference' is bad. It is however, the user's responsibility to assess which lines have been unintentionally removed and whether their removal renders the coverage unsuitable for further analysis.

- **Balancing the books; costs and benefits.**

Dirty data as a consequence of digitising techniques is a fact of GIS life. It is the way in which the data are dealt with that accentuates or reduces the inherent errors. The common errors of undershoots and overshoots must be removed, but at the same time coverage integrity should be preserved. Eliminating node errors can be done manually, with the user making a value judgement about each case. This ensures that only the correct lines are removed and also that lines are matched at the correct meeting points. Unfortunately such particular attention to detail is time consuming. The full coverage subset had 250 node errors, and it was a relatively small coverage consisting of only 827 lines. As the coverage size increases, there is a likelihood that a greater number of errors will emerge. Editing each node problem by hand can take up to two minutes, even for an experienced user. The human capital required to prepare large map sheets for analysis by hand is therefore phenomenal, and very few organisations will be in a position to meet that demand, or even to require it.

The optimal situation would be to delegate the usual and easily dealt with undershoots and overshoots to the computer algorithms, and to deal with previously defined problem areas individually. This again would entail an assessment of all the node errors and a classification of their potential complexity and would be costly in terms of human input. Potential trouble areas could be flagged and dealt with using differing tolerances, but this would provide further problems of edge matching, and would require a substantial amount of additional data management. Furthermore 'problem' areas are often dependent on the idiosyncrasies of the individual coverage and the way in which the information has been internally stored.

In many respects therefore, it is down to the user to assess the extent to which the removal of such large numbers of errors by machine defined algorithms justifies the subsequent change in coverage geometry. User decisions will in turn be based upon the intentions of the organisation and the use to which the data will be put. Large thematic land use studies for example are less likely to require detailed information as are market research companies looking for the optimum siting for a new cash dispenser. The costs of manual node error eradication would be beyond the reach of most organisations, but as an alternative some cut off point should be drawn, beyond which the reduction of more node errors will seriously undermine the geometry of the coverage in areas not necessarily connected to the nodes in question.

- **A threshold value?**

Possible threshold values could include the total amount of distance moved in terms of ground units, provided the user is in a position to perform a comparison. This type of method would allow the user to set thresholds on the basis of the data, and of the intended use. Alternatively a threshold value could be found as a result of rigorous testing, using an iterative process of tolerances and visual quality checks, however, assessing the margin of error associated with a particular coverage visually is no easy task, since the user is unaware as to how much the geometry of the entire coverage has been modified. Unfortunately, such a method is unable to deal with lines that have been completely eradicated, or the cumulative effect on an area that has had many of its nodes displaced.

Ideally, the number of *dangle*s removed should be played off against the widening difference between the test coverage and the original coverage. This can be done graphically, and a cut off point can be delimited.

Two graphic scenarios suggest themselves in the light of the above experiments (figure 5.87). In the first case there is an optimistic curve which occurs as the gradual increase in the *match node* and *dangle* values no longer reduce the number of *dangle* errors, but continue to augment the difference in line length with respect to the original coverage. The second case is a more depressive curve, and would describe the situation that occurs when gradually increased *weed* tolerances are employed. In this case *dangle* errors are continually reduced, but the difference in the feature lengths or position with respect to the original coverage escalates dramatically.

In addition to relative guidelines, which are data dependent, absolute guidelines could also be considered. These would centre around the use of the *weed* tolerance. As has been demonstrated in all the experiments, it is the high *weed* tolerance that significantly changes coverage features, therefore any use that is made of this tolerance should be very carefully monitored. All small islands or features in close proximity should be checked and verified each time the technique is applied.

- **The need for accuracy.**

In vanquishing the errors introduced at the digitising stage, the user needs to choose between accuracy of the data and the time taken to eliminate those inaccuracies. The easy, quick method to the removal of *dangle* errors and gaps between arc endpoints is to set high tolerances and quickly remove a large proportion of the obstacle to the creation of consistent polygons. Many institutions are quite happy to operate on broad error margins and for them dirty data is not a problem. Other users may rank accuracy to be of paramount importance to their particular application, and therefore an awareness of the potential dangers that automatic rectification may inflict is vital.

- **The need for benchmarking details.**

The evolution of a propriety GIS frequently includes a benchmarking process. This may involve the processing of a complex artificial coverage where certain results are expected from certain processes. Alternatively a 'real' coverage can be used and the results assessed in the light of the original 'ground truth'. Both these processes tend to be regarded as proprietary knowledge by GIS consultants, and not much detail exists in the literature.

- **The importance of scale.**

Scale also plays an important role in determining the degree of feature displacement. The same distance in centimetres on two coverages of different scales will relate to different distances on the ground. The greater the scale, the less large tolerances will alter the positional accuracy of the features, the smaller the scale, the more drastic similar tolerances will be. Therefore due to the accuracy of the operator, digitising table and cursor the accuracy of the map diminishes with a decrease in its scale. In effect lines on the ground are allowed a freer licence to be moved on a small scale map than on a large scale one.

- **Legal aspects.**

As an increasing amount of geo-referenced data becomes incorporated into information systems, the question of legal liability raises its ugly head. Who is responsible if geo-referenced data is inaccurate.? Will it be possible for the public to sue those responsible for producing faulty information? If so, then the reduction of error in digital cartographic coverages will be critical.

**Chapter Six - Errors associated with overlaying.**

**6. 1. Introduction.**

Geographic Information Systems claim to offer the user the potential to merge information in a variety of ways. For example, individual map sheets often published as individual geographic elements, can be combined with the aim of creating a 'seamless' database. Yet this is not an exclusive feature of GISs, as the representation of cartographic information in any digital form encourages the data to be viewed as contiguous, rather than being made up of a variety of components. Of greater importance, is the ability a GIS posseses to both link and integrate information.

Linkage allows cartographic information to be associated with attribute values. The association is made by allowing the positional feature to act as a 'spatial hook', onto which information that pertains to that particular point in space is attached. For example, initial associations might denote a particular line as a river and another as a road, whilst a secondary association might be the cargo a particular river carries, or the class of a particular road. Further associations might be the number of pleasure cruisers that use the river in August, or in the case of the road, the number of accidents due to drunk drivers at Christmas.

> "The crucial advantage offered by the GIS is the ability to link together seemingly disparate information provided it is spatially referenced ... This increases the accessibility of the data and extends their influences by allowing information to be used for purposes other than those for which it was originally intended."
>
> Singh (1990, 385)

Integration on the other hand implies more than this. For rather than dealing with an increasingly complex single geography, integration looks at multiple layers of geographical representation, and tries to combine the data from them into a single, unified and easily queried information base.

Data suitable for input into such a system, will usually be derived from various sources and in some cases the data may actually consist of different 'geographies'. For example, different enumeration units may have been utilised for the acquisition of socio-economic information,[1] or enumeration areas may give a large area the characteristics of a smaller zone within it, by making assumptions of equal density distribution throughout the area.[2] The general tendency however, is for the same geographical phenomena to be represented inconsistently in different sources as a consequence of the inherent error processes present in data collection and representation.

Before such data can be utilised for spatial analysis, its reliability must be verified and any errors arising due to the integration process must be removed. If the results of the integration are to be effective, then minimising mismatches is a vital priority. This chapter seeks firstly to investigate the nature of such errors and the consequences of their propagation within coverage, followed by an assessment of the popular methodologies for error removal.

Depending on the type of analysis to be performed, such data may consist of time series data, for use in the analysis of change, or alternatively a variety of themes may be integrated for an area in question, allowing suitability analysis to be performed for a predefined area.

**6. 2. Sources of data for input.**

Data for digital cartographic integration is derived from many sources, both in terms of its type and its origin (table 6.1).

Table 6.1. Basic division of information for integration analysis

| SOURCES | Digital Mapping Use | GIS database Use |
|---|---|---|
| Maps/cartographic/ remotely sensed material | Geometry | Integration |
| Statistics | Numbers | Integration |
| Non-statistical 'soft' data | none | Domesday Disk |

[1]Flowerdew and Green 1989 discuss this problem for the integration of data collection units.

[2]Openshaw 1984 outlines this in his discussion of the modifiable areal unit problem.

Data may be in attribute form without associated graphics, for example housing profiles may be derived from the Census of Population data for particular enumeration districts, or literacy rates may be reported for administrative units. Such data will need to have a geographic identifier that will allow that particular attribute or value to be related to a particular location. Alternatively, data may be in a spatially referenced form, be it in the form of maps or remotely sensed data. In traditional digital mapping systems these were viewed as being the separate 'lines' and the 'numbers' components of geographic information (table 6.1). It is important to note however, that what both digital mapping systems and most GIS' lack is the ability to assimilate 'soft' or 'extra' data such as news reports on employment statistics, although an attempt was made at this type of data assimilation by the Domesday project in the 1980s. A GIS therefore receives a relatively restricted subset of information about the real world consisting of *only* geometry and hard statistics. This has forced 'analysis' to operate only on those items, offering the user a relatively limited set of tools.

However, not only is there a difference in the actual type of data involved. Each individual piece of information, be it even of the same type, is likely to have undergone different processing routines and will have followed different quality control procedures. Even in the same agency where procedures and quality regulations are consistently maintained human errors may influence the result, as may the varying quality of the surveyed material.

**6. 3. Cartographic data.**

Cartographic data are derived from either primary or secondary sources (table 6.2).

Table 6.2. Sources of cartographic information.

| |
|---|
| • Primary data source |
|     • survey material |
|         positioning |
|         baseline definition |
|         primary, secondary, tertiary triangulation |
|         local surveys (plane surveying & thematic surveys) |
|     • remote sensing |
|         orbital |
|         sub-orbital |
| • Secondary data source |
|     • map |
|         geometry |
|         features/attributes |
|         context |

### 6. 3. 1. Primary data sources.

- **Surveyed material.**

Surveying techniques rely on the procedures of measurement, registration, recording of data, processing of data, analysis of data and finally its presentation. Each of these stages will have an error component attached to it. The extent of such errors will be determined by the observational skill of the surveyor and interpreter, but also by the precision of the instruments and techniques used to acquire and analyse the data.

Initial surveys deal with positioning and the identification of points upon the earths's surface, which in turn will form the basis for further localised surveys.[3] Surveyors tools and techniques for achieving such aims have essentially remained the same for the last few millennia.

> "It is indeed remarkable that so few changes in land surveying practice took place between the time of the Mesopotamian who surveyed and drew the plans and designed the irrigation of agricultural land in the Tigris and Euphrates valleys and the late nineteenth century surveyor who produced maps and plans and set out construction works in industrialised Europe and North America."
>
> Cooper(1985, 1)

However, throughout this century, improvements in the available technology and a consequent improvement in techniques have served to elevate the potential speed and

[3] Once points have been accurately identified, baselines can be identified, and from these, using the processes of triangulation, trilateration and increasingly, satellite observations primary, secondary and tertiary networks can be set up.

accuracy with which positioning can be carried out. Such tools include Global Positioning Systems (GPS)[4], Electronic Distance Measurement (EDM)[5], total surveying stations and satellite observations[6].

Secondary surveying techniques look at portraying features on the earth's surface. Traditionally this has been carried out using plane surveying following a traverse. This

---

[4] GPS was conceived as a highly precise effective continuous positioning system. It consists of 18 satellites in high orbit, of which four should provide the ideal geometry for position fixing at any point on the earth at any time. Two signals are emitted from the system, and those agencies with Department of Defence (DoD) clearance are able to access one or both of the signals, for positioning. The primary advantages of GPS systems are a) continuous observations are possible, b) high frequencies eliminate higher order refraction effects, c) higher orbits reduce the effect of the short wave part of the earth's gravitational field and d) oscillator errors are reduced by high quality instruments. Relative positioning through GPS is usually carried out by 'white noise phase delay measurement'. Here data is time lagged and collected at two or more stations independently. These are later cross tabbed and the time delay difference measured. Observations on a number of satellites then enable the computation of the three components of the baseline vectors between the stations.

[5] Recently EDM techniques have been introduced. These measure distances with infra-red waves, microwaves or laser beams. A transmitter unit is placed at the start of a baseline, or link to be measured and a transmitter/receiver is placed at the end. The length of time for the sent signal to make the round trip is then recorded. Some commercial models are capable of reducing errors down to more or less 1-2cm only.

[6]These aid with simultaneous satellite observations and doppler measurements. The simultaneous method involves the use of a satellite that emits a precisely timed intermittently flashing light. Simultaneous photos are taken of these from the earth's surface. Stars which appear in the photograph are used to precisely position unknown observation points. The doppler method measures the change in frequency of a signal due to the movement of the sound source relative to the observer. Doppler satellites transmit continuous signals on fixed short wave radio frequencies. These are monitored on the ground and the result compared to the known frequency of the signals. The received frequency is higher than the satellite's frequency when the satellite is approaching and lower when it is moving away. Measurement of frequency shifts allows for angular calculation, which in turn is used to calculate the satellite position with respect to the recording stations. The position of an unknown station is then assessed by looking at its relationship with the satellite's position.

method is still utilised today, but with more precise theodolites, and sometimes with EDM.[7] An alternative to plane surveying has included inertial positioning systems, which allow a traverse to be more speedily measured[8]. Finally there are observational thematic surveys. These deal with themes such as land use, or soil type, and in many cases such direct observation has been replaced by remote techniques.

Errors in survey material arise either due to some aspect of the measurement of the data, or alternatively due to the interpretation of the data. Errors in primary positioning will affect all subsequent triangulation networks and these in turn will dictate how positionally accurate localised surveys are. The precision of the equipment utilised will call into question the viability of all data at all stages of acquisition, and the operational skills with which these tools are used, and their readings logged will also determine the accuracy of the material.

The accuracy of observational surveys relies primarily on the observer, and accuracy verification procedures. However, since much of this area has been replaced by remote techniques, the liability for accurate thematic and topographic information now tends to lie with the interpreter rather than the observer.

- **Remotely sensed data.**

The term remotely sensed imagery incorporates a wide range of mappable information, ranging from visible air photographs to satellite imagery responsive to various components

---

[7] This process can also be enhanced by those that can afford it, thus large organisations such as the Ordnance Survey has developed their own digital field up-date system (DFUS) to digitise change in field offices immediately after revision survey has taken place. Digital map files are transferred by telecommunications between the digital archives disc and the field office.

[8] Inertial systems rely on three gyroscopes and three accelerometers, linked by a computer. These are mounted on a truck or helicopter to allow for fast movement over long distances. Measurements start at a ground control point where the three instruments are aligned with the horizontal, vertical and azimuth. All are then moved over the area to be surveyed in a similar way to plane surveying a traverse. Any responses the instruments have to the changes are recorded.

of the electromagnetic spectrum. Air photographs have been used in various degrees of sophistication for most of this century, for military reconnaissance work[9] and then increasingly for both topographic and thematic mapping. In the last twenty years, more sophisticated satellite retrieval techniques have evolved, and these too have followed a similar route of military development followed by potential geographic and environmental use.

Information retrieval varies, depending on how the electromagnetic spectrum is being utilised for image acquisition. Sub-orbital platforms rely on retrieval using equipment placed on board aircraft. These may utilise the visible spectrum or alternatively they may take the form of multi-spectral scanners, side looking airborne radar (SLAR) or thermal infra-red scanners (Harris 1987).

Orbital platforms came into use with the launch of Landsat I in 1972. This was followed throughout the 1970s and 1980s by the Landsat series and in 1986, by the launch of SPOT (Systeme Probatoire de l'Observation de la Terre). Scanners aboard the Landsat series deal with multi-spectral imagery via the MSS (Multi-Spectral Scanner) and TM (Thematic Mapper), whilst the SPOT satellite is able to retrieve data in both multi-spectral and panchromatic form.

Given the continuously increasing volume of data generated,[10] and the potential for satellite remote sensing, its usage still remains relatively patchy. In many cases the data are simply not available for the area or resolution[11] required. Generally, however it is the prohibitive cost of purchasing both the data and the processing equipment that limits the use of the technology. Even if data and technology can both be secured, interpretation can still prove

[9]A heavily highlighted feature of the Gulf War in 1991 was the reconnaissance capability of the allied forces.

[10] "A complete channel of Landsat MSS imagery comprises 2,340 lines and 3,240 pixels per line, a total of around 7.5 million pixels per scene of four channels ... this awesome amount of data, which fills just one computer compatible tape (CCT) can be collected in just 25 seconds." Curren (1985,143)

[11] MSS data have a ground resolution of 79m x 79m. TM data have a ground resolution of 30m x 30m. SPOT panchromatic data have a resolution of 10m x 10m, whilst SPOT multi-spectral data have a resolution of 20m x20m.

difficult, as the tools of image analysis are not precise enough for many uses.

- **Problems associated with remotely sensed imagery.**

Problems equated with remotely sensed imagery, be it aerial photography or satellite imagery, depend on the skill and technology with which the image is both obtained and interpreted. In the case of sub-orbital platforms, flight operated scanners are prone to image distortion as a result of aircraft instability such as pitch roll and yaw, in addition SLAR data suffers from the intrusion of topographic effects. Problems of the photogrammetrist and aerial photograph interpreter have been extensively documented elsewhere[12], and in the most part rely on the quality of the image, the skill of the photogrammetrist and the precision of the equipment being used.

When satellite remote sensed data are utilised as either a topographic or thematic cartographic information source, the quality of the data will initially be determined by the conditions in which the image was obtained. This includes the weather conditions prevailing at the time, the angle of the satellite and the resolution of the satellite receiver. Furthermore however, quality will also be affected by the post processing techniques undertaken to produce 'satisfactory' data for information systems.

- **Post processing techniques.**

These techniques attempt to extrapolate the maximum amount of detail from the image, by positional re-adjustment and thematic interpretation. Restoration occurs in the form of skylight, haze and shade suppression in an attempt to clarify image features, and geometric correction takes place by altering predictable sampling errors, compensating for changes in satellite or aircraft angle and most importantly by manipulating the coverage to fit on another image or map using ground control points. As the effort is made to force these points to match predefined cartographic images, positional and topographic features attached to such ground control points can inadvertently become warped and distorted, and the geometry of the coverage becomes a composite of satellite observation and surveyed material.

---

[12] Burnside 1979, Moffitt and Mikhail 1980, Campbell 1991.

Other image processing techniques deal less with the topographic geometry of images, but rather, they attempt to enhance the thematic information contained within them. However, when applications involve transition zones, then the boundaries of thematic change effectively become positional information, represented within the GIS as a linear feature.

Enhancement techniques include contrast stretching and filtering. Here data are manipulated to enable as clear a picture as possible to be extrapolated from the image. Contrast stretching allows for the maximum contrast between the image and its surrounding area, and the span of spectral responses obtained within the image is stretched to fit between 0 and 255. More complex filtering can be carried out by suppressing or enhancing certain spatial frequencies directions and textures, either within the frequency domain or the spatial domain. For example low pass filtering tend to smooth out irregularities and banding and high pass filters tends to enhance edges. Thus the low pass filter favours large areas and the high pass filter favours small areas, whilst directional filters enhance edges in a particular direction. As these are all generalisation techniques, fine detail is likely to be eliminated in feature definition, but, at an acceptable cost to image augmentation.

- **Classification techniques.**

As remotely sensed imagery deals with both theme and topography, a large component of research has concerned the proficiency with which imagery can be classified into geographic regions[13]. As with ground surveyed material, classification involves a degree of generalisation, as no natural thematic feature will abruptly stop at a defined boundary of a specified width and scale. Here however, rather than using visually observed criteria, the user is reliant upon an ability to distinguish remotely sensed spectral responses using a variety of statistical techniques. These may be performed upon a density slice of one waveband or alternatively by supervised classification of several wavebands.[14]

---

[13]Fox, L., Brokhaus, J. A. and Tosta, N. Mw 1985, Kelly, R. 1985, Thomas, H. C, Scarpace, F. L and Lillesand, T. M. 1986, Toll, D. L 1985, Young, J. A. T. 1986, Remote Sensing Yearbook, 1986, 1987, Hilton, 1991.

[14] After suitable wavebands have been selected, test classifications are carried out on sub-divisional areas. The results of these are then extrapolated over the whole data set using a classifier. Examples of these are the box classifier and the parallelepiped classifier. The most accurate is however the maximum likelihood classifier where pixels are added to classes for

Accuracy can then be interrogated using sampling techniques which compare the result of the sampled points in the field, to the points that have been sampled in the image. Naturally such techniques can only produce an approximation to reality, ensuring that the unclassified pixels are at least not assigned to areas with which they have nothing in common. As with any classification procedure, the quality of such thematic classifications could be improved by increasing the information available to the classifier. This may take the form of increasing the spatial resolution of the data, (comparable to improving the scale at which information is depicted), alternatively, a decrease in the number of classes required would diminish the probability of mis-assigning uncertain pixels.

### 6. 3. 2. Secondary data sources.

Once survey material has been translated into cartographic form, these maps themselves form a suitable information source for digital cartographic data capture. For most users unable to afford the resources required for primary surveys, most digital cartographic information will be obtained from maps, and will therefore be prone to the further inaccuracies involved in the map making process, such as drafting errors, scaling, generalisation and symbology. Thus the representation of criteria on the map will be both real (in terms of positional locations) and fuzzy (in terms of attribute defined locations, such as the boundary between two soil types). The GIS database is capable of acquiring both these sets of information and storing them and their relationships for both geometry and features (table 6.3). A third level of information that a map contains is its contextual information, such as what the map was designed for and the message it is trying to convey. Current GIS database storage techniques do not allow the assimilation of such information, and therefore the map as such can never be fully incorporated into the database system.

Table 6.3. Levels of information held in a map and GIS database.

| | Real | Fuzzy | Behavioural/cognitive aspects |
|---|---|---|---|
| Geometry | map AND GIS | map AND GIS | |
| Features | map AND GIS | map AND GIS | |
| Context | | | map ONLY |

---

which they have the most probability of belonging.

Geometric data input into GISs can be subdivided into three main groups. Firstly there are those which deal with the representation of topographic features in the real world, such as the location of a triangulation point, the centreline of a road, or the shape of a football stadium. Similar to these cartographic elements are artificial features which have their boundaries rigidly defined. The most common usage of such data is for thematic operations, in which postcode sectors, administrative units, or most national borders may be delimited and used as the basis for attribute depiction. There may be a level of generalisation in the drawing of the areas themselves, but their location in space, although artificial, is an accepted one. The final group of cartographic material that may be used as input into a GIS, is data which deals with subdividing spatial information into categories. This is often termed categorial or fuzzy data (Chrisman 1982a). This type of data tends to be primarily areal in nature and usually deals with landscape descriptors, although it has a more general application in contour and isopleth mapping. Such data may represent features such as different vegetation zones, climatic zones, soil types, land use, purchasing power or spheres of influence. As these attributes are difficult to partition or demarcate in terms of a linear delimitation, their borders always remain fuzzy. The reality is that often the boundary between such areas does not exist as a clearly defined linear feature, but rather as an area of intermixing in which one of the neighbouring features might predominate, but not necessarily. Thus when the boundary between two zones needs to be defined from multi-source data input, problems arise. As most GISs are incapable of holding *fuzzy* data, some decision has to be reached as to which of the sources should be defined as having the accurate and accepted boundary. This will then form the 'baseline' onto which other material can be attached. If the wrong decision is made at this point, then all further analysis using the data, may be error-prone.

Such graphic data are defined by the areal attributes they delimit, so any mistake in the graphic, implicates the geography attached to the areas that are defined by them. The need for accurate data preparation and overlay is thus all the more heightened when such data forms an important component of integration.

The reliability associated with all cartographic data are those of survey quality and the pre-processing techniques associated with coverage preparation. In the case of fuzzy data, the methods used for interpolating the boundaries are also subject to question, as are the criteria for the classification system.

- **Features.**

The features of categorical zone maps have already been discussed in the previous section, and as with remotely sensed data, it is the changing feature that determines the geometry. For other data types however, the geometry is usually clearly defined and further information which may relate to the identification or class of the geometrical piece of data, simply has to be attached onto it. For example, roads will need to be distinguished as such and then they will require a class and a name.

- **Context.**

Identifying and obtaining the contextual information that is inherent to any map, is a subjective process, depending upon the cognition of the individual. Thus the errors associated with such information will be very much tied to individual perceptions. Furthermore, even if such information were to be easily obtainable, its storage as an integral part of 'cartographic' data would prove difficult, for whilst the geometrical component of cartographic material is held within the topographic data structure, and the feature attributes often held in separate structures defined by the type of database utilised by the GIS[15] there is no easily identifiable slot into which the contextual information can be added. Thus in many cases contextual information is lost or ignored.

## 6. 4. Attribute data.

Attribute data can be subdivided in the same way as the graphic data they pertain to. There are those attributes which are absolute, such as the names of features, the area contained within a state, the number of people in a particular district, the housing type of a particular estate. All of these attributes are liable to change, but at any point in time there can be an undisputed value assigned to that feature. In addition there are attributes which are derived from a sample of the total population which they represent. These attributes are dependent

[15]Object oriented systems such as Smallword are an exception to this, as they store information in total units, rather then as separable entities of geometry and feature.

on the survey techniques used to extract them and interpolate from them. Adequate sampling techniques are required both for surveys of natural phenomena and populace, as are viable interrogative techniques. The accuracy of the sample should be statistically verified whenever possible, and questionnaire based data can often reveal particular responses on the basis of the question structure or wording.

## 6. 5. Overlay.

The most common form of digital cartographic integration occurs as a geometric overlay. This involves the integration of data on two levels. Firstly there is the integration of the digital cartographic data itself in the form of points, lines and areas. This can be compared to overlaying a set of conventional traced cartographic outlines on a light table. A particular location can then be selected and its characteristics analysed by looking at the composite result of overlaid data. For example such a point may fall within a zone of high population density, in an area within a mile of the main highway, which has fewer then two banks per square kilometre. On a light table such areas can be detected by visual inspection and repetitive verification. In a GIS, such areas can be located on the basis of the attributes falling within algebraic equations. In effect, therefore, the overlay process does not form a composite picture with the information of several sources simply drawn on one map, but it creates new geographic areas as defined by having several criteria in common. Thus information and unique geographic areas can be accessed to suit the user's specifications.

Overlay operations within Geographic Information Systems are based on the principles of set algebra. Sets of information can be manipulated to produce new data in which ceratin conditions hold true. Most GISs offer the user the ability to use methods such as UNION, INTERSECTION and IDENTITY. However, as most GISs concentrate on an accumulation of information, few allow the user to retain areas that do not hold true for certain conditions as part of the overlay process (figure 6.1). Such operations may be performed by more complex querying. Other facilities provided often include buffering and 'cookie cutting', whilst some systems such as SPANS allows for equations representing models of spatial information to be input into the overlay processor.

## 6. 6. Topological overlay versus cartographic overlay.

A commonly confused issue in digital mapping and GIS is the difference between the ability of systems to merely draw one map on top of another, and to be able to spatially overlay one map on top of another. In the first case, the two sets of digital data are merely drawn using a similar projection and coordinate system, whilst in the second case, a third dataset has been created as a consequence of mathematically combining portions of space that fit the criterion defined in the overlay procedure. The first case is often referred to as cartographic overlay, as the process has merely been carried out for display. Admittedly a degree of analysis can take place, but this is primarily visual. The resultant image contains no interrogable information about the newly created areas. Contrarily, if information is overlaid topologically, then each newly created area will form part of the new spatial image in its own right, and as a consequence, will have a series of attributes attached onto it. These will include descriptive information relating the portion of space it depicts in the real world, such as its land use, as well as descriptors of its geometry, such as its perimeter and area. Furthermore, these newly created areas will be topologically related to all those areas which surround it, thus making geographical enquiry possible.

## 6. 7. ARC/INFO and geometric overlay.

### 6. 7. 1. Data types.

Data suitable for entry into ARC/INFO's overlay processor include points lines and areas. Although areal information may be overlaid with other areal maps, it is more common for points and lines to be combined with polygon data, than with coverage of their own genre. This is most commonly done in the operations of line in polygon and point in polygon. Buffers may be built around any of the datatypes, although the resulting coverage will always be an areal one.

### 6. 7. 2. Requirements of the data.

An initial requirement of the data is that topology has been built, and in the case of polygon data, that all polygons are fully closed. In addition polygon integrity should have been verified. Graphically, this entails ensuring that all the lines making up the polygon actually exist and that no lines have been missed out, whilst in terms of the information base, this involves checking that the correct attribute is attached to the polygon or other feature in question. Other checks include the verification of assigned user-ids, and the removal of any small polygons with zero ids as these are likely to have been created as the result of the triangulation effect of the error removal routines.[16]

**6. 7. 3. Operations available**

In addition to the buffer operations, all set theory operations are possible. The first of these is UNION which merges the two datasets in totality, thus the resultant dataset contains all the areas that were present in both of the original datasets. The more selective operations of INTERSECT and IDENTITY are also available. In the first case, only those areas which are present in both of the initial datasets are retained, whilst in the second case, all of one dataset is retained in addition to all that coincides with it in the second dataset. An enhanced overlay operation is that of 'cookie cutting', in which the extremities of the data region considered are determined by the extent of one of the data sets.

**6. 7. 4. The fate of newly created features.**

Once areal data sets have been overlaid, all the areas of overlap are defined as new polygons. Not only does this affect the graphic, but it also has implications for the information held within the database. The area held will now be divided up using a different set of dividing lines or polygon boundaries, and each of these will be documented anew in the database. Attributes associated with each tract of land will also have changed, either by superimposition or extension.

**6. 7. 5. The consequences of such processes.**

The positive consequence of such processes is to add value to the data already held, and to make the data more suitable for analysis. If an analysis of change is to take place then newly created areas will indicate an increase or decrease in a particular variable. However, the negative side of the coin is that the very process of overlay creates problems of imperfect linear feature matching, figure 6.2. When lines that are overlaid do not match perfectly, two possibilities present themselves. Firstly, the mismatch may be a result of some change in the dataset, secondly, the mismatch may be a result of two map outlines that have been subjected

[16]See Chapter Five.

to all the subjective processes described in the preceding two chapters, and the mismatches may simply be a symptom of the error processes involved. In many respects therefore, unless the data are perfectly matched by hand prior to the use of the overlay routines, the overlay process is inherently error prone.

**6. 8. What are sliver polygons?**

As sliver polygons actually exist within a coverage, they cannot be ignored in any analysis that might then be undertaken. The space they occupy is real world space, the extent of which will depend upon the size of the sliver and the scale at which the data has been represented. That space is therefore erroneously designated, and attribute and locational information obtained from the coverage will be lost. Where change occurs in urban land use, or areas of intensive resource usage, it often does so in small areal units. Yet these small units may easily be grouped in with sliver polygons by a standard GIS 'sliver removal' algorithm. The cumulative effect of the process can also be quite drastic, as the more a coverage is overlaid on another, the greater the opportunity for a gradual depreciation in the quality of the data and in the reliability of any subsequent results.

**6. 9. The origins of such polygons.**

Any integration of digital cartographic data from several sources will inevitably produce situations in which the same feature in space (such as natural barriers or administrative boundaries) will be represented by lines in different physical locations. The true location on the ground will undoubtedly lie somewhere in between the cartographically portrayed ones. Such errors or inconsistencies are inevitable because of the number of sources being used, and the widely differing range of processes that might have been used to create each one.

The origins of such polygons lie in all the error prone processes of data entry and input that have already been outlined in previous chapters. Particular problems include those associated with the map data itself, those associated with the transfer of the data into the system, and those identified with tidying up the inevitable repercussions of the data entry process. In the first group, scale and the effects of the generalisation reduce the quality of the data, as does the problem of edge matching. Edge matching may be achieved within most GIS packages, but only with a degree of error and generalisation that the user should be aware of before the operation is attempted. An option for error minimisation is to match the edges prior to data entry, although human error may of course occur here.

The second phase of error accretion occurs in the elimination of detail that may occur in the tracing of the map onto a stable media, and also in the digitising of the graphic features represented in that tracing into a GIS coverage. The digitising process, as outlined in Chapter Four, is susceptible to operator fallibility, unavoidable generalisation and node matching errors.

Lastly, sliver polygons may arise due to errors introduced into the coverage as the data is transformed from 'spaghetti' into the system's data model. As outlined in Chapter 5, the post digitising routines often cause the location of a linear feature in space to shift in order to permit perfect geometric node matching. A further consequence of these routines is to create tiny triangles in the linear coverage as small lines collapse. Unless all these errors are removed during post-digitising verification, they will have uncertain and potentially serious impacts during subseqent analysis.

