DACST
INNOVATION FUND

ACTIVITY A7
PACKAGING AND ROLL-OUT

Part 1
Theoretical Aspects of GIS and Geocoding

Crime Analysis and Decision Support in the South African Police Service: Enhancing capability with the aim of preventing and solving crime

INNOVATION FUND PROJECT NUMBER: 13102

Submitted by:
P M U Schmitz, M Potgieter, P du Plessis, A K Cooper and C Schwabe

March 2000

MRC

CSIR
Your Technology Partner

Executive Summary

Funded by the Department of Arts, Culture, Science and Technology (DACST), a consortium consisting of the CSIR, the Human Sciences Research Council (HSRC) and the Medical Research Council (MRC) have been involved in the project: *Crime analysis and decision support in the South African Police Service (SAPS): Enhancing capability with the aim of preventing and solving crime.* A component of this project is Activity A7, *Packaging and Roll-out*, and as part of this activity, this report is a user manual, describing how to implement the different applications developed for SAPS.

The deliverables of this activity were developed in consultation with SAPS officers from the Crime Information Analysis Centre (CIAC) at National Headquarters and with inputs from the HSRC. Part 1 (this document) of this activity deals with the theoretical aspects concerning Geographical Information Systems (GIS) ranging from what GIS is, vector and raster models (the two different options available to represent the real world in digital form), database theory, and spatial analysis to Web-enabled GIS.

This report also covers aspects of geocoding, the building block for crime analysis since it gives the location of incidents in police station areas of jurisdiction. The last section in Part 1 gives information on the two main GIS software packages used by police forces over the world as well as digital background data necessary for a successful roll out of GIS in SAPS.

1. Introduction

Funded by the Department of Arts, Culture, Science and Technology (DACST), a consortium consisting of the CSIR, the Human Sciences Research Council (HSRC) and the Medical Research Council (MRC) have been involved in the project: *Crime analysis and decision support in the South African Police Service (SAPS): Enhancing capability with the aim of preventing and solving crime.* A component of this project is Activity A7, *Packaging and Roll-out*, and as part of this activity, this report is a user manual, describing how to implement the different applications developed for SAPS.

The deliverables of this activity were developed in consultation with SAPS officers from the Crime Information Analysis Centre (CIAC) at National Headquarters and inputs from the HSRC. Part 1 (this document) of this activity deals with the theoretical aspects concerning Geographical Information Systems (GIS) and also covers aspects of geocoding, the building block for crime analysis since it gives the location of incidents in police station areas of jurisdiction. The last section in Part 1 gives information on the two main GIS software packages used by police forces over the world as well as digital background data necessary for a successful roll out of GIS in SAPS.

Part 2 is a user manual explaining how to use GIS to map and analyse crime incidents. It also provides information on other tools and techniques developed or investigated by the consortium such as shift and roster models, the use of multivariate analysis to explain certain aspects of crime, the CSIR product AccessMap used to analyse accessibility issues and a patrolling simulation model to optimise patrolling activities in SAPS.

Part 3 and 4 deals specifically with the issues of distributing GIS over the Internet or intranet. Part 3 deals with specific software that is used for distributing GIS over the SAPS intranet. It must be stressed that this particular software, AutoDesk MapGuide was used only to demonstrate the concept to SAPS and that CSIR does not promote the software. Part 4 gives more detailed information on GIS and its use over the Internet or intranet. Both Part 3 and 4 are referred to in the section on Web-enabled GIS in this document.

2. Introduction to GIS

2.1. Mind maps and what is a GIS?

The mind map.

Mind mapping is an easy method of recording your thoughts on paper without formally organising your ideas (Cornelius and Heywood, 1995). You only start to organise your

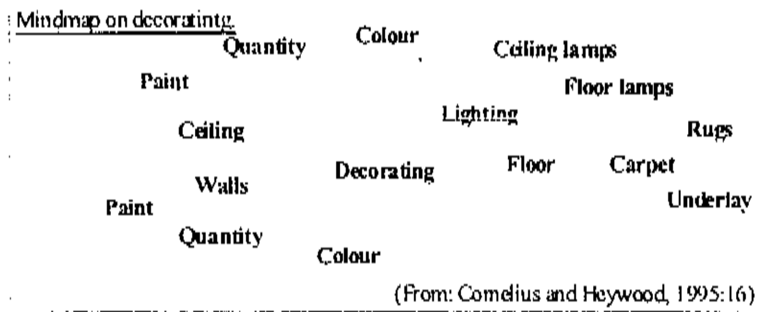


Figure 1. The Mind Map

thoughts and ideas **after** you have put them on paper. Mind maps are handy study tools, giving you an easy way to plan writing essays and reports. Mind maps are not only useful for the above but can also provide you with a tool to construct an outline to solve a particular problem with the aid of a geographical information system. A mind map on decorating a house (see Figure 1) is an example.

Another advantage of mind maps is that you start to see your problem or your solution to the problem in a spatial sense. Since we are busy with geography, we have to think spatially, start to see things in relation to each other in a spatial context, e.g. what is the shortest route from a fire station to a location where there is a fire, or how many other shopping centres are present within a 10 kilometre radius from the planned shopping centre, etc? In the next few sub-sections we are going to take a closer look at what GIS is all about.

2.2. Geographical Information Systems (GIS)

Geographic(al) Information Systems is a term that is used to describe a computer technology that has a geographic orientation. These are integrated systems that are used in a wide range of applications ranging from inventories, facilities management to complex geographic analysis, some of it in the form of simulations, "What if" games.

It is also a new discipline which attracts a lot of interest world wide, inside and outside academic institutions. GIS is used in research institutions as well as in the public and private sector. We have to take care that we do not see GIS as the answer to our problems, but as a highly sophisticated **tool** to help us find a solution or a range of solutions to our problems as well as a tool to do better management of our natural and cultural environments.

GIS has also a huge economic side to it. There is a wide variety of GIS products on the market ranging from the mediocre (revamped computer mapping systems or CAD) to very powerful GIS software systems/packages such as Arc/INFO, which are capable handling large data sets and are also able to do complex GIS operations. There are also numerous GIS consultants and vendors offering their services. GIS has a wide range of applications and brings together ideas developed in many fields such as agriculture, economics, surveying, cartography, mathematics, etc.

There is a wide range of GIS applications and one can broadly classify them into two main groups: those in the socio-economic arena such as market analysis, demographics, crime prevention, etc. and those in the natural arena such as forest management, environmental impact assessments, natural resources management, etc. The following two examples are from each of these two arenas.

The first example comes from the socio-economic arena. A city council opened two community workshops in order to enable the residents in the community to develop basic skills so that they can more easily find employment.

The person who runs these community centres needs to advertise their capabilities but she wants to ensure that the adverts reach the right people at minimum expense. Therefore to establish where the potential clients (the poor and the jobless) are staying in order to distribute the information to them.

You are the GIS person at an institution or consultancy given the job to find her an answer. What follows is a possible way of finding the information she needs. First you have to establish the target group.

They are the unemployed, the ethnic disadvantaged groups, the single parent families and the economically active women (meaning they are neither too young or too old to earn money).

Secondly the manageress wants people that live within walking distance of the workshops and she therefore uses suburb boundaries to specify the geographical region. What information do you need for the GIS to provide an answer?

Census data at enumerator level giving the following information:

- number of people who are heads of households of ethnic disadvantaged people;
- number of economically act've women; and
- number of single parents.

From the unemployment offices:

- the number of unemployed people.

The above is the data that will be needed to do the analysis in the GIS. These variables are imported into the GIS as a series of maps. One map showing the number of heads of households, the other the number of economically active women, etc. The overlay technique, meaning overlaying the different maps one on top of the other, is used to integrate the different map layers. Then the selected suburbs are imposed on the layered maps to determine the target areas.

A report is written using the relevant maps, advising the manageress about the location of potential clients (Cornelius and Heywood, 1995).
The second example is from the natural arena and concerns the blue swallow, a little bird from the Natal Midlands that is high on the red data list. The major habitat requirement for the blue swallow is rolling grasslands, which are currently being changed to commercial forests, causing a threat to the bird's survival. The aim of the exercise is to find suitable alternatives for this bird. Once these alternatives are identified one has to identify corridors between the areas to enable the bird to migrate between the habitats. To do this we need to set up a GIS to identify and model the landscape requirements for the blue swallow.

This is achieved by establishing the ecological requirements of the bird and translating it into a landscape model as well as mapping their existing territories. Once this is established the GIS is used to find other suitable areas and identify the corridors for the bird to survive.

We need the following:

- Topographical data of the Natal Midlands;
- Satellite data to provide details on forest cover and grasslands;
- Maps showing current blue swallow distribution, nesting sites and territories;
- Maps of natural landscape units.

The above data can be inputted via satellite data that is processed in a GIS using signatures from other sources, such as the Surveyor General and using a digitizer.

The following can now be generated by using a GIS:

- The satellite image reclassified to show the forested areas as well as the grasslands;
- The size of the forests and grasslands in hectares or square kilometres;
- The relief at 20m contour interval;
- The hydrology of the area;
- The blue swallow current territories;
- The land cover in these territories;

- Natural landscape units in these territories;
- A map showing the new areas by using the landscape model generated from the land cover and landscape units determined from the existing territories; and
- A map showing suitable corridors between the existing and new areas.

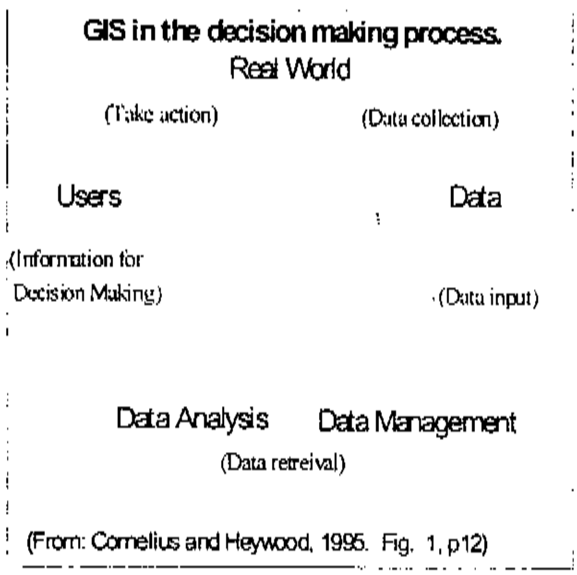


Figure 2. GIS in the decision making process

These two examples give a glimpse of what can be done with a GIS and how it helps in the decision making process. The figure shown above gives a visual understanding how GIS is used in the decision making process (Figure 2).

2.3. Definition of a GIS.

There are several definitions for a GIS but only two will be given here:

> A GIS is "a powerful set of tools for collecting, storing, retrieving at will, transforming, and displaying spatial data from the real world" (Burrough, 1986)

and

> "A geographical information system is a group of procedures that provide data input, storage and retrieval, mapping and spatial analysis for both spatial

and attribute data to support the decision-making activities of the organisation." (Grimshaw, 1994 as quoted in Cornelius and Heywood, 1995)

There are hordes of other definitions, but these give a fairly good description of what a GIS actually is.

- Most definitions of GIS have three things in common:
- Computer hardware and software;
- Geographically or spatially referenced data;
- Management and analysis procedures.

A note :

> "Geographically referenced (geo-referenced) information consists of all information that refers to the human-environment system and that can be localized in space and time."(Cooper, 1989a as referenced in Cooper, 1993)

A fourth element are the people that apply and /or study GIS. The following diagrams are from Burrough (1986). We will use them to explain all the components that are involved with GIS.

2.4. Components of a GIS.

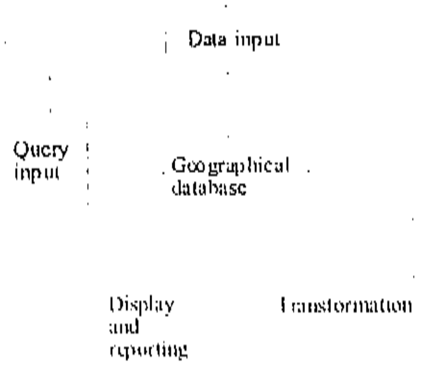


Figure 3: GIS software componentsS

GIS software: The following chief components of GIS software appear in the diagram.

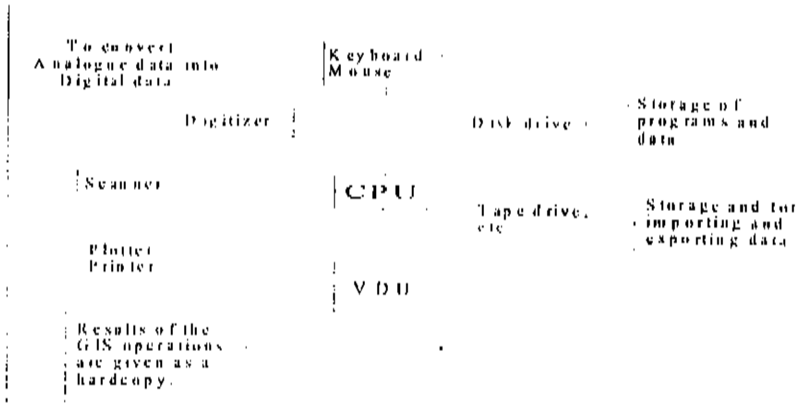


Figure 4. Hardware components

Consider these five components in more detail. Data input will be discussed first, followed by the geographical data base, data output and data transformation. Query input is done via the keyboard and /or the mouse. The query input will determine how the geographical database will be manipulated, what will be displayed, and what will be reported.

Data input:

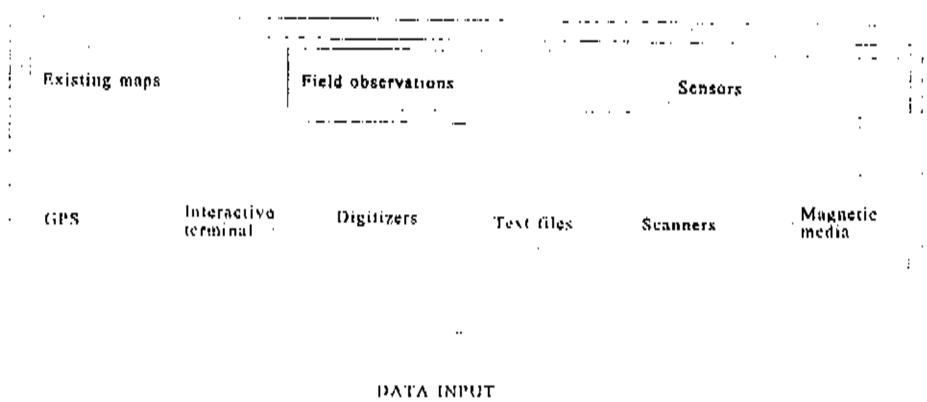


Figure 5: Data input

The geographical database:

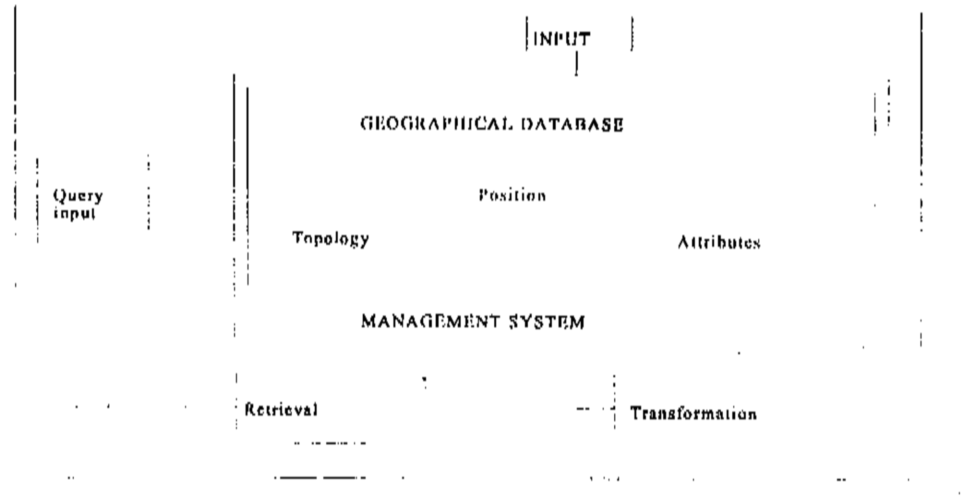


Figure 6: Geographical database

Position: It is the geographical location of the spatial elements, be it grid cells, points, lines or polygons.

Topology: This is the study of a geometric figure that is not dependent on position, having in general three characteristics namely **connectivity**, **containment** and **adjacency of features**.

Connectivity means that one node is connected to another node. This is important for network analysis.

Containment refers to polygons that contain something inside their boundaries. Eg. different shopping centres that occur within a 10 kilometre buffer around a proposed shopping mall.

Adjacency of features refers to features on the left or right hand side of a line or polygon.

Attributes: Non-spatial data explaining spatial elements, eg. what it is, the size of the element, type, etc.

Data output:

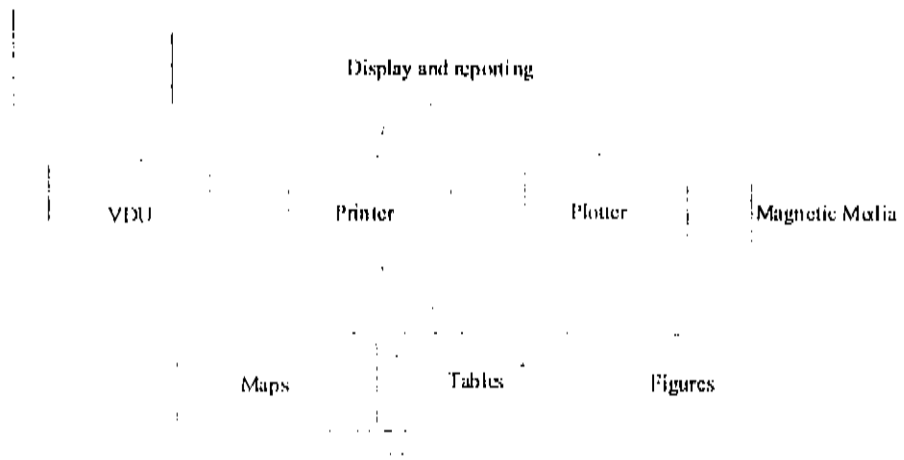


Figure 7: Data output

Data transformation

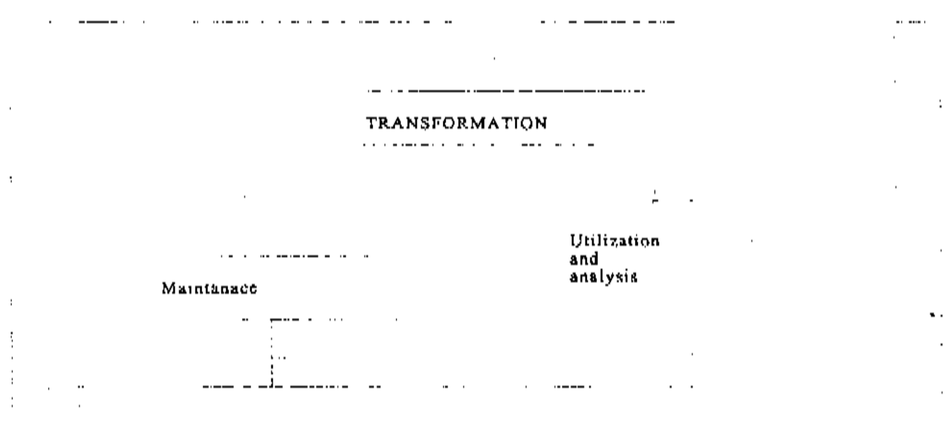


Figure 8: Data transformation

Maintenance: a) To remove errors from data;

b) Updating of existing data sets;

c) To match them with other data sets.

Analysis and There exist a huge array of analysis methods that can be utilization: applied to the data in order to achieve answers to questions asked of a GIS.

Both can be done on spatial and non-spatial data. Either seperately or on both combined.

Some examples of data transformations:

Scale changes;
fitting data to new projections;
logical retrieval of data;
calculation of perimeters and areas;
other manipulations that are application specific.

Some general questions a GIS person could expect:

1. Where is object A?
2. Where is A in relation to place B?
3. How many occurrences of type A are there within distance D of B? (Containment)
4. What is the value of function Z at position X?
5. How large is B? (Area, perimeter, count of inclusions, etc.)
6. What is the path of least cost, resistance or distance along the ground from X to Y along a pathway P?
7. What is at points X(1) and X(2), etc?
8. What is the result of intersecting various kind of spatial data?
9. What objects are next to objects having certain combinations of attributes?
10. Reclassify objects having certain combinations of attributes.
11. Using the digital database as a model of the real world, simulate the effect of process P over time T for a given scenario S (Burrough, 1986:9).

2.6. The relevance and importance of spatial data and why we use GIS?

From the above mentioned questions, and the usage of GIS to provide answers or alternatives to a problem, we realize that all of the above have one aspect in common namely geo-referenced data, to give us the answers. An example of how we use space occurs when we are invited to a function in another town and the route is explained with the aid of a map and a description of some outstanding spatial features, such as shopping centres, grog shop, robots, etc. These landmarks are seen by the host as unique and easily identifiable (not to mention that your host forgot that there are several shopping centres and also didn't mention the name of the shopping centre, where you have to turn right!).

What we actually do is continuously reference ourselves in space to find certain locations of interest.

In GIS we cannot do any applications without geo-referenced data. If we don't have any geo-referenced data we cannot determine spatial properties such as the length of a road centre line, the area or perimeter of a polygon, the volume of an object, the shape of an object, the indentation of an object (e.g. a coastline), the gradient (slopes) of a topographical feature or the aspect of the same feature and the centroid of a polygon. The same is true when we want to look at trends, distances, connections, flows or sequences. Thus we need to organize ourselves in space and Haggett, a prominent geographer, referred, in 1965, to the different categories of spatial organization as movement or flows, networks, nodes, hierarchies, surfaces and diffusion of ideas, diseases, etc in space. This we can replicate in a GIS provided we use geo-referenced data.

2.6. Why do we use GIS?

We use GIS as a tool to help make better decisions and to manage resources more effectively. With GIS we can do complex overlays with different map layers which would be extremely difficult if we would attempt it manually using tracing paper. Because we use GIS there are bound to be benefits as well as disadvantages.

Direct benefits of GIS are cost savings, staff savings, savings in storage space and to produce customized maps immediately on request, just to mention a few. Indirect advantages are the image enhancement of staff and the organization, improved quality in service provided and more improved flows of information.

Some benefits:

Savings in staff time;
savings in storage space;
fast data retrieval;
ability to keep maps up to date;
interactive map production;
ability to produce tailor made and standard maps;
quicker and cheaper map production; and
data sets can be shared, thus eliminating duplication of data.

Unfortunately there is a flip side to the coin and there are several kinds of disadvantages in using a GIS. These disadvantages are:

Initial costs to implement the system;
time consuming data input and data management;
time needed to learn to use the system;
lack of availability of commercial digital data; and
organisational changes required due to the introduction of a centralized digital database.

To conclude this section we need to understand the difference between a spatial data base and an attribute database. The following are extracts from each.

Database consisting of geo-referenced data:

```
Version 300
Charset "WindowsLatin1"
Delimiter ","
Index 1
CoordSys Earth Projection 1, 0
```

Columns 1

Pols_code Integer

[The above is a header of a file giving information that it is a spatial data set and the particular map projection it uses.]

Data

Point 24.90 -34.09

Symbol (35,0,10)

Point 29.80 -23.32

Symbol (35,0,10)

Point 26.23 -29.13

Symbol (35,0,10)

Point 26.23 -29.09

Symbol (35,0,10)

Point 27.94 -32.29 <----- longitude and latitude

Symbol (35,0,10) <------- a point symbol to show the location of the point on a map

Point 29.21 -29.84

Symbol (35,0,10)

An attribute database for the above spatial database:

STATION	AREA	South	East
PARADICE BEACH	UITENHAGE	-34.09	24.90
BANDOLLIERKOP	FAR NORTH	-23.32	29.80
BATHO	S FREESTATE	-29.13	26.23
BAYSWATER	S FREESTATE	-29.09	26.23
BEACON BAY	EAST LONDON	-32.29	27.94
BOSMANSNEK	BORDER POST UMZIMKULU	-29.84	29.21
BRIGHTON BEACH	DURBAN	-31.01	29.94
BRONVILLE	N FREESTATE	-28.84	26.75
BUFFALO FLATS	EAST LONDON	-33.02	27.09
CALCUTTA	GIYANI	-24.99	31.24
CAMBRIDGE	EAST LONDON	-32.97	27.90
CHRISSIESMEER	EAST HIGHVELD	-26.29	30.22

3. Spatial data modelling.

3.1. Maps - the primary source of data for the GIS.

Since maps are at present the primary source of data for a GIS, it should be appropriate to have a look at maps in general. Maps are scanned or digitised into machine readable form and we use this data to do our GIS operations, etc. Maps are used to represent the Earth on a flat surface to scale, to show specific characteristics of our environment, whether it is our natural or cultural environment or a combination thereof. Maps and their associated discipline, cartography, has been around for centuries and we need an understanding of the types of maps, the coordinate systems and projections they use to represent Earth.

The different types of maps depend on the use of the map. If we want to find our way from A to B we normally use a road map to guide us there. Road maps only contain the minimum of information necessary to convey it's message. Topographic maps are normally used as base maps in GIS and these maps give the third dimension, height above sea level, in the form of contour lines. Contour lines are lines of equal height. Lines that are used on maps to show lines of equal values are known as iso-lines. E.g. isohyetes are lines that show equal rainfall whether it is annual or seasonal rainfall. Isobars show us the areas of equal pressure on a weather map. Other examples of maps are dot density maps where a dot has a specific value and the number of dots in an area on the map are proportional to the data value for that region.

We use coordinate systems to locate ourselves in space or to locate a location on Earth. Coordinates "are based on measurements of displacement from a given location" (Goodchild and Kemp, 1990:26-3). There are a few different coordinate systems of which the plane and terrestrial (global) coordinate systems are the most commonly known. The plane Cartesian (x,y) coordinates are the coordinates that we use in mathematics at school and this is also the coordinate system that we use to draw graphs, etc. Terrestrial (global) coordinates are your longitudes and latitudes and is used to locate a position on the earth's surface.

We will look in more detail at the following three coordinate systems, namely Cartesian and Polar Coordinates (plane coordinate systems) and terrestrial coordinates. This information is based on the text by Goodchild and Kemp (1990:26-3, 26-8)

To determine a Cartesian coordinate we need an origin, from where we are going to determine a position, two axes that are perpendicular (90°) to each other where the horizontal axis is identified as X and the vertical one as Y. To link these two axes to geographical directions, X is east and Y is north, thus Y is anticlockwise from X.

The position is determined by measuring the displacement along both axes and the result is given as an ordered pair (X,Y) (see Figure 9).

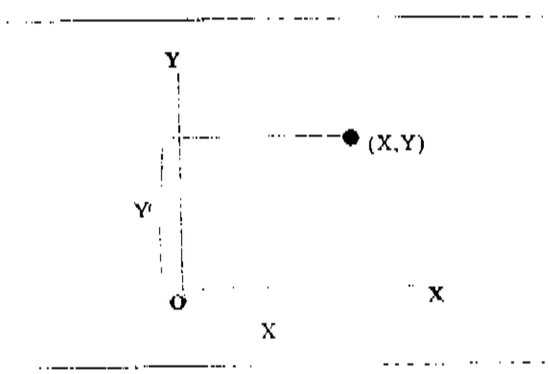


Figure 9. X and Y coordinates

We can measure distance between two points by using the Cartesian coordinates directly. There are two different methods that are used to determine distances between two points. The first method is the Euclidean (Pythagorean) distance and the second is the Manhattan Metric. The Euclidean distance is determined as follows:

$$D^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

the distance D is the square root of D^2

The Manhattan Metric is determined as follows:

$$D = |x_1 - x_2| + |y_1 - y_2|$$

The figure below illustrates both methods. The Manhattan Metric is based on three assumptions. The first is it assumes a rectilinear route that is parallel to the X and Y axes, secondly the alternate routes are of the same length and lastly it assumes that the travel is parallel to the axes. Figure 10 shows this in graphic form.

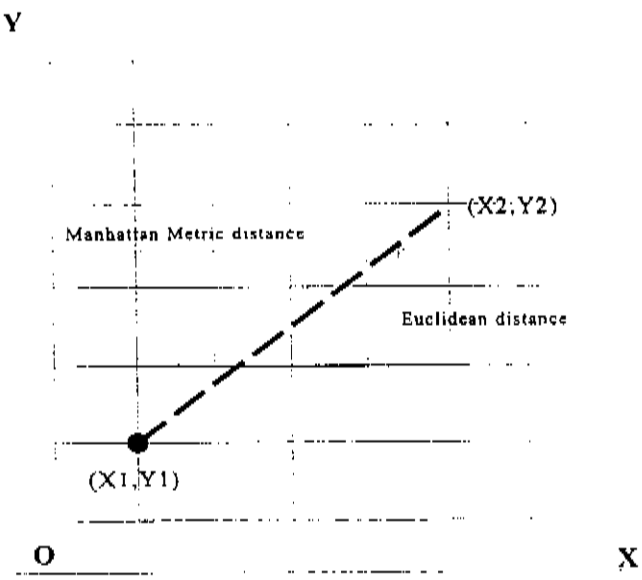


Figure 10. Euclidian and Manhattan distances

The second plane coordinate system is the Polar Coordinates (see Figure 11). Polar coordinates uses the distance r and the angle from a fixed direction Θ, normally from north (0°) turning clockwise. To convert from polar coordinates to Cartesian coordinates:

$$x = r.\sin(\Theta) \text{ and } y = r.\cos(\Theta)$$

and

$$r = \text{sqr of } (x^2 + y^2)$$

$$\Theta = \arctan(x/y)$$

(R = r in text)

Terrestrial (global) coordinates start of with a line that connects the South and the North Pole through a point on the earth□s surface. This line is known as the **meridian**, which is also a line of constant longitude. If we take the centre of the earth as a reference point, the **latitude** (φ) measures the angle between the point on the earth□s surface and the equator along the specified meridian. This is only true if we assume that the earth is a sphere and not an ellipsoid. An ellipsoid is roughly a sphere that is a bit flat at the top and the bottom. Latitudes range from -90° (the South Pole) to +90° (the North Pole). **Parallels** are lines of constant lattitude.

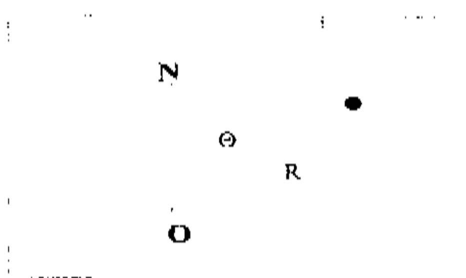


Figure 11. Polar coordinates

Longitudes (λ) □measures the angle on the equatorial plane between the meridian through the point (centre of the earth) and the central meridian (through Greenwich, England)□ (Goodchild and Kemp, 1990:26-7). Ranges from -180° (westerly) to +180° (easterly). Figure 12 gives a graphical representation of latitudes and longitudes.

Latitude and longitude are given in **degrees**, **minutes** and **seconds** where 60 seconds equals 1 minute and 60 minutes equals 1 degree (written as 23° 34' 45"). Unfortunately most GIS software does not understand this format thus we have to change it to decimal degrees (see also the spatial data set example in section 1).
The conversion is as follows:
from Degrees, Minutes and Seconds to Decimal Degrees,

Decimal Degrees = Degrees + ((Minutes + (Seconds/60))/60)

e.g. 42° 45' 30"

Decimal Degrees = 42 + ((45 + (30/60))/60)

= 42 + ((45 + 0.5)/60)

= 42 + (45.5/60)

= 42 + 0.758333

= 42.758333 (normally up to 6 decimal places)

Figure 12. Latitudes and longitudes

and from Decimal Degrees to Degrees, Minutes and Seconds

To calculate for 75.213458 decimal degrees to:

Degrees	= 75
Minutes	= (0.213458 * 60)
	= 12.807480
	= 12 (the 0.807480 is carried over to seconds)
Seconds	= (0.807480 * 60)
	= 48.4488
	= 48 (or 48.45 if you wish)

Thus: 75.213458 = 75° 12' 48"

(Source: MapInfo Corp. 1993:375)

3.2. Map projections

Since we want to represent the earth (which is a sphere of some kind) on a flat surface we need a projection to reduce the distortion caused by representing the earth on a flat surface. Thus a projection is an equation or a set of equations that contains a set of parameters. The number of parameters and the types of equations depend on the specific projection that is being used. □ A coordinate system is a collection of parameters that describe the coordinates. One of these parameters is the projection□ (MapInfo Corp, 1993:403). There are different map projections available such as the Transverse Mercator and the Gauss Conform projection. The latter is the projection that is used on the South African 1:50000 topographical maps. It is important that all the layers in a GIS are of the same projection otherwise you will get faulty results due to the fact that your positions (Lat/Long) don□t coincide with each other. When you digitize a map you have to use the same projection which was used to produce the map.

We have looked at map projections, etc. Since maps are models of the earth, lets have a closer look at what a model is and how it fits into the realm of GIS.

What is a model?

Haggett and Chorley (1967:21-22), as referenced in Heywood (1995), defined a model as follows:

> □..either a theory, a law, a hypothesis or a structured idea. It can be a role, a relation or an equation. It can be a synthesis of data. Most important, from the geographical view point, it can also include reasoning about the real world by means of translations in space (to give spatial models) or in time (to give historical models)□

3.3. Modelling the spatial dimension.

Modelling the spatial dimension addresses two dimensions of reality. The first dimension is the *spatial dimension* which gives the location of the object and the second is the *thematic dimension* which gives the characteristics of the location or the object occupying the location. If there are a number of them (in sequence) we can then add a third dimension namely the *temporal dimension.* With the third dimension we can do comparisons over time, to detect changes (Heywood, 1995). The spatial dimension can either be represented in a two dimensional framework (e.g. X,Y or latitude and longitude position) or in a three dimensional framework (the former plus a value that indicates the third dimension, e.g. height above sea level). According to Everest (1986) as quoted in Heywood (1995) □data are □facts□ represented by values, numbers, character strings or symbols which carry meaning in a certain context.□ In a spatial context the above will convey to the user information of the spatial location of the object as well as the attributes of the object at that location. Topological data is used to explain the relationships between spatial features. Topological data is important when you want to do network analysis in GIS.

3.4. Spatial data for use in the computer environment.

The data model that we use in GIS is basically a structure the computer will use to replicate the real world as a model. Cooper (1993) also mentions another world called the potential world, meaning the world as it might be or have been, whereas the real world is the world as it is or was. There are two data models that are used in GIS to represent the spatial component of the data (known as digital geo-referenced data). The first model is known as a **raster** model and the second is a **vector** model. The figure below illustrates both the models.

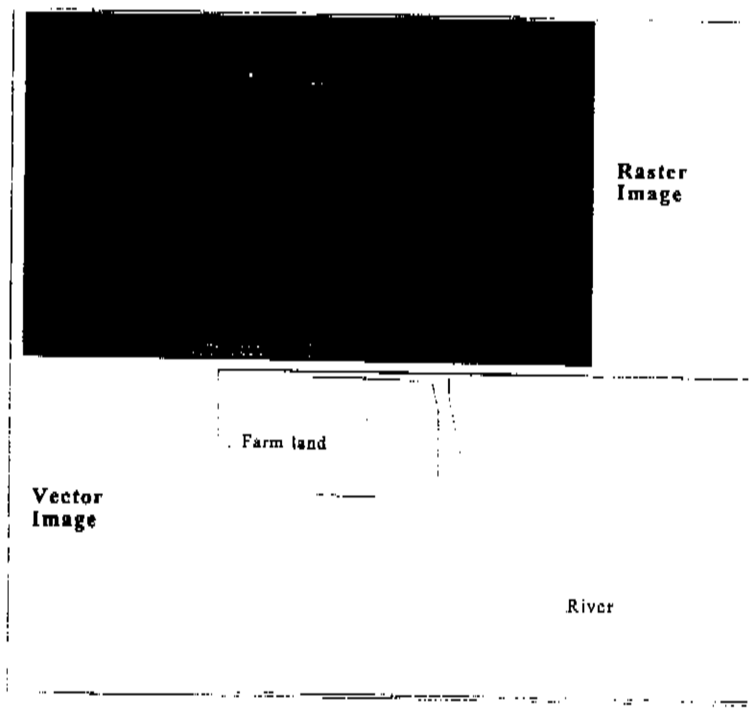


Figure 13: Raster and vector models

A data model in GIS can be defined as a general description of a specific set of entities and the relationship between these sets of entities. Peuquet (1990) states that data needs to be viewed at a number of levels. The first level is the real world, the next level is the data model that we will use, the third level is how are we going to structure the data and the last level deals with the file structure. Level 2 and 3 are abstracts from the real world and the last level deals with the machine structure itself. Schematically it will look as follows:

When taking the above into account we identify three stages that we have to do when we construct a GIS to give us an answer or answers to our problem.