**6. 10. Overlaying mapsheets of different scale.**

An experiment was undertaken to evaluate the success of integrating several topographical maps depicting the same area at different scales. In many cases, overlay uses different boundary features, such as land use and water usage, in which the user has no benchmark to assess the quality of the information that has been received from other sources, or which has been digitised within an agency. In other cases, especially in assessments of change (which may be very small), relatively similar areas will be overlaid on each other, and if change is to be correctly assessed it is important for these coverages to be integrated successfully and accurately. Therefore selecting maps that allegedly depict the same features, in the same spatial location, makes errors all the more apparent.

All the coverages were of the same projection, produced by the Ordnance Survey. The maps were topographic maps at 1:10,560 (appendix 6.1), 1:25,000 (appendix 6.2) and 1:50,000 (appendix 6.3). They were traced and digitised by the same operator, and were exposed to similar tolerance routines (appendix 6.6). Therefore theoretically, there should be very little difference between them except for the generalisation features due to scale. The area selected was Morton Fen in the Fenlands, Eastern England. A strict procedure was followed in the digitising of all the maps in an effort to eliminate all error parameters within the operator's control.

### 6. 11. Methodology.

#### 6. 11. 1. Map preparation.

The map sheets selected for the project were newly purchased and had no creases in them. Morton Fen at the 1:10,560 scale was however split into sheets TF SE and TF SW. Rather than digitise two sheets separately and include an extra error parameter of edge matching it was decided to match the two sheets manually prior to data entry. Each map was traced onto plastic tracing film, to counteract the problem of using paper which is an environmentally unstable medium. In addition areas or features that were unwanted in the final product were filtered out, (figure 6.3). The density of information available on the built up areas was reduced to include just the outlines of the house plots and bridleways and footpaths which were sometimes represented by parallel lines, or one straight line parallelled by one dashed line were generalised at this stage to single lines taken along the mid point of the parallel pair (figure 6.4). On the 1:50,000 scale map this method also had to be applied to main roads as these are depicted by very fine parallel lines with a strip of colour in between them to indicate what class of road the lines represent (figure 6.5). Originally, it was thought that such lines could be traced using a pen width that was the same as the width of the entire road but this soon proved problematic, as when it came to digitising such lines, the cross hair on the digitising tablet was so much finer than the line it was to trace. This prompted the consideration of the following procedures. Firstly, only one side of the large pen width would be digitised (figure 6.6), but this would have the effect of laterally displacing the true value of the line, resulting in the need to move junctions and intersections so that they could meet up with the displaced line (figure 6.6). Thus intersecting roads would have their lengths extended, resulting in an alteration to their accuracy. The second method was to approximate a centreline within the 0.7mm pen width. Although such a practice minimised the lateral movement of the line that was being digitised, it meant that intersecting lines would have to be slightly extended on both sides of the junction, thus causing two sets of alterations for each intersection (figure 6.6). In addition the effect of using a fine resolution cross hair on a thick line, whilst trying to approximate its centre often produces a bouncing line (figure

6.7). Sometimes the operator will digitise slightly to the left of the centreline and sometimes slightly to the right, thus the resultant digitised line will not possess the straight linear quality of the line on the tracing. As neither of these two procedures produced satisfactory results, subsequent maps were digitised with pen widths of no greater than 0.35mm, and any generalisation of lines that was to take place, was carried out prior to the digitising process itself (Appendix 6.4 and Appendix 6.5).

The 1:10560 maps were traced twice on two separate days by the same person; the same procedure was followed each time, but the draughtsperson had no access to the completed first copy whilst tracing the second, thus leaving the door open for the possibility of drafting error due to insufficient written instructions and subjective variability. It was hoped to demonstrate that even when the same individual is responsible for the drafting on two separate occasions a degree of inconsistency will occur, a phenomenon that can only be exaggerated if more than one person is responsible for such tasks as is often the case in commercial institutions. The 1:50,000 map was also traced twice, but the first version included some lines drawn with pen nibs of 0.7mm. The 1:25,000 map was traced only once.

For each map at least eight orientation points were identified. These consisted of the four outermost corners of the map and four points (usually at grid intersections) within the map. These points then form the basic reference points for transforming all the data contained in digitiser coordinates into a recognised grid. In addition during breaks in the digitising process, the map may have to be removed from the digitising table and replaced on a different portion of it; these points then form the basis for ensuring all the features to be entered are positioned in the correct relative locations to what already exists.

For each of the tracings the following pieces of information were noted down on the tracings themselves; the type of light source used and whether it came from above or below, the name of the operator, the date, the time of day, the pen nib width, the source data and the date of its survey as well as a list of generalisations that were carried out. The mylar sheets were then stored flat until they were required for the digitising process itself.

### 6. 11. 2. Equipment preparation.

The equipment used in the digitising experiments were an IBM PC-AT to communicate with the digitiser, a Summagraphics AO digitising table and the PC ARC/INFO ADS (Arc digitising System) software package. Prior to digitising the basic digitising table tests were carried out to ensure that the PC and the digitising equipment were communicating satisfactorily. The table was configured so that the operator was comforatble when digitising. The tracings were then firmly secured to the digitising table with masking tape.

**6. 11. 3. The digitising procedures used.**

As the ADS package records cartographic data for storage in a topological data structure the digitising procedure involves selecting chains and nodes or in the case of ADS, Arcs (consisting of segments) and nodes (start and endpoints of lines). After the coverages have been registered, several tolerances can be set. These include the snap tolerance, the edit tolerance and the weed tolerance. The snap tolerance allows nodes that fall within a specified linear distance of each other to be mapped together as the system assumes the nodes are meant to be one and the same[17]. Setting the snap distance too high will result in nodes that exist close together in reality being snapped onto each other, thus rendering the coverage more inaccurate than the digitising operator intended. As nodes can always be mapped together in post processing the snap tolerances were left at the default level. The edit distances were changed as needs arose. This distance indicates the radius of the search for a particular feature. The weed tolerance only allows points to appear if they are a certain distance apart from each other, thus eliminating clusters of points in small areas which may then take a heavy toll on storage capacity. As with the snap distance, this was also set to the system created default value. The first feature to be digitised on all the coverages were arcs. Three steps are involved in the digitising of arcs. Firstly identifying a start point, secondly electing representative points on the arc to retain as vertex co-ordinates and finally identifying an end point.

As the extent of the area in question made the lines that appeared on the PC's screen in response to the operators instructions quite small, each of the maps were divided into eight

[17] This works in much the same way as the match node tolerance discussed in Chapter Five.

windows, and each window was digitised as a separate section. The first map to be digitised was the 1:10560 map. During the digitising process, the aim was to try and match all the nodes as accurately as possible. Subsequently it was revealed that common ARC/INFO policy is to encourage users to 'overshoot' at node junctions rather than to aim for the junction exactly.[18] The reasoning behind such a philosophy is that overshoots are easier to deal with in the post processing routines than undershoots. Therefore actively aiming for an overshoot will simplify matters later on, rather than aiming for precision and falling short. The second 1:10,560 map and the second 1:50,000 were therefore digitised with a slight overshoot.

The coverage was saved approximately every thirty minutes, which also offered the operator a well earned rest. After all the lines were thought to be in, the result was checked and verified. Labels (denoting polygons) were added after the map had been fully cleared of all its node errors and transformed from digitiser co-ordinates to National Grid Co-ordinates.

**6. 11. 4. Removing digitising errors.**

The consequences of such errors in a topologically encoded data base is that they prevent the creation of true polygons as regions are left to spill into each other (figure 6.8). In addition, if it is the lines that are to be used for modelling or network analysis, they will need to be able to flow into each other, rather then being isolated in space (figure 6.8).

In this case the 1:10560 coverage digitised first was used (10560) as a test coverage to assess which tolerance values were most applicable for use on this occasion. Suitable match node values were seen as 0.05", whilst suitable dangle tolerance values were in the range 0.03 to 0.05. These values eliminated the majority of errors whilst limiting the degree to which the coverage features were compromised. When the minimum distances were measured for the

---

[18] This policy was first revealed by Green and Wiggins 1986 (Personal Communication), when ARC/INFO was mounted at Birkbeck College London as part of its beta release testing in the 1980's. Reference to this policy can also be noted in the Understanding GIS Manual published by ESRI:

> "When digitising arcs that connect to the outer boundary, you can extend them by only the slightest distance. This creates 'dangles' which are often easier to fix than 'undershoots', as you will discover when editing errors in lesson five."
>
> ESRI, (1990, 411)

smaller scale maps, it was found that similar tolerance values had to be used.

Any node errors that were not eradicated by blanket tolerances, were manually edited using the powerful graphic capabilities of the ARCEDIT module. This might have involved a series of operations associated with individual lines. For example, lines can be selected and then deleted or replaced, whilst nodes can be moved and vertices extended. A predominant use of the software was to extend lines beyond their current endpoint by adding either another line or vertex, or alternatively by moving the line or the node to which that line pertained. Adding extra features was generally more acceptable than moving features to a new location. This was primarily due to the fact that the majority of the lines that were being dealt with were straight, and therefore consisted of one vertex only. Moving such lines would cause the relocation to ripple all the way through the line, so even if the endpoint was now firmly matched with another line, the start point of the line might have been readjusted to the extent that it now wandered aimlessly in space.

### 6. 11. 5. Procedures employed in verifying the integrity of the coverages.

Other digitising errors such as digitising a line twice or missing out arcs may be dealt with interactively, using a manual editing procedure. The editing facility comprises an entire module of PC ARC/INFO and can therefore be expected to offer a wide range of options to ameliorate any single graphical problem. For example lines that don't quite match, even after match noding and dangle tolerancing, may be brought together, by extending the line with an extra vertex. Alternatively, the line might be deleted and the whole entity redrawn. Another option would be to allow an extra arc to be added to the end node of the line. Finally the last vertex of the arc might be moved so that it extends to meet the node it is supposed to intersect with (figure 6.9). With such a wealth of geometric manipulations available to the user, it is small wonder that even the same map digitised and edited by the same person will not appear to be entirely the same.

### 6. 11. 6. Adding topology to the coverages.

Once any missing arcs or geometric mistakes had been addressed, and the coverages were deemed to be free of any node errors, topology was created. The 'clean' procedure discussed in Chapter 5, creates topology on the assumption that the coverage is to be one containing polygon data, which may or may not be the case, as topology can be created for line or network based data as well as polygon based information from the same set of map features. If polygon topology is desired, then each polygon must be individually itemised so that it can be easily identified for analysis. Polygon labelling can be carried out graphically or as a database routine. For the most part, labels will be copied from the traced map image either at the time of digitising using a cursor, or at a later juncture via a manual editing routine.

Assigning areas identifiers is just as prone to error as digitising the boundaries of areas, especially if the number of polygons is large and the real extent of them relatively small. Common mistakes are often double identifers or missing identifiers, either of which will cause problems for analysis.[19]

Identifiers were assigned to all of the polygons within each of the test coverages using a graphic method[20]. This was verified visually, by checking each area on the mylar film. Following this, each of the coverages was built as an areal coverage with a full PAT to accompany it. The PAT was interrogated and all polygons with a zero identifier and were removed, as these were not present in the original mylar tracings, but had arisen as a result of tolerance routines such as those outlined in Chapter Five.

**6. 11. 7. Overlaying the coverages.**

Once all five maps had been digitised (two at 1:50,000, two at 1:10560 and one at 1:25,000) and topologically built, they were what most users would consider ready for integration through the process of geometric overlay. As all the data coverages pertain to the same area and deal with the same phenomena, any mistakes in coverage representation are easy to identify. In order to overlay all the data from three differing map scales, each coverage was transformed onto the Ordnance Survey Grid Coordinate System, as until this present time all

---

[19] If there is no identifer in the polygon then the system will leave the polygon as having a zero value. Zero values may then be used as a logical determinator by the user to eliminate polygons that are unwanted. If there is more than one identifier in each polygon, then the polygon may be suspended from any areal analysis.

[20] Identifiers were assigned using both the digitising table and a mouse within the post-digitising editing module in ARC/INFO.

processing has been carried out using the digitising coordinates.[21] Once all five data sources are operating on the same projection then the coverages can be overlaid.

### 6. 11. 8. Evaluating the errors of the data acquisition process.

As some of the coverages are quite detailed, the map extent has been divided into eight sub sections so that coverage variations may be inspected in greater detail (figure 6.10). If the map coverages themselves are then inspected, a gradual decrease in the provision of detail can be noticed as scale decreases (figure 6.11 to figure 6.15). This is to be expected as the smaller the scale of any map, the more generalised it becomes. If however the maps of the same scale are observed, certain variations also exist between these, which as both are of the same scale and from the same source can only be due to variations in the digitising process, be they in the area of map preparation, and digitising or post processing.

If the 1:10560 map in red (figure 6.11) and the 1:10560 map in green (figure 6.12) are compared, certain discrepancies become apparent. A prominent area of discord is on the boundary of section 1 and section 5; here the red map provides a greater degree of detail than the green. This is primarily true in the case of depicting the urban areas. If such digitised coverages were now to be used for an application such as assessing the level of landscape change for this area, with one map holding attributes for the year 1980 and the second holding attributes for the year 1989, the observer might be led to the false conclusion that a level of landscape change had emerged, with an increase in the number of dwellings on the outskirts of Bourne.

Another main area of discord between the coverages is that of the river network and the portrayal of some of the drainage ditches in sections four and eight. On the red map these are defined by parallel lines, whereas on the green map, single lines are used to indicate the same area. If geometric overlay were to be performed on these coverages the level of detail on the red map would have to be reduced so that a perfect match with the green map could be effective. Finally, there is also a difference in the portrayal of dwellings, with the number

[21] Coverages are transformed onto common co-ordinates using an affine transformation which relies on a minimum of four registration points.

of farmsteads and dwellings on the red coverage exceeding those on the green map despite identical source material.

In comparing the 1:25,000 map with either of the 1:10560 maps a depreciation in the level of detail can be noticed, especially in the representation of the urban areas in sections 1 and 5 (figure 6.13), in the river channel (sections 4 and 7) and in the general absence of farmsteads. Some lines portrayed as parallel lines in the 1:10560 map have also been reduced to single lines on the 1:25,000 map, especially the roads leading towards the urban area on the boundary between section 1 and section 5. There is also a considerable difference in the area of Morton Fen (sections 2 and 3) in favour of the 1:25,000 coverage. This strange phenomenon arises as the area was re-surveyed in 1986 and the new observed field boundaries were drafted onto a 1:25,000 map in the field.

Observing the two maps digitised at the 1:50,000 scale (figure 6.14 and figure 6.15) highlights many of the points already acknowledged in the comparison of the two 1:10560 maps. The yellow 1:50,000 map (figure 6.14) has a greater level of detail, especially on roads, drainage ditches, river channels and outlines of dwellings (section 5). Finally an overlay of the two maps (figure 6.16) reveals a degree of displacement between them, which may have been due to the thicker pen nibs used for the tracing on one of these coverages. Alternatively, it may be the result of operator digitising error or overlay transformation error. For example, if the two original maps (figure 6.14 and figure 6.15) are overlaid manually, a better fit can be obtained than occurs with the with the computer generated overlay (figure 6.16). This is perhaps due to the fact that the human eye looks at the totality of the overlay and makes many subtle adjustments in the comparison. ARC/INFO on the other hand looks at each polygon separately and can only function on the information provided for it. Thus the accuracy of the map overlay algorithm itself, may well depend upon the positioning and number of the tic registration points.

The degree of such mismatches only becomes apparent when the map layers are superimposed, and this is exactly the same process that comes into being when data layers are integrated in geometric overlay. In an overlay situation however, errors are likely to be further compounded as it is the norm for data to be gathered from several sources, rather than one mapping agency such as the Ordnance Survey. If all the five coverages are overlaid, total chaos ensues (figures 6.17 to 6.24). Variations are greater in sections four, seven and eight where data consist less of straight line segments, and a greater proportion of curved features. Trying to make sense of such data would prove a nightmare as so many spurious areas have been created. If these coverages were chronologically derived through time, such spurious polygons might be mistaken for areas of change, whilst in reality they are errors encountered in the process of digitising and overlay.

In map overlay analysis, such errors are quite likely to be greater than those illustrated in this example as the chances of having the same operator digitise all the inputs to the system are relatively small. Furthermore, in many 'overlay' situations the source data will normally relate to varying attributes and features, therefore actually discriminating mismatches and errors will prove to be nigh on impossible, and the user will simply assume the data to be accurate. Only by maintaining the highest of standards in data digitising and post processing can the user rest assured that errors have been kept to a minimum.

**6. 11. 9. An analysis of coverage differences.**

The degree to which such errors, or sliver polygons arise, can be assessed by comparing the distribution of polygon size in each of the overlaid coverages with those coverages from which the composite was derived.

As sliver polygons are known to be small in size, an increase in polygons of small dimension within a coverage could be a potential indication of sliver polygons. Alternatively, the change in the distribution of polygon size following overlay could also be due to some modification of the attribute that is being depicted. In the case of land use for example, an overlay of land use cover for two periods in time may produce small areas which are valid units of change. Therefore, to eliminate the possibility of veritable feature change causing the increase in small polygons, two coverages allegedly depicting the same features at the same scale were initially selected for comparison, followed by coverages depicting the same features at different scales.

Table 6.4 Number of polygons remaining in overlaid coverages, with the associated areal increase.

| Coverage name | Area ($m^2$) | No. of polygons |
|---|---|---|
| 1:50,000a | 49889678.181 | 168 |
| 1:50,000b | 49927337.147 | 190 |
| 1:50,000a + 1:50,000b | 50112443.117 | 882[1] |
| 1:10060a | 49429576.270 | 979 |
| 1:10560a | 50030194.344 | 1119 |
| 1:10560a + 1:10560b | 50045254.002 | 4923[2] |

[1] 364% increase in polygon number.
[2] 340% increase in polygon number.

In both the overlay of the two 1:50,000 maps and the two 1:10560 coverages, there has been a dramatic increase in the number of polygons held within the coverage as a whole. In the case of the 1:50,000 composite, this increase has exceeded 364% and in the 1:10560 composite, the increase is 340% (table 6.4). A more detailed portrayal of the quality of the change rather than merely its quantity can be ascertained from distribution histograms of the 1:50,000b (figure 6.25) coverage and the overlay of the two 1:50,000 coverages, (figure 6.26) and in the case of the 1:10560 scale, between the 1:10560b coverage (figure 6.27) and the 1:10560 composite (figure 6.28).

In the composite of the 1:50,000 case, (figure 6.26) the number of polygons in the first distribution class with a midpoint of 47820$m^2$ totalled 753, accounting for 85% of the total number of polygons in the coverage. In the 1:50,000b coverage (figure 6.25) however, the number of polygons whose area fell around a similar midpoint (50162$m^2$), totalled 55; a mere 28% of the total. A similar pattern was observed in the case of the 1:10560 overlay (figure 6.28). Here 4327 polygons fell around the first class midpoint of 15009$m^2$, giving it 88% of the total number of polygons. This compares with the 1:10560b coverage, (figure 6.27) in which 713 polygons are within the class having a midpoint of 24038$m^2$, thus accounting for only 64% of the total number of polygons held.

In both cases, the number of polygons with a smaller area has increased, and as the two pairs of overlays deal in exactly the same features, with their geometry being obtained from the same source, the increase cannot be attributed to feature change, but rather to incompatibilities between the coverages, presumably caused by differences in the tracing and digitising of the two coverages.

If the histograms of the original coverages are analysed (figures 6.25, 6.27 and 6.29) then a general trend of an increase in smaller polygons can be observed with an augmentation in scale of the coverage. Thus a greater number of smaller polygons becomes an indication of increasing accuracy. To reduce the misinterpretation that might be caused by the large number of 'accurate' small polygons contained within a large scale coverage when it is superimposed on a smaller scale coverage, comparisons involving a coverage made up of two scales with one of its original coverages, must be made with respect to the most detailed coverage held within the composite.

The increase in the total area for the 1:50,000 overlays and the 1:10,560 overlays is an interesting phenomenon, in that the smaller scale coverages are having their area increased by a greater amount than the larger scale coverages. Several interpretations present themselves. In the first case one might assume that given the way that the transformation and overlay algorithms function, the increase in area will be related to the best fit the algorithm can produce given the positioning, number and subsequent match of the tic points on the two original coverages. One might further assume that if the same coverages were to be digitised and overlaid an infinite number of times the areal change would probably tend to a normal distribution, with some overlays producing cases in which the total area was greater than the original map extent and a similar number of overlays in which the total area was smaller than the original map extent. However if the inputs were to change in some way, for example if the two maps to be used for overlay were not from the same scale, this would presumably invoke a skew, the direction of which could only be ascertained by empirical testing. Another interpretation would relate to the scale from which the material was obtained. For example, smaller mistakes relating to the boundary of polygons in the 1:50,000 scale coverages will produce larger deviations in terms of area than larger mistakes on the 1:10,560 coverages.

Four other combinations were overlaid (table 6.5). Here, although the features being dealt with are theoretically the same, variations exist due to generalisation associated with scale, so a greater number of sliver polygons can be expected. As in the previous cases, the total number of polygons held within the coverages increases with the number and scale of the coverages being incorporated. The increase, as before, comes in the form of small polygons. This is especially marked in the three-ply overlay, where the increase in the number of polygons held (as compared with the 1:10560 coverage) is 538% and in the overlay of the 1:10560 and 1:25,000 coverage where the increase is 327%[22]. As table 6.6 indicates, the number of smaller polygons held within these two coverages has also increased. In the three-coverage composite the number of polygons falling into the histogram category with a midpoint of 6530 corresponds to 88% of the total number of polygons within the coverage (figure 6.30), whilst in the case of the 1:10560 and 1:25000 composite, the percentage of polygons in the distribution group with a higher midpoint of 12677 equalled 87% (figure 6.31). Comparative percentages are shown for the original coverages which formed part of the input to the overlay, and as in the previous cases, the percentage of small polygons within a coverage has increased. Interestingly enough, this has not shown itself to be the case for the intersections involving the 1:50,000 coverage (figure 6.32). In both these cases (table 6.6) the percentage of smaller polygons has not greatly increased as compared with the original coverages, and yet the total number of polygons held has escalated. In these cases it might be safe to assume that the differences between the coverages are not of the tiny sliver variety, but in fact slightly larger, thus they are more evenly distributed among the distribution classes, rather than being clustered in one zone.

[22] The 1:25,000 coverage was not exactly the same as the Ordnance Survey depiction, as several new field boundaries were added from visual surveys. However the size of the new polygons depicting the reorganised boundaries were not comparable with the size normally encountered in sliver polygons.

Table 6.5 Overlay of the various coverages and the associated increase in polygon number and area.

| Coverage names | Area ($m^2$) | No. of polygons | % increase in area |
|---|---|---|---|
| 1:25,000 | 49958129.282 | 908 | - |
| 1:25,000 + 1:50,000b | 50076632.316 | 2111 | 133% |
| 1:10560b | 50030194.344 | 1119 | - |
| 1:10560b + 1:50,000b | 50122233.131 | 2539 | 126% |
| 1:10560b + 1:25,000 | 50105386.643 | 4777 | 327% |
| 1:10560b + 1:25,000 + 1:50,000b | 50176792.451 | 7144 | 538% |

Table 6.6 Coverage overlays and the associated number and percentage of polygons with a small areal extent.

| Coverage names | % change in total number of polygons | No. of polygons | Midpoint value ($m^2$) |
|---|---|---|---|
| 1:25,000 | 71 | 653 | 32317 |
| 1:25,000 + 1:50,000b | 75 | 1593 | 27425 |
| 1:25,000 + 1:50,000b | 64 | 1341 | 9141 |
| 1:10560b | 64 | 713 | 24038 |
| 1:10560b + 1:50,000 | 66 | 1665 | 6750 |
| 1:10560b + 1:25,000 | 87 | 4134 | 12677 |
| 1:10560b + 1:25,000 + 1:50,000b | 88 | 6299 | 6530 |

In summation therefore, whilst the overall distribution of the map remains the same, there has been an increase in the areas of smaller areal dimension, primarily due to the presence of erroneous sliver polygons.

### 6. 12. How can sliver polygons be dealt with?

Effectively removing sliver polygons entails accurate identification. The characteristics inherent to sliver polygons can be discriminated on the basis of their size, shape and feature class. In general, sliver polygons are very small in area, and can often be identified in the database by their relatively insignificant size as compared with other areal features. In addition to their small area, sliver polygons often have an extensive perimeter, as they are areas which have been created by marginal mis-alignment of lines (figure 6.33). If the two characteristics are combined, it can be seen that the average sliver polygon will have the properties of a small area combined with a large perimeter. There is therefore, the possibility of designating polygons on the basis of their area to perimeter ratio. A further characteristic that can be utilised is that of class or attribute. Particular combinations of classes can be isolated as being untenable in a particular coverage and therefore worthy of removal.

The difficulty in identification, arises when polygons that need to be retained display similar traits to the sliver polygons. For example roads, drainage ditches and other linear features often exhibit large perimeters in comparison to area, whilst small units such as farmsteads may fall below the allowable areal tolerance. The situation becomes further complicated when change is being investigated, as discriminating sliver polygons from change can often be a demanding process, given that both features will often display similar dimensions and shape. Change rarely arises in large uniform polygons, rather it manifests itself in small incremental zones very similar to the way slivers arise.

Various routines exist in each of the major systems to overcome the problems posed by sliver polygons, on the whole these rely on elimination rather than subjective reassignment. In ARC/INFO, the advised procedure for the removal of slivers relies on the removal of areas on the basis of features in the attribute table falling below a user-defined criterion.

> *eliminate* reduces the number of polygons or lines in a coverage by merging selected features with one of their neighbours...
>
> ARC/INFO PC overlay users Guide, section 4.

Advised methods for selection rely on the ratio between the area and the perimeter of sliver polygons, although the suggestion is that the area divided by the perimeter should be less than 10.[23]

### 6. 12. 1. Polygon size as an indicator.

Areas for elimination can be selected on the basis of some component of their size. The two options that are available are the area and perimeter of the polygon. To test the effectiveness of these options the test coverages were overlaid in the following combinations:

- 1:50,000b and 1:50,000a,
- 1:10,560b and 1:10,560a,
- 1:10,560, 1:25,000 and 1:50;000,
- 1:10,560b and 1:25,000,
- 1:10,560b and 1:50,000,

and their results subjected to elimination on the basis of their perimeter and area and a combination of the two.

### 6. 12. 2. Perimeter as an indicator.

Using these geometric attributes as the basis for polygon removal relies on the user being able to define a particular type of perimeter that would be associated with the sliver polygons. In this case perimeters can fall above or below a certain criterion. Large perimeters might indicate large areas, but they might also indicate long thin areas, or areas that have long complex boundaries. Small perimeters might relate to small sliver polygons, but they may also indicate polygons of simple and smooth shape.

### 6. 12. 3. Area as an indicator.

The areal value is perhaps not as unpredictable as the perimeter value as polygons with large areal values are unlikely to have been generated as a consequence of error or generalisation. Therefore, once a suitable threshold value is decided upon, all polygons below that value can be eliminated. In the same way as tolerance values operate, the areal tolerance that is set must always be less than the smallest retainable polygon. Any polygons which are unwanted, and which fall above this value should be removed manually, otherwise information will be lost and the quality of the final product will have undergone avoidable deterioration.

---

[23] This is a strange ratio to pursue, as it would result in the elimination of a lot of areas which do not display the properties of sliver polygons, as many robust areas display that type of ratio. A superior ratio would be to suggest that the perimeter divided by the area should be greater than 10.

Many sliver polygons however, can be quite large in areal extent, given that they may consist of long thin features, where the true boundary of the region has been slightly misplaced on one of the overlay coverages (figure 6.34). Thus the use of such tolerance values may eliminate polygons which do not have sliver characterisitics. Overcoming such polygon problems can be achieved by using a combination of perimeter and areal values.

### 6. 12. 4. Area and perimeter as an eliminator.

Using both the areal extent and the perimeter of the polygons in the coverage as a basis for sliver elimination allows long thin areas to be removed, by asking for those polygons in which the area divided by the perimeter falls below a predetermined threshold. The implication of removing features corresponding to such profiles, is that polygons may conform to the sliver polygon definition and yet they may in fact be veritable features that require retention.

## 6. 13. The geometrical effects of such threshold values.

Although several coverage combinations were overlaid,[24] two combinations were selected for further analysis. These were the two 1:50,000 coverages and the coverage consisting of one input from each of the scales. For the initial tests, the two 1:50,000 coverages were used, as these were more economical with both processing time, and disk space, than the larger three way combination. Tolerance values that were seen to be the most successful on the 1:50,00 combination were then enacted upon the three scale merge.[25]

Three tolerances based on the perimeter values of suspected sliver polygons were then applied to the combined 1:50,000 coverage. These values were selected on the basis of sliver polygon shape, in that sliver polygons tend to have large perimeters. Four values were

---

[24] The combinations consisted of the two 1:50,000 coverages, the 1:25,000 and the latter of the 1:50,000 coverages, the two 1:10,560 coverages, the latter 1:10,560 and the latter 1:50,00 coverage, the 1:25,000 and the latter 1:10,560 and finally a combination consisting of a coverage from each of the three scales.

[25] Implementing the UNION option on ARC/INFO uses a considerable amount of disk space to store the scratch files required to enable the process. This might be up to eleven times the size of the coverage itself. Furthermore the UNIONed coverages themselves are hungry on disk space, with the 1:10560 overlays requiring over a megabyte a piece.

selected and these are shown with their respective areas and total number of polygons in table 6.7.

Table 6.7 A comparison of the number of polygons and areal extent of the 1:50,000 overlaid coverages with the originals following the application of *eliminate*s based on the perimeter.

| Coverage originals | | Area ($m^2$) | % change in area wrt 1:50,000b coverage | % change in area wrt unioned coverages | No. of polys |
|---|---|---|---|---|---|
| 1:50,000a | | 49889678.181 | - | - | 168 |
| 1:50,000b | | 49927337.147 | - | - | 190 |
| UNIONed coverage | | 50112443.117 | 0.37 | - | 882 |
| | Perimeter tolerance | | | | |
| 1:50,000a + 1:50,000b | 1000 | 49876522.720 | -0.10 | 0.47 | 320 |
| 1:50,000a + 1:50,000b | 1500 | 49405883.221 | -1.04 | -1.41 | 214 |
| 1:50,000a + 1:50,000b | 2000 | 48673887.050 | -2.51 | -2.87 | 126 |
| 1:50,000a + 1:50,000b | 2500 | 48673384.131 | -2.51 | -2.87 | 94 |

If the number of polygons in each of these toleranced coverages is compared with the number of polygons held in the original 1:50,000a and 1:50,000b coverages as well as in the UNIONed coverage, (table 6.7) a pattern of initial increase followed by a loss in the number of units becomes apparent. As the coverages used for the UNION have not been perfectly matched, an increase in the number of units and also in the sum of polygon areas (table 6.7) arises. In an effort to temper this disparity, the elimination routines attempt to reduce the number of polygons and the sum total of polygonal area to more closely resemble the original. Table 6.7 also illustrates the percentage change in total areas held with respect to each of the original coverages and to the UNIONed coverage too. In all three cases there has been a reduction, implying that the elimination routines are removing more than they should be. Although the changes are slight in percentage terms, they represent 50,815$m^2$ to 1253,953$m^2$ units on the ground. Depending where the reduction occurred, losses of such magnitude could be vital. Figures 6.35 to 6.37, illustrate this point. Thus although feature reduction has taken place, it is not necessarily only intended polygons that have been withdrawn.

If the plots of the toleranced coverages are compared with the plot of the UNIONed coverages, the degree to which intentional polygons have been removed can be ascertained. In the first case, employing a perimeter tolerance of 1000 units (figure 6.35), has resulted in the deletion of some redundancy in the central section of the coverage, although mismatches still predominate in the south eastern areas and in parts of Morton Fen itself. A higher tolerance of 2000 units (figure 6.36) is successful in eliminating almost all the slivers and mismatches, but coverage detail begins to be eroded and simplified. An even larger tolerance of 2500 units, causes serious node collapse and results in excessive detail depletion (figure 6.37).

Although sliver polygons have the distinction of being long and thin, and therefore in the possession of large perimeters, they are not the only polygons exhibiting this criterion. Furthermore, the larger, or more intricate a polygon becomes the greater the extent of its boundary. The effects of this phenomenon can be seen in figure 6.37. Here many of the field boundaries in the centre of the coverage have disappeared due to their perimeters falling below that of the tolerance set.