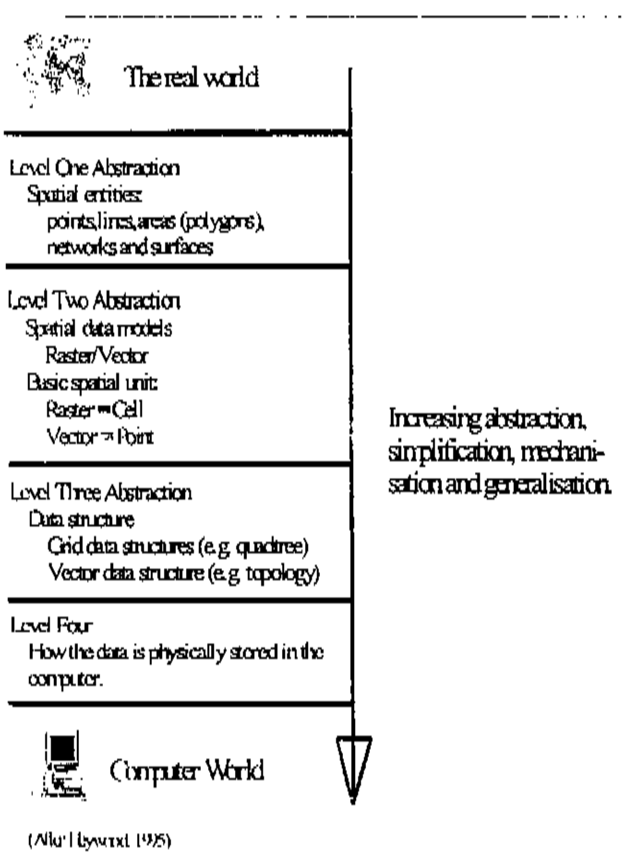


Figure 14. Stages of abstraction of real world data

Stage One:

Identify the spatial entities from the real world that are of importance to the project and decide **how** are you going to represent them. Either as a point, or a line, a polygon, a network or as a surface.

Stage Two:

Chose the data model (raster or vector) to represent those entities. The type of data model can be influenced by the available software and/or the type of applications you want to do. To determine trends might be easier in a raster model than a vector model, etc.

Stage Three:

This is the nuts and bolts stage where you instruct the computer how to recreate the entities you have identified using the spatial data model you have chosen (Heywood, 1995).

3.5. The big five of GIS - The five spatial entities.

We use five spatial entities to represent (model) the real world in GIS. These five entities are mentioned in Level 1 of abstraction. The first three spatial entities namely *point*, *line* and *polygon* (some GIS literature refer to it as area or region) are used to represent geographical phenomena or entities in a two dimensional framework. Points are used to give the location of a borehole, a shopping centre on medium scale map, crime incidents such as the hijacking of motor vehicles, etc. Lines are used to show rivers, roads, pipelines, etc. Polygons are used to show the areas of different land uses, census tracts, suburbs, police station areas of jurisdiction to name a few.

The other two spatial entities are *networks* and *surfaces*. A network can be seen as a series of interconnecting lines along which there is a flow of information. Here we see the flow of information at its broadest sense in this context. Examples of networks are telephone lines, electricity, rivers, rail networks and cable TV. You have noticed that some of the examples are also mentioned as line entities. This is true. It only becomes a network entity due to the specific problems that we want to address. If we want to use roads only as a line around which we want to create a buffer strip (e.g. 50m wide) for development purposes, we don't need it as a network entity, a line entity will be sufficient. If we want to determine the most cost effective route for a refuse collecting truck then the network entity is essential.

"A surface is a continuous entity for which at any point or location there is a measurement or a value which can be either quantitative or qualitative" (Heywood, 1995:24). Quantitative values are values such as height above sea level, temperature, mean annual rainfall, etc. An example of quantitative values is a friction surface that is generated from land use. This is done for example to determine a route that has the least impact on the environment. Since we are modelling reality there are some problems associated with the entity definition process. There are several of these but we will look at the three most common problems. These are:

The dynamic nature of the real world;

Scale;

Definition (Heywood, 1995).

We will discuss them separately.

3.6. The dynamic nature of the world.

As we all know the world changes continually, rivers change their course, forest get bigger or smaller due to exploitation, cities grow and land use change from natural to agriculture to urban over time. This dynamic world leads to two particular problems namely, how to select the correct entity type (one of the big five) that represent the feature appropriately when it is used for modelling purposes. To illustrate, do we need the suburbs of a city or do we actually need each land parcel (erven) for our specific project? The second problem is change over time, e.g. forest areas are converted into agricultural areas. This can have an impact especially if we are interested in forest species diversity and bio-prospecting. We record the locations of plants but when we want to use it a year later the area□s land use might have changed and the plants are lost.

Scale.

Durban on a world map with a scale of 1:20,000,000 will be represented as a point feature. On a 1:250,000 map we can identify Durban by its approximate outer boundaries using the area entity. If we use a 1:2000 map we will only see a part of Durban and we can represent Durban as follows:

points: e.g. street lamps and fire hydrants;
lines: e.g. road centre lines and pipelines;
areas: e.g. erven and outlines of buildings.

The art is to decide at which scale we want to run our project and then to make the correct choice of entities.

Definition.

□The definition of real world features into one of the five entity types is also plagued by the fact that many real world features simply don□t fit into the character of the models available□ (Heywood, 1995:27).

The representation of soil types in GIS is a good example to illustrate the above. In the GIS we show the different soil types as a feature with distinct boundaries. This is however not true in reality. In reality there is a gradual change from one type to another.

Fuzzy boundaries could be a problem, but on the other hand one could say that a 0,5mm thick line on a 1:50,000 map is equivalent to a line, 25m wide, on the ground. This could then account for the fuzziness of the boundary.

The decision to choose what entity type is very important. If you represent a road network simply as a series of line entities rather than a network entity you **won□t** be able to determine the shortest route between A and B. The same is true if you show a suitable building site as a point entity or as several line entities and **not** as an area entity, you will not be able to calculate the area of the site. We can represent the different entity types either as a RASTER or a VECTOR model.

When we use GIS we seldom look at a single feature, e.g. soil types, but we look at several different features simultaneously. The urban landscape can be represented as follows:

Layer 1: the street network (network entity)

Layer 2: the utilities network, such as storm drains, water mains, etc. (network entity)

Layer 3: the individual erven (area entity)

Layer 4: rivers (network entity)

Layer 5: parks and open spaces (area entity)

Layer 6: street lamps (point entity)

etc.

All these layers may be presented in one map. We use different information to solve particular problems and the most common method to handle the problems of multi faceted information by the GIS, is to divide the information into different layers as in above. This is known as the **layered approach**. MapInfo calls it layers and ArcView calls it themes that form a view (map). Each of these layers are independently stored as a file in the GIS, either as a raster or a vector or in both models. These layers can be selected, or deselected, in a map (view) in the GIS according to the need.

3.7. Raster and Vector Data Models

We had a brief introduction to raster and vector data models earlier in this section. In this subsection we are going to look, in more detail, at raster and vector structures, based mainly on the work of Aronoff (1989) Chapter Six. According to Aronoff (1989:166) there are several advantages and disadvantages for each of the models.

The advantages for the raster models are:

* it has a simple data structure;
* overlaying of layers are easy and also more efficient and time saving;
* high spatial variability is easily represented; and
* and we need raster models when we work with digital images such satellite imagery.

The disadvantages are:

* raster data structures are less compact (taking a lot of disk space), but there are several data compaction techniques, such run-length encoding, chain coding, etc, available to compact raster data;
* topological relationships are more difficult to represent (i.e. very difficult to do network analysis in the raster environment); and
* the graphical representation of raster maps are not as aesthetically pleasing as vector maps which make use of lines. Raster maps tend to be very blocky, if you using a finer resolution it will overcome your problem but will result in very large data sets.

The advantages of vector models are:

* the data structure is much more compact than that of a raster model;
* topological relationships can be established easily which enables analysis procedures dependent on topological information such as network analysis; and
* vector models are close representations of hand drawn maps.

The disadvantages are:

* the data structure is much more complex than those of the raster model since these structures also have to make provision for topological information;
* overlay operations are more difficult since you have merge polygons as well as dissolve lines which relies on complex algorithms (see figure on p29);

* high spatial variability is very difficult to represent in the vector format; and
* the manipulation and enhancement of digital images using methods such as filtering, can not be effectively done in the vector environment.

Overlaying of polygons in the vector environment

We mentioned, when discussing the disadvantages of raster models, that there are several methods to compact raster data structures. Although there are several methods we will only consider the two methods as discussed by Aronoff (1989). The first method is run-length encoding and the second method is quadtrees.

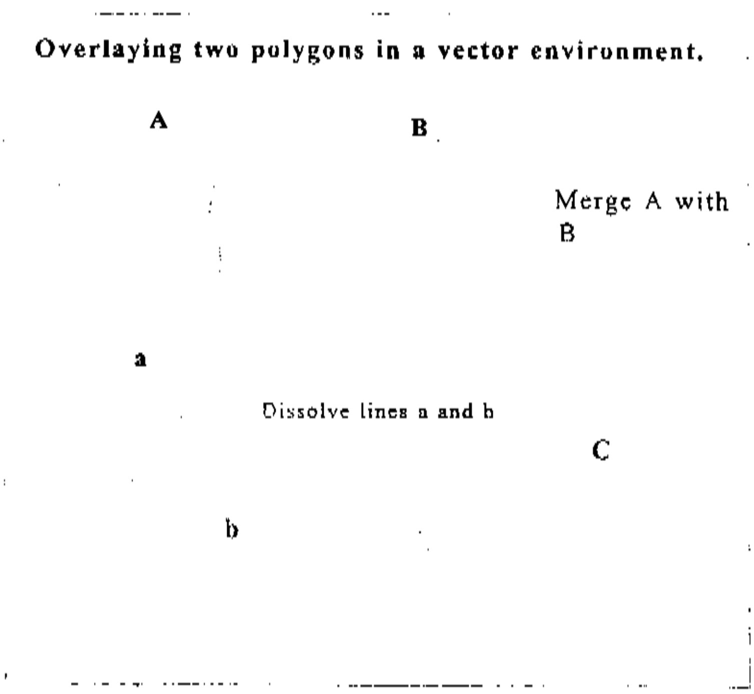


Figure 15. Overlaying of polygons in Vector-GIS

3.8. Run-length encoding.

"In run-length encoding, adjacent cells along a row that have the same value are treated as a group termed a run" (Aronoff, 1989:168). There are two major forms of run-length encoding which make use of the position of the rows. The first one is the standard run-length encoding which uses the value, the length (how many cells does this value occupy) and the row number. As a convention the rows and columns in a raster model starts in the top left hand corner and each begin with 0. If you have 10 rows and 10 columns then the rows and the columns will each be numbered from 0 to 9. The other compression technique is the value point encoding. Here each cell is given a number starting at the top left hand corner (starting with 0) proceeding from left to right for each row from top to bottom. The end position of each value in a run is given and stored in the point column (see figure below).

Run-length encoding of raster data

ROWS	COLUMNS 01234567
0	AAAAAAAA
1	AAABBBBB
2	BBBBBBBB
3	BBBBBCCC

Full raster encoding (32 values)

VALUE	LENGTH	ROW
A	8	0
A	3	1
B	5	1
B	8	2
B	5	3
C	3	3

Standard run-length encoding (18 values)

VALUE	POINT
A	10
B	28
C	31

Value point encoding (6 values)

Source: Aronoff, 1989:169 Figure 6.12

Figure 16. Coding of raster data

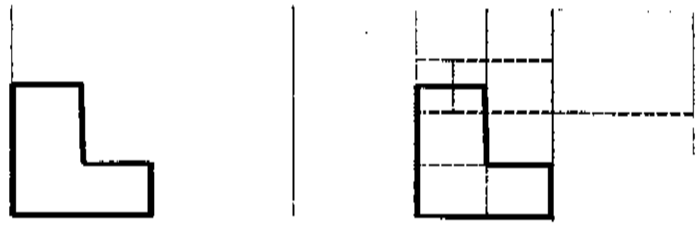
Run-length encoding is a useful compaction tool provided there is little variation in the number of different values. If a raster image is too complex, it might be better to use a different compaction method. Such a method is the quadtree method.

Quadtrees The quadtree is developed to compact raster data even more efficiently than the methods discussed above. It makes use of grid cells of variable sizes which are represented on various levels. The higher the level the smaller the grid cell and this higher resolution is only used where it is needed. The advantage of the quadtree method is that you can use limited topology to describe relationships between entities. The use of higher resolutions are only used in the vicinity of lines, points and polygon boundaries. Large areas of the same value can be accurately described by a much lower resolution. We will illustrate the above with a diagram on page 34.

The above diagram showed us how to code a quadtree. The bintree gives us a schematic representation of the quadtree and it starts off with the ROOT which represents the □MAP□. The first level (also known as a LEAF) represents the first division of the map into quadrants (hence the name quadtree). If the entity is not present in a quadrant (cell) or completely fills the cell then it terminates, i.e. there is no need for further subdivision. If the cell is only partially occupied by the entity, further subdivision is needed. Then we move to levels 2 and 3 (also known as NODES) until the whole entity is covered. The table gives us the attributes of each cell. In the example we only stated if the entity is present or not, but we could easily replace these with land use classes such as industrial, residential, etc. This type of a quadtree is known as a linear quadtree, meaning 031 is contained within 03 and 03 is contained within 3 (Aronoff, 1989).

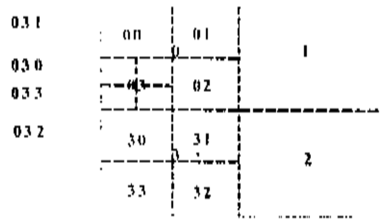
With a quadtree it is easy to determine the neighbouring cells by looking at the numbers, the neighbouring cells of 031 are 030, 032 and 033, thus clustering cells close together according to numbers as they are on the map. This also makes it easy for a computer to retrieve the information correctly. One of the big disadvantages of quadtrees is the time needed to set it up and to code it, as well as to modify the codes when necessary.

How to code a quadtree

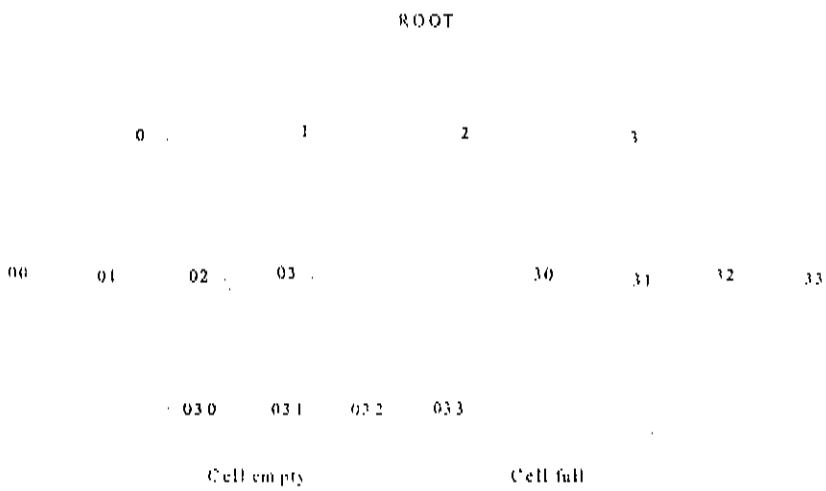


The 'Map'

Graphic representation of quadtree cells



The coding of the different cells indicating the different levels.



Bintree structure for quadtree representation

LEVEL 1	2	3	ENTITY
0			PRESENT
	00		ABSENT
	01		ABSENT
	02		ABSENT
	03		PRESENT
		030	ABSENT
		031	ABSENT
		032	PRESENT
		033	PRESENT
1			ABSENT
2			ABSENT
3			PRESENT
	30		PRESENT
	31		ABSENT
	32		PRESENT
	33		PRESENT

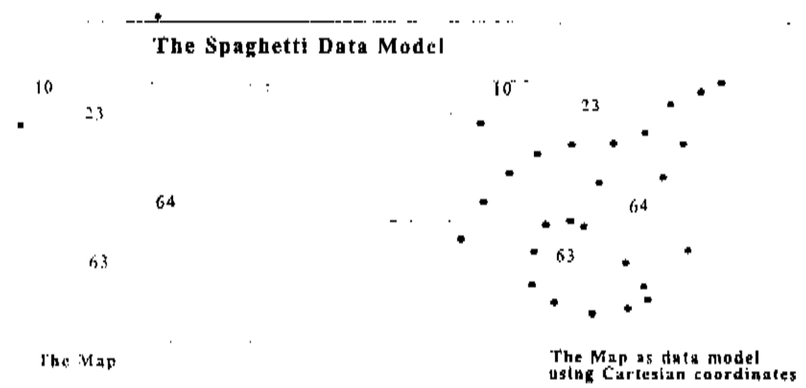
Tabular representation of the bintree

(From Heywood, 1995:60)

In this section we looked briefly at the raster model. In the next section we will have a closer look at the vector model.