A second parameter used as a basis for tolerancing was the polygon area. A variety of magnitudes were selected, ranging from 5000 units to 150,000 units. As with the perimeter tolerances, the more substantial these became, the greater the degree of coverage distortion. The smallest areal tolerance of 5000 units (figure 6.38) still ensured that a large number of mismatches were preserved. The use of an areal tolerance of magnitude 80,000 units was more successful in removing unwanted areas (figure 6.39), although some node collapse has occurred and smaller areas have been lost altogether.[26] The largest tolerance displays the extreme case, where polygon features approach distortion (figure 6.40).

The number of polygons present also indicates that the tolerance values are approaching the number of polygons held in each of the original coverages when it holds a value of between 168 and 190. Changes in total area (table 6.8) would indicate that the 50,000 tolerance leaves the coverage most akin to the original pair. It is however, important to note that maintaining an acceptable overall coverage magnitude is no guarantee that all desired areas will have the same areal magnitude, or even exist. As can be seen from figure 6.41, when it is compared with the either of the original coverages, many areas which could not be classed as errors in the south east and north east of the coverage have been removed.

The final parameter used on this UNIONed coverage, was that of area divided by perimeter, as this allegedly takes account of polygons with a high perimeter to area ratio. A wide variety of tolerances were chosen, ranging from 10 to 50. If the procedure only removed redundant polygons then, given the total number of polygons in each of the coverages (table 6.9), it could be hypothesised that a tolerance of between 20 and 30 units would result in a final coverage which compared favourably with the two originals. The areal comparisons (table 6.9) also indicate that a tolerance of such magnitudes would be suitable as percentage differences for these values are lower than the other tolerances tested.[27]

---

[26] The edge is maintained on this coverage by using an option that ensures no polygon with a border on the perimeter of the coverage is removed.

[27] In cases where there may have been change, the final coverage selected for geographical analysis, following error removal may have more polygons than the original, and may in addition have a greater total area. In the case under analysis in this study however, the coverages are allegedly portraying the same features and therefore all differences are errors and need to be removed, thus the lower the percentage change the better.

Table 6.8 A comparison of the number of polygons and areal extent of the 1:50,000 overlaid coverages with the originals following the application of *eliminate*s based on area.

| Coverage originals | | Area ($m^2$) | % change in area wrt 1:50,00b coverage | % change in area wrt the unioned coverages | No. of polys |
|---|---|---|---|---|---|
| 1:50,000a | | 49889678.181 | - | - | 168 |
| 1:50,000b | | 49927337.147 | - | - | 190 |
| UNIONed coverage | | 50112443.117 | 0.37 | - | 882 |
| | Area tolerance | | | | |
| 1:50,000a + 1:50,000b | 5000 | 50109839.552 | 0.37 | -0.11 | 462 |
| 1:50,000a + 1:50,000b | 25000 | 50059585.199 | 0.26 | -0.11 | 202 |
| 1:50,000a + 1:50,000b | 50000 | 49923184.684 | -0.01 | -0.38 | 156 |
| 1:50,000a + 1:50,000b | 80000 | 49727754.691 | -0.40 | -0.77 | 138 |
| 1:50,000a + 1:50,000b | 150000 | 49013208.750 | -1.83 | -2.19 | 109 |

Table 6.9 A comparison of the number of polygons and areal extent of the 1:50,000 overlaid coverages with the originals following the application of *eliminates* based on area/perimeter.

| Coverage originals | | Area ($m^2$) | % change in area wrt 1:50,000b coverage | % change in area wrt Unioned coverages | No. of polys |
|---|---|---|---|---|---|
| 1:50,000a | | 49889678.181 | - | - | 168 |
| 1:50,000b | | 49927337.147 | - | - | 190 |
| UNIONed coverage | | 50112443.117 | 0.37 | - | 882 |
| | Area/Perimeter tolerance | | | | |
| 1:50,000a + 1:50,000b | 10 | 50103370.208 | 0.35 | -0.02 | 383 |
| 1:50,000a + 1:50,000b | 12 | 50100276.612 | 0.35 | -0.02 | 324 |
| 1:50,000a + 1:50,000b | 15 | 50097837.967 | 0.34 | -0.03 | 266 |
| 1:50,000a + 1:50,000b | 20 | 50078727.969 | 0.30 | -0.07 | 217 |
| 1:50,000a + 1:50,000b | 25 | 49993954.975 | 0.13 | -0.24 | 176 |
| 1:50,000a + 1:50,000b | 30 | 49991804.992 | 0.13 | -0.24 | 160 |
| 1:50,000a + 1:50,000b | 50 | 49819378.418 | -0.22 | -0.58 | 142 |

The progression of these tolerances are shown in figures 6.42 to 6.48. The higher the tolerance becomes the more detail is eradicated. Both the tolerance of 10 units and 12 units do not achieve their aim of error reduction as many sliver polygons still exist in the main field areas, as well as in the south east of the coverage. With a tolerance of 15 units, the majority of the problems in the field boundaries are removed, but difficulties still exist in the south east. The tolerance of 20 units has improved the river channel, but has still failed to deal with the jumbling in the south eastern corner. The tolerance of 25 units is probably the most successful of those used, as few slivers exist and few areas worthy of retention have been removed. Features that have been unintentionally removed are generally those that share the same characteristics as the suspected sliver polygons, thus the areal path of the river channel becomes reduced to a line (figure 6.46 section 4) and further southwards the centreline disappears altogether (section 7). One feature within this section becomes coalesced with an area it is supposed to cut across. Thin dykes in section 8 also disappear, as do linear features in section 5. Further coverage depletion occurs in the coverages with tolerances of 30 and 50 units, especially of thin footpaths, roads or ditches (figure 6.47 and figure 6.48). In the final case, very few narrow linear features exist, as they have been swept away by the excessive tolerance value.

So much for the effects of such tolerance values on the 1:50,000 coverages. The question that follows is whether similar tolerances will have any effect on the combination of coverages from a variety of scales. A triple overlay was thus created, using one coverage from all three scales. The 1:10560 and the 1:50,000 inputs into the overlay were taken directly from the Ordnance Survey map sheets, whilst the 1:25,000 coverage contained additional information from ground surveys, although this was limited to the specific area of Morton Fen within the coverage. Three geometric tolerances were applied to the triple overlay. These were all of the area divided by perimeter variety, with values of 10, 20, and 30 units respectively. These results are shown in figures 6.49 to 6.51. The overlay of the three coverages highlights the problems associated with overlaying and integrating data from different scales, as small deviations in the 1:50,000 coverage are stretched out and enhanced in the overlay, making the sliver polygons that arise between the scales difficult to eliminate.

The first tolerance of 10 units resulted in the retention of many small unwanted areas, especially on the boundary of roads and water features. The boundaries of Morton Fen in section two and three have also experienced severe multi-representation. As with the 1:50,000 case therefore, this tolerance cannot be regarded as suitable for the elimination of sliver polygons in this coverage context. A higher sliver tolerance of 20 units was then applied to the coverage. This produced a more favourable result, removing the majority of parallel lines representing one feature, but as a consequence small thin areal features are beginning to disappear regardless of whether they are erroneous or not. Extreme examples of this can be found in section 8 of the coverage, although cases exist throughout the coverage. The tolerance of 30 units (figure 6.51) results in further coverage depreciation, with superfluous polygons being lost from section 8 and along the southern linear features. There has also been further node collapse throughout the coverage, and urban areas in the west of the coverage has suffered loss of detail.

Although similar tolerances produce similar effects in both the original test data of the 1:50,000 coverages and the combined three scale coverage, in the latter case, the tolerances have the effect of reducing the quality of the larger scale data to that of the smaller scale material. Thus if the level of quality inherent to the scale of each of the UNION inputs is to be maintained, the choice of a tolerance value becomes difficult when it is applied in such a blanket and geometric fashion.[28]

The accumulated areas of all the polygons within each coverage are compared with the largest scale coverage contained within the UNIONed threesome (table 6.10).[29] As expected, the number of polygons contained within each toleranced coverage does not match the number held within their original counterpart (table 6.10).

This is due to a combination of the UNION procedure, which will have increased the polygon number as a consequence of its functionality, and also to the subsequent use of the elimination routines, which serve to reduce polygon numbers. The closer the approximation to both the original number of polygons and to the areal total, then the closer the *eliminate*d coverages might be to resembling the original. As can be seen from table 6.10, the tolerances of 10 and 20 units bring the UNIONed coverage closest to the original (as defined by the 1:10560b coverage). The effect of the 30 unit tolerance is to remove the number of polygons

---

[28] An alternative method would be to attach importance to the geometry of one coverage. Thus areas of conflict could perhaps be addressed in the first instance by prioritising the various data inputs according to assumed quality, of which scale might be a factor.

[29] Of the three component coverages that make up the combined coverage, the 1:10560b coverage was selected as being the most suitable for comparison, as it has the largest original area and the greatest number of original polygons.

to less than the original number of polygons, thus implying that a degree of detail has been lost.

Table 6.10 A comparison of the number of polygons and areal extent of the triple overlaid coverages with the originals following the application of *eliminates* based on area/perimeter.

| Original value | | Area ($m^2$) | No. of polygons |
|---|---|---|---|
| 1:10,560b | | 50030194.344 | 1119 |
| UNIONed value | | 50176792.451 | 7144 |
| | Area/Perimeter tolerance | | |
| 1:10,560b+ 1:25,000+ 1:50,000b | 10 | 50102145.945 | 1364 |
| 1:10,560bb+ 1:25,000+ 1:50,000b | 20 | 49867801.680 | 993 |
| 1:10,560b+ 1:25,000+ 1:50,000b | 30 | 49786763.301 | 914 |

In terms of guidance for users, assessing the comparative difference between the areal dimensions of coverages and the number of polygons contained within each coverage, is a relatively simple task, and yet it could provide a rudimentary guide as to the suitability of a particular tolerance type and value. A time saving method which would allow the user to probe several scenarios without the processing time incurred in geometrical overlay could be obtained by experimenting with various tolerance sizes within the tables component of the system. This would give an indication of the number of polygons that would be removed and of the loss in area that would result, however such a method would give no indication of the geometrical changes that were taking place within the coverage and could therefore only be used as an exploratory tool.

**6. 14. Conclusions.**

This chapter has illustrated the wide range of spatial data types entering a GIS, and the errors associated with the collection and acquisition of each type. The consequences of overlaying data from a wide variety of sources has been highlighted using five coverages, each supposedly representing the same features in space. This example exemplified the common error problems associated with the overlay of digital cartographic data sets. In the first instance the number of polygons present within the coverage increases dramatically. Secondly, the distribution of polygon size showed that these new polygons were very small, resulting from slight inaccuracies between the coverages.

Methods for removing such sliver polygons were examined, and their effectiveness evaluated. The routines utilised were geometric in nature relying only on the size and the shape of sliver polygons. These routines were examined in the first instance as they are the primary way of sliver error elimination within the majority of GISs today.[30]

Geometric techniques are very crude, in that they take no account of the information held within the sliver polygons themselves. The removal of sliver polygons can be more intelligently determined by making use of such information. Chapter Seven will look at enhanced methodologies for sliver polygon removal and reassignment.

---

[30]This is not an opinion limited solely to ARC/INFO, but Prime's System 9, Sysscan, and IBM's GIFIS to name a few also advocates the same procedure.

## Chapter Seven - Some solutions to the sliver polygon problem.

### 7. 1. Introduction

Chapter Six looked at a particular overlay problem; that of overlaying two or more coverages that are supposed to hold the same spatial boundaries. No aspect of temporal change between coverages is involved. Differences between the coverages are merely put down to mismatches arising as a consequence of alternative cartographic production techniques, and thus the removal of small sliver polygons is very much a geometrical process. In this chapter, the context of overlay will be extended in two ways. Firstly to include the superimpostion of coverages in which spatial units are expected to change and secondly to look at the sliver polygons in terms of the small piece of space they represent.

The techniques used to effect such analysis will be limited to those available within the ARC/INFO software package, as for most users the tools their package comes with will be the only one afforded to them. In addition to looking at the way error in coverages incorporating temporal change can be removed, this chapter also looks at improving the way land use change can be evaluated through time, thus introducing new issues of versioning and the maintenance of a temporal GIS, within a system that in many respects was not designed to cope easily with such problems.

One of the problems of monitoring resource data is that the data are in effect a sub-optimal version of reality, relying on just geometry, attributes and geo-codes, to describe the circumstances they find. Thus in this version of the world, no fuzzy boundaries between categories are allowed to exist. Only absolute boundaries are allowed. This creates problems for data which might represent areas, but be interpolated from points, or for categories which share a 'mixed' zone with their neighbouring category.

In this chapter, an assessment of error alleviation is made in terms of sliver polygon reduction. This can be evaluated through the geometry of the features or the attributes they represent. The tools that can be used in ARC/INFO are severalfold. Firstly, there is the default operation (as was used in Chapter Six). This relies on the removal of small slivers on the basis of their area to perimeter ratio. Such removals are supplied and calculated by ARC/INFO, with the areas earmarked for removal being attached onto other zones of

ARC/INFO's choosing. Secondly, an enhanced analysis can be performed. This might consist of a tolerance set by the shape of the polygon (in this case the area to perimeter ratio), plus the added information provided by the attribute the polygon represents. Although this method might improve the selection procedure for the areas that are small and erroneous, it still suffers from the problem of what happens to areas once they are denoted as having failed the retention test, in that such methods still utilise the *eliminate* tools offered by ARC/INFO. One might argue that if areas have failed the test, they are no longer of any consequence, however, as these shapes are more than pieces of geometry, the space they represent may end up being wrongly classified.

An alternative approach is to ignore the standard tools offered by ARC/INFO and to develop methodologies which do not use the *eliminate* option. Such methods require a reorientation of the way slivers are viewed. Rather than seeing them as areas requiring identification and 'elimination', this approach deals with the fuzziness of data by recognising that sliver polygons should be merged with the most suitable of their neighbours. This methodology would utilise INFO routines and the *dissolve* option within ARC/INFO. This works by removing lines that have the same attribute on either side. In the case of classified data this would be the classification attribute. The drawback of this type of technique is that it requires prior knowledge about the nature and behaviour of the attributes.

Several options exist when using this type of approach. Firstly, a matrix of allowable movement can be created, thus change is only allowed in certain directions, which are defined by a set of predefined rules. Secondly, spurious polygons failing a geometric test can be merged with their neighbours on the basis of probability matches. Again these are defined by the user for the data at hand. Thirdly, if changes through time are being modelled, a best fit history approach can be implemented. This looks at data in terms of time, rather than in terms of the current or past classification only. Sliver polygons are merged with bounding areas that have 'acceptable' histories.

The implementation of such routines is not as simple as it might first appear. Aside from problems of data structure and the nature of the database query language, there is the implication of the way polygons are processed. The very nature of the system relies on the sequential processing of information. Thus, if one has n polygons then there are $n^2$-n possible results given any one start polygon. In choosing a starting polygon, for merging with any of its neighbours (some of which may require merging as well), the user effectively enforces a spatial priority which will affect the outcome of the analysis.

Coverage priority is a further problem facing the user. If coverages for overlay reflect different spatial units then one of them may be set as the accurate baseline, to which other coverage layers converge. Alternatively, if coverage layers reflect different versions of the same space at different times, one of the data layers will have to be set as the base time onto which changes are overlaid. Thus $t_0$ (where t represents time) might be set as the original time, whilst further time snaps $t_{(n+1)}$ might be regarded as less accurate in terms of bases geometry. Solving such problems requires the ability for spatio-temporal modelling. Yet at most, the existing GIS technology can only identify geometric change. It cannot easily identify *when* change occurred merely that it has. These problems will be discussed in greater length as they arise in various stages of the assessment.

**7. 2. Classification of land cover.**

Attributes based on a classification of space usually rely on a series of rules attached to classical classification/categorisation theory. The first is that categories have clear boundaries and are defined by common properties. Thus the categories will be distinctly different, and the boundary between them will be easily interpreted. This is true for some geographic data, but certainly not for all categorical geographical data. Secondly, properties defining the category should be shared by all members (Wittgenstein, 1953), thus all members of the set will have equal status as category members. In terms of geographical data this will depend on the classification systems used and on the perception and requirements of the user.

Deriving a categorisation system is essentially a subjective operation, depending initially on the purpose to which the classified material is to be put. Classification techniques will also depend upon culture, as differing cultures will utilise different semantics and schemes for categorisation. Lakoff (1987) suggests that some schemes may be derived by analogy and the perceived relationships features have to each other, thus properties relevant to the description of categories are interactional properties, in that they are only characterised by the interaction of human beings as part of their environment.

Alternatives to the classical set or category are those which do not have clearly defined boundaries. In a classical set, features are either in the set or outside it, thus they either have a membership value 1 (in the set) or they have a membership value of 0 (outside the set). Zadeh (1965) defined fuzzy sets in which additional values that extend between 0 and 1 are allowed. This system of categorisation acknowledges that some categories have clear boundaries, but that within those boundaries there are graded effects, thus some category members are better members than others. In many ways this applies to much of the interpolated geographical data collected for natural categorical features such as soils or vegetation, although the mechanisms currently available for the storage of such data forces the user to classify them in classical terms.

Once a classification scheme has been decided upon, the land cover needs to be sampled to assess how the real world might be sub-divided into that scheme. Land cover can be assessed in terms of a series of point samples which are then interpolated, or by zones. Each method will have its associated problems, thus point sampling will leave the classification prone to sampling and interpolation errors, whilst attaching an attribute to a zone will assume that a particular land use is homogeneous to the entire zone, which may not be the case in reality. Zones take on several forms. Natural zones are those defined by the extent of category they represent. Artificial zones include pre-determined collection zones such as cells or administrative units which remain constant and are used for the ease of surveys. In both cases, homogeneity will be linked to the resolution of the data and the type of classification that is used.

Before a classified land cover map can be used, some assessment of its accuracy will be required.

> "unless some examination of the likely errors has been made and vigorous checking has taken place, the data are worse than useless - the discovery of a single error will then render the rest of the data unreliable except at the most gross and superficial level."
>
> Rhind and Hudson (1980, 51)

The accuracy of land cover classifications depends upon the resolution of the sample, the frequency of the sample, the date of the categories and the scale at which the data are to be held. Some errors are much worse than others. For example, the misclassification of a school as a university is far more acceptable than classifying it as a manufacturing concern.

Once obtained, land cover data can be put to several uses. Rhind and Hudson (1980, 16) cite the most common ones, and these can be modified for GIS applications to produce the following major typification:

- To provide area and or volume descriptions of the land uses within defined regions, usually on a statistical basis, or to define the land use at a particular point in space.

- To carry out statistical analysis to observe relationships within an area.

- To produce comparisons between areas at the same moment in time.

- To overlay data for an area at different moments in time and to produce rates of change in land use. Thus spatial units remain more or less the same and time varies.

- To overlay different sets of data for the same area to determine co-incidence of land use and/or to observe the relationships between several sets of data. Thus temporal units remain the same and spatial units vary.

This analysis will focus on the way the last two types of land cover assessment can be carried out within GISs. The problems associated with the uses of categories within GISs have also been investigated. These include the problem of time freezes, in which the real world is isolated within specific time constraints. For instance in terms of an agricultural classification of land use, the time of the year in which the survey is carried out will determine the classes portrayed on the map. This raises the question of which classification is actually correct for the year as a whole. A method of circumnavigating this problem is to change the resolution and merely classify zones as containing grain, and ignore the fact that the land use in that particular zone may have two different grain crops per year, in addition to being classified as bare, stubble and ploughed for some of the year. Such problems become more acute as satellite imagery, with its ability for continuous update, begins to provide more data for inclusion in land cover assessments.

**7. 3. From categorisation to GIS storage.**

The way a GIS stores the classifications of individual zones will depend upon the type of data structure the system employs. Traditional relational data base structures will represent the category a particular polygon falls into as an extra field within its tabular structure. This will be stored in the same way as other attributes relating to the polygon such as its area, perimeter or identification number. This 'class' attribute can also be used and manipulated in the same way as other attributes. Newer object oriented systems such as Smallworld, utilise a different approach to classifying categorical zones in which the classifications are an integral part of the object. This facilitates versioning and temporal issues. In ARC/INFO feature attributes may be used as a prerequisite for the elimination or retention of particular polygons holding certain classes, or alternatively they can form the basis for selective merging.

**7. 3. 1. Using feature attributes.**

Thus, in addition to the area and perimeter items in the polygon attribute table, user identifiers may also be added. These may then be used as the basis for logical operations to eliminate unwanted sliver polygons. When coverages are overlaid on top of each other using geometric overlaying techniques, both the graphic geometry and the feature attributes merge. In the newly created feature attribute table, there exists a lineage of the attributes attached to particular points in space. If, for example, one coverage depicts population density and another land use, then any particular areal unit within the resultant coverage will have a value for its population density and also for its land use, both derived from the two input coverages to the overlay process. This combination of attributes can then be queried using logical expressions, and the polygons corresponding to undesirable subsets can be removed.

**7. 3. 2. Attributes as a basis for sliver polygon removal.**

In a combined coverage representing change between two time periods, making the elimination of an error polygon successful requires three factors. Firstly, that the polygons of error are removed. Secondly, that small original polygons remain, and thirdly, that polygons of change remain. To ensure that the correct polygons are retained and their rivals removed the characteristics of each of the groups need to be identified and searched for within the data set.

Error polygons can be identified by their size, shape, new identification class and their old identification class, whilst small original polygons can be identified by their attribute, in that they will have the same class value in both the original and overlay coverage. In this way they are mutually exclusive from the error polygons.

Variations in polygon characteristics can be elucidated from the class attributes and certain rules can be attached to specific class values allowing them to remain after there has been an elimination performed on the basis of a geometric component of the data. The use of attributes will depend upon the degree of mutual exclusivity that exists between the features for retention and those for removal. In a case where the two coverages for overlay were intending to represent the same geometry, some feature combination may be inconsistent, and any areas defined as such should be removed. For example, if an area is defined as being a road and an agricultural area, a discrepancy has arisen, or if a lake is defined as having a particular forest type then it is likely that there has been a boundary error and the inconsistent area needs to be removed. Alternatively, if the coverages to be combined consist of the same polygon information, with each containing different attributes, as might be the case for fixed boundaries such as administrative units or field boundaries, then each unit should have an output polygon user identifier equal to both the two input values (figure 7.1). If any areas exist in which for example the names of two adjacent districts appear then the polygon needs to be regarded as a potential sliver error.

## 7. 4. Testing the use of categories for error elimination.

To investigate the use of an eliminator based on polygon type rather than polygon size, a simple test coverage was created. This was used in preference to the Fenland coverage for several reasons. Firstly, to have individually coded and verified the authenticity of all the polygons in the Fenland coverages would have been extremely time consuming. Secondly, land use classifications were not available for maps of such large scale, and the only easily available source would have been an interpolation of the land use maps of the 1960s. Unfortunately, the area in question is not covered in the University of London's printed maps. The only way to have obtained the information would have been to derive it from the survey material or to have conducted a survey, neither of which were practical. Finally, it was felt

## Table 7.1 Polygon dimensions of test coverage

| AREA | PERIM | TS-ID | CLAS |
|---|---|---|---|
| 9.600 | 13.600 | 1 | 3 |
| 14.600 | 16.102 | 2 | 2 |
| 0.600 | 3.200 | 3 | 1 |
| 23.800 | 21.902 | 4 | 3 |
| 14.400 | 19.600 | 5 | 4 |
| 6.000 | 12.400 | 6 | 2 |
| 0.180 | 1.800 | 8 | 1 |
| 2.820 | 10.600 | 7 | 2 |
| 3.000 | 30.400 | 9 | 5 |
| 25.000 | 20.000 | 12 | 6 |
| 2.960 | 30.000 | 13 | 7 |
| 1.840 | 6.200 | 11 | 2 |
| 20.700 | 22.600 | 10 | 3 |
| 11.750 | 14.600 | 14 | 3 |
| 0.250 | 2.000 | 15 | 2 |
| 12.000 | 15.000 | 16 | 4 |
| 0.500 | 3.000 | 17 | 1 |
| 150.000 | 243.005 | | |

## Table 7.2 Transformation shift for the test coverage

| IDTIC | XTIC | YTIC |
|---|---|---|
| 1 | 0.000 | 0.000 |
| 2 | 15.000 | 0.000 |
| 3 | 0.000 | 15.000 |
| 4 | 15.000 | 15.000 |

| IDTIC | XTIC | YTIC |
|---|---|---|
| 1 | 0.200 | 0.200 |
| 2 | 15.200 | 0.200 |
| 3 | 0.200 | 15.200 |
| 4 | 15.200 | 15.200 |

that a smaller coverage could be more easily processed and monitored than a larger and more complicated one.

The test coverage was created in ARC/INFO containing all the major components of a problematic coverage for overlay. This included several long thin features (represented in reality by roads, ditches, canals or footpaths) and in addition, several features that were small in magnitude, without being long and sinuous in shape. In reality these would be represented by housing, small agricultural investments, and in many cases small zones of change (figure 7.2 and table 7.1). The coverage was created using a series of x,y coordinates, in preference to digitising, as this ensured that there would be no need to accommodate overshooting or undershooting arcs following coverage creation. After being topologically built, the coverage was then labelled and appropriate land use values were assigned to each polygon (figure 7.3).

Two suitable coverages for overlay were then created. The first consisted of the same features that were present in the original map, but with a transformation shift in two axes (table 7.2). Thus in this case there was no change in the features held within the coverages, only in their relative positions, with the transformation shift creating areas of error which were identical in size and shape to areas of value. This was overlaid with the original test coverage in ARC/INFO, using the intersect command, and the graphic result is shown in figure 7.4. Table 7.3 illustrates the tabular result of the combined attributes, which contain both the new polygon identifiers and the identifiers of that space prior to integration. This test overlay was seen as suitable for illustrating the situation that might arise if two maps intending to represent the same areas suffered from slightly differing representations. On occasion it might be necessary to perform such an overlay, if for example the first coverage had four attributes pertaining to each polygon (A, B, C and D) and the second coverage had three attributes pertaining to each polygon (B, F, and G) the overlay would be necessary for an integration of all the available data held on each zone. In this case we have a situation where the spatial units do not change, but the categories they hold within their bounds might. Another situation in which this might occur is in the case of constant spatial units being observed at differing times.

## Table 7.3 Results of the combined attributes

| Area | Perimeter | Overlaid coverage id | 1st coverage id | 2nd coverage id | 1st coverage class | 2nd coverage class |
|---|---|---|---|---|---|---|
| 0.400 | 4.400 | 1 | 0 | 1 | 0 | 3 |
| 0.900 | 9.400 | 2 | 0 | 2 | 0 | 2 |
| 0.120 | 1.600 | 3 | 0 | 3 | 0 | 1 |
| 2.140 | 21.800 | 4 | 0 | 4 | 0 | 3 |
| 0.960 | 10.000 | 5 | 1 | 0 | 3 | 0 |
| 8.280 | 12.800 | 6 | 1 | 1 | 3 | 3 |
| 0.600 | 6.400 | 7 | 2 | 1 | 2 | 3 |
| 13.039 | 15.312 | 8 | 2 | 2 | 2 | 2 |
| 0.280 | 3.200 | 9 | 3 | 2 | 1 | 2 |
| 0.320 | 2.400 | 10 | 3 | 3 | 1 | 1 |
| 0.160 | 2.000 | 11 | 4 | 3 | 3 | 1 |
| 21.660 | 21.102 | 12 | 4 | 4 | 3 | 3 |
| 0.381 | 4.395 | 13 | 4 | 2 | 3 | 2 |
| 0.961 | 10.010 | 14 | 2 | 5 | 2 | 4 |
| 0.639 | 6.790 | 15 | 4 | 5 | 3 | 4 |
| 0.960 | 10.000 | 16 | 4 | 6 | 3 | 2 |
| 0.240 | 2.800 | 17 | 0 | 6 | 0 | 2 |
| 0.320 | 3.600 | 18 | 5 | 1 | 4 | 3 |
| 12.480 | 18.800 | 19 | 5 | 5 | 4 | 4 |
| 0.240 | 2.800 | 20 | 6 | 5 | 2 | 4 |
| 4.800 | 11.600 | 21 | 6 | 6 | 2 | 2 |
| 0.060 | 1.000 | 22 | 6 | 8 | 2 | 1 |
| 0.900 | 9.400 | 23 | 6 | 7 | 2 | 2 |
| 0.120 | 1.600 | 24 | 0 | 7 | 0 | 2 |
| 0.080 | 1.200 | 25 | 8 | 5 | 1 | 4 |
| 0.040 | 1.000 | 26 | 8 | 8 | 1 | 1 |
| 0.080 | 1.200 | 27 | 7 | 8 | 2 | 1 |
| 1.800 | 9.800 | 28 | 7 | 7 | 2 | 2 |
| 0.360 | 4.000 | 29 | 1 | 9 | 3 | 5 |
| 1.600 | 16.400 | 30 | 5 | 9 | 4 | 5 |
| 0.060 | 1.000 | 31 | 8 | 9 | 1 | 5 |
| 0.940 | 9.800 | 32 | 7 | 9 | 2 | 5 |
| 0.040 | 0.800 | 33 | 0 | 9 | 0 | 5 |
| 0.040 | 0.800 | 34 | 9 | 0 | 5 | 0 |
| 1.000 | 10.400 | 35 | 9 | 12 | 5 | 6 |
| 0.040 | 0.800 | 36 | 9 | 13 | 5 | 7 |
| 0.160 | 2.000 | 37 | 9 | 11 | 5 | 2 |
| 1.760 | 18.000 | 38 | 9 | 10 | 5 | 3 |
| 0.460 | 5.000 | 39 | 0 | 10 | 0 | 3 |
| 1.960 | 20.000 | 40 | 12 | 0 | 6 | 0 |
| 23.040 | 19.200 | 41 | 12 | 12 | 6 | 6 |
| 0.960 | 10.000 | 42 | 13 | 12 | 7 | 6 |

Table 7.3 continued

| Area | Perimeter | Overlaid coverage id | 1st coverage id | 2nd coverage id | 1st coverage class | 2nd coverage class |
|---|---|---|---|---|---|---|
| 0.580 | 6.200 | 43 | 11 | 13 | 2 | 7 |
| 1.260 | 5.400 | 44 | 11 | 11 | 2 | 2 |
| 0.420 | 4.600 | 45 | 10 | 11 | 3 | 2 |
| 18.480 | 21.800 | 46 | 10 | 10 | 3 | 3 |
| 1.800 | 18.400 | 47 | 10 | 13 | 3 | 7 |
| 0.040 | 0.800 | 48 | 0 | 13 | 0 | 7 |
| 0.040 | 0.800 | 49 | 13 | 13 | 7 | 7 |
| 0.860 | 9.000 | 50 | 13 | 14 | 7 | 3 |
| 0.100 | 1.400 | 51 | 13 | 15 | 7 | 2 |
| 0.800 | 8.400 | 52 | 13 | 16 | 7 | 4 |
| 0.160 | 2.000 | 53 | 13 | 17 | 7 | 1 |
| 0.100 | 1.400 | 54 | 0 | 17 | 0 | 1 |
| 0.460 | 5.000 | 55 | 14 | 13 | 3 | 7 |
| 10.330 | 13.800 | 56 | 14 | 14 | 3 | 3 |
| 0.160 | 2.000 | 57 | 15 | 14 | 2 | 3 |
| 0.090 | 1.200 | 58 | 15 | 15 | 2 | 2 |
| 0.060 | 1.000 | 59 | 16 | 15 | 4 | 2 |
| 10.540 | 14.200 | 60 | 16 | 16 | 4 | 4 |
| 0.260 | 3.000 | 61 | 17 | 16 | 1 | 4 |
| 0.240 | 2.200 | 62 | 17 | 17 | 1 | 1 |
| 0.400 | 4.400 | 63 | 16 | 14 | 4 | 3 |
| 0.400 | 4.400 | 64 | 0 | 16 | 0 | 4 |
| 0.040 | 0.800 | 65 | 13 | 0 | 7 | 0 |
| 0.960 | 10.000 | 66 | 14 | 0 | 3 | 0 |
| 1.000 | 10.400 | 67 | 16 | 0 | 4 | 0 |

**Table 7.4 Polygons dimensions of the coverage of change**

| Polygon Change-id | Class-id |
|---|---|
| 0 | 0 |
| 21 | 7 |
| 1 | 2 |
| 29 | 4 |
| 2 | 5 |
| 3 | 1 |
| 4 | 2 |
| 18 | 6 |
| 20 | 7 |
| 36 | 6 |
| 22 | 7 |
| 28 | 6 |
| 5 | 2 |
| 27 | 3 |
| 6 | 3 |
| 8 | 7 |
| 9 | 6 |
| 12 | 9 |
| 13 | 6 |
| 11 | 9 |
| 10 | 8 |
| 19 | 7 |
| 23 | 2 |
| 14 | 2 |
| 15 | 7 |
| 16 | 2 |
| 17 | 7 |
| 24 | 2 |
| 25 | 8 |

The second overlay coverage consisted of similar features to those present in the original map, but with slight additions and alterations as might be found as a consequence of change (figure 7.5, table 7.4). For example long thin areas 18, 26 and 28 have been added as have small zones 22, 20, 19 and 25. Furthermore the road and footpath features, whilst remaining in the same location have taken slightly different shapes, as would occur due to poor digitising, the effects of routine tolerancing or the effects of some transformation algorithm. Thus 'errors' and 'change' are both built into the system.

As in the first case, the two coverages were overlaid in ARC/INFO, using the intersect command, and the result is shown in figure 7.6. This might be compared to a situation where boundaries between the two coverages under consideration are different. One instance of this might be two completely differing sets of information. For example an overlay of soil categories and population density. In this case spatial units vary, but temporality remains constant. Alternatively, the situation could apply to the evolution of a particular spatial feature, for example the land use. In this case, both the spatial units and the temporal units may change. In this set of experiments, the sliver polygons have been deleted using the standard ARC/INFO *eliminate* tool, with certain user extensions. Thus the methodology is simply an enhanced version of what was carried out with the crude areal tolerances in Chapter Six.

### 7. 4. 1. Using overlay one: spatial units constant.

In the first test the land use classification schema was used as a basis for error removal by suggesting that all areas in which the original and overlay coverages did not share the same attribute should be eliminated. This produced the result in figure 7.7. This can be compared with a geometric *eliminate* performed on the basis of the AREA/PERIMETER ratio (figure 7.8). Although less prone to the movement of nodes than the geometric case, the thin road features overlaying each other with a high degree of error have still been eradicated completely. Thus irrespective of whether geometry or feature is used, the *eliminate* option allows the user to select what should be removed, but still provides no means of control as to where the unwanted polygons should go.

### 7. 4. 2. Error versus change.

If the intention of the analysis is to investigate change using an overlay of two coverages from different time periods, then assuming change has occurred, the two maps will undoubtedly differ, with change often taking on the appearance of small polygons, not too dissimilar to sliver polygons. Consider for example an investigation into land use change, where two coverages from different time periods are overlaid in an attempt to gain some insight into the temporal variation of land use. Coverages of essentially the same outlines will be overlaid on top of each other. How will the user be able to differentiate between veritable areas of change and spurious polygons thrown up as a side effect of using the computerised techniques? In situations when hundreds of such areas may be created, distinguishing each case individually would be an impossible task, as it would nullify the benefits of using the software in the first place. Some computerised procedure is therefore desirable. Although it would be difficult to pre-determine the qualities that authentic small areas of change would exclusively possess, there is potential for the user to contribute some intelligence to the re-assignment of areas on the basis of the attribute qualities attached to them.

### 7. 4. 3. Using overlay two: spatial units and temporal units changing.

So far the test coverage has dealt with the overlaying of two supposedly identical coverages. However, in the majority of analyses, it is more likely that the two input coverages will differ in some way, and the extent of that variance will be the mainstay of the analysis. In the case of land use, areas may have simply changed their attribute, but it is also likely that some areas will disappear, whilst other new ones will be have been created.

The mutual exclusivity of feature attributes can be applied in this context using a variety of methods. All such techniques work in tandem with an areal tolerance, but with the added proviso of removing some areas from consideration for automatic rearrangement. For example if an area is less than the areal tolerance set, but has a particular feature code, then it can remain as a polygon within the coverage. Such simple rules can be extended to a more complex set.

Consider the following overlay shown in figure 7.9. For those areas which pass the areal tolerance set, then the classification in the new coverage is adopted as being the more accurate of the two, and any dissimilarities between that and the original can be regarded as change. Thus both case A and case B show either a change (from B to C and from C to B respectively), or alternatively an error. For the sake of argument, it can be assumed that classification B is a nuclear power station and that classification C relates to sheep grazing. According to prior knowledge of the data, it could be said that it is unlikely that any polygon which shows change to or from a nuclear power station is depicting true change, any such areas of change are more likely to be errors. If change is unlikely, then the old classification (that belonging to the first coverage) can hold true, in preference to the new one. Alternatively, if one change is more likely than another, then for those matches, the likely class will always predominate. If in the hypothesised example, it is believed that new extensions may have been added to nuclear power stations, but no power station premises have been torn down, then in all cases the area pertaining to the B classification will predominate. Thus if the old class = B and the new class = C, the old class will be retained. If alternatively, old class = C and new class = B, then the new class will be retained. To ensure that only areas of change are enacted upon, only areas in which the geometry of the area takes it below the threshold can be selected and furthermore of these, only those areas in which the old class does not equal the new class should be considered. To evaluate the effectiveness of such rules, several trials were performed on the overlaid coverage illustrated in figure 7.6.

The first test looked primarily at the problem of polygons with a small areal extent and also those with a narrow shape. In the test coverages such features are either error polygons, or polygons depicting housing, roads or footpaths. To try and force the software to concentrate on polygons which display the characteristics of error polygons, but which have not been identified as housing or road features, the test reselected all those polygons whose features were not equal to the attribute for road or housing and whose areal extent fell below the value defined by the tolerance, as shown in table 7.5. Three tests were carried out. Initially, the 'class' of the polygon was seen as the old class value. This produced results illustrated in figure 7.10. As can be seen however, any of the changes which fall within the areal tolerance are removed. Using the new class value as the attribute for consideration would have reversed the position by not allowing any of the selected features in the old coverage that fall below the tolerance to remain. If, alternatively, the 'either or' option is used, then very little of the coverage is altered with respect to the overlay coverage (figure 7.11).

If it is accepted that roads are allowed to move their absolute position, and that this may entail a new road or a wider road, then only retaining the location of the old roads would be unjustifiable. The alternative is to retain only the new roads, or a combination of both old and new and to accept that a certain amount of error will occur. This is demonstrated in figure 7.12 where a *dissolve* has been performed on the coverage shown in figure 7.11 using only the new class value. In this case therefore polygons 19,24 and 28 will represent changes in the road system, which in effect have not occurred. Such a methodology is however still preferable to a purely geometrical one, as at least some areas worthy of protection may be saved as a result of user intervention.

Table 7.5 Rules for elimination of sliver polygons.

| |
|---|
| ELIMINATE RULES<br>Rule 1: If the area to perimeter ratio of the polygon is less than 0.1 AND the old class of that polygon is NOT housing OR road OR footpath, then eliminate the area.<br><br>Rule 2: If the area to perimeter ratio of the polygon is less than 0.1 AND the new class of that polygon is NOT housing OR road OR footpath, then eliminate the area.<br><br>Rule 3: If the area to perimeter ratio of the polygon is less than 0.1 AND the new OR the old class of the polygon is NOT housing OR road OR footpath, then eliminate the area. |

To assess the effectiveness of an attribute operation, as opposed to a geometrical one, a comparison between the two operations needs to performed. Thus four further coverages were created using the same datasets, in which the errors were toleranced out using values corresponding to the ratio of the AREA/PERIMETER and the ratio of the PERIMETER/AREA.

The first three coverages involved area/perimeter techniques carried out in accordance with the advice of the suppliers, and these produced the results in figures 7.13, 7.14 and 7.8. The tolerances used were 10, 0.5 and 0.1 respectively. These tolerances either did not remove all of the sliver polygons (figure 7.14), or alternatively all the slivers were removed, but at considerable expense to all of the intentional features that resembled slivers such as the roads (figure 7.13).[1] Similar results can be obtained by using the perimeter of a polygon divided by the polygon's area as an indicator of sliver shape. In this case (figure 7.15), the ratio of perimeter to area was given as 10:1 and anything below that was deemed to be a sliver error. If this is compared with figure 7.8, then the results are comparable in error removal and detail reduction.

If these coverages, along with those derived using the attribute method (figure 7.12) are compared to the change that has occurred (figure 7.5) then the combination of attribute rules is still more successful than the geometric routines, as in situations of doubt information is retained rather than being removed.

### 7. 5. The fate of *eliminate*d polygons.

In the previous experiments, basic rules to govern the direction that some combinations of classes below a particular threshold could follow were investigated. However, once the changes were made, the re-assignment of the area in question to one of its neighbours always assumes the same pattern, with the largest line making up the boundary of the area in question being *dissolve*d, and the polygon in question being merged with one of its counterparts. This chapter seeks to look at that merging operation in greater detail, and to assess possible alternatives to the random geometric coalition that currently abounds.

The methods and concepts behind the elimination of spurious sliver polygons have their foundation in regarding such areas as geometric abnormalities. If such anomalies can be swept away, then somehow, the coverage will become accurate. Rather than furnishing the user with the expectation that such intrusions into cartographic purity are to be expected, sliver polygons are often seen as irritating by-products of a powerful analysis procedure, and the

[1]When the results of the eliminate operations were looked at in detail, it was noticed that polygons within the tolerance value still existed. This can only be attributed to the sequential nature of the polygon elimination routine, as not all of the polygons are processed in unison, but rather each relies on the result of its predecessor.

attitude towards them in proprietary GISs is quick removal, usually involving geometric solutions.

The sliver polygon is rarely viewed as an areal entity in its own right. Instead, it is the boundaries defining the polygon that are seen as the problem, and the solutions to the elimination of such slivers rely on the removal of one, or some, of the bounding lines. In removing one of the bounding lines of the polygon, the area that was once held within becomes exposed and the small sliver will merge its geography with one of its adjacent neighbours. The method for reassigning such areas within GISs usually relies on a merge with one of the surrounding areas based on a chosen geometrical property, for example in ARC/INFO ...

"The longest arc and label points of each selected polygon are eliminated."

ARC/INFO PC overlay users Guide, section 4.

Therefore the fate of such polygons is in many respects random, as they are not assigned to the most suitable neighbouring area, but on the magnitude of one of its boundaries. In a geometrical sense, with no other option possible, this type of routine is preferable to a random assignment as it at least relies on a modicum of intelligence. The basis for this is that the sliver polygon will have most in common with the region with which it shares the longest boundary. Unfortunately this method does not take account of the fact that one boundary may be only fractionally longer than the other, or that the longest boundary may not be the most suitable.

The most significant problem associated with the nearest largest neighbour approach is that areas become mis-assigned. These are usually borderline areas, which are often the ones of the most interest to geographers. It is the designation of the boundaries that usually causes the most conflict, especially in categorical or 'fuzzy' data, such as soil maps or vegetation distribution.

Examples of the neighbourhood effect have been drawn from the test coverages previously used for feature attribute analysis. These coverages were chosen as their size and simplicity make them capable of easily illustrating the way in which polygons failing the tolerance test are re-assimilated into the coverage.

In all of the above cases areas have been lost. In the case of reductions on the basis of feature attributes the user has exercised a greater degree of choice over which areas were to be removed, but in terms of deciding as to where these unwanted polygons should go, the user has as little control as with the purely geometric cases, as unwanted areas are lost to the neighbour with which they bear the longest shared boundary.

An example of how this routine works is given with respect to the coverage in figure 7.8. If this coverage is compared with the raw overlaid coverages (figure 7.6) and the original coverage of change (figure 7.5) an alteration in some of the field boundary shapes becomes apparent. For example the boundary between what was polygon 22 and 29 in the overlay coverage (figure 7.6) has been removed. This has come about in the following manner. As polygon 22 fell below the tolerance value, it was selected for elimination facilitated by the dropping of one of its boundaries. As the boundary between polygon 22 and polygon 29 was the largest that made up the outline of polygon 22, it was removed and polygon 22 joined with polygon 29. Another more complex example can be seen in the removal of polygon 24. The obvious choice here would have been polygon 20, even by the longest shared boundary criterion. Instead polygon 24 has been merged with polygon 25 and consequently polygon 26 has now lost its boundary with the road. This matching would have originated with polygon 20 merging with polygon 19, but as far as polygon 24 is concerned it does not recognise its new relationship with polygon 19, thus of its other neighbours, namely polygons 23,25 and 26, the largest border occurs with polygon 25. This inability to accumulate new relationships as they are formed is due to the sequential nature of the processing routines, which does not update the topological base of neighbouring polygons. A similar situation occurs with polygon 28, as it no longer finds its link with polygon 22, which has reemerged as polygon 29.

In figure 7.12, the longest shared boundary routine is in action again, but here a different result has ensued. Polygon 24 (figure 7.11) has dropped its largest boundary and joined up with polygon 20. Polygon 28 on the other hand has joined polygon 27, rather than polygon 22. Presumably, this was a choice based on the internal ordering of the polygon structures of the polygons. As polygons are not numbered by the user to take account of this, the process can be viewed as relatively random and archaic.

### 7. 6. Alternatives to random geometric merges.

These methods throw away the traditional ARC/INFO tool of *eliminate* and rely on alternative methods for area removal. If one of the bounding arcs that make up an areal feature is 'dropped', the feature will merge with whichever of its neighbours the line borders. Thus, if a suitable neighbour can be found for the suspected sliver polygon to merge with, the line separating them simply needs to be removed. This ensures that the error polygon will merge with a neighbour of the user's choosing rather than one with which it merely shares the longest boundary. This dropping of arcs is carried out by another ARC/INFO tool called *dissolve*, which is used in conjunction with manipulations carried out by the INFO database. The way that this procedure is applied for error minimisation depends upon the use that is to be made of the data. If the user is simply after the most up-to-date coverage given changes through time, the procedures will be different to that enacted if the user wished to specifically analyse the difference between two time periods.

#### 7. 6. 1. Database error removal.

In the first of the above cited cases, differentiating between small polygons of error and change, might involve the initial assumption that sliver polygons do not exist, and any mismatches between the coverages are due to the fact that the second coverage is more accurate than the first. This would hold true for the entire coverage, with the exception of certain pre-determined cases governed by a series of rules, that limit the changes polygons can make through time. If for example, it was desired to freeze certain features from the original dataset, these could be considered to be outside the analysis.

This method is illustrated using the same overlay as before, the contiguity and compatibility of the polygon attributes were used to formulate a series of hypothetical rules which gave precedence to roads, footpaths, grazing and housing as defined in the original coverage, unless they had been altered in certain allowable directions. Otherwise, the overlay coverage was deemed to be the more accurate of the two. These rules are shown in table 7.6.

Table 7. 6. Rules which could be applied to a classified coverage to preserve certain features.

| DATABASE RULES | |
|---|---|
| Rule 1: | Anything that used to be a road and which is not now, let it become a road again. |
| Action 1: | Reselect polygons in which the old class [CLASS] is equal to 5 and copy this into the *dissolve* class item [NCLASS]. |
| Rule 2: | Anything that used to be a footpath, and which is not now, let it become a footpath again. |
| Action 2: | Reselect polygons in which the old class [CLASS] is equal to 7 and copy this into the *dissolve* class item [NCLASS]. |
| Rule 3: | Anything that used to be grazing, which is not housing, road or grazing now, let it be grazing again. |
| Action 3: | Reselect polygons in which the old class [CLASS] is equal to 6 and in which the new class [CLASS2] is not equal to 1 or 5. Calculate the *dissolve* class [NCLASS] to be equal to the old class [CLASS]. |
| Rule 4: | Anything that used to be a house which is not a road, let it be a house again. |
| Action 4: | Reselect polygons in which the old class [CLASS] equals 1 and the new class [CLASS2] is not equal to 5. Calculate the *dissolve* class [NCLASS] to equal the old class [CLASS]. |
| Rule 5: | Accept the other changes as viable. |
| Action 5: | For the rest of the polygons calculate the *dissolve* class [NCLASS] to equal the new class [CLASS2]. Perform a *dissolve* on the item [NCLASS], thus removing any boundaries between polygons of the same class. |

This method allows the user to exercise subjective control over the resulting dataset, as in some cases the original dataset will prove to be the more accurate, whilst in others the overlay dataset will be the more reliable. Without such a mechanism, it is always the new, overlaying dataset that is deemed to be the more accurate of the two, with any discrepancies between the two datasets being attributed to update. In employing such a methodology, the user is in effect creating a matrix of allowable change. In simple cases this might be a simple yes/no matrix, but in more complex cases, probabilities could be assigned to the various change options. In the above example a matrix of allowable movement would look similar to that shown in table 7.7.

Table 7.7 A matrix of allowable movement.

| To<br>From | House | M'ket G'dn | Wheat | Barley | Road | Grazing | Foot-path |
|---|---|---|---|---|---|---|---|
| House | Y | N | N | N | Y | N | N |
| Market Garden | Y | Y | Y | Y | Y | Y | N |
| Wheat | Y | Y | Y | Y | Y | Y | N |
| Barley | Y | Y | Y | Y | Y | Y | N |
| Road | N | N | N | N | Y | N | N |
| Grazing | Y | N | N | N | Y | Y | N |
| Footpath | N | N | N | N | Y | N | Y |

Thus some features would become fixed once they had been implanted into the system using the original coverage. This series of rules were applied to the test coverage in the following way. Firstly the PAT of the overlay (figure 7.10) was selected and the rules were applied to reselected portions of the database, then a *dissolve* was performed. The results are shown in figure 7.16.

If this is compared with the original coverage (figure 7.2), the change coverage (figure 7.5), the purely geometrical *eliminates* (figure 7.8 and figures 7.13 to 7.15) and the combined attribute and geometrical *eliminates* (figure 7.11), then some differences emerge. Firstly, areas have been removed which are much larger than a pre-defined tolerance, as the user believes change in that direction is impossible. Thus polygons 12, 25 and 24 in figure 7.5 have been disallowed because polygon 12 in figure 7.2 represents grazing, which in the user defined matrix of allowable movement cannot reemerge as any other agricultural attribute.[2] As with figure 7.12, the original roads and footpaths have been retained, thus falsely

[2]As with the first eliminate using feature attributes polygons 10 and 23 which were both wheat have been joined, as have polygons 26 and 5 which are both barley.

increasing the amount of land under these uses. Although this is inaccurate, it may still be deemed preferable to the more randomly based geometric wipe out.

### 7. 6. 2. Selective neighbourhood matching.

In adopting a methodology of selective neighbourhood coalescence, the user acknowledges that small suspected error polygons will have to be eliminated using a variety of techniques, but the emphasis, is on *where* those areas go, rather than *whether* they should go. If the user can direct the fate of areas to be removed, then their elimination does not induce as a great a measure of concern and uncertainty as in the previous cases, given that they will be merged with suitable areas. Admittedly, some degree of detail will be lost, but at least the user has the satisfaction of knowing that it will not be redirected into false, or random avenues. Thus this approach deals with error minimisation on 'fuzzy' boundaries, rather than error removal. Damage is minimised by relocating areas of uncertainty to probable polygon partners and neighbours.

### 7. 6. 3. Using probabilities to determine best neighbour matches.

If one considers a land use cover consisting of several classified categories, a probability matrix could be defined which would determine the probabilities of a particular land use being merged with all other land uses in that analysis. Probabilities may be given in terms of percentages which are independent of each other; or alternatively, probabilities for all land uses may be given in terms of proportions, which when summed total one (or percentages which when summed equal 100%). If a polygon fails to pass an areal tolerance test (such as an area/perimeter ratio), then the matrix will suggest which of its neighbours the polygon should merge with given the available options.

Implementing such a procedure within ARC/INFO involves several stages. Firstly the matrix has to be set up by the user within the INFO data base management system. This will require as many rows and columns as there are land use categories. The areal tolerance defining what is a sliver can then be specified and each polygon in the overlaid coverage can be tested to see whether it falls below the tolerance. Those polygons that do fall below the tolerance then need to be organised so that the relationship each one has with its neighbours is explicitly defined within a user accessible file. One possible way of doing this is to redefine the slivers as 'inside' polygons and to list a possible set of 'outside' neighbours. Each neighbour can then be referenced to the best fit matrix and the most suitable neighbour found on the basis of the predefined probabilities. Once this is ascertained the class of the sliver polygon can then be changed to match that of its best fit neighbour[3]. Using the polygon identifier of the sliver will then allow the same change to be made within the PAT file. Once the same process has been carried out for all the slivers, the PAT file of the overlaid coverage can be *dissolved* on its new category value. Thus all boundaries dividing areas of the same category will be removed. The basics of this process are outlined in figure 7.17.

Such a system however, can only be applied to polygons that do not share any common categories with their neighbours. If they did then the whole process of assessing the best fit neighbour would be a time consuming and pointless exercise, as it could be easily achieved by a *dissolve* in the first place. In removing slivers from an overlay in which an updated version of a classified land cover was desired, this process would be of little use as any sliver would be related to either the old state of the space it represents or the new state, and thus

---

[3]To take into account the possibility that the sliver polygon might be merged with another sliver polygon, thus increasing its area and perimeter and making the new combined polygon less likely to be defined as a sliver, the procedure should check each new 'inside' polygon against the tolerance again before proceeding to merge it. Thus when a sliver polygon is with one of its neighbours the new polygon will have its area and perimeter defined according to the following equations:

$$P_{np}=\sum(P_s+P_m)-2dl_{sm}$$

Where
P = perimeter of polygon
m = polygon to be merged with
s = sliver polygon
dl = length of dividing line
np = new polygon

$$A_n=A_s+A_m$$

Where
A = area of polygon

This allows for the new polygon to be re-tested against the areal tolerance.

it could be removed by choosing which coverage is the more accurate and dissolving on the classification associated with that coverage.

This system might be usefully employed in an analysis of land use change, in which the aim was to define zones of change rather than update existing material. As each zone would not be defined as 'maize' or 'grassland', but as 'maize to maize' or 'meadow to moor', slivers do not automatically share the same parent class as one of its neighbours. The problem the user would face in a large coverage is the large number of potential changes that might exist. With a coverage of only 4 classes, for example, the potential number of changes through time is 16, which would then have to be matched with each other bringing the total number of neighbour match assessments required to $16^2$. For a classification system with 10 classes the matrix would require $10^4$. Not only would this be time consuming to build, but it would also take time to check and access. The amount of effort such a system might require may well counteract any benefits to be gained from its application. A further problem associated with this type of approach is that it is sequential in nature. Consider for example the case of three slivers next to each other (figure 7.18). Sliver 1 is best matched to area 4 (class K) and is merged with it. This then denies sliver 2 the opportunity of merging with sliver 1, even though class A might have been most suitable for it. Sliver 2 is now offered a sub-optimal choice, when if it had been first in the sequence it might have merged with sliver 1 and resulted in the creation of a new polygon that was above the sliver tolerance (figure 7.18).

### 7. 6. 4. Using histories to determine best neighbourhood matches.

An alternative approach, which would reduce the size of the matrix required for reassignment, involves making use of the histories attached to each polygon. If the land use cover is being assessed in terms of its change through time, then each polygon will have a series of land uses attached to it. Instead of viewing these as separate entities, they could be combined to form an historical record of land use within that defined polygonal space. Thus, following overlay, when small zones of change are suspected of being slivers, their histories can be referred to as an indicator of which neighbouring polygon such zones should be merged with.

This type of approach was applied to the test coverage used in the previous experiments using the following methodology. Firstly, all the polygons in the original and overlay coverage were labelled, and classification classes were associated with them. The classification classes were taken from two surveys of land use carried out in the Fenlands during the months of December 1985 and August 1986. Thus polygons in the test coverage were assigned to polygons in the December 1985 survey and their classes duplicated, and the same polygons were then identified in the August 1986 coverage and their classes duplicated in the coverage for superimposition. The coverages are shown in figures 7.19 and 7.20, and the tabular information is illustrated in table 7.8 and table 7.9.

Table 7. 8. The original coverage with its class number and category

| Poly-id | Class-no | Category |
|---|---|---|
| 1 | 1 | Beets |
| 2 | 2 | Wheat |
| 3 | 7 | Building |
| 4 | 1 | Beets |
| 5 | 2 | Wheat |
| 6 | 4 | Rapeseed |
| 7 | 3 | Bare |
| 8 | 7 | Building |
| 9 | 6 | Road |
| 10 | 3 | Bare |
| 11 | 5 | Stubble |
| 12 | 3 | Bare |
| 13 | 6 | Road |
| 14 | 2 | Wheat |
| 15 | 7 | Building |
| 16 | 5 | Stubble |
| 17 | 7 | Building |

Table 7.9 The coverage of change with its class number and category

| Poly-id | Class-no | Category |
|---|---|---|
| 1 | 2 | Wheat |
| 2 | 5 | Stubble |
| 3 | 7 | Building |
| 4 | 2 | Wheat |
| 5 | 2 | Wheat |
| 6 | 3 | Bare |
| 8 | 7 | Building |
| 9 | 6 | Road |
| 10 | 8 | Beans |
| 11 | 9 | Potatoes |
| 12 | 9 | Potatoes |
| 13 | 6 | Road |
| 14 | 2 | Wheat |
| 15 | 7 | Building |
| 16 | 2 | Wheat |
| 17 | 7 | Building |
| 18 | 10 | Footpath |
| 19 | 7 | Building |
| 20 | 7 | Building |
| 21 | 7 | Building |
| 22 | 7 | Building |
| 23 | 2 | Wheat |
| 24 | 2 | Wheat |
| 25 | 8 | Beans |
| 26 | 10 | Footpath |
| 27 | 3 | Bare |
| 28 | 10 | Footpath |
| 29 | 4 | Rapeseed |

The next stage involved setting up a matrix of potential histories. In this all the land use categories from $t_0$ are plotted against all the land use categories from $t_1$. The matrix is then filled in with either a 0 or a 1. A zero denotes that given the choice between the land use classification at $t_0$ and at $t_1$, the most probable present class for the polygon will be that existing at $t_0$. In the same way a 1 indicates that the most suitable present class for the polygon will be that existing at $t_1$. For the classes present in this example, the following matrix was designed; table 7.10.

Table 7.10. Matrix of probable histories.

| | To<br>From | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Beets | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| Wheat | 2 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| Bare | 3 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 |
| Rapeseed | 4 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| Stubble | 5 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| Road | 6 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| Building | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| Beans | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 |
| Potatoes | 9 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| Footpath | 10 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 |

The two coverages were overlaid giving each polygon both an old class and a new class value. In order to create a composite class the first value was multiplied by a 100 and the second class value added to that. This gave a three digit number denoting the combined class value. The combined class or history values for each of the overlay polygons are shown in figure 7.21. A tolerance value was then set at AREA/PERIMETER = 0.05 units. Thus anything below this would be regarded as a potential sliver. Slivers were then tested against the matrix to see which of their two histories was the most acceptable. If for example, the previous state was deemed to be the most acceptable, then the neighbouring polygons of the sliver were searched and the first one found to be sharing a common previous state was flagged as the polygon most suitable for merging with the sliver. The merge was carried out by replacing the 'history' of the sliver polygon with that of the polygon it is to be merged with (figure 7.22), and a *dissolve* operation was then carried out on the PAT of the overlaid coverage on the basis of polygon 'history' (figure 7.23).

In comparing the results with the geometric case the system produces a superior result, with zones of true change clearly defined, and dubious zones of change removed and combined with their most appropriate user defined neighbour. The problems with trying to fully automate such a system within the ARC/INFO package are however severalfold. Firstly, one is tied to the system and the limited operational tools provided within it. The INFO component of ARC/INFO is clumsy and less than user friendly to operate and its programming facilities offer little scope for easy implementation of this algorithm. ORACLE or DBASE as the data base component of the GIS might make the implementation easier. The cost at which such an improvement in quality can be achieved, is primarily one of time and additional processing. The example cited uses only a test example; running the process on a larger coverage would require a larger matrix and a longer processing time. Another potential problem is defining the matrix. How sure can the user be about particular changes through time? To what extent can user preconceptions preclude changes in previously unexpected directions? In terms of land use change such deterministic delimitations are probably less likely to run into problems than they would in the field of resource management. Furthermore, given that the alternatives are geometrically biased, most cases of operator knowledge intervention should at least produce a considered result, rather than a randomly determined one.

## 7. 7. Conclusion.

To clearly show the change in polygon movement as a response to various sliver removal routines, this chapter has used small illustrative examples, consisting of simple stylised coverages. Larger coverages with more complex data relationships may yield different results, but it is not the focus of this study to benchmark differences between coverage types and dimensions.

A variety of methods have been suggested to deal with the sliver polygon problem relying on the use of feature attributes, combined with a series of user defined rules. Although all of the error elimination options are successful in their own right, none of them provides the ideal solution for all problems.

The use of geometrical routines succeeds in removing small thin polygons or polygons answering a particular attribute value, but irrespective of how sophisticated the rules become for determining which polygons represent change and which are merely erroneous, if the elimination routines always rely on re-assignment for eliminated polygons to the neighbour which merely happens to have the longest shared boundary, then the user will never have control over the final result. An alternative to the elimination routine is to use a matrix of allowable attribute movement. However this approach relies on the user having an in depth knowledge of the likely attribute change through time and the emphasis is still on error removal on the assumption that an accurate coverage is attainable.

As the presence of error within GISs is inescapable, if the user continues to view the removal of inaccuracies as one of error elimination, in which problems of mismatches are conclusively solved, then the user will always be left in an uncertain position in which a proportion of the coverage is erroneous. An alternative philosophy is to accept that a degree of uncertainty will persist within coverages that have been created from less than perfect datasets and that in dealing with this, emphasis should be placed on the reassignment process, rather than on the elimination process itself. Other alternatives to allay overlay mismatches can be derived using methods which stress selective neighbourhood coalescence, as opposed to forced neighbourhood matching. This was suggested in the case of probability matching, and demonstrated in the case of history matching.

Attributes can therefore form the basis of a more intelligent polygon removal methodology, albeit a still imperfect one, as they give the user some control in the form of an ability to determine the result of polygon reassignment. Unlike other methods of polygon removal however, it also requires a greater prior knowledge of the data, forcing the user to cover every eventuality. Furthermore, assuming a superior accuracy for one coverage over another, or of one time state over another can be dangerous, as coverages often have varying accuracies within their extent. Therefore in using such a methodology, the onus for accuracy assessment is transferred from the software to the user.

# Chapter Eight - Concluding comments.

## 8. 1. Summary of findings.

The issue of error within GISs is a relatively new one, and before solutions can be offered to diminish its impact, the extent and nature of the problem require investigation. For many users this begins with a simple admission that inaccuracies actually exist, and that most geographical data are 'fuzzy'. Beyond this stage, the next step demands empirical measurement. Thus one of the roles of this thesis has been to measure the extent to which inaccuracies are introduced in various stages of GIS database preparation. Some of the errors have arisen as a consequence of data quality, whilst others are a result of the tools GISs provide for dealing with problems of data input. Sliver polygon accumulation following data overlay and the tools provided to deal with them were also assessed. In addition to measurement, this thesis has attempted to offer some solutions to the sliver polygon problem. The major findings of the thesis can thus be summarised as follows:

- Errors exist at all stages of data accumulation and these will affect the results of any subsequent analysis on that material.

- Errors are present in digital cartographic data before they reach the GIS input stage, and assessing the extent of these depends very much on user-based decisions.

- Errors in digitising are a result of both imprecision and generalisation. Neither can be satisfactorily predicted because manual digitising is an essentially subjective process. Both can be measured to give an indication of the digitising profile associated with a particular individual. Measures of precision can be made with respect to time, or with respect to the position of the feature on the digitising table. Measures of feature representation can be made in terms of the number of points used to represent a line and the length of the line following digitising.

- Routines which allegedly 'clean' digital cartographic data do not make the data more accurate as they might imply. Rather, they mould the information into a form that is suitable for the system to handle. This may be done at a cost to coverage integrity and may use techniques that are not clearly documented, but which appear hidden from the user as 'black box' methodologies.

- Such routines when used to excess can be quite destructive to coverage morphology, and users should be aware of the consequences of using them without question.

- The amount of coverage change or displacement resulting from such 'cleaning' and 'building' routines can be measured in terms of the change in line length and the movement of coverage features.

- If tolerances are not carefully selected, unintended changes in line lengths and node shifts can occur which leave the coverages displaying different properties to the originals.

- Overlay following the use of data digitised from different scales at different times will result in a large number of sliver polygons, which then have to be removed if meaningful analysis is to be carried out. The number of slivers is likely to be higher if the data have been obtained from different capture techniques or from different cartographic sources.