3.9. The vector model

The vector model is used to represent the locations of entities in space using a set of coordinate pairs which are mathematically exact. It makes use of a Cartesian coordinate system to represent the locations on the earthís surface on a two dimensional space. The basic three entities used in a vector model are points, lines and polygons. A point is represented with a single coordinate pair whereas lines are represented by several coordinate pairs. The order of the coordinate pairs give an indication of the direction of the line (this is important for building topology). A polygon is actually a special form of a line feature that is closed, meaning the first and the last coordinate pairs are exactly the same. In GIS the positions of features are stored in standard geographic coordinate systems such as the Universal Transverse Mercator (UTM) coordinate system, State Plane or Latitude and



Feature	Number	Location
Point	10	X Y (Single point)
Line	23	X1 Y1. X2 Y2. ...... Xn Yn , (String)(Closed Loop)
	63	
Polygon		
	64	

Source: Aronoff (1989:174) Figure 6.16

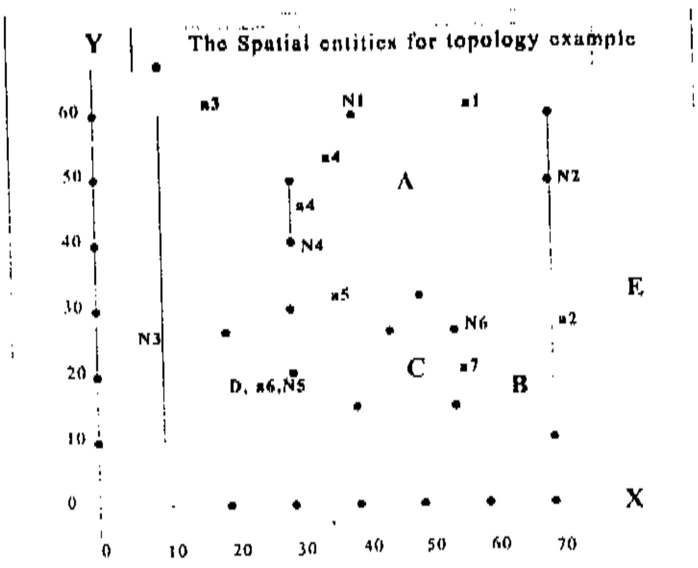
Figure 17: Spaghetti Data Model

Longitude.When working with line or polygon features the more complex the features are (irregular in shape) the more coordinate pairs are needed to represent them. There are two specific vector data models, the first is the **spaghetti data model** and the second is the **topological data model**. The spaghetti data model is a model where there is no topology involved. Points, lines and polygons are stand alone entities with no relationships between one another. This data model records only the spatial features, i.e. the coordinates that are involved to represent these features as entities. According to Aronoff (1989) this model is sufficient to plot maps. An example of a spaghetti data model is given below.

3.10. Topological data model

The topological data model has all the characteristics of the spaghetti data model plus the spatial relationships amongst the spatial entities. To illustrate topology we will use the work of Aronoff (1989: 174 - 177). Aronoff (1989) uses an Arc-Node data model to illustrate topology. In this model the basic entity is an arc (another word for a line) which consists of a series of interconnected points. The first and last points act as nodes.

Nodes are intersection points where two or more arcs meet, although a node may occur at an end of an arc that represents a dead-end street. Isolated nodes (there are no arcs that meet each other there) represent points. Arcs that are connected to each other at nodes and follow a certain directions are known as chains. A polygon is formed when a chain of arcs form a closed loop and these arcs form then the boundary of the polygon. In the figure below, that will be used as an example, we can record topology for each of the three different entities. Each will be represented in a data table and the fourth table will represent the coordinate data for each of the arcs. In a GIS points, lines and polygons are stored in separate layers, but for this demonstration we will use only a single layer.



Source: Aronoff (1989:175) Fig. 6.17

Polygon Topology	
Polygon	Arcs
A	a1,a5,a3
B	a2,a5,0,a6,0,a7
C	a7
D	a6
E	area outside map coverage

Node Topology	
Node	Arcs
N1	a1,a3,a4
N2	a1,a2,a5
N3	a2,a3,a5
N4	a4

N5	a6
N6	a7

Arc Topology				
Arc	Start Node	End Node	Left Polygon	Right Polygon
a1	N1	N2	E	A
a2	N2	N3	E	B
a3	N3	N1	E	A
a4	N4	N1	A	A
a5	N3	N2	A	B
a6	N5	N5	B	B
a7	N6	N6	B	C

Arc Coordinate Data			
Arc	Start X,Y	Intermediate X,Y	End X,Y
a1	40,60	70,60	70,50
a2	70,50	70,10; 10,10	10,25
a3	10,25	10,60	40,60
a4	40,60	30,50	30,40
a5	10,25	20,27; 30,30; 50,32	70,50
a6	30,20		30,20
a7	55,27	55,15; 40,15; 45,27	55,27

When we work with arcs in topology, the arcs, that make the boundary of a polygon, move as a rule in a clockwise direction. Polygons may have islands in them (polygon B has an island called polygon C) and points can by seen as polygons without an area and they also have an arc (a6 in our example). E denotes the area outside our interest, it may contain other arcs, points and polygons. With the aid of topology we can do spatial analyses such as contiguity (spatial relation of adjacency) and connectivity (arcs are connected to each other).

In this section we had a brief overview of spatial data modelling in GIS. In the next section we look at data and also have an introduction to database theory.

4. Data and an introduction to database theory

In this section we look at data in general and where do we get our data for the GIS and data accuracy. This is followed by a discussion on generalization of features in a GIS. Generalization is one of the serious problems a GIS user will encounter. Generalization in the realm of GIS is the change of geographical features due to a change in map scale. An example is a coastline, the coast line at a 1:2,000 map will show all the little nooks and crannies, but this will be lost if the same coastline is represented on a 1:50,000 map. The last section deals with the introduction to database theory and some thoughts on object oriented GIS.

4.1. Data and data acquisition

We use two different data sets in GIS:

1) Spatial data which is also termed positional data, is used to define where the cartographic or real world features occur in geographical space; and
2) attribute data which tells us what this feature is all about.

We use different kind of tools to input data into a GIS (see section 1) and is sometimes a major project in GIS. Data acquisition can, due to the unavailability of digital data, take up to 80 percent or more of the project costs and time. If you have to input your own data via digitizers, scanners, etc. it can be very labour intensive, tedious and error prone. We also have to take care that the acquisition of data does not become the project itself, so that we lose sight of the analysis part of the project. Since time costs money we use different methods to economise, but this will bring its own problems which we have to address. In short data input is as follows:

1) entering (e.g. digitizing) of spatial data;
2) entering of non-spatial (topological) data and associated attributes; and
3) linking the spatial data with the non-spatial data.

All data entries through all stages of the project must be verified and checked for accuracy.

Data sources in GIS can be classified into two broad classes namely, primary data and secondary data. Those can be either digital or non-digital.

Primary data

Point locations that represent data from recording instruments, such as rain gauges, thermometers, etc., are either digitized from a topographical map or obtained directly from a global positioning system (GPS), and are termed primary data sources. Another example of primary data sources are the point locations of herbal plants determined with the aid of a GPS. Primary data is data that has not been interpreted by someone else and then represented in a map form. Examples of non-digital primary data are field maps, hand recorded data (readings from a rain gauge, etc), analogue logs such as charts from autographic temperature recorders, borehole logs, etc. Digital examples are field digitized maps, geophysical data (seismic recordings) and remotely sensed images.

Secondary data

Secondary data sources are established as soon as someone interprets, edits and processes primary data. Examples of non-digital secondary data are maps and tables. These have to be converted into digital formats. Digital databases can also be secondary data. When we deal with secondary data an "audit trail" (the process description, checks and balances for accurateness of data) is necessary to ensure users that the data used is accurate. This data about data is sometimes referred to as **metadata.** If you are interested to read more about metadata read Cassettari (1993), Chapter Two. In the next sub-section we look at data quality and the accuracy of spatial databases.

4.2. Data quality and accuracy of spatial data

When the data is in the computer we tend to lose sight of errors and we assume that the database is accurate. There are errors in any GIS project and the different errors are discussed below:

1) errors that are obtained from the original source due to errors made during the development of the original source. These types of errors are known as source errors and one of the most common examples of source errors are errors caused due to distortions at the edges of maps;

2) errors made during data capture and storage (processing errors);

3) errors made when data is extracted from the computer or when data is converted from one database system or GIS to another; and

4)errors that are made during GIS operations, such as the overlaying of layers.

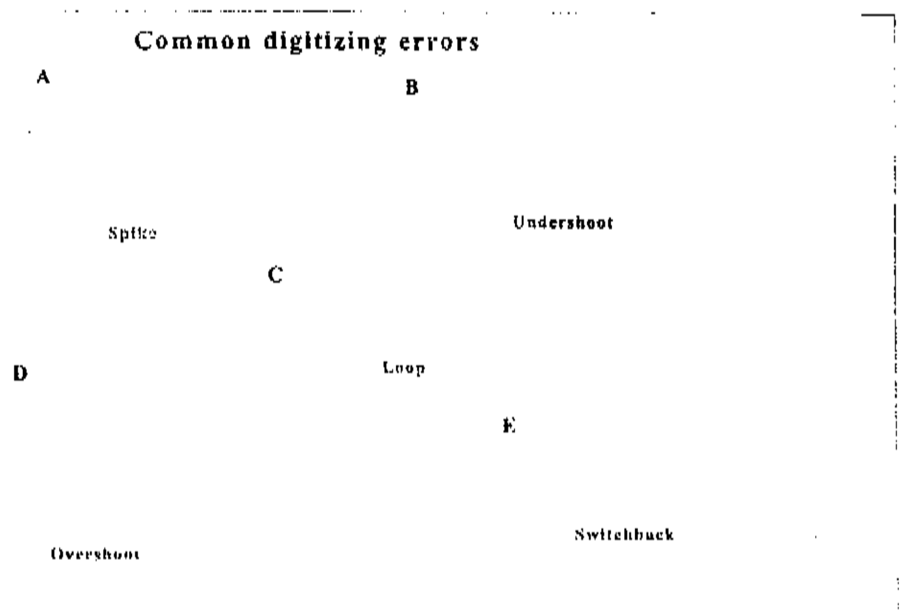


Figure 18. Common digitising errors

Some common errors that we may encounter are boundaries, e.g. lakes fluctuate and the border of the lake might have been established during a dry period and not during a wet period and may lead to developmental problems. Classification errors are another example of errors. These occur when we deal with complex classifications of features. On the other hand a too simplistic classification can also lead to further problems. Other errors are data capture errors during the digitizing process such as over-shoots, under-shoots, switchbacks, loops and spikes. These errors may occur due to user fatigue which causes involuntary muscle spasms. The figure below gives a visual example of these errors. Other errors are caused by psychomotor problems in line following (NCGIA, 1990).

The following paragraphs deal with the accuracy of spatial databases and is based on the work discussed in Unit 45 of the NCGIA Core Curriculum (NCGIA, 1990).
Before we look into the accuracy of spatial databases we have to understand what is meant by accuracy and precision in the context of GIS. Accuracy is... □defined as the closeness of results, computations or estimates to true values (or what is accepted to be true)□ and precision is... □defined as the number of decimal places or significant digits in a measurement□ (NCGIA, 1990:45-3). The latter does not mean that the data or measurement is accurate.

Standards are set to ensure data quality that provides a common language, identifies good practice and saves time and cost once developed and accepted (AGI, 1989 as referenced in Cooper, 1993). Standards in GIS in terms of data quality looks at the following components, namely positional accuracy, attribute accuracy, logical consistency, completeness and lineage (NCGIA, 1990) and currency (Cooper, 1993).

With positional accuracy we look at the closeness of locational information (e.g. latitude and longitude, X,Y coordinates, etc.) to the true position on the earth□s surface. Attribute accuracy is linked to how closely it describes the actual values. If the attribute describes surfaces (continuous attributes) the accuracy is expressed as a measurement error, e.g. elevation accurate to 1m. For categorical attributes we look if the categories are appropriate and are they sufficiently detailed and defined. Logical consistency looks at the internal consistency of the data structure and is important when we deal with topological data. Completeness refers to the inclusion of all possible objects into the database. Lineage refers to the recording of data sources and how the database was created (metadata) (NCGIA, 1990). Currency refers to when was the data captured, as the world around us is a dynamic world and the positional and attribute accuracy might change over time (Cooper, 1993). An example is that a road is downgraded from a national road (N3) to a route (R103) as happened in Natal due to the construction of a toll road.

We also use standards when we want to exchange geo-referenced data. The South African National Exchange Standard (NES) is such a standard and is used to ensure that information on data quality (as discussed above) are exchanged together with the digital geo-referenced data. The data set in the format of an exchange standard is not a database but only a set of data extracted from one database to be incorporated into another database (Cooper, 1993). For more information on NES see Cooper (1993) □ Standards for Exchanging Digital Geo-referenced Information□.

Generalization is one of the serious problems that is encountered in GIS and it will be discussed in the next sub-section.

4.3. Generalization

□Generalization is an inherent characteristic of all geographical data□ (João, Herbert, Rhind, Openshaw and Raper, 1993:64), meaning that when we change scale we lose distinguishable shapes due to the lack of physical space on a map. Generalizations are done by experienced cartographers and with the explosion in the use of GIS and the lack of experienced cartographers, some researchers developed mathematical algorithms to automate some of the generalization operations.

Why do we use automated generalization? Firstly it is to create a level of detail appropriate for the scale that we use in our GIS application, secondly to permit analysis of data at different levels of resolutions and thirdly to minimize storage and input/output operations (João, *et al*, 1993). The result of the generalization process is the generalization effect. Generalization effects inherent in the data come from two main sources, (a) digitizing maps that have been generalized during the map making process and (b) the use of generalization algorithms in the GIS itself.

Generalization effects manifests itself e.g. in the change of the length and angularity of the feature, the elimination or displacement of features. These will have an effect on the GIS analysis that we want to do. Other examples of generalization effects are the exaggeration of features (cartographic licence), such as road widths in a Map Studio road map. The effect of generalization on GIS analysis is well illustrated in João, *et al*, 1993:66-69 using the example of finding a suitable area that can be developed into a park using three different map scales. The table below explains the effect of generalization on the size (m^2) of the potential sites.

Area of the zones (m^2) that resulted from an overlay to determine the potential location of a new park for the region of Canterbury and East Kent, using three different map scales

Zone	**1:50 000**	**1:250 000**	**1:625 000**
1	712 963	529 378	1 526 598
2	1 379 760	1 491 389	1 635 338
Total	2 092 723	2 020 767	3 161 936

Source: João, *et al*, 1993:68. Table 1

Thus we can see that the right scale and the amount of generalization is crucial to do effective GIS analysis. There are different techniques on how to implement effective generalizations, but it falls outside the scope of this guide. However the following may give you a good idea how features change when we generalize and might give you a guideline to which scales are suitable for specific GIS analysis you might be doing. Generalization occurs when we change scale. However there are moments where features change abruptly and change form before further generalization can take place. These abrupt changes are "catastrophic change" (Fisher, 1991).

To illustrate the above we will take two types of generalization into account. The first one is geometrical generalization where features are simplified, enlarged and displaced.
The second generalization is conceptual generalization which influences the selection, classification, typification and symbolization of features. We will only use changing map scales.

Up to a map scale of 1:10 000 the world can be mapped without much generalization of features.

At 1:20 000 a threshold for abrupt change occurs. At this level street widths are being exaggerated, buildings are simplified, combined and displaced. Land parcels are regrouped and classified according to land use.

Here the "first catastrophe" occurs where we change from isomorphic to generalized mapping.

From 1:20 000 up to 1:250 000 further generalizations occur, e.g. rivers and contours are further simplified (losing angularity), etc.

At 1:500 000 the "second catastrophe" occurs. Geometric shapes of spatial objects vanish, e.g. towns are now replaced by a symbol. Here the size and/or shape of the symbol indicated the size of a town or its political significance. Airports are now shown with an aeroplane symbol showing its general location. Displacement of certain features, such as the airport example, is very common at this level. Here we changed from a realistic to a highly symbolized representation.

Up to now we looked at certain aspects of data and the effect of generalization on features. In the next sub-section we look at database theory as such, to give us some insight how GIS deals with data. We will base our discussion on the NCGIA Core Curriculum Units 43 and 44 (NCGIA, 1990).

Introduction to database theory

In the early days of GIS they didn't make use of existing database systems. Today most GIS's are build around existing database systems. According to NCGIA (1990) there are two approaches, the one is to use an existing database management system (DBMS) as is and the GIS must fit the assumptions build into the DBMS and the second is make partial use of DBMS and the rest of the data (mostly spatial data) is handled by the GIS itself. When working in the Windows environment one can even use the open database connectivity (ODBC) to link an existing DBMS such as Oracle or MSAccess with a GIS.

4.4. Definition of a database

There are two definitions that we can look at:

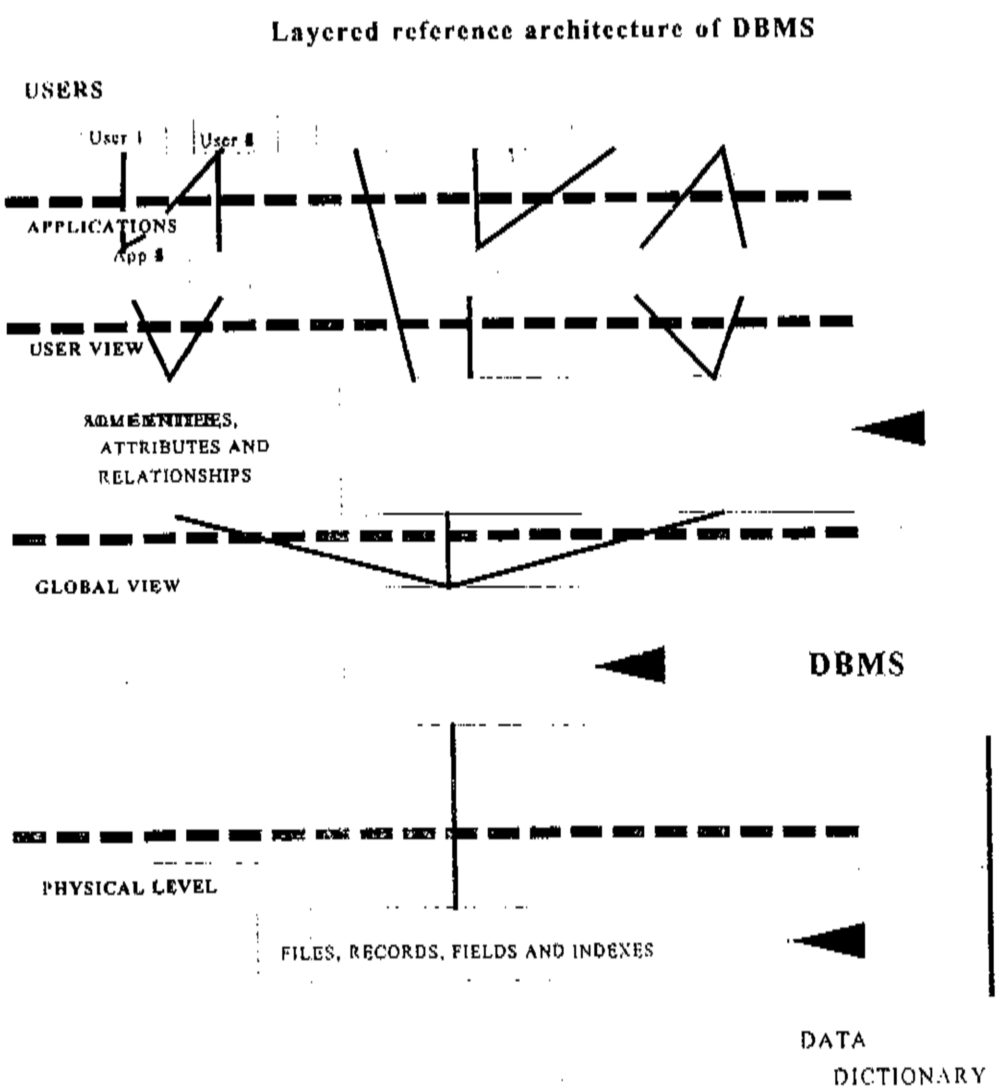
> "A DBMS is a software system with capabilities to define data and their attributes, establish relationships among data items, manipulate and manage the data".
>
> (Stamper and Price, 1990:56 as quoted in Reeves, 1996)

or

> "A DBMS is a pool of shared facilities used to access and maintain a database. A DBMS acts as an interface between end users, application programs and the database. It allocates storage, provides security and handles all the traditional demands of file processing".
>
> (Benyon-Davies, 1991:8 as quoted in Reeves, 1996)

4.5. Why do we use DBMS?

It reduces data redundancy, meaning data is shared between different users and is of the same standard, it maintains data integrity and quality, metadata about the database is saved with the data. When using a DBMS certain models, rules and standards has to complied with, thus ensuring that inconsistencies are avoided. In



From Reeves, 1996: [illegible], Fig.5

Figure 19. Layered reference architecture of a DBMS

the long run it reduces cost in software although it might be expensive to install and to maintain in the beginning. Lastly it has security restrictions which will only allow legitimate users to access certain parts of the database (NCGIA, 1990).
DBMS make use of a layered structure to achieve the fundamental objective of data independence, meaning the data as such is not affected by what the users do with it. Figure 19 shows the layered structure of a DBMS.

The Global View in a DBMS contains the conceptual design of the database expressed in the terms of the chosen DBMS software. This is achieved by using the data definition software. In the Global View all the entities, attributes and relationships of the data are present.

At the Physical Level, the physical structures which are used to implement the logical structures are established. Here the files, indexes, lists and algorithms exist. At this level only the database professionals would need to be aware of the structure.

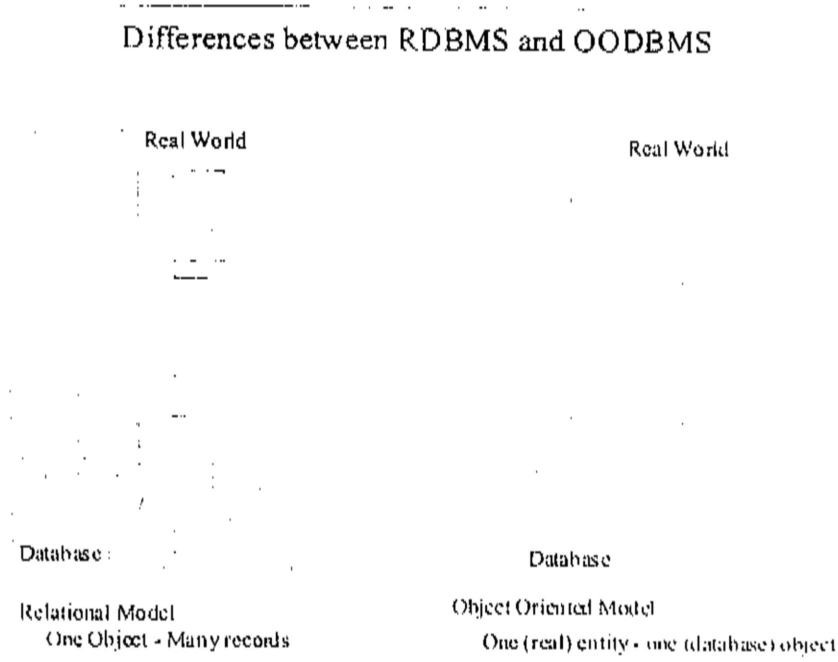
Since there are different users with different needs, the User View allows the user to use only that part of the database which is of use to him or her. The user (User 1 to 6) can either directly access the User View or can access it via specific written application programs (Application Level). The three levels (User View, Global View and Physical Level) make the database management system and each DBMS has its own data dictionary (metadata), which contains the definition of the variables it uses in the different application programs. It stores the data dictionary in a central place thus forming part of the database. This allows for standardization of definitions thus eliminating duplications.

There are four different databases. The first type or model is the hierarchical data model which is a tree structure similar to a bintree (root and leaf setup). A disadvantage of the hierarchical data model is that you have to duplicate it to represent all the different ownerships of the data.

The second model is the network data-model. Similar to the hierarchical data-model the only difference is that the leafs can belong to more than one root to overcome the problem of multiple ownerships. GIS does not make use of these two data-models since they are difficult to handle and to understand.

The third data-model is the relational data-model, which developed from the two previously mentioned data-models. The relational model is easy to understand and the user does not need any knowledge of the realities of data storage, which is a pre-requisite for the previous two data-models. To the user the data appears to be held in the conventional 2D tables and the links between these tables are established by the sharing of identical columns in the different tables. This is the data-model used by GIS

since it makes use of the relations between its spatial and attribute data. The shared



From: Reeves, 1996:235 Fig. 28

Figure 20. Differences between RDBMS and OODBMS

column is the unique ID to which the spatial as well as the attribute data belongs.

The last data-model is the object-oriented models. The object-oriented model is designed to take into account those entities of the real world that cannot be easily represented by simple tables. The difference between an object-oriented database model and the relational model is as follows: If we want to represent an entity (object) of the real world in a relational data model we need many records (see Figure 20),

(RDBMS - Relational database management system

OODBMS - Object oriented database management system)

the same entity will be represented as one (database) object (see figure above) (Reeves, 1996).

In the next subsection we discuss the relational model in more detail since this is the basic model used in GIS to handle data. We will base our discussion on the work by NCGIA (1990:43-8 to 43-10).

4.6. The relational model

The relational model allows for flexible links between records and each record has a set of attributes. For each attribute you can define the range of possible values, e.g. for DAY the domain will range from 1 to 31, meaning if you put in 0 or 32 the database won't accept your input. The records of each type form a table of records known as a **relation**, where each row that consists of a record is known as a **tuple** and each column is known as an **attribute**. "The **degree** of a relation is the number of attributes in the table" (NCGIA, 1990:43-9). If there is only one attribute in a relation then it is known as an **unary** relation, e.g. COURSES(SUBJECT) where COURSES is the relation and SUBJECT is the attribute. A **binary** relation consists of two attributes, e.g. PERSONS(NAME,ADDRESS) and relation with **n** amount of attributes is as an **unary** relation, e.g.

PROPERTY(ERFNUMBER,NAME,ADDRESS,.....,VALUE).

To link relations with each other we make use of **keys**. A key is a subset of attributes which has to fulfil the following criteria:

It MUST have a unique identifier, meaning it cannot be repeated in the same column; and

it must be non-redundant, meaning if we delete it we will lose all the other information linked to it as well.

There are two keys, the first key is a primary key which acts as the unique ID for the relation and the second key is a foreign key which links the relation with another relation. The primary key in a relation is the same as the foreign key in a different realtion.

Normalization is an action that you do to find the simplest structure for a given set of data. It deals with the dependence between attributes and avoids loss of general information when records are inserted or deleted (NCGIA, 1990).

The next two tables will give us an example of a non-normalized and a normalized relation.

In HOMES1 (the primary key is underlined) this relation is not normalized since PRICE is uniquely determined by STYLE and can lead to problems when we want to delete or insert records. If we delete the last entry of ranch we will lose the relationship between ranch and its price value of $50,000. If we enter a new STYLE e.g. triplex (costing $75,000) we have to enter a whole new relationship, when the first triplex record occurs.

Non-normalized relation

HOMES1		
BUILDER	STYLE	PRICE
Cadillac	Duplex	65,000
Delzoto	Duplex	65,000

Howlett	Bungalow	45,000
Joint	Ranch	50,000
Metro	Bungalow	45,000
Monza	Duplex	65,000
Terex	Ranch	50,000
Wimpey	Ranch	50,000

The next table gives us a normalized relation. Here we split the non-normalized relation into two normalized relations namely COST and HOMES2. HOMES2 is created to establish the STYLE for each BUILDER and COST is created to establish the PRICE for each STYLE.

HOMES2		COST	
BUILDER	STYLE	STYLE	PRICE
Cadillac	Duplex	Bungalow	45,000
Delzoto	Duplex	Duplex	65,000
Howlett	Bungalow	Ranch	50,000
Joint	Ranch		
Metro	Bungalow		
Monza	Duplex		
Terex	Ranch		
Wimpey	Ranch		

STYLE acts as a foreign key in relation HOMES2 and as a primary key in relation COST.

In the next section we are going to look at Structured Query Language (SQL), which is a language that will enable to create, delete and query relational tables.

4.8. Structured Query Language (SQL)

Structured Query Language (SQL) has been developed to aid users to create tables, populate tables, query the data in these tables, alter the data as well as deleting tables without any major programming skills. SQL has the following advantages:

Completeness: It is a comprehensive database language which gives us a full set of database tools in a consistent format;

Simplicity: Despite its comprehensiveness it only uses close to 30 commands to do everything, thus enabling the user to learn the different commands quickly;

Declarative: It states clearly what it wants, but it is up to the computer how to achieve the answer. The software determines how this is done;

Pseudo English Command Structure: SQL uses English words such as CREATE, DELETE, SELECT, etc. to specify the action. Remember an SQL sentence ends with a semi-colon (;) and NOT with a full stop. Full stops are for decimals (Reeves, 1996).

SQL commands can be placed into three different categories. We will briefly look at each category and give an example or two. The first category are the SQL Database Definition Commands. There are five commands which allows us to undertake database definition tasks, namely:

CREATE TABLE;
ALTER TABLE;
CREATE INDEX;
CREATE VIEW;
DROP TABLE/ VIEW/ INDEX.

A view is a table which shows only a part of the database to a user. An index is a small file which is used to keep track of the position of the rows in a table to enable fast extraction of data when a query is run, by going directly to the records in question rather that going through each row sequentially.

Example:

CREATE TABLE <table-name>

(<column-name><data_type> [NOT NULL]

[,<column-name><data_type> [NOT NULL]....]);

The data type can be either real, integer, float, logical, character (string), etc. NOT NULL means there must be something in the row, it must not be empty. NULL is not a zero or a blank (Reeves, 1996).

DROP is a command that is used to delete the table, view or index that was created. If you use the DELETE command you will only delete the contents in a row and not the complete table. When you use the CREATE command you only create the table, view or index but it is still empty. The next group of SQL commands are used to populate these tables.

The second category are the SQL Data Manipulation Commands. This allows data to be inputted, updated, queried and deleted.

INSERT INTO;

UPDATE;

DELETE FROM;

SELECT.

INSERT INTO allows data to be inputted into a table by creating a new row each time. UPDATE allows one to change the content of these entries. DELETE FROM is the command to delete a row that was inserted incorrectly into a table, view or index. The SELECT command allows us to query the contents of the table or view.

Example:

SELECT <column_list>

FROM <table_list>

[WHERE search_conditions]

Multiple search conditions are linked with AND or OR.

The third category of SQL commands are the two commands that allow people to use or not to use a database. GRANT allows those users on the list to access the database. Once the user does not use the database anymore then the REVOKE command is used to restrict the user to the database.

The above is a glimpse into database theory as well as a short introduction to SQL. The next section will look at spatial operations in GIS.

5. Spatial Analysis and other GIS Analysis Functions

GIS has several analysis functions ranging from maintenance to spatial analysis. It is the spatial analysis function of the GIS that distinguishes it from other systems such as Computer Aided Design software packages. □ These functions use spatial and non-spatial attribute data in the GIS data base to answer questions about the real world□ (Aronoff, 1989:189).

Since data in a GIS is a model of the real world we use models to mimic certain aspects of the real world. But when we create models of reality we only select a few features of the real world, thus when we mimic the real world, we mimic only those selected parts. We use spatial analysis to play □what if?□ games or we try to establish □what is best?□ to aid us in our decision making process, etc. □In order to find usefull answer one must ask the right question□ (Aronoff, 1989:190). To find out if we have the right answer, we reverse the process, via analysis, to establish if the answer that we got, gives us the right question.

To do effective spatial analysis we have to organize the geographic data in such a way that it enables us to do analysis. There are several ways of doing this.

One of the methods is to separate the data into data layers which consist of layers of logically related geographic features with their attributes. An example is that the railway lines and roads are represented in a transport data layer whereas rivers and lakes are represented in the hydrology data layer.

If we deal with a huge database but are only interested in a section (spatially) of the database, we can partition the data into smaller sections called **tiles**. These tiles use a stable boundary (meaning it does not change over time) such as a grid determined by the latitude and longitude of a UTM Grid. If available, map library software is used to keep track of and manage all the tiles in the GIS database.

5.1. Classification of GIS analysis functions

GIS analysis functions can be classified into four main groups. The first group is the maintenance and analysis of spatial data, the second is the same as in the first group, but it only applies to attribute data. The third group is the integrated analysis of spatial and attribute data and the last group deals with output formatting (Aronoff, 1989). The last group concerns itself with the different options of map production and reporting.

5.2. Maintenance and analysis of spatial data

Maintenance and analysis functions are used to transform spatial data files, edit them and to assess their accuracy. We have previously looked at tiling and if we want to connect adjacent tiles into one area of interest we call it mosaicing.

Format transformation functions are needed when we import a database from another source, which might be a different GIS software package with it's own file formats and data structure. This procedure is relatively easy when working with a raster data structure, but it is more complex when it involves vector based data with its topology. We normally have to recreate or "build" the topology from the coordinate data once it is transformed from its original source into the required format.

Geometric transformations are used to assign ground coordinates to a map or a data layer within a GIS or to adjust one data layer□s map projection so it can be correctly overlayed over another data layer (Aronoff, 1989). This procedure is known as registration and there are two different methods of doing it.

The first method is to register the data layer by using relative positions. Here the features of a layer is placed relative to the same features (e.g. the boundary of Natal) of the other layer. The first layer is known as the SLAVE and the second layer as the MASTER. This type of registration is also known as **rubber sheeting**. The second method of registration is by using the absolute geographical position of the same features. Here we correct the absolute position of the features in each data layer separately by using the same map projection for each layer. This method prevents the propagation of positional errors. When we use the first method the positional errors that occur in the MASTER data set will be propagated to the SLAVE.

Conflating according to Aronoff (1989:201) □ is the procedure of reconciling the positions of corresponding features in different data layers□. This method is used for example when you have maps of different vintage, where certain areas were stable (no positional change over time) and you want to overlay these maps to detect change in other areas. After the overlaying procedure the stable areas may show small polygons (slivers), these are is due to small input errors during data capture. To eliminate this error is to create a template that acts as a base map from where all the other relevant features are redrafted.

Due to the characteristics of the paper used for the maps that we need, as well as other factors such as humidity, the edges of maps are sources of errors as shown in the figure below.The procedure to correct these discrepancies is known as edge-matching.

Editing functions are used to correct spatial features after digitizing them from a map. This means cleaning the data set of spikes, switch backs, etc. that is part of the digitizing process, but has a profound influence on the outcome of our analysis procedure, if left unattended. With editing functions we can add and delete spatial features or change their geographic positions. Aids such as line-snapping are tools that are designed to minimise digitizing errors and are also used when the spatial features need editing.

Line coordinate thinning is a form of generalization (discussed earlier). Its function is to decrease the number of coordinate pairs that defines a line or the boundary of a polygon without losing the feature□s general form of the line or boundary.

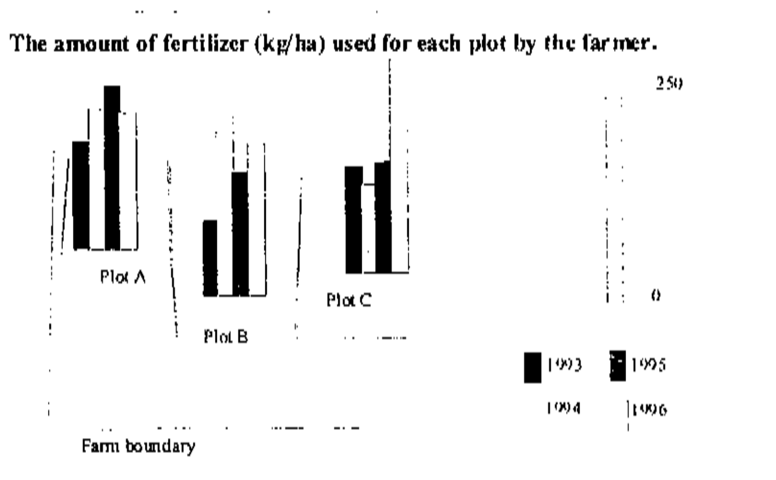
5.3. Maintenance and analysis of non-spatial attribute data

In this sub-section we discuss two different functions. The first are the attribute editing functions. Since many GIS analysis functions depend on attribute data alone, it is necessary to be able to edit the attribute data set. Some GIS software allows easy access to the attribute data set by using the □info-button□ and clicking it on the spatial feature that needs to be changed. The relevant information appears on the screen and the incorrect information can be changed. Other commands such as table maintenance allows the changing of the data type (e.g. text to integer) or the deleting of irrelevant columns from the table. Another form of editing is called **file matching** or **address matching** where the attribute data is matched to a spatial data file. This is done to display the attributes spatially. Another term for this function is **geo-coding**.