- The current methods for removing sliver polygons rely primarily on geometric tools which are relatively crude and which either remove as many genuine polygons as erroneous ones or which leave the error removal job unfinished. Finding a suitable tolerance figure is a matter of trial and error.

- Alternatives to the areal tolerance routines rely on utilising a user's knowledge of the data; but also require a system that is designed to easily manage time and versions.

**8. 2. Evaluation of methodologies.**

The methodologies used for measurement and assessment have been relatively easy to implement, and have proved successful in providing a basis for error evaluation in terms of the operator, the data and the system. An assessment of GIS ability to deal with 'fuzzy' data was carried out in both Chapter Five and Chapter Six using simple benchmarking techniques. In Chapter Five, the creation of a test coverage that replicated problems the tolerance routines might have to remove, provided a visual illustration of how system algorithms can distort coverage data. The careful monitoring of such a coverage in response to various error removal routines is within the scope of most users, and can quite painlessly be applied. Obtaining the results of such tests will firmly establish in a user's mind the degree of crudity with which such routines operate and the appropriate amount of caution that such routines should be treated with. In a similar way, testing the effectiveness of the sliver polygon removal techniques, (as was done in Chapter Six) is easy to perform, once an overlay coverage has been created and the user is aware of how the features are supposed to appear.

Extensions of these simple techniques were used in Chapter Five to assess the relative strengths and weaknesses of the linear error removal techniques. These were more time consuming to perform as they required the data to be manipulated into a certain format. For example, in evaluating the changes in coverage line lengths, the internal relabelling of lines following linear tolerancing had to be overcome, so that lines in the toleranced coverage could be correctly compared with lines in the original coverage. Whilst in the comparison of node movements, data referring merely to the endpoints of linear features had to be extracted from all of the x y coordinates of the coverage. In both cases however, simple programs can quite easily solve these problems. The actual comparisons also involve a degree of programming, but these are by no means difficult to write.

The methods for assessing the extent of the differences between coverages used simple descriptive statistics and comparative techniques. Admittedly more analysis could be carried out on this material, which would enable the differences to be illustrated to a finer degree. For example, the Mann-Whitney U test might have been applied as a non-parametric alternative when some of the coverages were proved to be unsuitable for the parametric t-test. In addition, the measurements of linear length change and node shift are not the only factors that can be cited as being indicative of coverage change. Other measures might include the changing shape of features, or loss in sinuosity or a change in line characteristics. Alternatively, the areal difference between the original and the toleranced coverage could be used as a measure of coverage change. It might also be argued that the use of a coverage with very few curved lines would prejudice the results in favour of distortion. This may be true, but such a coverage also serves to clearly illustrate the effects of line movement. Furthermore, such coverages are not inventions, but actually exist.

Measuring the extent of the sliver polygon problem, and the degree to which the GIS is able to counteract it, could have been heightened by including a coverage derived from satellite imagery. This would have perhaps provided a different type of problem, and differing solutions might have suggested themselves. Initial tests were undertaken with Thematic Mapper data, but these showed the PC ARC/INFO system to be unsuitable for dealing with imagery containing many classes, due to problems of space and processing. A further problem concerned actually inputting the material into ARC/INFO, since at that time many of the input formats were not compatible with those that image processing systems were currently using. Transferring the data therefore required quite extensive programs. It was hoped that panchromatic SPOT data would be less cumbersome to use and would require less image processing, but unfortunately although the data were acquired, the pre-processing necessary could not be undertaken due to equipment problems. Thus it was decided to use coverages of the same area at various scales, given that this would cover GIS applications adhered to by many users.

Operator precision and generalisation in terms of data input was measured in Chapter Four. Assessing generalisation and individual variability in linear representation is a difficult task for a geographer to undertake, as so many perceptual and psychological factors are undoubtedly responsible for the final decisions. However, what the experiment did clearly highlight was the degree of generalisation present in the cartographic source material due to the confines of scale, and the way in which such generalisations can be compounded by operator generalisation. Operator precision is an easier factor to evaluate as it is less subjective in nature. Using points to assess operator deviations focuses the user on precision only and requires little more than a willingness on the part of the operator to perform the test, and some means of displaying the results graphically with respect to time or the spatial extent of the digitising table. It is a simple test, but an effective and revealing one, which could point to problems digitising operators may have that might not otherwise have been foreseen. It could also form the basis of a statistical measure of the operator's precision.

The various tests that were carried out in Chapter Seven were all applied to a small sample coverage so that changes could be effectively highlighted. A larger and more complicated data set would really be required to test these approaches properly in terms of both effectiveness and speed of processing. Furthermore, due to the difficulties of programming with a very limited language such as INFO, some of the methodologies could not be fully automated. Using a different database language such as DBASE or ORACLE might have made the task easier. Or, alternatively, gaining access to some of the 'unseen' files which store the relationships between features in an explicit way, would have reduced the load on the program and data preparation.

The methodologies used in this thesis for measurement could easily be replicated by users interested in exploring the error issue with respect to their own data, and their own approaches to building a GIS data base. Alternatively, such techniques could easily be incorporated by GIS vendors into simple routines. Adoption will however be very much determined by the degree of time and expense users are willing to go to, which will in turn be dictated by the intended use of the data.

**8. 3. Implications for the user.**

It is therefore misguided to assume that error can be removed entirely from GIS methodologies or from GIS products. However, a reduction of error which leaves the user with a quality product is feasible, yet it relies on a willingness by the user to recognise the presence of inaccuracies and also on an ability to measure their extent. Unfortunately, many users are unaware of how error arises and what form it might take. In part this is due to the fact that many GIS users are reliant on the system vendors for advice relating to the use of both their information and the system. Established large organisations will undoubtedly have checking procedures to maintain the quality of their data, but many inexperienced smaller users are unlikely to exercise as great a control. The gradual introduction and adoption of digital cartographic data standards will go some way to reducing this problem, but its heavy reliance on 'fitness for use' will still put most of the responsibility in the hands of the users.

## 8. 4. Implications for the vendor.

Although this thesis has looked exclusively at the ARC/INFO GIS, the issues raised are valid for many of today's systems, as they all utilise similar techniques for 'making the data fit' and for sliver polygon removal. These techniques are undoubtedly crude and as users increase in number and sophistication one would hope that alternatives which place a greater onus on user familiarity with the data will be made available. One of the improvements that could be carried out relatively easily would be making the results of internal benchmarking tests available, as this would assist users in decisions as to the applicability of certain operations. There is hope here as ESRI (the ARC/INFO vendors) are very keen to respond to user demands, although most of ESRI's improvements have been made in the field of user interfaces, as easy to use systems are easier to sell. One assumes however, that as users become more familiar with the system and more discerning, there will be a demand for more attention to be paid to the error issue.

## 8. 5. Implications for ARC/INFO.

Although this thesis has not intended to be a critique of the ARC/INFO GIS, it has highlighted some of the shortcomings of the system and the way in which the system's techniques are presented. For example some of the tolerancing routines are very crude in the way they respond to the data, yet they are advocated in the manual without any attached 'health warning'. Only the positive benefits of removing node mistakes are mentioned, and these are couched in terms of 'clean' data; which for many users will be synonymous with 'accurate' or 'correct' data. Some of the approaches to digitising are also dubious. Suggesting that users intentionally 'overshoot' when they digitise because this ensures the subsequent use of one algorithm in favour of another encourages a technologically deterministic attitude to error which should be avoided at all costs. The methodologies suggested in Chapter Seven are relatively simple and work within the confines of the system, thus automating some of the methodologies suggested in that chapter as a macro, or optional routine for sliver polygon removal, would be a relatively easy task for ARC/INFO programmers who have direct access to the files such as the PAL (Polygon Arc List).[1] If ARC/INFO is to maintain its standing as a world leader in the GIS field issues such as error measurement, tracking and

[1] Otherwise known as the polygon topology file.

understanding need to be fully met in the near future, whether it is by the adoption of simple methodologies or sophisticated new approaches to data structure and modelling.

## 8. 6. Areas for future research.

What hopes are there for both the awareness and reduction of error in GISs? Most observers fall into two camps; those who see the solutions lying in technological improvements, and those who regard the problem as solvable only through an increase in awareness and training. Possible technologies that might be applied to the problem include Artificial Intelligence (AI) and a more acceptable data model for data storage. Despite a promising welcome into the GIS field in the mid-1980's, very little has been achieved in the area of AI short of pattern recognition. If AI solutions are to be found and adopted by vendors, it is unlikely to occur in the short term. An alternative data model to the currently widely used relational based data model is one which relies on object oriented data storage. Currently INTERGRAPH's TIGRIS system professes to have object oriented components, but the only true object oriented GIS on the market today is the Smallworld system. In an object oriented environment, although error at the digitising stage cannot be removed, the use of tolerancing can be more safely applied. As each linear feature is regarded as an entity in its own right, complete with its own attribute, the user can attach allowable movements or feature changes to each element in the data base. These pieces of information are then part of that feature and can be called upon whenever an issue of contention arises.

A further technological improvement concerns the cartographic data that are actually used as the basis for GIS analysis. One could argue, that as the mechanisms for high speed accurate surveys become more widely available, GISs in the future will probably not have the same kind of error inputs as GISs at present. Although this may well be the case for positional information, such as material derived through GPSs, thematic information will always embody some degree of generalisation, and as such will lead to an element of 'fuzziness' within GISs.

In the short term, however, it seems unlikely that a deterministic technological approach will alter system methodologies. In fact it could be argued that systems should not bear the responsibility of error reduction as it is better addressed as a factor of education and training.

Admittedly, there is a productive responsibility, but in pragmatic terms given the large investment in GIS it would be ridiculous to suggest that vendors discard methodologies overnight. It is likely therefore that the error issue (like three dimensional GISs) will be adopted in an evolutionary fashion.

Other important issues that need to be addressed by the GIS community in the near future include those of time and versioning. So far the discussion has revolved primarily around errors in space, but methodologies for dealing with errors in time, and in 'space and time' are becoming equally pressing. Entwined with this, is how time and change should be stored within GIS data bases. At present there are very few systems which have sophisticated methods for dealing with this problem, Smallworld being a notable exception.

The present therefore becomes a holding operation in which all efforts should be made to ensure that systems are not misused, by providing users with adequate training in both techniques and underpinning methodologies. This in turn will provide them with a suitable base for carrying out error assessment and minimisation procedures. Furthermore, users should be encouraged to become familiar with the 'fitness of use' concept in assessing data suitable for integration. They should be made to realise that error cannot necessarily be predicted, neither can it be entirely prevented; at best it can only be coped with. In portraying the extent to which error can arise in GISs, this thesis has sought to form a basis from which solutions to error limitation and awareness can be achieved.

## Bibliography


Abler, R. F. 1987: The National Science Foundation National Center for Geographic Information and Analysis. International Journal of Geographic Information Systems, 1, 303-326.

Adams, S. L. and McLeod, R. G. 1984: Comparison between multispectoral classification accuracy of Landsat-4 MSS and TM, in Klepfer, M. M. and Morrison, D. B. (eds) Machine processing of remotely sensed data, p49. Purdue Research Foundation: Indiana.

Ahn, J. and Freeman, H. 1984: A program for automatic name placement. Cartographica, 21, 2&3, 101-109.

Andrus, J. F. Campbell, C. W. and Jayroe, R. R. 1975: Digital image registration method using boundary maps. IEE Transactions on Computers, 24, 9, 935-940.

Archer, H and Crosswell, P. L. 1989: Public access to geographic information systems: an emerging legal issue. Photogrammetric Engineering and Remote Sensing, 55, 11, 1575-1581.

Armstrong, M. P. 1988: Temporality in spatial databases. Proceedings of GIS/LIS, San Antonio, pp880-889.

Armstrong, M. P. and Densham, P. J. 1990: Database organisation strategies for spatial decision support systems. International Journal of Geographical Information Systems, 4, 1, 3-20.

Aronoff, S. 1982: The minimum accuracy report: A user's view. Photogrammetric Engineering and Remote Sensing, 48,8, 1309-1312.

Aronoff, S. 1985: The minimum accuracy value as an index of classification accuracy. Photogrammetric Engineering and Remote Sensing, 51,1, 99-111.

Aronson, P. 1987: Attribute handling for geographic information systems. Chrisman, N. R. (ed) Proceedings, Auto Carto 8, pp346-355.

ASPRS 1990: ASPRS accuracy standards for large-scale maps. Photogrammetric Engineering and Remote Sensing, 56, 7, 1068-1070.

Atly, J. L. and Coombs, M. J. 1984: Expert systems: concepts and examples. NCC Publications: Manchester.

Baker, J. R. and Drummond, J. E. 1984: Environmental monitoring and map envision using integrated Landsat and digital cartographic data. ITC Journal, 1-10-19.

Balchin, W. G. V. and Coleman, A. M. 1967: Cartography and computers. The Cartographer, 4, 2, 120-127.

Baskerville, R. 1988: Designing information systems security. John Wiley: Chichester.

Baugh, I. D. H., and Boreham, J. R. 1979: Measuring the coastline from maps: a study of the Scottish mainland. The Cartographic Journal, 13, 2, 167-171.

Baxter, R. S. 1978: Algorithm 4: Affine transformations. Environment and Planning A, 10, 105-110.

Beard, K. 1987: How to survive on a single detailed database. Chrisman, N. R. (ed) Proceedings Auto-Carto 8, pp211-220.

Beard, K. 1989: Use error: the neglected error component. Proceedings, Auto-Carto 9, pp808-817.

Beck, C. 1977: Information systems and social sciences. American Behavioural Scientist, 20,3,427-448.

Beck, F. J. and Olsen, R, W. 1986: Quality control and standards for a national digital cartographic database. in Moellering H. (ed) Digital cartographic data standards: a report on evaluation and empirical testing. National Committee for Digital Cartographic Data Standards. Columbus: Ohio.

Bedard, Y. 1987: Uncertainties in land information system databases. Chrisman, N. R. (ed) Proceedings Auto Carto 8, pp175-184.

Benson, W. 1982: An application of fuzzy set theory to data display. Yager, R. R. (ed) Fuzzy set and possibility theory. Pergamon Press: NY.

Bickmore, D. 1964: Experimental cartography. A report on the Oxford Symposium October 1963. Clarendon Press: Oxford.

Bickmore, D. 1966: Experimental cartography project. Interim report to the Natural Environment Research Council. Royal College of Art: London.

Bickmore, D. P. 1970: Exact Automatic Cartography. Journal of the British Interplanetary Society, 23, 363-368.

Blakemore, M. J. 1984: Generalisation and error in spatial data bases. Cartographica, 21, 2-3, 131-139.

Blakemore, M. J. 1985a: High or Low Resolution? Conflicts of Accuracy, Cost, Quality and Application in computer mapping. Computers and Geosciences, 7, 3, 345-348.

Blakemore, M. J. 1985b: From lineprinter maps to geographic information systems: a retrospective on digital mapping. SUC Bulletin, 19, 2, 65-78.

Blakemore, M. J. 1990: The UK Regional Research Laboratory Initiative: regional centres of excellence. Proceedings, GIS for the 1990s, Ottawa, Canada, 1172-1184.

Blakemore, M. J. 1990: NOMIS Dynamic Local Labour Market Information. Mapping Awareness, 4, 10, 26-30.

Blakemore, M. J. and Harley, J. B. 1981: Concepts in the history of Cartography: A review and perspective. Cartographica, 17, 4.

Blakemore, M. J. and Rybaczuk, K. 1987: Digital mapping. in Parry, R. B. and Perkins, C. R. World Mapping Today, pp46-52. Butterworths: London.

Blakemore, M. J. and Dewdney, J. 1990: Census Mapping. in Perkins, C. R. and Parry, R. B. (eds) Information Sources in Cartography. Bowker-Saur: London.

Board, C. 1981: Cartographic communication. Cartographica, 18, 2, 42-78.

Bolstad, P. V., Gessler, P. and Lillesand, T. M. 1990: Positional uncertainty in manually digitised map data. International Journal of Geographical Information Systems, 4, 4, 399-412

Borgeson, W. T., Batson, R. M. and Kieffer, H.H. 1985: Geometric accuracy of Landsat-4 and Landsat-5 thematic mapper images. in Photogrammetric Engineering and Remote Sensing, 51, 12, 1893-1898.

Bossler, J. D., Finnie, T. C., Petchenik, B. B. and Usserman, T. M. 1990: Spatial data needs: the future of the National Mapping Program. Cartography and Geographic Information Systems, 17, 3, 237-242.

Boudriault, G. 1987: Topology in the TIGER file. Chrisman, N. R. (ed.) Proceedings Auto-Carto 8, pp258-263.

Bouille, F. 1982: Actual tools for cartography today. Cartographica, 19, 2, 27-32.

Boyle, A. R. 1970: The quantised line. Cartographic Journal, 7, 91-94.

Boyle, A. R. 1974: Geographic information systems in hydrography and oceanography. The Canadian Cartographer, 11, 2, 125-141.

Boyle, R. 1980: Scan digitising and drafting. ACSM-ASP Convention Technical Papers, pp136-142.

Boyle, A. R. et al. 1983: Review and synthesis of problems and directions for large scale GIS development. NASA-CR- 166563.

Bracken, I. and Webster, C. 1989: Towards a typology of geographical information systems. International Journal of Geographical Information Systems, 3, 2, 137-152.

Bracken, I. and Higgs, G. 1990: The role of GIS in data integration for rural environments. Mapping Awareness, 4, 8, 51-56.

Brassel, K. 1977: A survey of cartographic display software, International Yearbook of Cartography, 17, 60-77

Braund, M. 1980: An analysis of the effects of temperature and humidity on a variety of drawing media. Bulletin, Society of University Cartographers, 14, 2, 25-35.

Bribiesca, E., and Guzman, A. 1979: Shape description and shape similarity measurement for two dimensional regions. Geo Processing, 1, 2, 129-144.

Bridges, E. M. and Davidson, D. A. 1982: Principles and applications of soil geography. Longman: London.

Brod, R. 1979: Generalisation. A Fundamental process in cartography. Bulletin, Illinois Geographical Society, 21,1, 24-29.

Brooks, A. A. and Upperman, J. V. 1986: Interim NCDCDS testing report. in Moellering, H. (ed.) Digital cartographic data standards: a report on evaluation and empirical testing, pp33-43. National Committee for Digital Cartographic Data Standards. Columbus, Ohio.

Broome, F. 1986: Mapping from a topologically encoded data base: the US Bureau of the census example in: Proceedings, Auto-Carto London, Vol 1, 402-411.

Brujin, C. A. De 1986: Integrating emerging low-cost technologies in a mapping and geo-information system for urban planning in developing countries: the RUMADS approach. Technical Papers, pp152-161. ACSM: Falls Church Va.

Brusegard, D. and Menger, G. 1989: Real data and real problems: dealing with large databases, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Bryant, N. A., Zorbist, A. L. Walker, R. E. and Gickhmann, B. 1985: Landsat thematic mapper geodetic accuracy: implications for geo-coded map compatibility. Advanced technology for monitoring and processings global environmental data, pp 257-273. Remote Sensing Society: London.

Bundock, M. 1987: An integrated DBMS approach for geographic information systems. Chrisman, N. R. (ed) Proceedings Auto-Carto 8, pp292-301.

Burnside, C. D. 1979: Mapping from aerial photographs. Wiley and Sons: New York.

Burnside, C. D. 1982: Electromagnetic Distance Measurement 2nd edition. Granada Books: St. Albans.

Burrough, P. A. 1984: Fakes, facsimiles and facts: fractal models of geophysical phenomena. Proceedings of IBM conference on science and uncertainty. London.

Burrough, P. A. 1986: Principles of Geographical Information Systems for land resources assessment. Clarendon Press: Oxford.

Burrough, P. A. 1987: Multiple sources of spatial variation and how to deal with them. Chrisman, N. R. (ed) Proceedings Auto-Carto 8, pp145-154.

Buttenfield, B. P. 1985: Treatment of the cartographic line. Cartographica, 22, 2, 1-26.

Buttenfield, B. P. 1987: Automating the identification of cartographic lines. The American Cartographer, 14, 1, 7-20.

Buttenfield, B. P. 1989: Scale dependence and self-similarity in cartographic lines. Cartographica, 26, 1, 79-100.

Calkins, H. W. and Tomlinson, R. F. 1977: Geographic information systems, methods and equipment for land use planning. USGS: Reston, VA.

Campbell, J. 1991: Introductory cartography. (2nd Ed.) WCB: London.

Campbell, W. G. and Mortenson, D. C. 1989: Ensuring the quality of geographic information system data: a practical application of quality control. Photogrammetric Engineering and Remote Sensing, 55, 11, 1613-1618.

Cao, Tho Trang, 1978: GIMMAP: An example of a geographic information and mapping system. ACSM Fall Convention pp57-69.

Carcknell, A. and Hayes, L. 1987: Remote Sensing Yearbook 1987. Taylor Francis: London.

Card, D. H. 1982: Using known map category marginal frequencies to improve estimates of thematic map accuracy. Photogrammetric Engineering and Remote Sensing, 48, 3, 431-439.

Carruthers, A. W. and Waugh T. C. 1988: GIMMS Reference Manual, Release Edition 5. GIMMS Ltd: Edinburgh.

Carstensen, L. W. Jr. 1990: Angularity and capture of the cartographic line during digital data entry. Cartography and Geographic Information Systems, 17, 3, 209-224.

Carter, J. R. 1984: Computer mapping progress in the '80s. AAG: Washington.

Carter, J. R. 1988: A typology of geographic information systems. Technical papers. ACSM-ASPRS Annual Convention, pp207-215.

Cerny, J. W. 1972: Use of the SYMAP comupter mapping program. The Journal of Geography, 74, 167-174.

Chen. Z - T. 1989: A spatial low pass filter working for triangular irregular network (TIN) and restricted by breaklines. Proceedings, Auto-Carto 9, pp138-145.

Chorley, R. 1988: Some reflections on the handling of geographic information. Int. Jnl. GIS, 2, 1, 3-10.

Chrisman, N. R. 1977: Cartographic data structures. Laboratory for Computer Graphics and Spatial Analysis Report. LCGSA: Harvard, USA

Chrisman, N. R. 1982a: Methods of spatial analysis based on error in categorical maps. Ph.D. thesis, University of Bristol.

Chrisman, N. R. 1982b: Beyond accuracy assessment: correction of misclassification. Proc. International Society of Photogrammetry and Remote Sensing Commission, IV 24-IV, 123-132.

Chrisman, N. R. 1982c: A theory of cartographic error and its measurement in digital data bases in: Foreman, J. (ed): Proceedings, Auto-Carto 5, pp159-168. American Congress for Surveying and Mapping: Falls Church Va.

Chrisman, N. R. 1983: Epsilon filtering: a concept for automated scale changing. Proceedings of American Congress on Surveying and Mapping 43rd Annual Meeting, pp322-331.

Chrisman, N. R. 1984a: On storage of coordinates in geographic information systems. Geo-processing, 2, 259-270.

Chrisman, N. 1984b: The role of quality information in the long-term functioning of a geographic information system. Cartographica, 21, 2-3, 79-89.

Chrisman, N. 1985: An interim proposed standard for digital cartographic data quality: supporting documentation. in Moellering, H (ed.) Digital cartographic data standards: on interim proposed standard, pp113-132. National Committee for Digital Cartographic Data standards: Columbus, Ohio.

Chrisman, N. R. 1986a: Obtaining information on quality of digital data. Proceedings, Auto-Carto London, Vol. 1, 350-358.

Chrisman, N. R. 1986b: Testing the Interim proposed standard for digital cartographic data quality. in Moellering, H. (ed.) Digital cartographic data standards: a report of evaluation and empirical testing, pp53-55. National Committee for Digital Cartographic Data Standards: Columbus, Ohio.

Chrisman, N. R. 1987a: The accuracy of map overlays: a reassessment. Landscape and Urban Planning, 14, 427-439.

Chrisman, N. R. 1987b: Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. International Journal of GIS, 1, 3, 265-278.

Chrisman, N. R. 1987c: Fundamental principles of geographic information systems Chrisman, N. R. (ed.) Proceedings Auto-Carto 8, pp 32-41.

Chrisman, N. R. 1987d: The design of geographic information systems based on social and cultural goals. Photogrammetric Engineering and Remote Sensing, 53, 10, 1367-1370.

Chrisman, N. R. 1988: The risks of software innovation: a case study of the Harvard Lab. American Cartographer, 15, 3, 291-300.

Chrisman, N. R. 1989a: Error in categorical maps: testing versus simulation. Proceedings, Auto-Carto 9, pp521-529.

Chrisman, N. R. 1989b: Modelling error in overlaid categorical maps, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Cimon, N., Cole, P. K. and Quigley, T. M. 1990: Technical note. A regression technique for estimating the time required to digitize maps manually. International Journal of GIS, 4, 1, 51-54.

Clarke, K. C. 1986: Advances in geographic information systems. Computers, Environment and Urban Systems, 10, 175-184.

Colvocoresses, A. P. 1984: Mapping of Washington DC and vicinity with Landsat 4 thematic mapper. Technical Papers, 50th Annual Meeting, pp757-763. American Society of Photogrammetry: Falls Church, Va.

Colwell, R. N. and Poulton, C. E. 1985: SPOT Simulation Imagery for Urban Monitoring: a comparison with Landsat TM and MSS imagery and with high altitude colour infrared photography. Photogrammetric Engineering and Remote Sensing, 51, 8, 1093-1101.

Comeau, M. A. and Holbaek-Hanssen, E. 1984: Compression and compaction of binary raster images. Cartographica, 21, 2 & 3, 140-147.

Cooke, D. F. 1989: Spatial data systems beyond GIS. Proceedings Auto-Carto 9. ACSM/ASP: Falls Chruch, VA.

Cooke, D. F. and Maxfeild, W. I. 1967: The development of a geographic base file and its uses for mapping. Urban and regional information systems for social programs. Proceedings of 5th Annual Conference of the urban and regional information system association. pp207-218.

Cooley, R. 1990: National Center for Geographic Information and Analysis. Mapping Awareness, 4, 10, 14-15.

Cooper, M. A. P. 1985: New instrumention: an end to survey observation. Paper presented to 'Surveying in Tomorrow's world'. Newcastle University: Newcastle.

Coppock, J. T. 1988: The analogue to digital revolution: a view from an unreconstructed geographer. American Cartographer, 15, 3, 263-275.

Coppock, J. T. and Rhind, D. W. 1991: The history of geographical information systems. in Maguire, D., Goodchild, M. F. and Rhind, D. W. Geographical Information Systems. Longman: London. (in press)

Couclelis, H. 1986: Artificial intelligence in geography: Conjectures on the shape of things to come. Professional Geographer 38, 1, 1-11.

Cowen, D. 1990: GIS versus CAD versus DBMS: what are the differences? in Peuquet, D. J. and Marble, D. F. (eds) Introductory readings in Geographic Information Systems. Taylor and Francis: New York.

Crain, I. K. and Macdonald, C. L. 1984: From land inventory to land management. The evaluation of an operational GIS. Cartographica, 21, 2&3, 40-46.

Cress, J. J. and Deister, R. R. P. 1990: Development and implementation of a knowledge-based GIS geological engineering system. Photogrammetric Engineering and Remote Sensing, 56, 11, 1529-1535.

Cromley, R. G. and Morse, S. J. 1988: Some comments on numerical line generalisation. Geographical Analysis, 20, 3, 263-269.

Crone, D. R. 1953: The accuracy of topographical maps. Empire Survey Review, 12, 88, 64-70.

Cross, P. A. 1985: Prospects for satellite and inertial positioning methods in land surveying. Paper presented to 'Surveying in Tomorrow's world'. Newcastle University: Newcastle.

Csillag, F. 1987: Mapping the spatial variability of soils: a cartographer's approach. Chrisman, N. R. (ed) Proceedings Auto-Carto 8, pp155-164.

Curren, P.J. 1985: Principles of remote sensing. Longman: London.

Dangermond, J. 1983: A classification of software components commonly used in geographic information systems. Proceedings, US/Australia workshop on design and implementation of computer based geographic information systems. pp70-91

Dangermond, J. 1984: Description of techniques for automation of regional natural resource inventories. in Hopkins P. F. (ed.) Extraction of Information from Remotely Sensed Images. ACSM: Falls Church, Va.

Dangermond, J. 1988: A review of digital data commonly available and some of the practical problems of entering them into a GIS. Technical Papers ACSM-ASPRS Annual Convention, pp1-10.

Dangermond, J. 1990: The future of GIS technology. The 1990 GIS source book. pp7-12. GIS World Inc: Fort Collins, Colorado.

Dangermond, J. and Morehouse, S. 1987: Trends in hardware for geographic information systems. Chrisman N. R. (ed.) Proceedings, Auto-Carto 8, pp380-385.

Daratech, 1989: Geographical Information Systems Markets and Opportunities, Daratech Inc, Cambridge, Mass.

Davies, C. G. 1982: An investigation of line generalisation in Computer-assisted cartography. MSc. Thesis: University of Manchester, Institute of Science and Technology.

Davis, W. A. and Hwang, C. H. 1986: Organizing and indexing of spatial data. Proceedings, second international symposium on spatial data handling. ICA: Seattle, Washington

De Bruijn, C. A. 1984. MASMAP, design for a project orientated geo-information program package for urban upgrading schemes. ITC Journal, 2, 95-103.

De Jong, S. M. 1990: Integration of remotely sensed and GIS data to determine SPOT classification accuracy. Proceedings EGIS '90, pp517-525

DeMan, W. K. E. (ed.) 1984: Conceptual framework and guidelines for evaluating Geographic Information Systems. UNESCO: Paris.

Demko, G. J. 1988: Geography beyond the ivory tower. Annals of the Association of American Geographers, 78, 4, 575-579.

Denegre, J., Deschamps, J. C. and Galtier, B. 1986: Landsat and SPOT high resolution satellite images: a new component for geographic data bases. Proceedings, Auto-Carto London, Vol.1, 527-537.

De Vorsey, L. 1973: Florida's seaward boundary: a problem in applied historical cartography. Professional Geographer, 25, 3, 214-220.

Dickinson, H. J. 1989: Techniques for establishing the value of geographic information systems. Proceedings GIS/LIS '89, pp412-420.

Dickinson, H. J. and Calkins, H. W. 1988: The economic evaluation of implementing a GIS. International Journal of Geographical Information Systems, 2, 4, 307-327.

Dickinson, H. J. and Calkins, H. W. 1990: Comment on 'concerning "The economic evaluation of implementing a GIS"'. International Journal of Geographical Information Systems, 4, 2, 211-212.

Dicks, S. E. and Lo, T. H. C. 1990: Evaluation of thematic map accuracy in a land-use and land-cover mapping program. Photogrammetric Engineering and Remote Sensing, 56, 9, 1247-1252.

Digital Cartographic Data Standards Task Force 1988: The proposed standard for digital cartographic data. American Cartographer, 15, 9-144.

Distefano, J. 1986: Report to the National Committee for Cartographic Data Standards. in Moellering, H. (ed.) Digital cartographic data standards: a report on evaluation and empirical testing, pp44-51. National Committee for Digital Cartographic Data Standards: Columbus, Ohio.

Dobson, M. W. 1985: The future of perceptual cartography. Cartographica, 22, 3. 27-43.

Dodani, M. H., Hughes, C. E. and Moshell, J. M. 1989: Separation of powers. Byte, March, 255-262.

DOE, 1987: Handling geographic information. Report of the committee of enquiry chaired by Lord Chorley. HMSO: London.

Donelson, W. C. 1978: Spatial management of information. Computer Graphics, 12, 203-290.

Dougenik, J. 1980: Whirlpool; a geometric processor for polygon coverage data. Proceedings of the Int. Symposium on Cartography and Computing: Applications in Health and Environment, pp304-311. ACSM: Falls Church, Va.

Douglas, D. and Peuker, T. K. 1973: Algorithms for the reduction of the number of points required to represent a digitised line or its caricature. The Canadian Cartographer, 10, 2, 112-122.

Dowman, I. J. 1990: Remote sensing in topographic mapping. Mapping Awareness, 4, 5, 26-27.

Doyle, F. J. 1984a. The economics of mapping with space data, in ITC Journal, 1, 1-10.

Doyle, F. J. 1984b. Surveying and mapping with space data. ITC Journal, 4, 314-321.

Drummond, J. E. 1989: Models and data quality handled in a dBASE/AutoCad GIS. ITC Journal, 3/4, 205-211.