The second set of functions are the attribute query functions. These functions are used to query or to retrieve parts of the attribute data set, using search conditions, as discussed at the end of Section C. If you have different attribute tables then you can join these (called a relational join) when they share a common column. See discussion on primary and foreign keys in the previous section.

You can also create an open database connectivity (ODBC) between the GIS and a DBMS, such as Microsoft Access, and then make use of limited SQL commands to query the attributes and create a new layer in the GIS (if spatial data is available) or join it to an existing spatial data set.

Integrated analysis of spatial and attribute data



This capability is what makes GIS such a power tool as an aid in decision making. There are four different categories which will be discussed in more detail below. These are:

retrieval, classification and measurement functions,

overlay,

neighbourhood functions, and

connectivity or network functions (Aronoff, 1989).

This discussion is based on the work of Aronof (1989) Chapter 7.

Retrieval, classification and measurement functions

Retrieval involves the extraction, manipulation and the displaying of selected data. The geographical position of the entities in question does not change and we also do not create new data layers. Thematic mapping falls under this sub-category.

Thematic mapping is used to display selected data in a spatial format, either as a mosaic, or as pie charts or bar graphs at the centroid of a polygon or at a line or point on the map (see Figure 21).
Figure 21. Thematic mapping using bargraphs

Classification and Generalization

If we want to group entities together into several groups, then we classify them. E.g. individual plots classified either as residential, commercial; agriculture or industrial, etc. When we classify entities we may have to reclass them by assigning them new values

and thus creating a **new** data layer from the existing data layer. We use classification to help us to recognise patterns. Patterns can be used to identify areas of high crime or areas that are suitable for logging trees (Aronoff, 1989).

Another form of generalization is called **map dissolve** by merging polygons of the same class together by dissolving the common boundaries between them. This brings not only a simpler map but it can also high light underlying patterns which were not visible before.

Measurement functions

Since we work in space using coordinates and scales we are able to calculate the perimeters, distances and areas of the spatial features in question. These are done in both the data models in GIS. Using measurement functions we do not change the geographical position but we use the spatial information to create new attributes which can be used in queries such as: Show all the forested areas larger the 120 square kilometres.

5.4. Overlay operations

We have mentioned previously the overlay functions of a GIS. With overlay we mean putting one layer on top of another layer using different mathematical functions as well as layering the one image over the other. We use overlay functions to identify suitable areas for development, or showing environmentally sensitive areas, or showing areas of suitable habitat and corridors for the blue swallow in the Natal Midlands. The mathematical functions are the easiest to use in a raster data model since we work with a single value in each grid cell. One can add two layers together or subtract, multiply or divide them with each other. You can take the one layer to the power of the other layer.

In the vector environment you use the area of e.g. polygons as a variable and when you overlay the two polygons you divide the two polygons into smaller polygons. Each of these polygons will have their own area, but the sum of these areas are equal to the area of the two original polygons. This is also known as polygon clipping (Aronoff, 1989). Since overlaying is fairly easy in the raster data model, some GISI is incorporate both data models (the hybrid approach) to alternate between the two models when the need arises. ArcINFO is such a GIS.

5.5. Neighbourhood operations

If we want to understand the characteristics of the surrounding area of a feature that are of interest to us we use neighbourhood operations. An example is the proposed closure of a primary school. We want to determine how many children between the age of 6 and 13 within a 6 km radius from the school will be affected by this proposal. According to Aronoff (1989) there are three requirements that have to be fulfilled when we want to do neighbourhood operations, they are:

i) one or more target locations such as the school;

ii) the neighbourhood must be specified, e.g. a 6 km buffer; and

iii) a function that has to be performed on the elements with in the neighbourhood, such as find the total number of children between the age of 6 and 13 that will be affected and determine the teacher to child ratio when 26 teachers are employed.

Other neighbourhood functions are the search function where ..i la function is applied to a neighbourhood to generate the neighbourhood valuel l (Aronoff, 1989:212).

Search functions are either applied to numeric data, e.g. showing the different property values within a 5 km radius from a fire station, and thematic data, e.g. the number of different classes in a neighbourhood based on different land uses. Search areas can be either regular in shape or irregular (known as a window) and both can be defined by the user. In the vector environment we can use the line-in-polygon or the point-in-polygon functions. A simple point-in-polygon example is to find all the supermarkets within a town and a line-in-polygon example is to show all the National Roads (N1, N2, etc) that cross KwaZulu Natal.

The other set of functions deals with the topography of the area and are commonly known as topographic functions (Aronoff, 1989). A topographic model of the landscape showing only the elevation above mean sea level is normally known in the GIS community as a Digital Elevation Model (DEM). From a DEM several calculations can be made, they are:

i) the **slope** of the area which is the rate of change of the elevation;

ii) the **aspect** of the surface in the area, meaning the direction the surface faces; e.g. North or Northwest, etc or 25^0, 270^0, etc.

iii) or the maximum slope of the surface (the **gradient**)

The following functions are those that are used to convert point data into surface data. Thiessen polygons or Voronoi polygons are used to define areas of influence around individual points. Meaning that any other point within such a polygon is the closest point around which the polygon was generated. Thiessen polygons are mostly applied to climatic data, such as rainfall. The value of the polygon is the same as the value of the point around which the polygon was generated. The boundary of a Thiessen polygon is exactly halfway between two adjacent points.

The second function in this group is known as interpolation. Interpolation is used to predict unknown values using known point values at adjacent locations, using mathematical procedures, ranging from a simple linear functions to fourier series and kriging. The last function in this group is contour generation. Contour generation is used to connect points of equal value, be it elevation, rainfall values, crime rates, time zones, etc. It involves also interpolation functions to predict the unknown values between neighbouring points, once these values are established then the points of equal value are then connected.

The last group of analysis functions are known as connectivity functions and will be discussed in more detail in the next sub-section.

5.6. Connectivity Functions

According to Aronoff (1989) connectivity functions are those functions that accumulate values over the area when traversing it step by step, e.g. every 100 m interval, etc. Each connectivity function must include the following: firstly it needs a specification of how the different spatial elements (such as rivers) are interconnected (normally it is derived from the topology of the spatial data), secondly it needs a set of rules which specify what is allowed to move along these interconnections and sometimes even the direction of movement is needed, and lastly it needs a unit of measurement (Aronoff, 1989).

Contiguity measures is a type of connectivity functions that is used to establish areas of shared values and combine them in a single unit. Some contiguity functions do not allow gaps between the areas, whereas some allow gaps between areas such as roads or rivers which cross the areas of interest. The common measures of contiguity are the following: the size of the contiguous area, and the length of its longest axis and its shortest axis. An example where we will use such a function is to establish a new development site which is larger that 20 hectares and not narrower than 1 km.

Proximity functions are used to measure the distance between features of interest and must include four parameters to enable it to do so, they are:

i) the target location, such as a shopping centre;

ii) a unit measure such as the distance in kilometres;

iii) a function to calculate the distance such at the straightline (as the crow flies) distance from source to target ; and

iv) the area that is going to be analysed (Aronoff, 1989).

An example is the school example discussed previously in this section.

Network functions

"A network is a set of interconnected linear features that form a pattern or framework" (Aronof, 1989:225). According to Aronoff (1989) a GIS is used to predict network loading, resource allocation such as the strategic location of police vehicles to effectively combat crime in areas where the crime is high, and route optimization, e.g. the planning of bus routes that will give the maximum coverage at the least cost in a neighbourhood. When we want to do network analysis, we need the following:

i) a set of resources such as passengers that need to be transported;

ii) one or more locations where the resources are located, e.g. existing bus stops;

iii) an objective, as given for the bus route planning in the introduction; and
iv) a set of constraints that limits the criteria that has to be met to achieve the objective (Aronoff, 1989).

Spread functions make use of some of the characteristics of network and proximity functions to evaluate phenomena that accumulates with distance, e.g. the further you move away from a source the higher the travelling costs. This is done in a step-by-step fashion calculating the value (travel cost) at each successive step. Constraints such as barriers can be placed to influence the calculations of travelling costs. The output is normally known as an accumulation or friction surface. Friction surfaces are also used to retard movement across a surface or to place constraints when establishing the shortest route between two points. An example is to use a friction surface showing various levels of environmental sensitivity when planning a route for a powerline that has the minimum impact on the environment. This function is used in the raster data model environment using each neighbouring cell to progressively accumulate the required value.

To establish a stream□s pathway over a terrain, such as a digital elevation model (DEM), a seek or spread function is used. A seek or spread function uses a specific decision rule (e.g. find the next lowest elevation adjacent to this one) that is applied step-by-step until it can not apply this rule anymore.

Intervisibility functions or viewshed modelling is used to map out areas that are visible from a specific location taking into account the elevation above sea level and the height of the specific location above the ground. Viewsheds are used to establish □line-of-sights□ from scenic lookouts or the area that are covered by a specific radio antenna. Viewshed mapping is also used in landscape architecture to hide unseemly sites such as rubbish dumps or electricity pylons in the vicinity of a natural area.

Perspective views are 3-D generated images, mainly from DEMs, to display the topography. Other layers, such land use, may be draped over them to give a 3 dimensional perspective of the land use. It is mostly used for display purposes only.

The last group of GIS analysis functions is **output formatting** which is used to present the previously discussed analysis functions in a hardcopy format. It deals with map annotation, such as the North arrow, scale bars, etc., the placing of text labels on maps or graphs, the different symbols that are used to show different types of point locations, etc. These are the functions that are needed to produce high quality and easily understandable hardcopies for presentation purposes.

6. Web-enabled GIS

6.1. Introduction

This section draws on the literature by Plewe, 1997. This section deals with the distribution of geographic information over the Internet or intranet. This technology is going to be used more and more in the future.

6.2. Distributed Geographic Information

Distributed geographic information (DGI) applications range from simple, pre-draw2n maps, such as pin maps showing hijacking incidents, on a Web page to network-based collaborative GIS, where for example individual police stations contribute to the central database spatial data as well as sharing data with other stations and role players with the South African Police Service. (SAPS). The advantage of this is that one can communicate in real time with each other, meaning the neighbouring police station has simultaneous access to the data to help with crime prevention activities.

The following technologies have been developed to make DGI applications possible. These are: servers, which store the data and applications; clients , which use the data and applications; and network communications, which control the flow of information between servers and clients. Autodesk MapGuide is an example of DGI software. It consists of three entities namely Autodesk MapGuide Server (stores the data and applications as well as manages who is allowed to see what data or maps), Autodesk MapGuide Author allows the user to create applications and to store them on the server and Autodesk MapGuide Viewer allows the user to view the data and maps (the client).

6.3. Why DGI?

There are several reasons for DGI, ranging from sharing data within an organization to sharing data that belongs in the public domain, such as the TIGER files used in the US that contains census data up to street level, up to giving "teasers" on the Web to sell data. The SAPS route will be sharing data within an organization.

6.4. What types of servers and client combinations are available?

In this section we look at the different types of servers and clients available for DGI. The server can either be heavy or light and the client either thick or thin. In the next few paragraphs we will discuss these different options available and how the server type interacts with the client type. The combinations of the different types depend of computer processing needs, type of searches that has to be done, GIS operations that needed to performed and maps that have to be drawn again and again.

6.5. Heavy or light server?

The main principle behind client/server is to store everything on a single machine, from which the clients can access data and maps. This same principle applies when using the Web. If the GIS data is kept on a single machine it is much easier to update and maintain the data as well as to control the access to the data. This concentrating the data and processing load on a primary server is known as a **heavy server**. Due to the heavy processing needs and high data load associated with GIS it is very important to have a machine that is capable of handling the traffic. Response times should be far less than 3 seconds. This involves high installation costs which includes the interface of GIS processes and data with a Web server to process live, dynamic requests. Another short coming of this approach is that even for the smallest task you need Internet communication and server processing. This leads to an crease in network traffic, which slows down the response time significantly.

A **light server** is where most of the processing is done at the client's machine and not on the server. The server acts only as repository for data as well as some functionality on who is allowed to access what data. The choice between the two depends on what is expected from the GIS in a particular organization.

6.6. Thick or thin client?

In general Web browsers fall under the **thin client** variety, meaning that the server is handling the processing, with the browser processing the display only. The alternative is a **thick client**, which does most of the processing of the information and/or data. The latter does zooming in or out, panning and spatial queries at the client without bogging down the network or taxing the server. The use of Java or ActiveX applets as well as "plug-ins" enables the GIS user to increase his/her browser processing capabilities. These applets or "plug-ins" are available from the server the next time a user visits the site. The disadvantage of applets is that has to be downloaded every time when a user visits the site, but a "plug-in" allows for permanent installment on the client'smachine. The Autodesk MapGuide Viewer is an example of such a "plug-in" and has a wide variety of map viewing capabilities. The choice between the two depends on what your needs are and how popular your site is.

6.7. The answer to both questions!

A thick client, which interacts with a light server, allows for flexible and powerful analysis. Unfortunately it cannot handle a lot of network traffic and also limits your audience. It is also difficult to maintain, updating a whole bunch of clients can be a nightmare. On the other hand when using a **heavy server – thin client** option it uses a lot of bandwidth and is normally restricted to simple applications, such as simple map viewing capabilities. The latter has the advantage that a wide audience can use it.

6.8. Several DGI applications

The following applications can be done:

- Raw data download for further use in GIS. In this case the GIS is a stand alone GIS. The data is down loaded onto your machine's own disk and then is available for own use.
- Static map display. In this case maps are created using a GIS but then saved as a bitmap or as a JPEG and commonly forms part of an HTML document. This is used when a wide audience is targeted.
- Metadata search. This used to search data on GIS data, which contains all the relevant information on a data product that is available. In South Africa it can be accessed at **http://www.nsif.org.za**.
- Dynamic Map Browser. This is the most popular way of displaying data. It allows for the client to browse a map, such as zooming in or out, panning, basic query or when moving the cursor over a map feature it will display the attribute of the feature and switching layers on and off.
- Data preprocessing. This is a DGI application where the server hosts a high end GIS such as ArcInfo to do high en preprocessing of the data. The new data is stored on the server and is then down loaded onto the client for further processing using a lower end GIS such as AtlasGIS, MapInfo or ArcView.
- Web-based GIS query and analysis. This is a thin client – heavy server application which allows the client to have a wide range of GIS functionality available ranging from strait forward GIS applications to more complex functions such as overlaying layers or creating buffers.

6.9. Further reading

These few sections gave an insight to what Web-enabled GIS is about. For further reading about how to set up Web-enabled GIS and the different type architectures available for establishing Web-enabled GIS, the reader is advised to read the following textbook and reports:

- B Plewe, 1997: *GIS Online: information retrieval, mapping and the Internet.* Published by OnWord Press, Santa Fe, USA. SAN 694-0269
- Crime Analysis and Decision Support in the South African Police Service: Enhancing Capability with the Aim of Preventing and Solving Crime. Activity 7: Packaging and roll-out. Part 3: Web-enabled GIS - Using Autodesk's MapGuide to distribute information of geographical nature over the Internet or intranet. Schmitz, PMU.
- Crime Analysis and Decision Support in the South African Police Service: Enhancing Capability with the Aim of Preventing and Solving Crime. Activity 7: Packaging and roll-out. Part 4: The roll-out of a Geographical Information System (GIS) for crime pattern analysis in the South African Police Service. Schmitz, PMU

The next section deals specific with geocoding, an important GIS function available to plot crime incidents

7. Geocoding

7.1. Introduction

Traditionally, crime was mapped by sticking coloured drawing pins into wall maps to show the location of different crimes. Computers and Geographical Information Systems (GIS) using geocoding are increasingly replacing the placing of crime incidents on a map or pin mapping. Automated geocoding hasbeen successfully used in many police forces throughout the world including the South African Police (SAPS). However, there is an increasing realization within the SAPS that the implementation of automated geocoding is fraught with difficulties and that practical solutions are urgently required. The primary intention of this report is therefore to make a modest contribution towards achieving this goal. The second purpose is to describe research on defining the requirements of the SAPS for an automated geocoding system for crime analysis at an operational level. From this work a user requirement statement/document will be produced. Part of the work would be to define an acceptance test protocol for automated geocoding in the SAPS. An acceptance test protocol for the purposes of this report, is defined as tests the SAPS needs to implement in order to ensure that the correct accuracy levels are being maintained while geocoding crime incidence.

Utilising a qualitative methodological approach, unstructured inteviews were conducted with a small number of senior management as well as personnel that have been exposed to this technology within the SAPS. From the interviews valuable information on various perceptions and views regarding the utilization of automated geocoding were obtained. From the data it is very clear that there is a desperate need for automated geocoding of crime in the SAPS. This need is confirmed by the fact that it is those international police forces that are using information as the core of their fight against crime that are succeeding in the tide against this epidemic. This can be effectively achieved only by the use of management information systems including GIS. Ultimately, what is required in the SAPS is a Management Information System (MIS) that gives the police access to a wide range of information needed to prevent crime in South Africa.

The police also indicated that it was imperative that whatever system is developed to enable automated geocoding it must be simple to use, accessible and integrate with existing systems in the SAPS. It also became clear that the strategy for implementing automated geocoding in the SAPS should be comprised of short, medium and long-term goals.

Identifying several areas for further research concludes the report that forms a natural extension of the work completed in this project.

- Develop a seamless street address/property database for a selected police station area in South Africa and integrate it with other required GIS data sets for crime analysis and the development of appropriate crime prevention strategies.
- Research the viability of integrating or interrelating information from different databases and sources into a Management Information System for the SAPS (e.g. credit bureau, vehicle register, gun register, population register, voters roll, etc).
- Through focus groups with members of the SAPS gain an understanding of their perceptions of the SAPS and policing in South Africa.
- Evaluate existing training programs in the SAPS and present training programs to the SAPS on crime analysis as outlined by Steven Gottlieb, an international expert in this field.
- Evaluate the methods used to report crime in South Africa as well as the methods used by the SAPS to collect information on reported crimes.

The next section looks at what geocoding is all about and the two different approaches to geocoding available to SAPS

7.2. What is meant by automatic geocoding?

In a geographical information system (GIS), geocoding data involves linking a record from a database to a geographical feature in the GIS, such as a point, a line or a polygon. To geocode the record, one needs a unique key field in the data base that provides a spatial reference or geographic identifier (an indirect spatial reference [ISO/TC 211 1998]), such as the name of a magisterial district, the name of a local authority, the name of a farm, the name of a suburb, an erf number or a street address. This is illustrated in Figure 22, below, in which the data base record for a magisterial district *Riversdale* is geocoded through its key field, the name of the magisterial district, which in the map is a label for the polygon representing the district.