Duffield, B. S. 1975: The delineation of recreational landscapes: the role of a computer based information system. Transactions of the Institute of British Geographers, 66, 141-148

Dunn, R., Harrison, A. R. and White, J. C. 1990: Positional accuracy and measurement error in digital databases of land use: an empirical study. International Journal of Geographical Information Systems, 4, 4, 385-398.

Dutton, G. H. 1981: Fractal enhancement of cartographic line detail. The American Cartographer, 8,1, 23-40.

Dutton, G. H. 1984: Truth and its consequences in digital cartography. Technical papers, ACSM 4th annual meeting, Washington, pp278-283.

Dymon, U. J. 1989: Do we really know our map users? Cartographica, 26, 3-4, 49-58.

Eason, K. 1988: Information technology and organisational change. Taylor Francis: London.

Easterfield, M. E., Newell, R. G. and Theriault, D. G. 1990: Version management in GIS: applications and techniques. Proceedings of EGIS '90. p288-297.

Egeinhofer, M. J. and Frank, A. U. 1989: Object orientated modelling in GIS: inheritance and proportion. Proceedings, Auto-Carto 9, pp. 588-598.

Ehlers, M., Edwards, G. and Bedard, Y. 1989: Integration of remote sensing with geographic information systems: a necessary evolution. Photogrammetric Engineering and Remote Sensing, 55, 11, 1619-1627.

Elgin, R. L. 1982: Legal princples of boundary relocation for Arkansas - a study summary. Technical papers ACSM/ASP, Fall Convention. Falls Church, Virginia

Epstein, E. F. 1988: Legal and institutional aspects of global databases. in Mounsey, H. and Tomlinson, R. Building databases for global science. Taylor and Francis: London.

ESRI, 1987: ARC/INFO Starter kit manual. ESRI: California.

ESRI, 1987: ARC/INFO ARCEDIT manul. ESRI: California.

ESRI, 1987: ARC/INFO Overlay manual. ESRI: California.

ESRI, 1987: ARC/INFO Starter kit manual. ESRI: California.

ESRI, 1990: Understanding GIS. The ARC/INFO method. ESRI: California.

ESRI, 1990: The ARC/INFO Rev.5.1 release: Major enhancements to the world's leading GIS software. Mapping Awareness, 4, 10, 57-63.

Fagan, G. L. 1986: Routing of municipal service vehicles in the town of Oyster Bay, Long Island, New York. Technical Papers ACSM/ASPRS Annual Convention, Vol 3, pp20-30.

Faintich, M. B. 1986: Digital cartographic data bases: advanced analysis and display technologies. Mapping from modern imagery: Proceedings of an International Symposium of the International Society for Photogrammetry and Remote Sensing and the Remote Sensing Society. pp480-499.

Fisher, P. F. 1989: Knowledge-based approaches to determining and correcting areas of unreliability in geographic databases, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Flowerdew, R. and Green, M. 1989: Data integration: statistical methods for transferring data between zonal systems. Proceedings of the 29th Regional Science Association European Congress Cambridge. UK.

Flowerdew, R. and Green, N. 1989: Statistical models for inference between incompatible zonal systems, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Foster, M. J. and Shand, P. J. 1990: The Association for geographic information yearbook 1990. Taylor and Francis: London.

Fox, L., Brockhaus, J. A. and Tosta, N. D. 1985: Classification of timberland productivity in Northwestern California using Landsat, topographic and ecological data. Photogrammetric Engineering and Remote Sensing, 51, 11, 1745-1752.

Frank, A: 1987: Overlay processing in spatial information systems. Chrisman, N. R. (ed.) Proceedings Auto-Carto 8, pp16-31.

Franklin, W. R. 1984: Cartographic errors symptomatic of underlying algebra problems. Proceedings, International Symposium on Spatial Data Handling. University of Zurich: Zurich.

Fulford, M. 1981 The Fastrack automatic digitizing system. Pattern Recognition, 14, 1, 6, 65-74.

Gaits, G. 1969: Thematic mapping by computer. Cartographic Journal, 6, 50-68.

Gardels, K. 1987: The expert geographic knowledge system: applying logical rules to geographic information Chrisman, N. R. (ed.) Proceedings, Auto-Carto 8, pp520-529.

Gatrell, A. C. 1989: On the spatial representation and accuracy of address-based data in the United Kingdom. International Journal of Geographical Information Systems, 3, 4, 335-348.

Germsmehl, P. J. 1985: The data, the reader and the innocent by-stander. The Professional Geographer, 37, 329-334.

Giermam, D. M. (et al). 1975: Remote sensing and Canadian geographic information systems for impact studies. in Thompson, G. E. (ed.) Third Canadian Symposium on Remote Sensing, pp235-241. Canadian Aeronautics and Space Institute: Edmonton, Alberta.

GIMMS. 1988: GIMMS Reference Manul Version 5. GIMMS: Edinburgh.

GIS World Inc. 1990: The 1990 GIS source book. GIS World Inc: Fort Collins, Colorado.

Glanzer, M. N. 1980: A Lawyer's review of the Dolphin Lone Case. Technical papers, ACSM Fall Technical Meeting, section MS-3-B.

Goldberg, E. 1970: Urban information systems and invasions of privacy. Urban Affairs Quarterly, 249-64.

Goodchild, M. F. 1978a: Statistical aspects of the polygon overlay problem in Vol. 6. Dutton, G. (ed.), Harvard papers on Geographic Information Systems, Addison Wesley.

Goodchild, M. F. 1978b: Approaches to the estimation of geographical measures: a fractal framework. Department of Geography, University of Western Ontario.

Goodchild, M. F. 1980a: Fractals and the accuracy of geographic measures. Jnl. Int. Assoc. of Mathematical Geology, 12,2, 85-98.

Goodchild, M. F. 1980b: The effects of generalisation in geographical data encoding. in Freeman, H. and Pieroni, G. G. (eds.): Map Data Processing, pp191-205. Academic Press: NY.

Goodchild, M. F. 1983: A classification of software components commonly used in geographic information systems, Proceedings of US/Australia Workshop on Design and Implementation of Computer Based Geographic Information Systems. IGU: Amherst, NY.

Goodchild, M. F. 1987: Application of a GIS benchmarking and workload estimation model. in Frazier, J. F. et al. (eds.) Papers and proceedings of Applied Geography Conferences Volume 10, 1-6. SUNY-Binghampton: Binghampton, NY.

Goodchild, M. F. 1988a: Stepping over the line: technology and the new cartography. American Cartographer, 15, 311-319.

Goodchild, M. F. 1988b: The issue of accuracy in global databases. in Mounsey, H. and Tomlinson, R. Building databases for global science. Taylor and Francis: London.

Goodchild, M. F. 1989a: The National Center for Geographic Information and Analysis: Year 1. ACSM Bulletin, 123, 35-40.

Goodchild, M. F. 1989b: Modelling error in objects and fields, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Goodchild, M. F. and Rizzo, B. 1986: Performance evaluation and workload estimation for geographic systems. Proceedings, Second International Symposium on Spatial Data Handling, pp437-509. International Geographical Union: Williamsville, New York.

Goodchild, M. F. and Rizzo, B. R. 1987: Performance evaluation and workload estimation for geographic information systems. International Journal of GIS, 1, 1, 67-76.

Goodchild, M. F. and Dubuc, O. 1987: A model of error for choropleth maps, with applications to geographic information systems. in Chrisman, N.R. (ed.) Proceedings Auto Carto 8, pp165-174.

Goodchild, M. F. and Mark, D. 1987: The fractal nature of geographic phenomena. Annals of the Association of American Geographers, 77, 2, 265-278.

Goodchild, M. F. and Gopal, S., eds, 1989: Accuracy of Spatial Databases. Taylor and Francis: London.

Goodchild, M. F. and Min-hua, W. 1989: Modelling errors for remotely sensed data input into GIS. Proceedings Auto Carto 9, pp530-537.

Goodchild, M. F. and Wang, M. 1989: Modelling error for remotely sensed data input of GIS. Proceedings, Auto-Carto 9, pp530-537.

Goos, M. 1990: The global positioning system and its GIS applications. The 1990 GIS source book. pp254-259. GIS World Inc.: Fort Collins, VA.

Gould, M. D. 1989: Human factors research and its value in GIS user interface design. Proceedings GIS/LIS '89, 542-550.

Grant-Head, C. 1984: The map as a natural language: a paradigm for understanding. Cartographica, 21, 1, 1-32.

Green, N. P., Finch, S. and Wiggins, J., 1985: The 'state of the art' in geographic information systems. Area, 17, 4, 295-301.

Green, N. 1990: Statistical models for areal interpolation. Proceedings of EGIS '90, pp392-401

Greenland, A., Socher, R. M. and Thompson, M. R. 1985: Statistical evaluation of accuracy for digital cartographic data bases. Proceedings, Auto-Carto 7, pp212-221.

Greulich, G. H. 1980: Where the grass stops growing. Technical papers ACSM, Fall Technical Meeting, section MS-3-c-1 - MS-3-8.

Guelke, L. 1977: Cartographic communication and geographic understanding. Cartographica, Monograph 19, 129-145.

Guevara, J. A. 1989: On the design of geographic information systems procedures. Proceedings, Auto-Carto 9, pp789-797.

Gupta, A. K. and Rao, V. R. 1984: Mapping of Wolfamile region in the Sirchi district in India from different digitally enhanced data products of Landsat. in Klepfer, M. M. and Morrison, D. B. (eds.) Machine Processing of remotely nystsensed data, pp54-61 Purdue Research Foundation: Indiana.

Guptill, S. C. 1986: A new design for the US Geological Survey's national digital cartographic data base. Proceedings Auto-Carto London, Vol. 1, 10-18.

Guptill, S. C. 1987: Desirable characteristics of a spatial database management system. Chrisman, N. R. (ed.) Proceedings Auto-Carto 8, pp278-281.

Guptill, S. C. 1989a: Evaluating geographic information systems technology. Photogrammetric Engineering and Remote Sensing, 55, 11, 1583-1587.

Guptill, S. C. 1989b: Inclusion accuracy data in a feature based, object-oriented data model. in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Guptill, S. C. 1990: How to benchmark your geographic information system. ACSM Bulletin, 128, 49-55.

Haemers, P. B. M. 1990: Integration of GIS and remote sensing data accuracy assessment. Proceedings of EGIS '90, pp428-436.

Hakanson, L. 1978. The length of closed geomorphic lines. Journal of the International Association of Mathematical Geology, 10, 2, 141-167.

Hallam, C. A. 1979. The USGS geographic information retrieval and analysis system: overview. Proceedings of the American Congress of Surveying and Mapping. Annual Meeting, pp229-246.

Hansen, D. B. 1984: Generic geographic information systems: on evaluation of FWS and BLM Geographic data processing capabilities. ASP-ACSM Tech. Papers 50th Annual Meeting, Volume 2, 712-727.

Hardy, D. D. 1985: Today the earth is 100 000 000 000 000 bits. Advanced Technology for monitoring and processing global environmental data, pp1-4. Remote Sensing Society: London.

Harley, J. B. 1975: Ordnance Survey Maps: a descriptive manual. Southampton: Ordnance Survey.

Harley, J. B. 1966: Error and revision in early Ordnance Survey maps. The Cartographic journal, 5, 115-124.

Harmrin, R. D. 1981: The information economy. The Futurist, 15, 4, 25-30.

Harris, R. 1987: Satellite Remote Sensing. An introduction. Routledge and Kegan Paul: London.

Harrison, A. and Bell, S. 1983: Recent database developments of the ECU. Prepared for the Royal Society Meeting, 'Cartographic Data Structures and GIS' London, July 4.

Hartley, R. and Higgins, M. 1987: Developing a GIS for the European Community within the Corine programme. Paper presented at the May 1987 SORSA Workship, Durham.

Haywood, P. 1986: The differing requirements of producers and users of data and the design of topographic cartographic databases. Cartographic Journal, 23, 131-134.

Haywood, P. 1987: The Ordnance Survey topographic data base study. The Ordnance Survey, Southampton.

Healey, R. G. and Desa, G. B. 1989: Transputer based parallel processing for GIS Analysis: problems and potentialities Proceedings, Auto-Carto 9, pp90-99.

Heivly, C. G. 1986: Using expert systems concepts to fix USGIS digital boundaries. Proceedings, Second International Symposium on spatial data handling. ICA: Seattle, Washington.

Hemenway, D. 1985: Interview: David Rhind in Photogrammetric Engineering and Remote Sensing, 51, 10, 1651-1658.

Herbert, G. 1990: The use of knowledge-based techniques for map design and generalisation. South Eastern Regional Research Laboratory working report, 25. SERRL, Birkbeck College: London.

Herot, C. F. 1980: A prototype spatial data management systems. Computer Graphics, 14, 3, 63-70.

Herring, J. R. 1990: TIGRIS: a data model for an object-oriented geographic information system. Proceedings of GIS design models and functionality. Leicester 21-22 March.

Hessdorfer, R. W. 1976: An interactive Geog. Info. Sys. for the exploration and display of demographic data utilising the thematic map as its medium. Progress Report, May 1976.

Heuvelink, G. B. M., Burrough, P. A. and Stein, A. 1989: Propagation of errors in spatial modelling with GIS. International Journal of Geographical Information Systems, 3, 4, 303-322.

Heuvelink, G. B. M., Burrough, P. and Leenaers, H. 1990: Error propagation in spatial modelling with GIS. Proceedings of EGIS '90, pp453-462.

Hilton, K. 1991: Spaceship Earth. Geographical Magazine Supplement. LXIV, 3, 1-32.

House of Lords 1983: Select Committee on Science and Technology: Remote Sensing and digital mapping. HMSO: London.

Ibbs, T. J. and Stevens, A. 1988: Quadtree storage of vector data. International Journal of GIS, 2, 1, 43-56.

ICA. 1987: Cartogrtaphic activities in the United Kingdom, 184-1987. Cartographic Journal, 24, 97-116.

Ihemadu, S. O. 1985: Satellite remote sensing as a mapping tool. ITC Journal, 3, 187-191.

INTERGRAPH. 1989: The next generation of GIS. Mapping Awareness, 3, 5, 3-8.

Issacs, I. K. 1987: Looking for digital data errors - a misconception. in Rochon, G. and Cornut, J. Geomontics applied to municipal management, pp47-54. Canadian Institute of Surveying and Mapping.

Jackson, M. 1984: Image analysis and digital cartography within the NERC - thematic information service: The development of a unified spatial model Photogrammetric Record, 11, 64, 383-393.

Jackson, M. J. and Mason, D. C. 1986: The development of integrated geo-information systems. International Journal of Remote Sensing, 7, 6, 723-740.

Jackson, M. J. Thomas I. L. and Steward, N. J. 1986: Report of a workshop on geographical information systems held in London on 13 September 1985. International Journal of Remote Sensing, 7, 6, 745-755.

Jaske, R. 1984: FEMA's Integrated Emergency Management Systems (IEMIS). Federal Emergency Monitoring Agency, Unpublished paper.

Jenks, A. and Caspell, F. 1971: Error on choropleth maps: definition, measurement, reduction. Annals of the Association of American Geographers, 61, 217-244.

Jenks, G. F. 1979: Thoughts on line generalisation. Proceedings, Auto Carto IV, pp205-220.

Jenks, G. F. 1981: Lines, computers, and human frailties, Annals of the Association of American Geographers, 71, 1, 1-10.

Jenks, G. F. 1989: Geographic logic in line generalisation. Cartographica, 26, 1, 27-42.

Joao, E., Herbert, G. and Rhind, D. 1990: The measurement and control of generalisation effects. South Eastern Regional Research Laboratory working report, 19. SERRL, Birkbeck College: London.

Joao, E., Rhind, D. W., Openshaw, S. and Kelk, B. 1990: Generalisation and GIS databases. Proceedings EGIS '90, 504-516.

Johnannsen, T. 1978: Testing the accuracy of cartographic equipment. Nachrichten Auds bom Karton-und vermessungswessen, 2, 35, 43-54.

Jones, C. B. and Abraham, I. M. 1987: Line generalisation in a global cartographic database, Cartographica, 24, 3, 32-45.

Kainz, W. 1987: Co-ordinage transformations in map digitizing. in Chrisman, N. R. [ed]. Proceedings Auto-Carto 8, pp107-111.

Karson, A. J. 1980: The artistic elements in map design. The Cartographic Journal, 17,2, 124-127.

Keats, J. S. 1982: On the nature of maps and map logic. in Guelke, L. (ed) The nature of cartographic communication. Cartographica, Monograph 19, 1-7.

Keats, J. S. 1984: The cartographic art. Cartographica, 21, 1, 87-43.

Kelly, R. 1985: Landsat data used in water use surveillance in Arizona in Landsat Data Users Notes, 33, 1-2.

Kennedy, S. 1989: The small number problem and the accuacy of spatial databases, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Kennie, T. J. M and Mathews, A. C. (eds). 1985: Remote sensing in civil engineering. Surrey University Press: Glasgow.

Kennie, T. J. M. 1985: Remote sensing scanning systems in Kennie, T. J. M. and Mathews, M. C. eds Remote sensing in civil engineering. pp 41-86. Surrey University Press: Glasgow.

Kinnear, C. 1987: The Tiger structure. Chrisman, N. R. [ed] Proceedings, Auto Carto 8, pp249-257.

Kishimoto, H. 1968: Cartometric measurements. Juris Druck and Verlag: Zurich.

Kjerne, D. and Deuker, K. J. 1986: Modelling cadastral spatial relationships using an object oriented language. Proceedings, Second International Symposium on spatial data handling, pp142-157.

Koch, E. I. and Strurz, H. 1986: Geo-support information system: concepts and facilities, Department of City Planning, NY.

Kolancny, A. 1977: Cartographic information - fundamental concepts and terms in modern cartography. Cartography, 19, 28-45.

Krishnamuri, R. 1980: The arithmetic of shapes. Environment and Planning Series B, 7, 463-484.

Krogulecki, M. 1990: A prototype decision guide and audit log for preparation of spatial databases. Photogrammetric Engineering and Remote Sensing, 56, 4, 505-509.

Kubo, S. 1986: The basic scheme of TRINITY: A GIS with intelligence. Proceedings, second International Symposium on spatial data handling.

Laboratory for Computer Graphics and Spatial Analysis. 1985: Lablog 4. LCGIA: Harvard, USA

Lai, Poh-Chi. 1988: Resource use in manual digitization. A case study of the Patuxent basin geographical information system data base. International Journal of GIS, 2, 4, 329-346.

Lakoff, G. 1987: Women, fire and dangerous things. What categories reveal about the mind. University of Chicago Press: London.

Lambert, J. 1984: Getting human responses. Practical Robotics. May/June, 20-24.

Lambert, J. 1984: Putting our trust in an expert system. Practical Robotics, Sept/Oct, pp51-52.

Lamsac, 1987: Digital Mapping in local authorities Detailed analysis of 1986 survey Report #2. Lamsac: London.

Land, J. 1986: Integration of data from a map information system and a remote sensing system. Mapping from modern imagery: Proceedings of an International Symposium of the International Society for Photogrammetry and Remote Sensing and the Remote Sensing Society, pp93-98.

Land, N. 1989: Transfer standards. Mapping Awareness, 3, 1, 42-43.

Landsat Data Users Notes 1984: Commercialisation legislation passes in Congress, 31 1-4. NOAA: Sioux Falls, South Dakota.

Landsat data users notes: 1985. Congress transfers Landsat to private sector. Landsat Data User Notes, 34, 1-2.

Landsat data users notes, 1986: Expectations of Landsat commercialisation. Landsat Data User Notes, 35, 15-17.

Langefors, B. 1977: Information Systems Theory. Information Systems, 2, 207-219.

Langeraar, W. 1984: Surveying and Charting of the seas. Elsevier: Amsterdam.

Langran, G. E. 1988: Temporal GIS design tradeoffs. Proceedings GIS/LIS '88, pp890-899.

Langran, G. E. 1989a: Accessing spatiotemporal data in a temporal GIS. Proceedings, Auto-Carto 9, pp191-198.

Langran, G. E. 1989b: A review of temporal database research and its use in GIS applications. International Journal for Geographic Information Systems, 3, 3, 215-232.

Langran, G. E. and Poiker, T. K. 1986: Integration of name selection and name placement. Proceedings Second International Symposium on spatial data handling. ICA: Seattle, Washington.

Langran, G. E. and Chrisman, N. 1988: A framework for spatiotemporal information. Cartographica, 25, 3. 1-14.

Lauzon, J. P., Mark, D. M., Kikuchi, L. and Guevara, J. A. 1985: Two dimensional run encoding for quadtree representation. Computer Vision, Graphics and Image Processing, 30, 36-69.

Ley, R. G. 1981: An examination of various cartometric vertical accuracy tests. The Cartographic Journal, 18, 1, 25-31.

Libault, A. 1961: Les measures sur les cartes et leur incertitude. Editions geographies de France: Paris.

Lichtner, W. 1979: Computer-assisted processes of cartographic generalisation in topographic maps. Geo-Processing, 1, 183-199.

Lienberg, E. 1976: SYMAP: its uses and abuses. Cartographic Journal, 13, 1, 26-35.

Lo, T. H. C. Scarpace, F. L. and Lillesand, T. H. 1986: Use of multispectoral spectral profiles in agricultural land-cover classification. Photogrammetric Engineering and Remote Sensing, 52, 4, 535-544.

Lodwick, W. A., Monson, W. and Svoboda, L. 1990: Attribute error and sensitivity analysis of map operations in geographical information systems: a suitability analysis. International Journal of Geographical Information Systems, 4, 4, 413-428.

Loffler, E. and Margules, C. 1980: Wombats detected from space. Remote Sensing of the Environment, 9, 47-58.

Loudon, T. V., Wheeler, J. F., and Andrew, K. P. 1980: Affine transformations for digitised spatial data in geology Computers and Geosciences 6, 397-412.

MacDougall, E. B. 1975: The accuracy of map overlays. Landscape Planning, 2, 23-30.

MacEachren, A. 1985: Accuracy of thematic maps: Implications of choropleth symbolization. Cartographica, 22, 1, 38-57.

McGranaghan, M. 1987: Human interface requirements for vehicle navigation aids. Chrisman, N. R. Proceedings Auto-Carto 8, pp396-402.

McLening, M. 1984: A. Solution looking for a problem. Software, 12-14.

McMaster, R. B. 1986: A statistical analysis of mathematical measures for linear simplification. The American Cartographer, 13, 2, 103-16.

McMaster, R. B. 1987a: The geometric properties of numerical generalisation. Geographical Analysis, 19, 4, 330-346.

McMaster, R. B. 1987b: Automated line generalisation. Cartographica, 24, 21, 74-111.

McMaster, R. B. 1988: A reply to R. G. Cromley and S. J. Morse's " Some comments on numerical line generalisation". Geographical Analysis, 20, 4, 340-342.

McMaster, R. B. 1989a: The integration of simplification and smoothing algorithms in line generalisation. Cartographica, 26, 1, 101-121.

McMaster, R. B. 1989b: Introduction to 'numerical generalisation in cartography'. Cartographica, 26, 1, 1-6.

McMaster, R. B. and Shea, K. S. 1988: Cartographic generalisation in a digital environment: a framework for implementation in a geographic information system. Proceedings, GIS/LIS '88, pp240-249.

Maffini, G., Arno, M. and Bitterlich, W. 1989: Observations and comments on the generation and treatment of error in digital GIS data. in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Maguire, D. J., Worboys, M. F. and Hearnshaw, H. M. 1990: An introduction to object-oriented geographical information systems. Mapping Awareness, 4, 2, 36-39.

Maling, D. H. 1968: How long is a piece of string? The Cartographic Journal, 5, 2, 147-156.

Maling, D. H. 1977: Cartometry - the neglected discipline? in: Kretschmer, I. (ed) Studies in Theoretical Cartography, pp229-246. Franz Deuticke: Vienna.

Maling, D. H. 1989: Measurements from maps: principles and methods of cartometry. Pergamon Press: Oxford.

Mandelbrot, B. B. 1967: How long is the coast of Britain? Statistical self-similarity and fractal dimension. Science, 156, 636-638.

Mandelbrot, B. B. 1977: Fractals: form chance and dimension. Freeman: San Franscisco.

Mantey, P. E, Bennett, J. L. Carlson (eds). 1973: Information for Problem Solving: The development of an interactive GIS. Proc. IEEE Comm. Conf. Seattle, Washington, June 1973.

Marble, D. F. 1979: Integrating Cartographic and Geographic Information Systems Education. ACSM 39th Annual Meeting, pp493-499.

Marble, D. F. 1984: Geographic Information Systems: an overview. Proceedings, Pecora 9 conference, pp18-24. Sioux Falls: South Dakota.

Marble, D. F. 1990: Performance analysis of geographic information systems. Proceedings, GIS for the 1990s, Ottawa, Canada, pp577-586.

Marble, D. F. and Peuquet, D. (eds.): 1983: Geographic Information Systems and remote sensing. in R.N. Colwell (ed.): Marvel of Remote Sensing, pp923-958. ASP: Falls Church, Virginia.

Marble, D. F. and Sen, L. 1986: The development of standardised benchmarks for spatial database systems. Proceedings, Second International Symposium on spatial data handling. pp488-496. International Geographical Union: Williamsville, NY.

Marino, J. S. 1979: Identification of characteristic points along naturally occurring lines; an empirical study. The Canadian Cartographer, 16, 1, 70-80.

Mark, D. M. 1986: The Use of quadtrees in geographic information systems and spatial data handling. Proceedings Auto Carto London, pp517-526.

Mark, D. M. 1987: On giving and receiving directions: cartographic and cognitive issues. Proceedings Auto-Carto 8, pp562-571.

Mark, D. M. 1989: Conceptual basis for geographic line generalisations. Proceedings Auto Carto 9, pp68-77.

Mark, D. M. and Csillag, F. 1989: The nature of boundaries on 'area-class' maps. Cartographica, 26, 1, 65-78.

Mark, D. M. and Frank, A. U. 1989: Concepts of space and spatial language. Proceedings, Auto-Carto 9, pp538-556.

Marsden, L. E. 1960: How the national map accuracy standards were developed. Surveying and Mapping, 20, 4, 427-439.

Marx, R. W. 1986a: The Tiger system: Automating the geographic structure of the United States Census. Government Publications Review, Vol 13, 181-201.

Marx, R. W. 1986b: Computer mapping and the census bureau; the future and you. Paper presented to AM/FM International Eastern Regional conference, Williamsburg, Va. May 5-7.

Mason, R. W. 1990: Data conversion - British Standards. in Foster, M. J. and Shand, P. J. (eds) The Association of Geographic Information Yearbook 1990. Taylor and Francis: London.

Masser, I. 1988: The Regional Research Laboratory Initiative. A progress report. International Journal of Geographic Information Systems, 2,1, 11-22.

Meine, K. H. 1977: Cartographic communication links and a cartographic alphabet. Cartographica, 19, 72-91.

Miller, O. M. and Voskuil, R. J. 1964: Thematic map generalisation. Geographical Review, 54, 1, 13-19.

Miller, R. and Robinson, V. B. 1988: Development of an expert system polygon overlay processor for land-related information systems (LRIS). Proceedings of Annual Conference of the Canadian Society for Civil Engineering, May 23-27, Calgary, Alberta.

Mitchell, W. B. (et al). 1977 GIRAS: A geographic information retrieval and analysis system for handling land use and landcover data, USGS Professional Paper 1059, 16pp.

Moellering, H. 1986a: Developing digital cartographic data standards for the United States. Proceedings, Auto-Carto London, vol 1, 3012-322.

Moellering, H. (ed). 1986b: Digital cartographic data standards: A report on evaluation and empirical testing. (No.7 in Issues in Digital cartographic data standards). NCDCDS: Columbus Ohio.

Moellering, H. and Rayner, J. N. 1981: The harmonic analysis of spatial shapes using dual axis fourier shape analysis (DAFSA). Geographical Analysis, 13, 1, 64-77.

Moffit, F. H. and Mikail, E. M. 1980: Photogrammetry (3rd ed). Harper and Row: New York.

Mohan, J. and Maguire, D. 1985: Harnessing a break-through to meet the needs of health-care. Health and Social Service Journal, 9, 580-581.

Monmonier, M. S. 1980: The hopeless pursuit of purification in cartographic communication; a comparison of graphic arts and perceptual distortions of graytone symbols. Cartographica, 17, 1, 24-39.

Monmonier, M. 1982: Computer Assisted Cartography. Principles and Prospects. Prentice Hall: New Jersey.

Monmonier, M. 1989: Interpolated generalisation: cartographic theory for expert guided feature displacement. Cartographica, 26, 1, 43-64.

Monmonier, M. 1990a: Strategies for the visualisation of geographic time-series data. Cartographica, 27, 1, 30-45.

Monmonier, M. 1990b: Geographic Information Systems (GIS). in Parry, R. B. and Perkins, C. Information sources in cartography. pp214-231. Bowker-Sauer: London.

Morehouse, S. 1985a: Geo-relational data structure for GIS, Proceedings, Auto-Carto 9.

Morehouse, S. 1985b: ARC/INFO: A geo-relational model for spatial information. ESRI: California.

Morehouse, 1989: The architecture of ARC/INFO. Proceedings, Auto Carto 9, pp266-276.

Morrison, J. 1977: The science of cartography and its essential processes. Cartographica, 19, 58-71.

Morse, B. 1987: Expert system interface to a geographic information system. Chrisman N. R. [ed] Proceedings, Auto-Carto 8, pp535-541.

Mounsey, H. and Tomlinson, R. 1988: Building databases for global science. Taylor and Francis: London.

Muller, J. C. 1985: Mental maps at a global scale. in: Cartographica, 22, 4, 51-59.

Muller, J. C. 1986: Fractal dimension and inconsistencies in Cartographic line representations. Cartographic Journal, 23, 123-130.

Muller, J. C. 1987a: The concept of error in cartography. Cartographica, 24, 21, 1-15.

Muller, J. C. 1987b: Fractals and automated line generalisation. Cartographic Journal, 24, 1, 27-34.

Muller, J. C. 1987c: Optimum point density and compaction rates for the representation of geographic lines. Chrisman N. R. [ed] Proceedings Auto-Carto 8, pp221-230.

Muller, J. C. 1989: Theoretical considerations for automated map generalisation. ITC Journal, 3/4, 200-204.

Muller, J. C. 1990: The removal of spatial conflicts in line generalisation. Cartography and Geographic Information Systems, 17, 2, 141-149.

Muller J. C., Johnson, R. D. and Vanzella, L. R. 1986: A knowledge-based approach for developing cartographic expertise. Proceedings, Second International Symposium on spatial data handling. ICA: Seattle, Washington.

Muller, J. C. and Zeschen, W. 1990: A knowledge based system for cartographic symbol design. Cartographic Journal, 27, 1, 24-30.

Nagy, G. 1983: Optical scanning digitizers. Computer, 16, 5, 13-24.

Naithani, K. K. 1990: Can satellite images replace aerial photographs? A photogrammetrist's view. ITC Journal, 1990-1, 29-31.

Nale, D. K. 1990: Photogrammetric solutions for GIS databases. The 1990 GIS source book. pp246-252. GIS World Inc.: Fort Collins, VA.

Nelson, R. and Blakemore, M. 1986: NOMIS - a national geographic information system for the management and mapping of employment, unemployment and population data. Technical Papers 1986 ACSM ASPRS Annual Convention; Cartography and Education, pp20-29. ACSM: Fall Church, Virginia.

Newcomer, J. A. and Szajgin, J. 1984: Accumulation of thematic map errors in digital overlay analysis. The American Cartographer, 11, 1, 58-62.

Newell, R. G., Theriault, D. and Easterfield, M. 1990a: Temporal GIS - modelling the evolution of spatial data in time. Proceedings of GIS design models and functionality. Leicester 21-22 March.

Nickerson, B. G. and Freeman, H. 1986: Development of a rule-based system for automatic map generalisation. Proceedings, Second International Symposium on Spatial Data Handling, ICA: Seattle, Washington.

Norris, F. M. 1965: The nature of paper and board. Pitman: London.

Nystrom, D. A. 1986: Geographic information system developments within the US Geological Survey. Proceedings, Auto-Carto London Vol. 1, pp33-42.

Nystuen, J. D. 1967: Boundary shapes and boundary problems. Peace Research Society Papers, VII Chicago Conference. pp107-128.

O'Neill, M. and Mark, D. 1987: The psi-s plot: a useful representation for digital cartographic lines. Chrisman [ed] Proceedings, Auto-Carto 8, pp231-240.

Onsrud, H. J. 1989: Legal and liability issues in publicly accessible land information systems. Proceedings, GIS/LIS '89, 295-300.

Openshaw, S. 1984: The modifiable areal unit problem. Catmog, 38.

Openshaw, S. 1989: Learning to live with errors in spatial databases, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Openshaw, S. 1990: Why are we still waiting for spatial analysis functionality? The 1990 GIS source book. pp13-16. GIS World Inc: Fort Collins, Colorado.

Openshaw, S. 1990b: A spatial analysis research strategy for the regional research laboratory initiative. RRL Discussion Paper, 3. University of Sheffield Department of Town and Country Planning: Sheffield.

Openshaw, S., Cross, A. and Charlton, M. 1990: Building a prototype geographical correlates exploration machine. The International Journal of GIS, 4, 3, 297-312.

Parker, H. D. 1990: Introduction. The 1990 GIS source book. pp1-6. GIS World Inc: Fort Collins, Colorado.

Pavlidis, T. 1978: A review of algorithms for shape analysis. Computer Graphics and Image Processing, 7, 243-258.

Pearson, M. and Gareau, R. 1986: Exchanging digital base mapping data Proceedings. Second International Symposium on spatial data handling. ICA: Seattle, Washington, ICA.

Pegg, R. K. and Norris, G. 1985: Temperature and humidity variations in map storage areas. Cartographic Journal, 22, 20-22.

Perkal, J. 1966a: An attempt at objective generalisation. Michigan Inter-University Community of Mathematical Geographers, Discussion Paper, 40, 18pp.

Perkal, J. 1966b: On the length of empirical curves. Michigan Inter-University Community of Mathematical Geographers, Discussion Paper, 10, 32pp.

Petchenik, B. 1975: Cognition in cartography. Proceedings of the International symposium on computer assisted cartography. ACSM: Falls Church Virginia.

Petchenik, B. 1985: Maps, markets and money: a look at the economic underpinnings of cartography. Cartographica, 22, 3, 7-19.

Petrie, G. 1985: Remote sensing and topographic mapping. in Kennie, T. J. M. and Mathews, M. C. (eds.): Remote sensing in civil engineering pp119-161. Surrey University Press: Glasgow.

Peuker, T. K. 1976: A theory of the cartographic line. International Yearbook of Cartography, 16, 134-143.

Peuker, T. K. and Chrisman, N. R. 1975: Cartographic data structures. The American Cartographer, 2, 1, 55-69.

Peuquet, D. J. 1984: A conceptual framework and comparison of spatial data models. Cartographica, 21, 4, 66-113.

Peuquet, D. 1986: The use of spatial relationships to aid spatial database retrieval. Proceedings, Second International Symposium on spatial data handling. ICA: Seattle, Washington.

Peuquet, D. J. and Boyle, A. R. 1984: Raster scanning, processing and plotting of Cartographic documents. Spad Systems: Williamsville, NY.

Peuquet, D. J. and Marble, D. F., editors, 1990: Introductory readings in Geographic Information Systems. Taylor and Francis: London.

Phillips, R. 1974: Computer graphics in urban and environmental systems, Proceedings of the IEE, 62, 4, 437-452.

Piazza, P. A. and Pessaro, F. 1990: A cognitive model for a "Smart" GIS. The 1990 GIS Source Book. pp273-284. GIS World: Fort Collins, VA.

Pierce, R. J. 1982: Evolution of the Defence Mapping Agency's digital data specifications: Cardwell, F. S., Black, R., and Cold, B. M. Technical Papers, American Congress on Surveying and Mapping - American Society of Photogrammetry. Fall Convention, ACSM/ASP: Falls Church, Va.

Radwan, M. M. and Suharto, P. 1984. Integrated digital techniques for development of a geographic information system in Indonesia. ITC Journal, 3, 200-205.

Rase, W. D. 1986: The evolution of a graduated symbol software package in a changing graphics environment. Proceedings, Second International Symposium on Spatial data handling. ICA: Seattle, Washington.

Reinhard, C. 1984: State line boundary markers. Wisconsin Mapping Bulletin, 10, 2, p4.

Rhind, D. W. 1973: Generalisation and realism within automated cartographic systems. The Canadian Cartographer, 10, 1, 51-62.

Rhind, D. 1981: GIS in Britain. in Wrigley, N. and Bennett, R. J. (eds.) Quantitative Geography, pp17-31. Routledge and Kegan Paul: London.

Rhind, D. 1987: Recent development in geographical information systems in the UK. International Journal of GIS, 1, 3, 229-242.

Rhind, D. 1988a: A GIS research agenda. International Journal of GIS, 2, 1, 23-28.

Rhind, D. 1988b: Personality as a factor in the development of a discipline: the example of computer-assisted cartography. American Cartographer, 15, 3, 277-289.

Rhind, D. 1990a: The ubiquitous geographical information system. Sci. Publ. Affairs, 5, 2, 57-66.

Rhind, D. W. 1990b: Topographic data bases derived from small scale maps and the future of Ordnance Survey, in Foster, M. J. and Shand, P. J. The Association for Geographic Information Yearbook 1990, pp87-96. Taylor and Francis: London

Rhind, D. W. 1991: The next generation of Geographical Information Systems and the context in which they will operate. Paper presented to ESF Meeting in Davos, January 1991.

Rhind, D. and Hudson, R. 1980: Land use. Methuen: London.

Richardson, L. F. 1961: 'The problem of contiguity', an appendix to statistics of deadly quarrels. General Systems Yearbook, Volume 6, pp139-187.

Rix, D. 1989: Quality control of digital data. Mapping Awareness, 3, 4, 33-36.

Robinson, A. H. et al 1978: Elements of Cartography, New York: Wiley.

Robinson, A. H. and Petchenik, B. 1977: The map as a communication system. Cartographica, 19, 92-110.

Robinson, V. B. et al: 1985: Natural Language in Geographic Data Processing Systems. Paper presented to Int. Conf. on Advanced tech and monitoring and processing global environmental data. University of London, Sept. 1985 pp67-75.

Robinson, V. B. and Frank, A. U. 1985: About different kinds of uncertainty in collections of spatial data. Proceedings, Auto Carto 7, pp440-449.

Robinson, V. B., Frank, A. U. and Blaze, M. A. 1986: Expert systems and geographic information systems: review and prospects. Journal of Surveying Engineering, Vol. 112, 2, 119-130.

Robinson, V. B., Blaze, M. and Thong, D. 1986: Representation and acquisition of a natural language relation for spatial information retrieval. Proceedings, Second International Symposium on Spatial Data Handling, pp472-487. International Geographical Union: Williamsville, NY.

Robinson, V. B., Frank, A. U. and Karimi, H. A. 1987: Expert Systems for geographic information Systems in Resource Management AI Applications, 1, 1, 47-57.

Robinson, V. B. and Frank, A. U. 1987a ; Expert systems for geographic information systems. Photogrammetric Engineering and Remote Sensing, 53, 10, 1435-1441.

Robinson, V. B. and Frank, A. U. 1987b: Expert systems applied to problems in geographic information systems: introduction, review and prospects. Chrisman, N. R. [ed] Proceedings, Auto-Carto 8, pp510.

Rochon, G. 1986: Spot a new cartographic tool. Mapping from modern imagery: Proceedings of an International Symposium of the International Society for Photogrammetry and Remote Sensing and the Remote Sensing Society, pp192-265.

Rollin, J. R. A method of assessing the accuracy of cartographic digitising tables. Cartographic Journal, 23, 144-148.

Root, R. R. (et al). 1986: Vegetation and fire fuel models mapping of North Cascades National Park. Technical Papers, pp78-85. ACSM-ASPRS Annual Convention: Falls Church, VA.

Rowe, C. 1990: People and chips: the human implications of information technology. Blackwell: Oxford.

Rugg, R. and Schmidt, W. 1986a: Testing the interim proposed standard for cartographic features in Moellering H. [ed]. Digital cartographic data standards: a report on evaluation and empirical testing, pp90-96. National Committee for Digital Cartographic standards: Columbus, Ohio.

Rugg, R. D. and Schmidt, W. E. 1986b: Digital cartographic data feature standards in the United States. National Committee for Digital Cartographic Data Standards: Columbus, Ohio.

Ryden, K. 1987: Environmental Systems Research Institute. American Cartographer, 14, 3, 261-263.

Samet, H. 1984: The quadtree and related hierarchical data structures. in Computing Surveys, 16, 2, 187-260.

Samet, H., Shaffer, C. A., Nelson, R. C., Huang, Yuan-Geng., Fuijmura, K. and Rosenfield, A. 1986: Recent developments in quadtree based geographic information systems. Proceedings Second, International Symposium on spatial data handling. ICA: Seattle, Washington.

Sandberg, G. 1981: A primer on relational data base concepts. IBM Systems Journal, 20, 1, 23-41.

Sander, L. M. 1987: Fractal growth. Scientific American, pp82-88.

Schenk, T. and Zilbertson, O. 1990: Experiments with a rule-based system for interpreting linear map features. Photogrammetric Engineering and Remote Sensing, 56, 6, 911-917.

Scholten, H. J. and Stillwell, J. C., editors, 1990a: Geographical information systems for urban and regional planning. Kluwer: Dordrecht.

Scholten, H. and Van der Vlugt, M. 1990b: A review of GIS applications in Eurpoe. in Worrall, L. (ed) Geographic information systems: developments and applications. pp13-40. Belhaven Press: London.

Serpell, D. 1979: Report of the Ordnance Survey Review Committee, HMSO: London.

Shand, P. and Moore, R. 1989: The Association for Geographic Information Yearbook 1989. Taylor and Francis: London

Shapiro, L. and Haralick, R. M. 1980: A spatial data structure. Geo Processing, 1, 313-337.

Shea, K. S. and McMaster, R. B. 1989: Cartographic generalisation in a digital environment: When and how to generalise. Proceedings, Auto-Carto 9, pp56-67.

Shelberg, M. C. and Moellering, H. and Lamb, N. 1982: Measuring the fractal dimensions of empirical cartographic curves. Proceedings, Auto Carto 5, pp481-490.

Singh, 1990: The future of GIS. in Foster, M. J. and Shand, P. The Association for Geographic Information Yearbook 1990. Taylor and Francis: London.

Sleigh, R. W. 1976: Measurment - the key to topographic mapping. Geographical Association of Rhodesia Proceedings, 10, 58-68.

Smallworld Systems 1990: GIS in Europe - summary report. Geodetical Info Magazine, 4, 4, 28-29.

Smith, N. 1987: Data models and data structures for the Ordnance Survey. Ordnance Survey: Southampton.

Smith, T. R. and Pequet, D. J. 1985: Control of spatial search for complex queries in a knowledge based GIS. Paper presented to Advanced technology for monitoring and processing global and environmental data. Remote Sensing Society. pp439-452.

Smith, T. R., Menon, S., Star, J. L. and Estes, J. E. 1987: Requirements and principles for the implementation and construction of large-scale geographic information systems. International Journal of GIS, 1, 9, 13-32.

Snyder, J. 1984: Minimum-error map: Projections banded by Polygons. The Cartographic Journal, 21, 2, 112-121.

Southard, R. B. 1983: The Development of U.S. National Mapping Policy. American Cartographer, 10, 1, 5-15.

Sowton, M. and Haywood, P. 1986: National standards for the transfer of digital map data. Proceedings, Auto-Carto London, Vol. 1, 298-311.

Sowton, M. 1991: The development of digital mapping and support to GIS in the British national mapping agency. in Maguire, D., Goodchild, M. F. and Rhind, D. W. (eds) Geographical information systems. Longman: London.

Sperry, S. L. 1988 The ERDAS and ARC/INFO GIS Workstation - an application for the Chesapeake Bay Critical Area Program. Technical Papers, ACSM - ASPRS, Annual Convention, pp226-231.

Steinhaus, H. 1954: Length, shape and area. Colloquium Mathematician III, Wroclaw, Str. pp1-13.

Steinke, T. R. 1987: Eye movement studies in cartography and related fields. Cartographica, 24, 21, 40-73.

Steward, H. J. 1974: Cartographic generalisation; some concepts and explanation. Cartographica, Monograph 10, 1-77.

Stuheit, J. 1990: GIS and artificial intelligence. GIS World, February, pp37-38.

Taylor, D. R. F. [ed] 1983: Graphic communication and design in contemporary cartography. John Wiley: Chichester.

Teng, A. P. 1986: Polygon overlay processing: a comparison of pure geometric manipulation and topological overlay processing. Proceedings Second International Symposium on spatial data handling, pp102-119.

Thomas, D. 1989: What's in an object? Byte, March, 231-240.

Thomas, H. C., Scarpace, F. L., and Lillesand, T. M. 1986: Use of multitemporal spectral profiles in agricultural land-cover classification. Photogrammetric Engineering and Remote Sensing, 52, 4, 534-544.

Thompson, M. M. 1960: A current view of the national map accuracy standards. Surveying and Mapping, 20, 4, 427-457.

Thompson, M. M. and Rosenfield, G. H. 1974: On map accuracy specifications. Surveying and Mapping, 31, 1, 57-64.

Tobler, W. R. 1965a: Automation in the preparation of thematic maps. Cartographic Journal, 2, 1, 32-38.

Tobler, W. R. 1965b: Computation of the correspondence of geographical patterns. Papers, Regional Science Association 15, 131-139.

Tobler, W. R. 1978a: Comparison of plane forms. Geographical Analysis, 10,2, 154-162.

Tobler, W. R. 1978b: Comparing figures by regression Computer Graphics, 12, 193-195.

Toll, D. L. 1985: Effect of Landsat thematic mapper sensor parameters on land cover classifications. Remote Sensing of Environment, 17, 129-140.

Tolson, J. 1978: Developments in digitising and voice recognition to meet the requirements of hydrography and cartography. International Hydrographic Review, 57, 1, 109-122.

Tomlinson, R. F. (ed). 1972: G1S. The proceedings of the Unesco/IGU first symposium on GIS. pp1152-1281. International Geographic Union Commission on Geographic Data Sensing and Processing.

Tomlinson, R. F. 1987: Current and potential uses of geographic information systems. The North American experience. International Journal of GIS, 1, 3, 203-218.

Townsend, A. R., Blakemore, M. J and Nelson R. 1987: The NOMIS data base availability and uses for geographers. Area, 19, 1, 43-50.

Townshend, J. R. G., 1981: The spatial resolving power of earth resource satellites. Progress in Physical Geography, 5, 1, 32-55.

Trachsler, H. W. 1974: Aerial photographs and orthophotos as a source of data for geographic information systems. Zurich University, Ph.D. thesis.

Traylor, C. J. 1979: The evaluation of a methodology to measure manual digitisation error in cartographic data bases. Unpublished PhD Thesis, University of Kansas.

Trevett, J. W. (1986). Imaging radar for resources surveys. London: Chapman and Hall.

Tsurutani, T., Kashara, Y. and Nawiwada, M. 1980: ATLAS: A geographic database system - Data structure and long design for Geographic Information. Computer Graphics, 14, 13, 71-77.

Tuan, Yi-Fu. 1979: Space and place: the perspective of experience. Arnold: London.

Tukey, J. W. 1979: Methodology, and the statistician's responsibility for both accuracy and relevance. Journal of the American Statistical Association, 74, 786-793.

U.S. Department of Commerce Bureau of The Census, 1985: Tiger Tales. A presentation of the automated geographic support system for the 1990 census. USBC: Washington.

Vahala, M. 1986: Data standardisation of the integrated LIS in Finland. Proceedings, Auto-Carto London, Vol. 1, 323-332.

Van Horn, E. K. 1985: Generalising cartographic databases. Proceedings Auto Carto 7, pp532-540.

Veigner, P. 1989: Learning the language. Byte, March, 245-253.

Veregin, H. 1989: Error modelling for the map overlay operation, in Goodchild, M. F. and Gopal, S. (eds) Accuracy of spatial databases. Taylor and Francis: New York.

Vickers, D. W. and Robinson, G. J. 1986: Generalisation from large to medium and small scale Ordnance Survey maps using expert systems techniques. Proceedings Auto Carto London, pp267-275.

Vonderhoe, A. P., and Chrisman, N. R. 1985: Tests to establish the quality of digital cartographic data; some examples from the Dane County land records project. Proceedings, Auto-Carto 7, pp552-555.

Vonderhoe, A., Melra, D and Crossfield, J. 1984: Technological trends: the Odyssey software system. Wisconsin Land Information Newsletter, 2, 1, 3-4.

Wagner, D. F. 1988: A method of evaluating polygon overlay algorithms. Technical papers, ACSM-ASPRS Annual Convention, pp173-183.

Walsh, S. J., Lightfoot, D. R. and Butler, D. R. 1987: Recognition and assessment of errors in geographic information systems. Photogrammetric Engineering and Remote Sensing. 53, 10, 1423-1430.

Ward, J. R. and Phillips, M. J. 1987: Digitizer technology: Performance characteristics and the effect on the user interface. Computer Graphics and Applications, 7, 4, 31-34.

Waugh, T. C. 1978: GIMMS- An overview. Paper presented to the Harvard Computer Graphics Conference. 22pp.

Waugh, T. 1986: A Response to recent papers and articles on the use of quadtrees for geographic information systems. Proceedings, Second International symposium on spatial data handling. ICA: Seattle, Washington.

Weber, W. 1987: Mass digitisation of topographic cartographic information. Nachrichten aus dem Karton-und Vermessing-swesen Series II, 46, 101-122.

Welch, R. 1984: Merging Landsat 4 digital images with cartographic data bases in: Klepfer, M. M., and Morrison, D. B. (eds.) Machine Processing of Remotely Sensed data with Special Emphasis on Thematic Mapper data and geographic information systems, pp19-20. Purdue Research Foundation: Indiana.

Welch, R. 1985: Cartographic potential of SPOT image data. Photogrammetric Engineering and Remote Sensing, 51, 8, 1085-1091.

Welch, R., and Usery, L. 1984: Cartographic accuracy of Landsat 4 MSS and TM image data. Transactions of Geoscience and Remote Sensing, 22, 3, 281-288.

Weller, B. S. 1985: The significance of automated cartography to an information society. Cartographica, 22, 4, 38-50.

White, E. R. 1983: Assessment of line generalisation algorithms using characteristic points. The American Cartographer, 12, 1, 17-27.

White, M. S. 1978: A geometrical model for error detection and correction. Preceedings, Auto-Carto III, pp439-457.

White, M. S. 1984: Technical requirements and standards for a multipurpose geographic data system. American Cartographer, 1, 1, 115-126.

Wiggins, J. C., Hartley, R. P., Higgins, M. J and Whittaker, R. J. 1986: Computing aspects of a large geographic information system for the European Community. Proceedings, Auto-Carto London, Vol. 2, 28-43.

Wilcox, D. L. 1990: Concerning 'The economic evaluation of implementing a GIS'. International Journal of Geographical Information Systems, 4, 2, 203-210.

Williams, C. M. 1986: The geometric modelling and compression of terrain data. Proceedings, Second International Symposium on spatial data handling. ICA: Seattle, Washington.

Wittgenstein, L. 1953: Philosophical Investigations. MacMillan: New York.

Wood, B, and Douglas, D. 1984; Cartographic feature coding. Cartographica, 21, 2 & 3, 62-72.

Wood, D. 1984: Cultured symbols/thoughts on the cultural context of cartographic symbols. Cartographica, 21, 4, 9-37.

Wood, D. and Fells, J. 1986: Designs on signs/myth and meaning in maps. Cartographica, 23, 3, 54-103.

Worrall, L. (ed) 1990: Geographic information systems. Developments and applications. Belhaven: London.

Wright, J. K. 1942: Map makers are human. Geographical Review, 32, 527-544.

Yan, J. Z. 1982: Geo-cartographics technology transfer and exchange: levels, quicklines and checklist. in Douglas, D. H., Boyle, A. R. (eds.) Computer assisted cartography and geographic information processing, hope and realism. University of Ottawa.

Yan, J. Z., Parker, J. P. 1985: A framework for coordinating the development and applications of street network files for Canada. Statistics Canada: Ottowa.

Young, J. L. 1978: Some facts on the history of surveying. Proceedings of American Congress of Surveying and Mapping. Fall Technical Meeting, pp452-463.

Young, J. A. T. 1986: Remote sensing and an experimental geographic information system for environmental monitoring, resource planning and management. International Journal of Remote Sensing, 7, 6, 741-744.

Zadeh, L. 1965: Fuzzy sets. Information and control, 8, 338-353.

# Appendices

## Appendix 4.1

### Software used

**The hardware coinfiguration.**
The hardware configuration for this project consisted of an enhanced IBM compatible PC/IBM PS2 with mouse, an A0 digitiser and a desk top eight pen plotter. (Appendix 1) This compares favourably with the standard configuration that a small user would find affordable given limited purchasing power. The software that the majority of the analysis has been carried out is ARC/INFO. This is a vector based GIS system and is arguably one of the world leaders in the field at present.

**Software systems available.**
Originally, it was envisaged that investigations would be carried out using the ODYSSEY system marketed by the Harvard Graphics Lab which was resident on the Durham AMDHAL mainframe computer, thus ensuring adequate processing power, unlimited storage facilities, easy transfer of data for analysis to other software packages on the system and finally high quality plotting and output of data. Initial tests were carried out as to the suitability of the WHIRLPOOL polygon overlay processor contained within ODYSSEY. Although the system was not particularly user friendly or well documented WHIRLPOOL performed the basics of polygon overlay. At the same time however the Department of Geography at Durham was given the award of an IBM Institute by the IBM Education Trust. Included in the equipment donation were ten PC-AT's incorporating Professional Graphics and an 8 pen plotter. To fully utilise the potential of such hardware the Geography Department decided to invest in an A0 digitiser and also probed into the possibility of purchasing ARC/INFO, which at the time was only available on mainframes or minicomputers. As a consequence the department became a beta release site and it seemed advantageous to utilise this state of the art technology, for an analysis of integration in GIS, rather than what was rapidly being recognised as the ODYSSEY dinosaur.

**ARC/INFO - that state of the art PC-based GIS.**
ARC/INFO is a vector based GIS, and originally was available for large machines only. In the early 1980s, with the gradual rise of the power of the PC in terms of both processing and the ability to attract potential customers who previously would never have had the spending power to purchase mainframe software let alone the machine to run it. In many respects therefore the systen suffers from being adapted for the PC from mainframe/mini version, rather than having been coded with the limitations of the PC in mind. The PC version was on beta release until February 1987 and has since undergone a variety of enhancements and updates. Unlike its more powerful relatives, the PC version is available in a series of modules allowing the user to purchase only those units applicable to the needs in question. Six modules are currently in existence and these are; STARTER KIT, ARCEDIT, OVERLAY, ARCPOLT, NETWORK and GRIDCONVERSION. This study utilises all but the networking package.

**ARC/INFO's ability to integrate.**
As is to be expected from any GIS worth its salt, ARC/INFO is able to accept data from a variety of different sources and can handle a range of different data formats. Data may be entered from the digitiser manually, or it may be fed in from an image analysis system or

ternatively standard data types such as Digital Line Graph (DLG) can also be accepted. Such ta spatial data sets can be maintained and updated through a variety of management chniques. Analysis is primarily performed by the OVERLAY option as this allows for the perimposition of data sets in several different ways. Two data sets or coverages may be erlaid to find which areas are common to both coverages or alternatively to create a new overage containing all those areas and attributes present in both original coverages. Other nctions include the creation of buffer zones around selected features and the ability to direct arch enquires to such zones.

**he relational data base within the ARC/INFO architecture.**

side from its graphical capabilities ARC/INFO also possesses a very user friendly and ansparent data base management system. This is held within the INFO part of the system. NFO, is in its own right a DBMS and retails as a separate entity to ARC/INFO. The merging a vector based graphical system with a DBMS allows the user to interrogate every aspect of aphic data contained within the data base. Every point, line and area carries its own id, onto hich every piece of information pertaining to that feature may be attached and subsequently arched for and manipulated by the user.

iven that most of the algorithms within the system that dealt with integration and data query ere quite crude, the system was able to provide for its users a suitable backbone onto which ore exacting routines may be perched. In this way, despite its many short comings, the system as viewed as a suitable vehicle for assessing the problems associated with digital cartographic ta integration and flexible enough to allow for user defined modifications to enhance existing erations. The structure of ARC/INFO is based on the concepts of a relational database, thus formation is stored as flat files and may be joined of referred to other sets of information by lating these data sets. ARC/INFO uses the notion of a coverage to group together all the data rtaining to a particular geographic surface.

"a coverage contains both locational data (which define points, lines and polygons) about ch feature. Locations are described by coverage features. Descriptive information about atures is stored in feature attribute tables."

Arc/Info Starter Kit Manual (1988, 3-2)

**he coverage - the basis for data storage within ARC/INFO.**

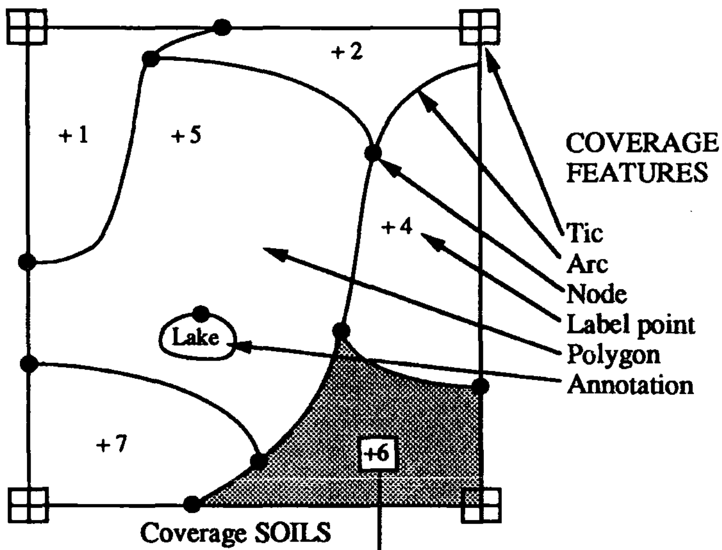
hus the coverage will have both a graphical element and a tabular element. The graphics will stored by a series of x y coordinate points. These may then be aggregated to form lines, hich in turn may form polygons. For example streets may form the boundary of a city block hilst the block will be defined by the streets comprising its border. In addition to their graphic mponent which will define their particular location in space, geographic entities are also made of some sort of descriptive information that give value and quality to mere lines in space. hus each street, for example will have a name, and a length attached to it. It may also have formation on the type of surface it has, whether it is one way or not and what the address nge of its dwellings are. Such data is termed attribute data and are reconciled to the features ey describe by a 'user-id' (diagram a.1).

**he graphic component of the coverage.**

ne graphic components that are dealt with within ARC/INFO are termed arcs, nodes, label

**Diagram a.1**

## THE FEATURE CLASSES OF A COVERAGE WITH RELATED FEATURE ATTRIBUTE TABLE




Feature attribute table SOILS. PAT

| RECNO | AREA | PERIMETER | SOILS | SOILS - ID | SOIL | CLASS | SUITABILITY |
|---|---|---|---|---|---|---|---|
| 1 | -36.0 | 24.0 | 1 | 0 | —— | —— | —— |
| 2 | 3.0 | 9.0 | 2 | 1 | A3 | 113 | HIGH |
| 3 | 2.5 | 8.5 | 3 | 2 | C6 | 95 | LOW |
| 4 | 15.0 | 15.0 | 4 | 3 | B7 | 212 | MODERATE |
| 5 | 4.0 | 8.5 | 5 | 4 | B13 | 201 | MODERATE |
| 6 | 2.0 | 4.5 | 6 | 5 | Z22 | 86 | LOW |
| 7 | 5.5 | 12.0 | 7 | 6 | A6 | 77 | HIGH |
| 8 | 4.0 | 7.0 | 8 | 7 | A1 | 117 | LOW |

**Source: ESRI**

Example of polygon topology

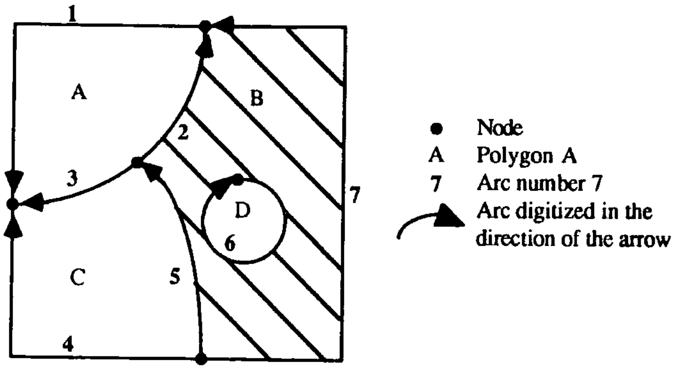
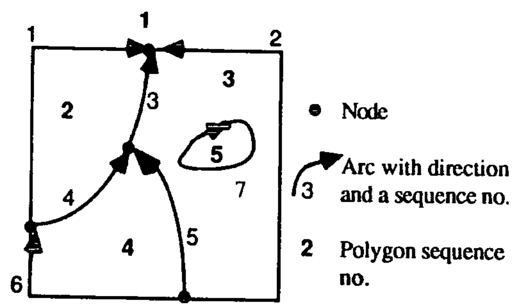



Diagram a.2

| Polygon | No. of Arcs | List of Arcs |
|---|---|---|
| A | 3 | -1, -2, 3 |
| B | 4 | 2, -7, 5, 0, -6 |
| C | 3 | -3, -5, 4 |
| D | 1 | 6 |

Example of contiguity

Diagram a.3




| Arc # | Left Poly # | Right Poly # |
|---|---|---|
| 1 | 1 | 2 |
| 2 | 3 | 1 |
| 3 | 2 | 3 |
| 4 | 2 | 4 |
| 5 | 4 | 3 |
| 6 | 1 | 4 |
| 7 | 3 | 5 |

Example of arc-node topology

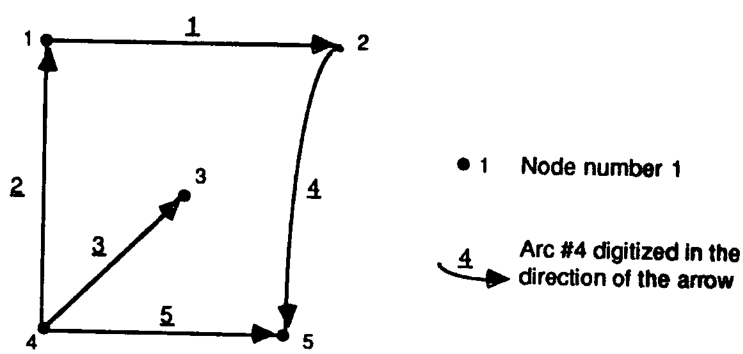



Diagram a.4

| ARC # | FNODE # | TNODE # |
|---|---|---|
| 1 | 1 | 2 |
| 2 | 4 | 1 |
| 3 | 4 | 3 |
| 4 | 2 | 5 |
| 5 | 4 | 5 |

**Source ESRI**

points and polygons. Arcs refer to segments that define networks and linear features such as drainage networks or street plans. Nodes represent the x y coordinates that make up the arcs, whilst label points are coordinates which may either represent the location of an attribute in its own right, such as an historic site or a telephone box. In addition this component may also be used to assign a user id to a polygon feature and also to indicate where labelling text is to be placed. Finally polygons are areas that are fully closed by bounding arcs. These may represent features such as fields, soil types and administrative units as well as any other geographical entity common to either categorical or choropleth mapping. Other geographical features used relate to registering the coverage and maintaining its boundaries and projection. Tics are citied at the time of coverage creation and are points on the source data which can be easily identified. These then form the basis for registering successive layers of data for coverage overlay. A BND file defines the minimum and maximum extent of the coverage.