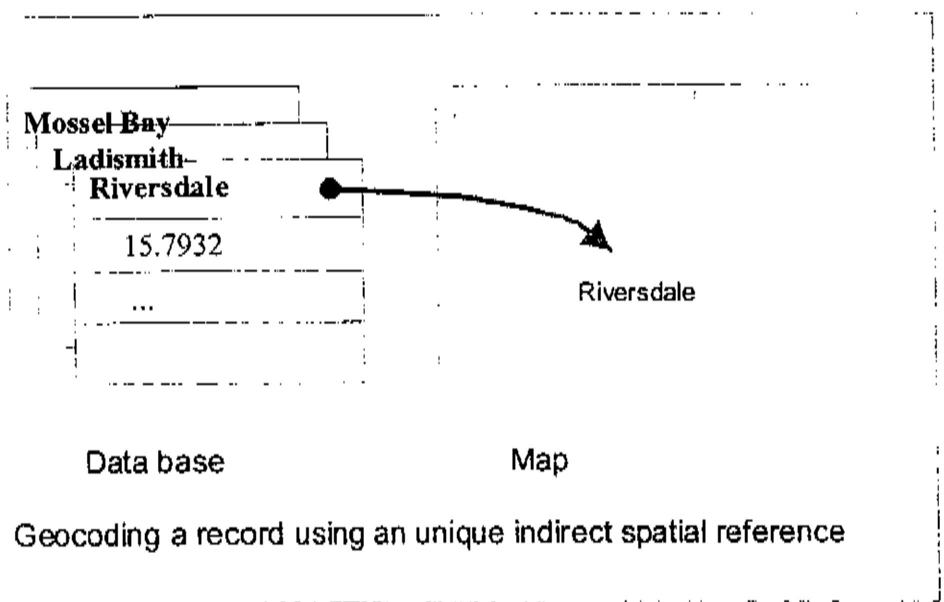


Figure 22. Geocoding a record using a unique indirect spatial reference

Generally, crime data are geocoded using street addresses. This process is also known as pin mapping, to reflect the traditional analogue method of sticking coloured pins in wall maps to map crimes. It is manual if an operator has to geocode the record, and automatic if the computer can do it by converting the street address into a pair of coordinates. A geocoded street address is a street name and number combination that has been linked to a pair of geographical coordinates. In South Africa, this could be done through the erf (property), though in other countries interpolating along the edge of a street block is normally done.

7.3. Research into the issues that will affect automatic geocoding.

During this phase two reports were produced. The first report "***On geocoding crime in South Africa***" gives a theoretical overview of issues involved in geocoding crime in South Africa. It starts with a background to geocoding and to the situation in South Africa with regard to the reporting systems used by the South African Police Service, such as the fields in the Crime Administration System (CAS) pertaining to geocoding, the geographical units used by the police, the quality of address data in CAS, etc. The situation with regard to geocoding is first sketched in other countries and then applied to the situation in South Africa. Possible ways of geocoding in South Africa are fully described in five recommended steps, namely: 1) Identification of the syntax of the address field, 2) reparation of the address, 3) replacing aliases with the standard address, 4) validating the address, and 5) reformatting the address field

The second report "***An investigation into approaches for geocoding of crime in South Africa***" reflects on definitions of geocoding and provides a summary of a literature study, done mainly on the Internet, of the various available technical methods of geocoding and the white papers ofvarious software vendors for the purpose of obtaining more information on the expected features regarding geocoding in the future. Examples are:

> When using the *Zip Centroid* method all the data is linked to a centroid, thus clustering all the locations around the centroid. The second method is the *Random Method* where the locations are assigned randomly within a

polygon. The advantage is that locations are not clustered around the centroid, but the disadvantage is that it does not show actual locations within the polygon. The third method is the *Representative Method*, where the locations are assigned based on the population density within the polygon. It places the locations were there are concentration of people and is thus a useful tool when no street addresses are available. The *Zip+4 Method* is a refined version of the representative method and the only disadvantage is that it does not place the locations at their exact positions. These four methods were developed to be used with Zip Codes, but one can apply the same techniques to other polygons such as suburbs, etc. The fifth method is the *Address-based Method*.

The report also provides a short summary of the various options that are available with regard to issues such as whether stand alone PC's, client server technology or server-only geocoding technologies should be considered. In this report possible methodologies are summarized and the five steps for geocoding in South Africa are again listed, but with additional information.

Various role players have attended a geocoding workshop given by the CSIR, the HSRC and members of SAPS on the 26th February 1999. During this workshop the following issues were fully explored: i) The rationale for geocoding from a project perspective, ii) the rationale for geocoding from a crime management perspective, iii) sociological perspectives on crime geocoding, iv) approaches to geocoding, v) layers of information needed for geocoding, vi) issues relating to CAS, and at that time vii) the way forward.

Investigation into available software tools that would be needed for:

1) parsing (i.e. different spellings and syntax need to be recognized as having the same meaning) the information in the free text fields from CAS, and

2) software with soundex (i.e. words with the same sound and therefore probably sharing the same meaning) capabilities, etc. One of the requirements is that the software should be compatible with the systems used by the police.

Conclusions from this investigation were that:

1) the free text in the fields pertaining to street addresses should be standardized and that standardization should be communicated to the data capturers in the police station.
2) Parsing in Oracle (the database the police uses) alone would not be possible. An external parsing tool would also beneeded.

7.4. Geocoding of incidents using data from the Johannesburg Area Pilot

During this phase vendors of data or possible data sources of street address dictionaries were identified and meetings were arranged with various possible vendors such as Africon, Intact Solutions, GIMS and Metropolis. This was done after the Johannesburg Municipality was contacted to find out whether they have already finished digitising the street addresses in the hope that we would be able to link the data with information from the servitude roll similar to what has been done with Pretoria. The data was, however too expensive, ranging from R 800 000 to 1,2 million. A decision was taken to obtain sample data from vendors. They would be willing to provide sample data for the following two reasons:

1) That it was too expensive to buy the data, and
2) the quality of the data needed to be assed first before buying the data.

Sample data was obtained from GIMS and from Metropolis which gave us the opportunity to compare two different models of geocoding, namely 1): the American model where street numbers are interpolated along the edge of a street block and 2) where the street address are linked to the centroid of the erf.

7.5. Detail of the automatic geocoding with data supplied by GIMS

GIMS asked for feedback on this sample data. Therefore a paper was prepared and presented during a conference in September. The data is road centre lines with address ranges on the left and the right of each road segment. If the number equal to 9999 or 0 for both the minimum and maximum left/right ranges, it indicates that there are no addresses available showing for example open spaces such as parks or that the land is undeveloped.

Several people from GIMS worked on the project and the street types were standardised, e.g. Avenue or Ave was standardised to Av, Close to Cl and Street as St. Spelling of street names were also checked. Once this was done the street data was displayed as a theme in ArcView. The second set of data received from GIMS was a sample set of cadastral data. The area covered is the area around Johannesburg Zoo, Parktown and Hillbrow. The projection used is the Gauss Kruger SALO 27 and the spheroid is Clark 1880 (see Figure 23).

Thus the data from the Anti-Hijacking Initiative (a project aimed specifically by SAPS to combat hijackings in the northeastern and southern parts of Johannesburg) is used since the addresses were more accurate. Data relating to hijackings specifically is collected on a daily basis on a separate questionnaire that provides for detailed victim and suspect details as well as detailed information on *modus operandi.*

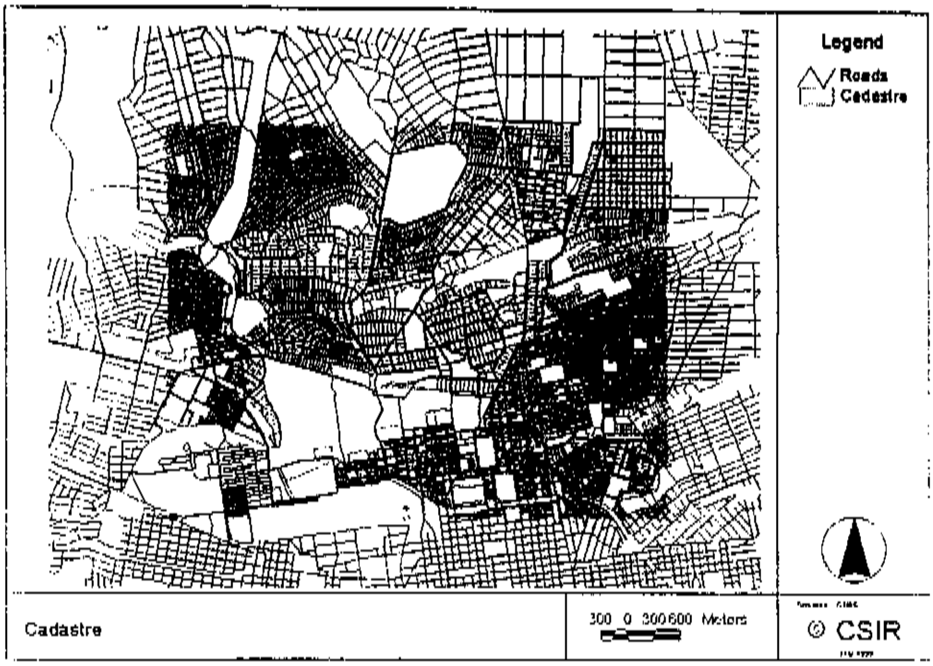


Figure 23. Road data with street address ranges and cadastral data

To facilitate geocoding the data was cleaned of all the records where street names were absent. Due to the design of the database the address was divided into two fields namely one field for the street name and type and one field for the number. These two fields were concatenated to form a field showing the number, street name and type. The concatenation was done using a tool within ArcView. This column was then edited to change the street type to LN, CR or AV and street numbers such as 44a was changed to 44, to facilitate geocoding of the address. The geocoding process will be discussed in more detail in the next section.

Geocoding

In this section we will discuss the geocoding procedure in ArcView. The data used for geocoding is the hijack data for the period 1 October 1998 to 8 February 1999.

Batch

For geocoding you need a table containing the data that you want to geocode and a theme to which you want to geocode. The theme can either be polygon data such as Zip+4, or a line feature with address ranges. Point themes can also be used for geocoding purposes.

You add the table to the project and make the view in which you want to geocode active and click on "Geocode Addresses" [ESRI, 1996]. The first option is to geocode the addresses using "US Streets without Zone" option. This was done in batch mode. Batch mode is an automated geocoding process using ArcView default settings for spelling sensitivity. The hit rate is then displayed to the user in table format.

The second option is to geocode the hijacking database using "US Streets with Zone". This method of geocoding allows for more accurate location of incidents since the address with the suburb is used. When using this option a left hand and right hand polygon should be given to indicate the zone. The reason is that a road might be the boundary between different zones. Zones can either be numerical (ZIP codes or postal codes) or a string (suburb name, etc). Since the geocoded reference database did not include information on left and right polygon, an assumption was made that the suburb occurred on both sides of the road, using the same column for both. The first geocoding run was also done in batch mode and the results were displayed in a table format.

Interactive

Based on the results after the batch mode run, the user has the ability to either rerun the batch mode and set the sensitivity at a lower level to improve the hit rate, or to run the geocoding process in the "Interactive Mode". The lower spelling sensitivity level allows ArcView to select several possible candidate street names and numbers for the misspelled address. The user then clicks on the street name and left/right range combination that is the closest to the misspelled address.

Using the interactive mode the user is then able to accept an address or not by clicking on the "Match" or "Unmatch" button. When the "Match" button is clicked the address will be geocoded and added to the number of matched addresses. If a street name and left/right range is selected for a misspelled address, the "new" address will be geocoded.

7.6. Results

Figure 24 shows the distribution of hijacking incidents using the "with Zone" option. There is a difference in the distribution of incidents. We managed to geocode 101 incidents out of 419 entries. The low number of geocoded addresses can be attributed to two reasons namely: that some of the entries did not have street names, type and number, and that some of the addresses which qualified fell outside the study area. It has to be kept in mind that a large part of the data in the database consists of hijacking incidentsin the Sandton area just north of the study area.

During a meeting with Metropolis it was decided to arrange a sample set of data from the police and test this with GeoPlus, a parsing tool used by Metropolis. The police agreed that Metropolis receive some data to see if it could be geocoded. Although the police data was of poor quality, Metropolis managed to perform rather well. A sample set of street address dictionary data was also obtained from Metropolis. We tested the same hijacking data used on the data from GIMS with this street address data for Johannesburg.

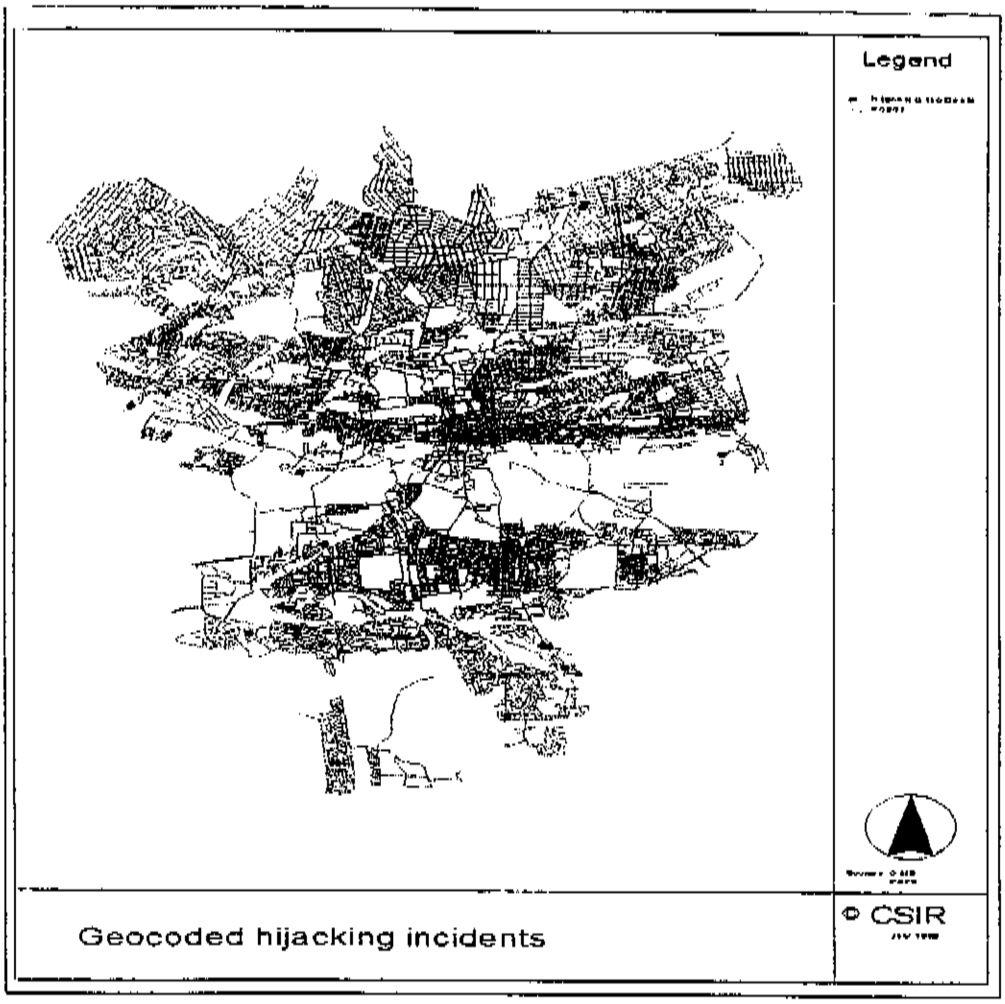


Figure 24. Geocoded hijackings using 'address with zone'

1. **Comparison between the American system and Land parcel**

Figure 25 shows the difference between the US Street address range method and the land parcel method.

There are inadequacies when using the US Street address range method, it divides the street centre line into equal length based on the number of addresses left and right of the centre line. These locations might differ from the actual location of the address, when compared to position of the land parcel in question. For the South African situation the land parcel option is the most viable route to go. The reasons are:

- It is costly to develop road centre lines with left and right ranges;
- There is land parcel data (cadastre) available for the whole country (Project Miracle);
- A private sector company succeeded in linking the addresses and erf number to the SG's unique land parcel number;
- The same company also wrote algorithms to match addresses taking into account local issues, including the formal black townships' numbering systems; and
- The US Street address range does not cater for street numbers such as 44a and 44b, these are addresses that are common when dealing with panhandle properties. For crime analysis land parcel geocoding is a better option, especially when dealing with *modus operandi* of burglars, etc. A specific gang may, for example, prefer to burgle panhandle properties since they are normally out of sight from the streets and patrolling police vehicles.

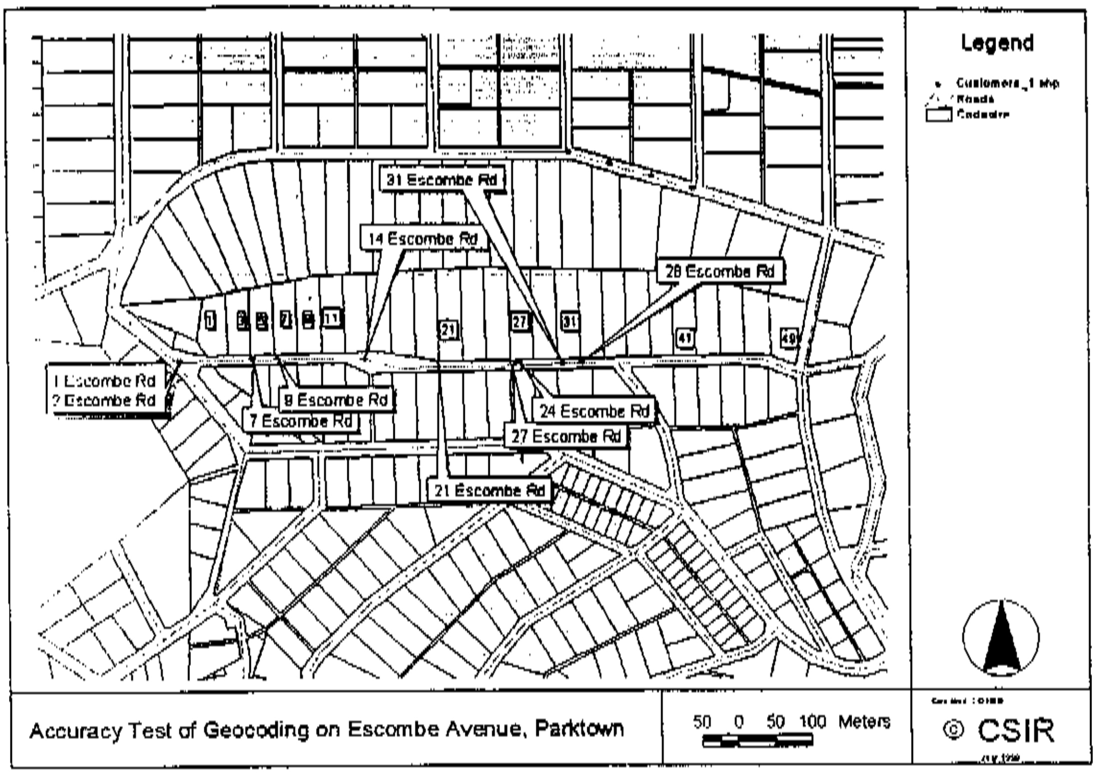


Figure 25: Difference between address ranges and land parcel data

7.7. After Geocoding

In this section we look at two examples of using geocoded crime incidents. The first example shows hot spot analysis (see Figure 26) the second example deals with aggregating point data into polygons.