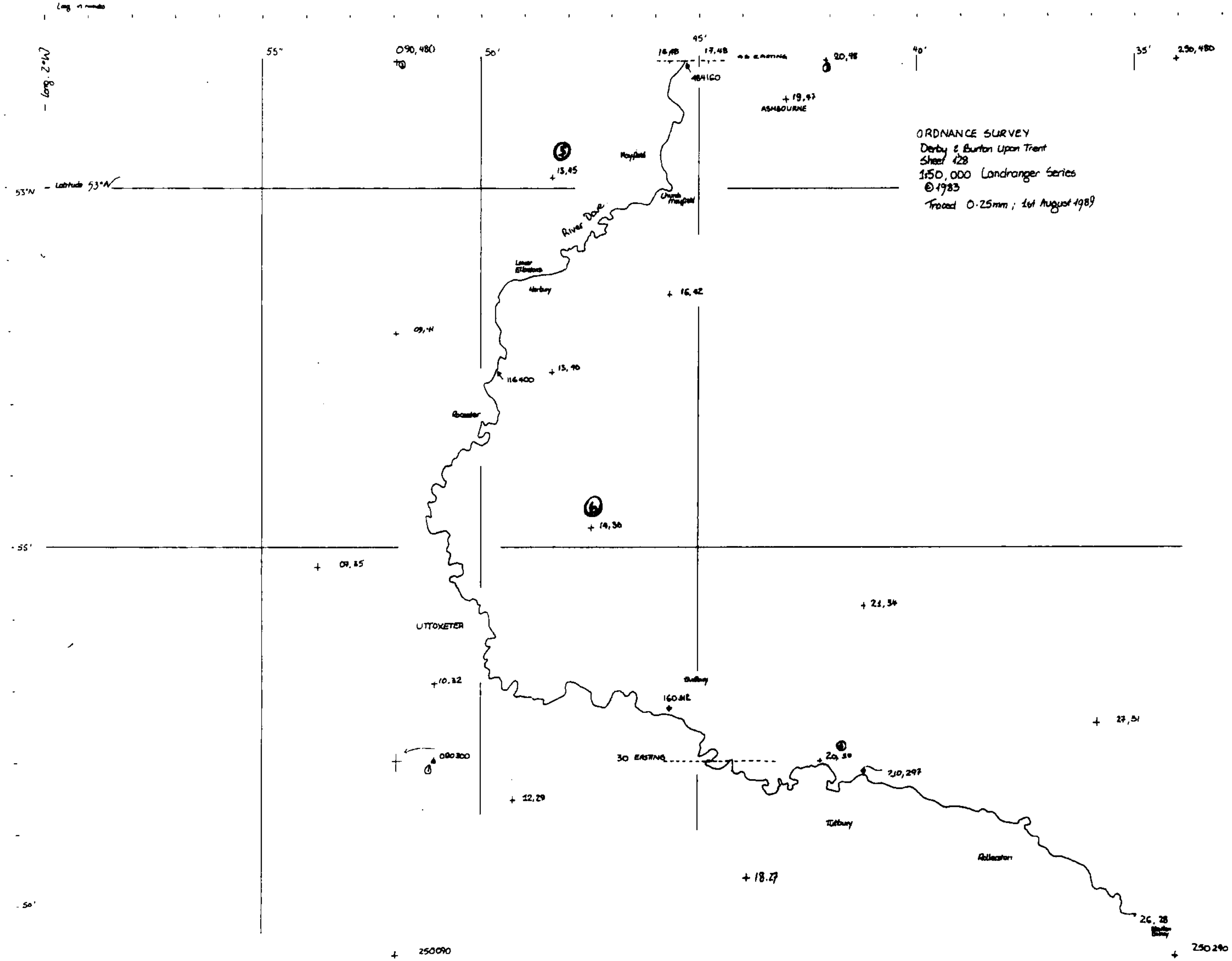
**The tabular component of the coverage.**
The tabular data held in relation to the graphical data is organised on the basis of topology. Some of the data is available to the user for manipulation whilst other sections that are vital to the integrity of the system are stored internally. Internal files contain information on polygon topology as defined by its bounding areas (diagram a.2), on contiguity between areas as defined by the sequence of arcs that define them (diagram a.3) and on arc to node topology. (Diagram a.4). Five types of externally accessible files are created by ARC/INFO as part of the coverage creation process. These are the Point Attribute Table (PAT), the Arc Attribute Table (AAT) the Polygon Attribute Table (PAT), the TIC file and the BND file. Of prime importance in all of the AT files is the cover-id item, as it is with this that all further attributes must be matched. Attribute tables can be set up within the INFO section of the system by defining a set of items to create a template and then filling in that template with attribute information. User defined attribute tables may contain factual information that relates directly to the feature itself, such as the number of people resident in area x, or the degree of pollution contained within a river section represented by line y, in addition attribute tables may also be used for classification for mapping, or for creating a pool of display options such as line widths or shading colours.

The ability of the software to handle the data in this manner enables the user to easily keep tabs on the information that is stored within each coverage feature. Furthermore, searching and manipulation of features can be easily carried out fron the databse. These qualities make the system well suited to asessing the problems associated with looking at land use change and data integration.

# Appendix 4.2 Pascal programs used in Chapter Four

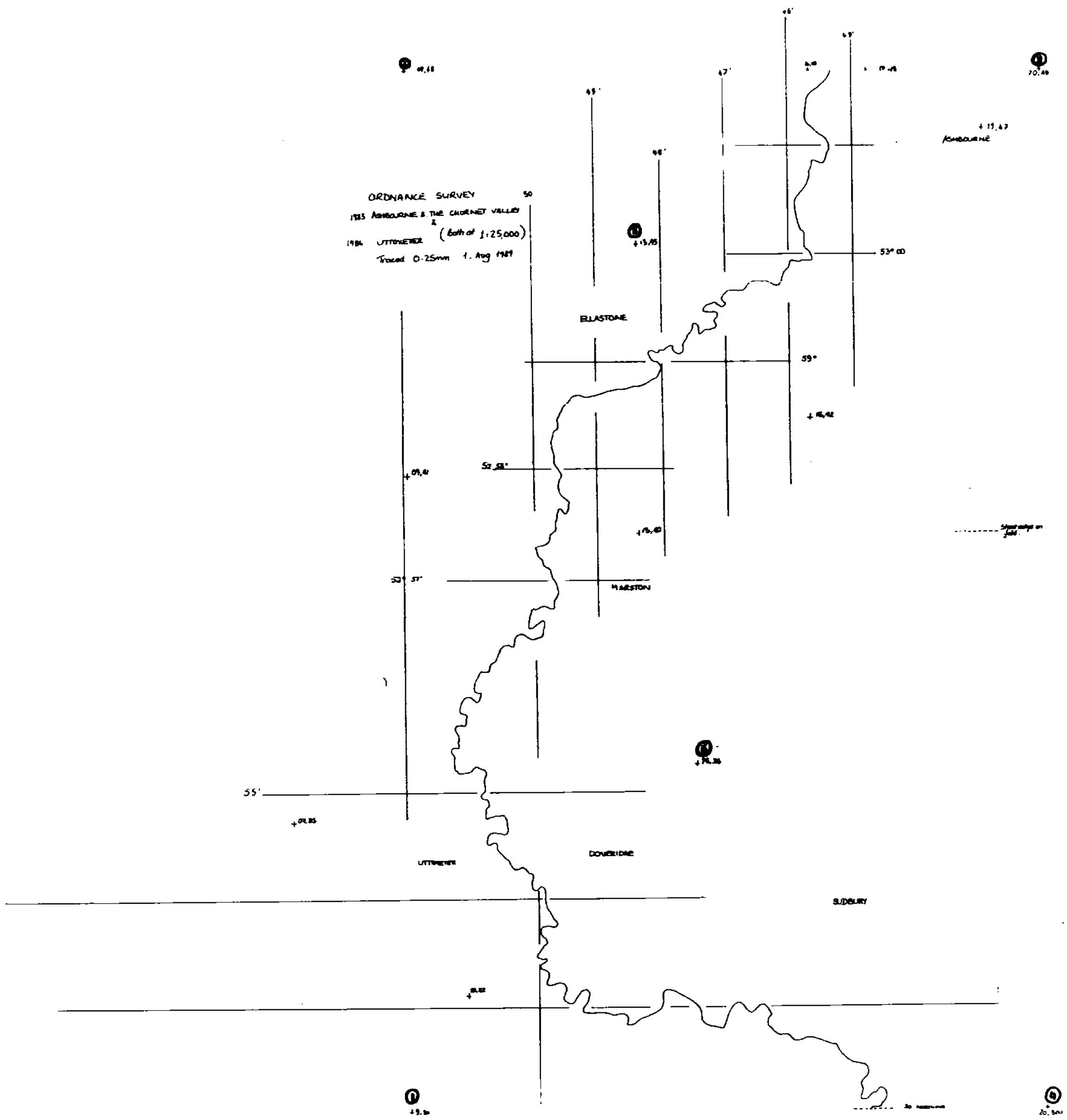
```
program residuals(infile,outfile);
(*Finds the total absolute value (in terms of both the x,y coordinates)*)
(*from the original x,y grid coordinates as compared with those digitised*)
(*by the operators*)

var
 infile,outfile : text;
 truenox,truenoy,index : integer;
 residx,residy,xcood,ycood,totres : real;


begin
 reset(infile,'FILE=indat');
 rewrite(outfile,'FILE=outdat');

 while not EOF(infile) do
   begin
     READLN(infile,index,xcood,ycood);
     residx:=0; residy:=0; truenox:=0; truenoy:=0;
     truenox:=ROUND(xcood);
     truenoy:=ROUND(ycood);
     residx:=xcood-truenox;
     residy:=ycood-truenoy;
     totres:=(abs(residy)+abs(residx))*10;
     WRITELN(outfile,index,totres :10:5)
   end;
 CLOSE(infile);
 CLOSE(outfile)

end.
```

```
program residuals(infile,outfile);
(*Finds the absolute residual value from the original x,y ground coordinates*)
(*as compared with those digitised by the operators*)

var
 infile,outfile : text;
 truenox,truenoy,index : integer;
 residx,residy,xcood,ycood : real;


begin
 reset(infile,'FILE=indat');
 rewrite(outfile,'FILE=outdat');

 while not EOF(infile) do
   begin
     READLN(infile,index,xcood,ycood);
     residx:=0; residy:=0; truenox:=0; truenoy:=0;
     truenox:=ROUND(xcood);
     truenoy:=ROUND(ycood);
     residx:=ABS(xcood-truenox);
     residy:=ABS(ycood-truenoy);
     WRITELN(outfile,index :4,residx :8:5,residy :8:5)
   end;
 CLOSE(infile);
 CLOSE(outfile)

end.
```



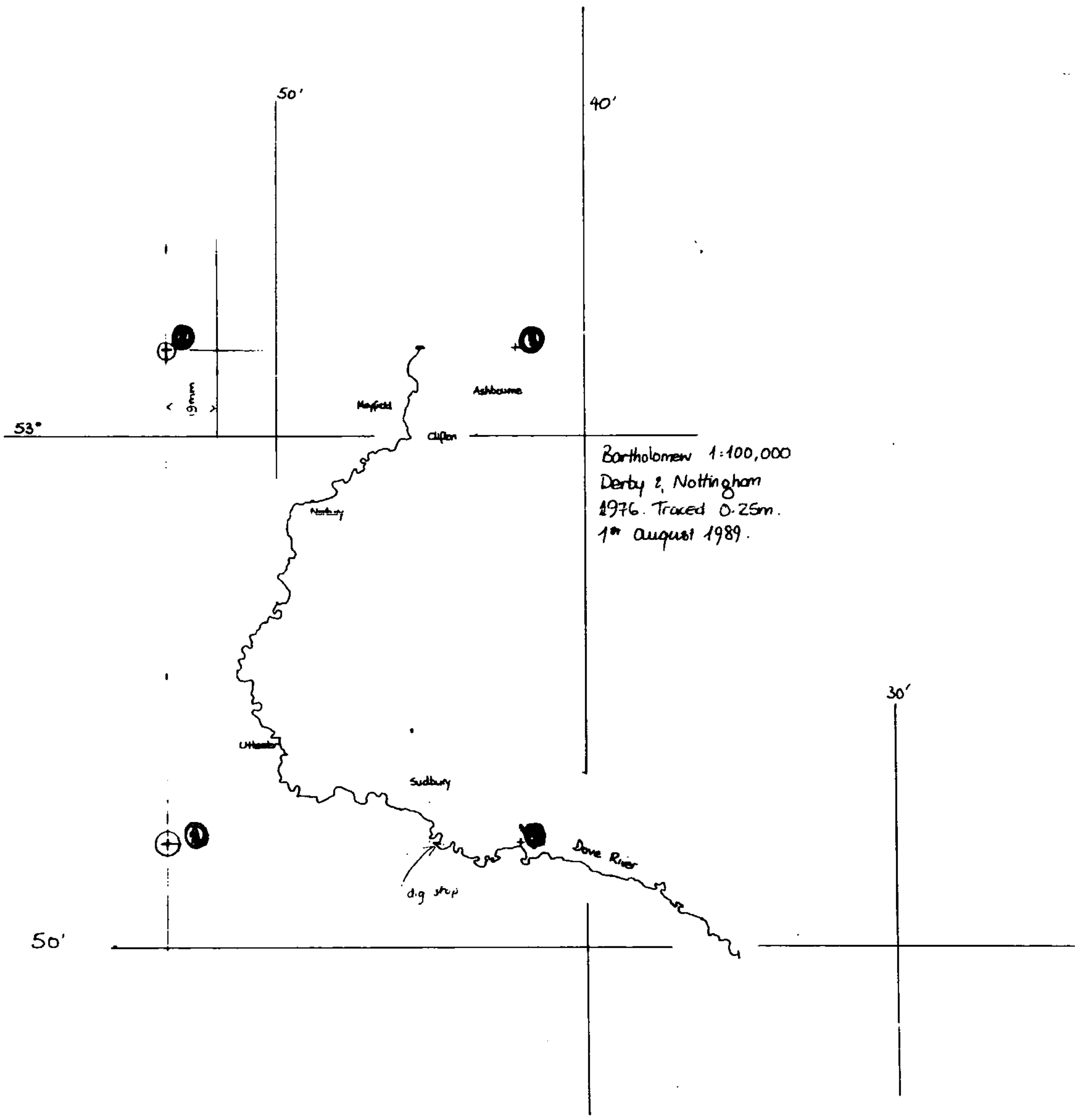

# Appendix 4.3

A reduced copy of the tracing taken from the Ordnance Survey 1:25,000 scale map




**Appendix 4.4**

**A reduced copy of the tracing taken from the Ordnance Survey 1:50,000 scale map**




Appendix 4.5

A reduced copy of the tracing taken from the Bartholomew 1:100,000 scale map

## ppendix 5.1 Pascal programs used for calculations of line differences

```
rogram format(infile,outfile);

*reads in data from the ARC/INFO dump files and rewrites it into a suitable*)
*format for analysis*)

ar
pointer : integer;
mlen,tlen,diff,pdiff  : real;
infile,outfile  : text;
*pointer is the identifier for each line*)
*mlen is the length of the line in the master file*)
*tlen is the length of the line in the test file*)
*pdiff = %diff; diff=tot diff*)

egin
*Set up the files*)
RESET(infile,'FILE=x.diffm');
RESET(outfile, 'x.dm');
WHILE NOT EOF(INFILE) DO
  begin
    READLN(infile,pointer,mlen,tlen,diff,pdiff);
    writeln(outfile,pointer:11,mlen:11:5,tlen:11:5,diff:11:5,pdiff:15:3)
  end;
nd.
```

```
program histogram(infile,outfile);
(*Assesses the number of lines with doubles/triples etc*)

const
  bignumber = 827;

type
  bigarray = array[1..bignumber] of integer;

var
  pointer,pointer2,counter1,counter2,counter3,i,j : integer;
  r1,r2,r3,r4 : real;
  invec,hist : bigarray;
  infile,outfile : text;

procedure resit;
begin
  reset(infile,'file=indat');
  rewrite(outfile,'file=outdat');
  pointer:=0; counter1:=1; counter2:=0; r1:=0; r2:=0; r3:=0; r4:=0;
  while (counter1 <= bignumber) do
      begin
         invec[counter1]:=0;
         hist[counter1]:=0;
         counter1:=counter1+1
      end
end;

procedure readinvec;
begin
   while NOT eof(infile) do
     begin
        counter2:=counter2+1;
        readln(infile,pointer,r1,r2,r3,r4);
        invec[counter2]:=pointer
     end;
   writeln(counter2)
end;

procedure sortit;
begin
 counter3:=1;
 while (counter3 <= counter2) do
    begin
       i := counter3;
       pointer2:=invec[i];
       hist[pointer2] := hist[pointer2]+1;
       counter3:=counter3+1
    end
end;

procedure writeit;
begin
  j:=1;
  while (j <= bignumber) do
    begin
       writeln(outfile,j :5,hist[j]);
       j:=j+1
    end
end;
```

```
egin
  resit;
  readinvec;
  sortit;
  writeit
nd.
```

```
rogram prepare(infile,outfile);
*Prepares the data by coalescing lines with repeated user ids, and also*)
*by counting their number*)

ar
oldpoint,newpoint,countlines : integer;
newlen,oldlen : real;
infile,outfile,outfile2 : text;


egin
 Set up the files }
RESET(infile,'FILE=X.KI');
REWRITE(outfile,'FILE=out.dat');
REWRITE(outfile2,'FILE=out2.dat');

newpoint:=0; oldpoint:=0; newlen:=0; oldlen:=0; countlines:=0;
READLN(infile,oldlen,oldpoint);
WRITELN(outfile2,'Points repeated');
WHILE NOT EOF(INFILE) DO
  begin
    READLN(infile,newlen,newpoint);
    IF newpoint = oldpoint THEN
      begin
         oldlen:=newlen+oldlen;
         countlines:=countlines+1;
         WRITELN(outfile2,oldpoint :6)
      end
    ELSE
     begin
         writeln(outfile,oldlen :6:3,oldpoint :6);
         oldpoint:=newpoint;
         oldlen:=newlen
     end
  end;
WRITELN(outfile,oldlen :6:3,oldpoint :6);
WRITELN(outfile2,'Number of repeated lines', countlines)
nd.
```

```
rogram ridzeros(infile,outfile);
*Reformats file to remove line lengths which were equal to zero,*)
*in preparation for comparison with the original*)

ar
pointer : integer;
mlen,tlen,diff,pdiff  : real;
infile,outfile  : text;


egin
*Set up the files*)
RESET(infile,'FILE=x.diffm');
REWRITE(outfile, 'file=x.mti');
WHILE NOT EOF(INFILE) DO
  begin
    READLN(infile,pointer,mlen,tlen,diff,pdiff);
       if diff <> 0 then
          writeln(outfile,pointer:4,diff:9:3)
  end;
nd.
```

```
Program leeqwe(infile1,infile2,outfile);
(*Compars the line length on the test coverage with the line lengths on *)
(*the original coverage*)

const
 no_columns = 2;
 no_lines = 827;
 no_lines2 = 827;

type
 bigarray = array[1..no_columns,1..no_lines] of real;


var
 counter1,counter2,pointer,KCOUNT,x  : integer;
 length,diff,rat : real;
 vec : bigarray;
 infile1,infile2,outfile : text;


begin
(*Set up the files*)
 RESET(infile1,'FILE=t009.kid');
 RESET(infile2,'FILE=x.pre');
 REWRITE(outfile,'FILE=out.dat');

(*initialise the counters for the next loop*)

 counter1:=1;
 counter2:=1;

(*initialise vec[] to zero*)

 FOR counter1:=1 TO no_lines DO
   begin
     FOR counter2:=1 TO no_columns DO
       begin
         vec[counter2,counter1]:=0.0
       end
   end;

(*reset counter for next loop*)

(*THIS LOOP MAY BE USED TO PROCESS THE MASTER FILE*)

 KCOUNT:=1;
 WHILE KCOUNT < no_lines + 1  DO
   begin
     readln(infile1,length,pointer);
     vec[1,KCOUNT] := length;
     KCOUNT := KCOUNT + 1
   end;

(*THIS LOOP MAY BE USED FOR THE AUXILIARY FILES*)

 KCOUNT := 1;
 WHILE KCOUNT < no_lines2 + 1 DO
   begin
     IF NOT EOF(infile2) THEN
       begin
```

```
        readln(infile2,length,pointer);
        IF pointer <> KCOUNT THEN
          begin
            WHILE pointer <> KCOUNT DO
              begin
                KCOUNT := KCOUNT + 1
              end;
            vec[2,KCOUNT] := length
          end
        ELSE
          vec[2,KCOUNT] := length;
        KCOUNT := KCOUNT + 1
      end
  end;

(*reset counters. Counter2 starts from 2 since column 1 is the master*)
counter1 := 1; rat:=0;
counter2 := 2;

(*work out absolute difference*)
FOR counter1 := 1 TO no_lines DO
  begin
    FOR counter2 := 2 TO no_columns DO
      begin
        diff := vec[1,counter1] - vec[counter2,counter1];
      (* diff := ABS(diff);  *)
        rat := (vec[counter2,counter1] / vec[1,counter1]) * 100;
writeln(outfile,counter1:4,vec[1,counter1]:10:3,vec[counter2,counter1]:10:3,diff:10:3
      end
  end

end.
```

```
Program total(infile);
(*Assesses the total difference in each of the test files*)

var
 newpoint : integer;
 mlen,tlen,diff,pdiff,count  : real;
 infile : text;


begin
{ Set up the files }
 RESET(infile,'FILE=X.KI');
 count:=0;
 WHILE NOT EOF(INFILE) DO
   begin
     READLN(infile,newpoint,mlen,tlen,diff,pdiff);
     count:=count+diff
   end;
 WRITELN('The total amount of difference is ',count:13:3)
end.
```

## Appendix 5.2 Pascal programs used for calculation of node differences

```
program verhard(infile,outfile);
(*prepares node file data for comparison from the standard ARC/INFO generate*)
(*format, into an x1,y1 x2,y2 format*)

var
 realx,realy,xdum,ydum,xm,ym : real;
 pointer : integer;
 infile,outfile : text;


begin
  reset(infile,'FILE=indat');
  rewrite(outfile,'FILE=-out');

  xdum:=0; ydum:=0;
  realx:=0; realy:=0; ydum:=0;
  while not EOF(infile) do
  begin
     xdum := 1;
     readln(infile,pointer);
     readln(infile,xm,ym);
     while xdum >= 0 do
     begin
        readln(infile,xdum,ydum);
        if xdum >= 0 then
        begin
           realx:=xdum;
           realy:=ydum
        end
     end;
     writeln(outfile,pointer :4,xm:11:5,ym:11:5,realx:11:5,realy:11:5)
  end;
  close(infile);
  close(outfile)
end.
```

```
Program findno(infile,outfile);
(*Identifies multiple occurances of particular node points on the basis*)
(*of thier identifier. Totals of their number and writes the values to a file*)

var
 histval,newpoint,count,threshold,bignumber : integer;
 infile,outfile : text;

procedure init1;
begin
 RESET(infile,'FILE=indat');
 count:=0
end;

procedure init2;
begin
 REWRITE(outfile,'FILE=out.dat');
 newpoint:=0; histval:=0; threshold:=0; bignumber:=0
end;

procedure readit;
begin
 READLN(infile,newpoint,histval)
end;

procedure testit;
begin
   if histval = threshold then
      begin
        count:=count+1;
        writeln(outfile,newpoint:5)
      end;
end;

procedure questions;
begin
 writeln('Enter in the lower threshold value');
 read(threshold);
 writeln('Enter in the upper threshold value');
 read(bignumber)
end;

begin
 init1;
 init2;
 questions;
 while threshold < bignumber do
 begin
   writeln(outfile,'The lines with ',threshold:2,' occurances are');
   while not eof(infile) do
     begin
       readit;
       testit
     end;
   writeln(outfile,'The total number of such lines is',count:5);
   threshold:=threshold+1;
   init1
 end
end.
```

```
rogram doublerid(inputf,outputf);
*matches node endpoints which have duplicate ids*)
*Takes into account all the possible ways the node points might appear as
*doubles*)

ar xs1,ys1,xe1,ye1,xs2,ys2,xe2,ye2 : real;
  pointer,pointer2,pointer3 : integer;
  inputf,outputf : text;

rocedure initf;
*initilaises the files*)

egin
 reset(inputf,'FILE=indat');
 rewrite(outputf,'FILE=-out')
nd;

rocedure init;
(*initialises all the data items to zero*)

egin
 pointer:=0; pointer2:=0; pointer3:=0;
 xs1:=0; ys1:=0; xe1:=0; ye1:=0;
 xs2:=0; ys2:=0; xe2:=0; ye2:=0
nd;

rocedure readit;
(*reads in the next item in the file*)

egin
 readln(inputf,pointer2,xs2,ys2,xe2,ye2)
nd;

rocedure reinit;
(*stores the old value in so that new value can be read and compared*)

egin
 xs1:=xs2;
 ys1:=ys2;
 xe1:=xe2;
 ye1:=ye2;
 pointer:=pointer2
nd;

rocedure writea;

egin
 writeln(outputf,pointer2:4,xs1:11:5,ys1:11:5,xe2:11:5,ye2:11:5)
nd;

rocedure writeb;
(*stores the values if the b condition holds true*)

egin
 writeln(outputf,pointer2:4,xs1:11:5,ys1:11:5,xs2:11:5,ys2:11:5)
nd;

rocedure writec;

egin
```

```
  writeln(outputf,pointer2:4,xs2:11:5,ys2:11:5,xe1:11:5,ye1:11:5)
end;

procedure writed;

begin
  writeln(outputf,pointer2:4,xe1:11:5,ye1:11:5,xe2:11:5,ye2:11:5)
end;

procedure writeit;
(*writes the value*)

begin
  writeln(outputf,pointer:4,xs1:11:5,ys1:11:5,xe1:11:5,ye1:11:5)
end;

procedure testit;
(*test the conditions for the matching of the nodes and invokes one of the *)
(*write procedures*)

begin
 if (xe1=xs2) and (ye1=ys2) then
   writea;
 if (xe1=xe2) and (ye1=ye2) then
   writeb;
 if (xe2=xs1) and (ye2=ys1) then
   writec;
 if (xs2=xs1) and (ys2=ys1) then
   writed
end;

begin
(*main part of program*)
initf;
init;
readit;
reinit;
while not EOF(inputf) do
  begin
     readit;
     if pointer2 = pointer then
        begin
           testit;
           readit
        end
     else
       begin
          writeit;
       end;
     reinit
 end;
 writeit
end.
```

```
ogram nodes(file1,file2,out);
reads in two files; one containing the master node file and the second*)
the node file to be compared. Differences are found using pythagoras*)
theorum for both the start and the end nodes of lines, as well as the total*)


r
diff,ediff,tdiff,sdiffb,ediffb,tdiffb,totdiff,x1,x2,x3,x4,y1,y2,y3,y4 : real;
oint,point2 : integer;
ile1,file2,out : text;


ocedure initf;

initialises the files*)

gin
 reset(file1,'file=gg18:m.nod');
 reset(file2,'file=indat');
 rewrite(out,'file=-oot')
d;

ocedure init;

initialises the variables*)

gin
 sdiff:=0; ediff:=0; tdiff:=0; totdiff:=0; sdiffb:=0; ediffb:=0; tdiffb:=0;
 x1:=0; x2:=0; x3:=0; x4:=0;
 y1:=0; y2:=0; y3:=0; y4:=0;
 point:=0;
 point2:=0
d;

ocedure readit;

reads in the data from both the files*)

gin
 readln(file1,point,x1,y1,x2,y2);
 readln(file2,point2,x3,y3,x4,y4)
d;

ocedure read1;
reads in the data from both the first file*)

gin
 readln(file1,point,x1,y1,x2,y2)
d;

ocedure read2;

reads in the data from both the first file*)

gin
 readln(file2,point2,x3,y3,x4,y4)
nd;

rocedure diff;
```

```
(*computes the difference between the two start and end points*)

begin
  sdiff:=sqrt(sqr(x3-x1)+sqr(y3-y1));
  sdiffb:=sqrt(sqr(x4-x1)+sqr(y4-y1));
  ediff:=sqrt(sqr(x4-x2)+sqr(y4-y2));
  ediffb:=sqrt(sqr(x3-x2)+sqr(y3-y2));
  tdiff:=ediff+sdiff;
  tdiffb:=ediffb+sdiffb;
  if tdiffb > tdiff then
    begin
      writEln(out,point2:6,sdiff:11:5,ediff:11:5,tdiff:11:5);
      totdiff:=totdiff+tdiff
    end
  else
    begin
       writeln(out,point2:6,sdiffb:11:5,ediffb:11:5,tdiffb:11:5);
       totdiff:=totdiff+tdiffb
    end
end;

begin
(*the main part of the program*);
initf;
init;
readit;
while not eof(file2) do
  begin
     if point2 > point then
        repeat
           read1;
        until point = point2;
     if point2 = point then
        begin
           diff;
           read2
        end
  end;
  writeln('The total amount of node movement in this file is',totdiff:11:5)
end.
```

```
program largnodes(infile,outfile);
(*selects node movements that are greater than a particular threshold value*)
(*snod=start node*)
(*enod=end node*)
(*tnod=total value*)
(*sum=user defined threshold*)

var
 point : integer;
 snod,enod,tnod,num : real;
 infile,outfile  : text;

procedure settup;
begin
  (*Set up the files*)
  RESET(infile,'FILE=x.dif');
  REWRITE(outfile,'FILE=out.dat')
end;

procedure init;
begin
 snod:=0; enod:=0; tnod:=0; point:=0; num:=0
end;

procedure nodtest;
begin
  IF (tnod > num) THEN
  begin
     writeln(outfile,point :6,snod :6:3,enod :6:3,tnod :6:3);
  end
end;

begin
(*main part of program*)
  settup;
  init;
  writeln('Enter in the threshold value');
  read(num);
  WRITELN(outfile,'Differences greater than ',num:4:2,' inches');
  WHILE NOT EOF(INFILE) DO
  begin
     READLN(infile,point,snod,enod,tnod);
     nodtest
  end
end.
```

Appendix **5.3** Number of lines with the same id which did not follow each other

| Coverage name | Number of lines |
|---|---|
| T026C02 | one case id=635 |
| T026C2 | one case id=635 |
| T09C02 | one case id=635 |
| T09C2W1 | one case id=454 |
| T009C2W1 | three cases id=430, id=454, id=471 |

## Appendix 6.6

COVERAGE HISTORIES

COVERAGE 50000

A suitable MNODE value for this coverage was selected using a distance parameter within the software system allowing the user to find out the lengths between two points. The largest gaps between the endpoints of lines that are supposed to meet are measured using this tool and a tolerance value is selected that encompases all of these cases. For coverage 50000, the largest gaps were 0.02015, 0.14, 0.15 and 0.07. Therefore a match node tolerance of 0.03" was selected. This is the equivilent of xmetres on the ground.

The CLEAN tolerance was selected in the same way, this time searching for lines that overlapped, rather than lines that fell short. A CLEAN dangle tolerance of 0.05" was selected and a fuzzy tolerance of 0.002" was used.

The coverage was then edited manually to remove any remaining redundant errors. Following this the coverage was CLEANed again using the the same tolerances as before. The coverage was manually edited again and the CLEAN procedure repeated once more. The coverage was then free of any graphical errors such as overshoot and undershoots.

Coverage process:
MNODE 50000 0.03
CLEAN 50000 50000-5 0.05 0.002
Manual edit 50000-5
CLEAN 50000-5 50000-5E5 0.05 0.002
Manual edit 50000-5E5
CLEAN 50000-5E5 50000c 0.05 0.002

COVERAGE 50000B

This coverage was MNODEd with a tolerance of 0.03 and CLEANed with a dangle tolerance of 0.03" and a fuzzy tolerance of 0.002. The coverage was then edited manually and the CLEAN process repeated. A further edit and CLEAN procedure was carried out before the coverage was fully free of system errors.

Coverage process:
MNODE 50000B 0.03
CLEAN 50000B 50000B-3 0.03 0.002
Manula edit 50000B-3
CLEAN 50000B-3 50000B-3E5 0.05 0.002
Manual edit 50000B-3E5
CLEAN 50000B-3E5 50000Bc 0.05 0.002

COVERAGE 25000

Coverage process:

COPY 25THOU 25THOUBK
MNODE 25THOU 0.03
CLEAN 25THOU 25THOU-5 0.05 0.002
Manually edit 25THOU-5

```
N 25THOU-5 25THOU-5E5 0.05 0.002
ME 25THOU-5E5 25THOU-MN3
ME 25THOUBK 25THOU-MN2
E 25THOU-MN2 0.02
ally edit 25THOU-5E5
N 25THOU-5E5 25000C 0.050.002

RAGE 10560

rage process:

 FENS2 FENS2MN4
E FENS2MN4 0.04
ally edit FENS2MN4 store results in FENS2MN4E
N FENS2MN4E FENS2MN4CD 0.05 0.002
N FENS2MN4E F2MN4ECD7 0.07 0.002
N FENS2MN4E F2MN4ECD9 0.09 0.002
N FENS2MN4E F2MN4ECD1 0.1 0.002
N FENS2 F2C5 0.05 0.002
 F2C5 F2C5MN4
E F2C5MN4 0.04
ally edit FENS2MN4CD
N 10560A 10560AC 0.050.002
```