Hot spot analysis

Hot spot analysis can be done using STAC (Spatial and Temporal Analysis of Crime), developed by the Illinois Criminal Justice Information Authority. This software package used the pin-mapped locations of crime incidents and calculates the hot spots and their centroids, using set parameters provided by the user. Figure 26 shows the hot spots for hijackings in Johannesburg, using STAC.

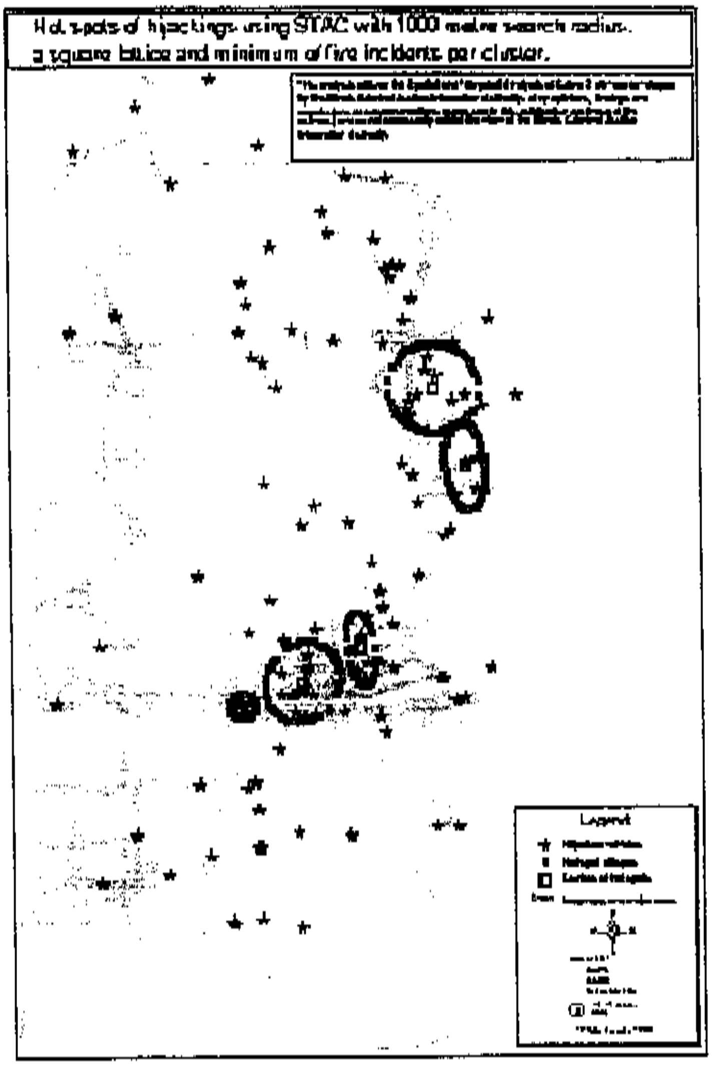


Figure 26. Hot spot analysis

CrimeView, an extension for ArcView, uses different algorithms to calculate hot spots.

Aggregate data to CAS Block level

With GIS it is possible to calculate the number of points within a polygon, such as a CAS Block, thus aggregating the data to CAS Block level and ultimately aggregating it to police station level. One can then use other data such as the area of a CAS Block or police station to express the number of crime per square kilometers. Since the EA's are available and there is currently a project done by the HSRC to align CAS Block and police station boundaries to EA boundaries, it is possible to express crime per 1000 of the population or per 100 000 of the population for comparative purposes.

8. Software Capabilities

8.1. Introduction

This section gives guidelines on how to evaluate the GIS software capabilities for use in SAPS. This framework can also be used for other types of software. This framework enables the user to make objective decisions about several GIS software that are available. Section 8.3 gives an overview of GIS software packages used by other law enforcement agencies over the world. This is included to give SAPS an idea of GIS that are in use at other agencies. The last part of this section deals with data needs that are necessary for SAPS when implementing a GIS solution.

8.2. Criteria table

This criteria table is a framework for software evaluation. These are guidelines only for use in SAPS when implementing a GIS solution. The aim with this framework is to help SAPS make an objective evaluation of GIS software available and to make objective decisions, on which GIS software package or packages to purchase as part of the GIS solution. Table 8.1 givesthe framework in detail. ArcView and MapInfo are only given as examples of software.

Table 8.1 Framework for software evaluation.

FRAMEWORK FOR SOFTWARE EVALUATION

	GIS	
	ARCVIEW	MAP INFO
WHERE DOES IT FIT?		
Was it designed to be a complete KM system or a specialist sub-system?		
Which parts of the Knowledge Management process does it support?		
not support?		
How well do the sub-systems and tools integrate?		
Does it help me capture and store the data?		

Is it feasible for fieldwork?		
Is it PC-based?		
(Some "for what purpose?" questions needed)		
Does it help me structure (organise) the information?		
Does it link disparate data sources?		
(Some "why?" and "for whom?" questions)		
Does it it help me find the information?		
Does it provide remote access to information?		
Does it rely on human intervention?		
(Some "Where can it find information?" questions?)		
Does it help me represent and present the information?		
What does the target audience look like?		
Will the presentation add insight?		
(Some "To whom?" and "How?" questions)		
Does it help me analyse the information?		
(For what purpose?)		
Does it help me make decisions about:		
Effective management of policing		
Responding to crime efficiently / develop good crime prevention practices		
Identifying the offenders		
Other		
Does it help me share the information and can we learn together through collaboration/sharing?		
Does it rely on human assistence?		
What does the target audience look like?		
("Why?" and "With whom?" questions?)		
Does it help me recycle my learning?		
(??)		
WHO WILL OPERATE IT?		
Technician		
Detective		
Other		

WHO WILL USE THE OUTPUT?		
Clerical		
Detective		
Prosecution		
Managerial - operations		
Managerial - middle		
Managerial - executive		
HOW DOES IT WORK?		
How does it help me capture and store the data?		
in a traditional way, through		
presentation tools		
Word processing docs		
spreadsheet info		
Digitiser		
e-mail		
with emerging technologies, through		
speech recognition		
Video or audio production		
optical character recognition		
remote sensing		
World Wide Web		
use of automated agents		
Can it store any information?		
Documents (Formats supported?)		
Multimedia (sound, video?)		
Does it store info in a searchable form (BLOBS, CLOBS,...)		
How well does it deal with dynamic information (such as e-mail, newsgroups, discussion forums, websites, etc)?		
For what purpose must it deal with dynamic information?		
How does it help me structure (organise) the information? (Classification schemes and navigation structures)		
Does it use hierarchies, taxonomies or networks to structure the information?		
Does it support the categorisation of documents according to similarity in content?		
Does it represent the classification scheme visually?		
Does it build electronic linkages from the classification structure to the underlying information?		
Does it enable multiple views of the classification scheme?		
Does the structure include human experts and their areas of expertise?		
Does it identify users and their community affiliations?		
Does it support manual and automatic information structuring?		
Does it use metadata to improve the structure and navigation of the knowledge resources?		
How does it deal with missing values		

How does it help me find the information?		
Do I have to be familiar with the information content?		
Do I have to understand the information structure?		
Does it transparantly link different sources?		
Efficiently by means of indexes etc?		
Can I access text, video, audio, images etc?		
Does it support the following searching functionality:		
boolean searching?		
fuzzy searching (pattern matching?)		
conceptual detection searches?		
contextual detection searches?		
feature detection searches?		
complex natural language queries?		
Does it filter the content?		
Does it allow for collaborative filtering?		
Can it search a thesaurus to find related information?		
Can it browse all information stores, including the knowledge map, thesaurus and threaded conversations?		
Can it do secondary searches against previously retrieved results?		
Can it store and retrieve previously defined searches?		
Does it push relevant information to the user?		
Does it go beyond retrieval to support extraction?		
How does it help me represent and present the information?		
Does it allow visual representation of trends and patterns?		
Does it support animation?		
Does it support 3D?		
Does it graphically illustrate the strength of links between objects?		
Output quality?		
How does it help me analyse the information?		
Does it have knowledge discovery techniques, e.g. data mining, skill mining, text mining?		
Does it have spatial analytical techniques?		
Does it allow downloading of data for user analysis and re-use?		
Does it encode information in a model rather than simply presenting raw data?		
How does it help me make decisions?		
By viewing maps graphs and charts?		
By viewing reports?		
By running a decision model in the background?		
Does it incorporate automatic suggestions to use decision models?		

Does it look at a number of alternatives and automatically select the best one?		
If so, does it evaluate the alternatives against a single objective or multiple objectives?		
Does it evaluate the alternatives and provide a number of outcomes, ranked in some order?		
Does it simulate the outcome of a user-defined scenario?		
Does it help a group of people to reach a decision quicker and/or easier?		
If so, does it incorporate all their inputs and mediate between the links?		
How does it help me share the information and can we learn together by collaboration/sharing?		
Can I send information to others?		
Does it notify users of new information?		
Does it notify users of old information?		
Does it notify users of irrelevant information?		
Does it allow users to share an electronic workspace?		
in real time?		
Does it allow users to do collaborative work?		
Does it allow me to have virtual meetings?		
Does it have chat-room functionalities?		
Does it support the identification of experts?		
How does it help me recycle my knowledge?		
Does it support "nuggetising" - capturing and grouping units of knowledge in shareable repositories?		
Does it provide an integrated repository for presentations, modelling results etc?		
WHAT IS THE TECHNICAL DETAIL?		
To what extent is systems integration supported?		
Links to databases		
Links to other repositories		
Links to investigative tools		
Links to GIS or other GIS's		
Links to other presentation tools		
Links to GPS, cellphones and other peripherals		
Can it be extended?		
Can it be customised?		
Just the user interface?		
Extensions / add-ons?		
What platforms are supported?		
Servers?		
Clients?		
Is it web enabled?		
Web servers?		

Browser clients?		
Cost		
What costing model is used with the software?		
How expensive is expert support?		
HOW DOES IT SATISFY THE MOST IMPORTANT USER NEEDS?		
Does it support the whole process from investigation through to prosecution?		
Does it help "ordinary" investigators become "super" investigators?		
Does it address problems of mistrust between investigators?		
Does it promote coordination/collaboration?		
Can security be maintained using this tool?		
Is there an audit track for the actions happening on the system (for example, can you check who searched through data)?		
Does it empower users to change/adapt the system to suit changed needs?		
Can it interact with legacy information and existing systems?		
Does it allow the user to access and search through a wide variety of data and information sources?		
Does it allow entry, and can it cope with, structured and unstructured data?		
Can it support both a large unstructured information pool and structured data in the form of (for example) a relational database?		
Does it allow the user to manually and/or automatically add structure to unstructured information contained in the system?		
Can it handle disparate data types?		
Can it handle disparate types of cases?		
Can it handle and integrate data and information in different languages?		
Can it handle inadequate data (wrong spelling, inaccurate OCR, etc)?		
Does it allow the investigator to test possible scenario's?		
Is it flexible and "open" enough to grow with the organisation (i.e. not locking in the user to a particular legacy solution)?		
Is it user-friendly?		
Does its "look and feel" correspond to other systems that investigators are used to?		
Can it be transported to, and setup in, a new location with the least effort possible?		
How good is the online help?		
What skills are needed to operate the software?		
Does it have an acceptable "pedigree":		
Was it specifically developed for investigative use?		
Was it designed + structured for the intelligence community?		
Do major investigative agencies use it?		
What is the declared market focus of the tool?		
How good is the local support?		

How good is the technical support?		
Is there training available (novice to expert) for the tool?		

8.3. GIS in use at other law enforcement agencies worldwide

To be able to do crime prevention and analysis effectively certain tools are needed. ArcView and MapInfo are two such GIS tools both with their strong and weak points. It is therefore better to use both tools and get the best of two "worlds". The CSIR use both these tools in doing their crime prevention research and application.

The following is an extract from the Crime Mapping News, which is published in the USA. In 1997 the Crime Mapping Research Centre (CMRC) of the National Institute of Justice completed a study of mapping technologies in use in police departments across the country (USA). The CMRC survey showed that MapInfo and ESRI's ArcView were the mapping software most commonly used in police departments, with MapInfo slightly in the lead (see Mamalian, Loveing, et al. available at *http://www.ojp.usdoj.gov/nij/rsrcdocs.htm* (Nelson L. and Kelly J., 1999, MapInfo or ESRI: The great debate. Crime Mapping News, Volume 1 (Issue 4), 1-8)

"Crime mapping came of age in the Implementation and Vendor period, when computing costs began to fall and soft-ware became more immediately useful. Over the past decade we have also seen more examples of police departments commissioning customized versions of software to meet their individual needs. A survey of police departments conducted in 1997-98 (Mamalian, La Vigne et al., 1999) showed that only 13 percent of 2004 responding departments used computer mapping. ..." Other information gleaned from the computer mapping survey showed that 88 percent of respondents used off-the-shelf GIS software, such as MapInfo® (about 50 percent), ArcView® (about 40 percent), ArcInfo® (about 20 percent), and others (about 25 percent). Some departments used more than one package. Approximately 38 percent of departments that used mapping had done some kind of customizing, and 16 percent were using global positioning system (GPS) technology." (Harries K., 1999, Mapping Crime: Principle and Practice. Chapter 4, 94)

It must be made clear that the CSIR do not favour or prescribe the one tool above the other, but merely wants to assist SAPS in the evaluation of these GIS tools.

Data is the one important aspect of GIS that cannot be left out when investigating the possible use of GIS for crime mapping. The following table shows different levels of GIS data that can be used by SAPS when implementing a GIS solution.

Data needed by SAPS

Macro Data	Micro Data
Topography / Cadastral	
Provincial boundaries	Towns, suburbs
Magisterial boundaries	Erven (if / when available)
Police area boundaries	Police stations
CAS blocks	Emergency services and all other important landmarks
Major towns	
Contour lines (± 100 m interval)	
Transportation	
Major roads	Roads and streets
Major railway lines	Railway lines
Environmental	
National Parks	Suburban greenbelts, parks
Major rivers	Rivers
Major dams	Dams

The following data is readily available

Macro Data	Micro Data
Topography / Cadastral	
Provincial boundaries	Towns, suburbs
Major towns	Police stations
	Emergency services and all other important landmarks
Transportation	
Major roads	Roads and streets
Major railway lines	Railway lines
Environmental	
National Parks	Suburban greenbelts, parks
Major rivers	Rivers
Major dams	Dams

The following data is not so readily available

Macro Data	Micro Data
Topography / Cadastral	
Magisterial boundaries	Erven (if / when available)
Police area boundaries	
CAS blocks	
Contour lines (± 100 m interval)	

References:

Aronoff S., (1989), *Geographic Information Systems: A Management Perspective*. Ottawa: WDL Publications

Burrough P.A., (1986), Principles of Geographical Information Systems for Land Resources Assesment. *Monographs on Soil and Resources Survey No. 12.* Oxford Science Publications. Oxford: Clarendon Press.

Cassettari S., (1993); *Introduction to integrated Geo-information Management.*

Cooper A.K., (1993), Standards for exchanging digita georeferenced information. Unplushed MSc Thesis, University of Pretoria.

Cornelius S. and Heywood I. Module 1. *GIS: An Overview.* Course notes from the Diploma Office, Department of Environmental and Geographical Sciences. The Manchester Metropolitan University. Manchester.

Fisher P.F., (1991), Spatial data sources and data problems. In Maguire D.J., Goodchild M.F. and Rhind D.W., (eds), *Geographical Information Systems Vol. 1.* Longman Scientific and Technical, Essex.

Goodchild M.F. and Kemp K.K., (1990), Technical Issues in GIS. *NCGIA Core Curriculum.* National Centre for Geographic Information and Analysis. University of California, Santa Barbara.

Heywood I., (1995), Module 2. *Spatial Data Modelling.* Course notes from the Diploma Office, Department of Environmental and Geographical Sciences. The Manchester Metropolitan University. Manchester.

João E., Herbert G., Rhind D.W, Openshaw S. and Raper J., (1993), Towards a generalization machine to minimize generalization effects within a GIS. In Mather

P.M., (eds), *Geographical Information Handling - Research and Applications.* John Wiley, Chichester.

MapInfo Corp. (1993), *MapInfo User's Guide*. MapInfo Corporation, Troy, New York.

NCGIA, (1990), *NCGIA Core Curriculum*. National Centre for Geographic Information and Analysis, University of California, Santa Barbara.

Peuquet D.J., (1990), A conceptual framework and comparison of spatial data models. In Peuquet D.J. and Marble D.F. (eds.), *Introductory Readings in Geographic Information Systems*. London: Taylor and Francis

B Plewe, 1997: *GIS Online: information retrieval, mapping and the Internet.* Published by OnWord Press, Santa Fe, USA. SAN 694-0269

Reeves D., (1996), Module 3. *Database Theory.* Course notes from the Diploma Office, Department of Environmental and Geographical Sciences. The Manchester Metropolitan University. Manchester.