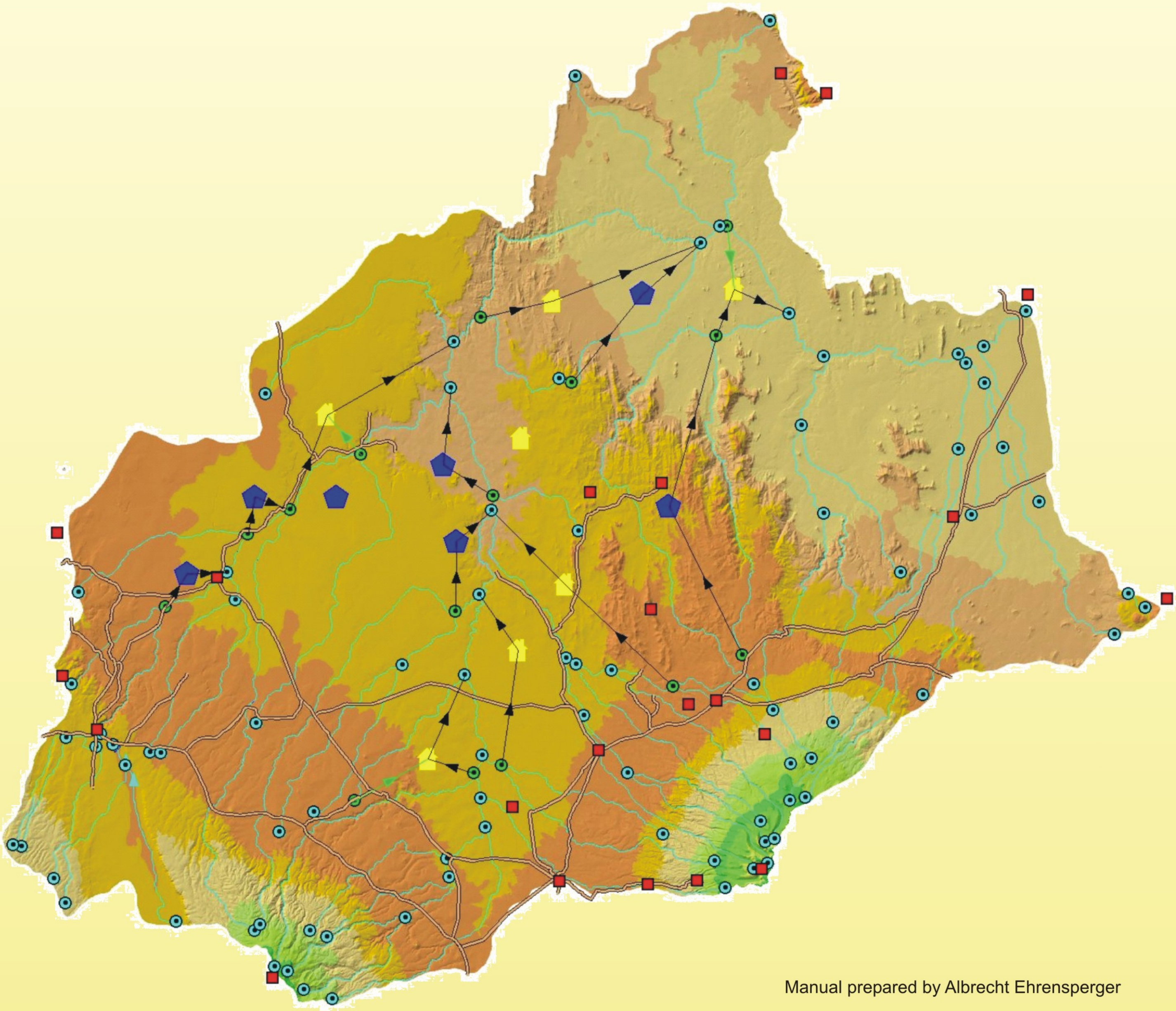
GIS Training Course
November 8th to 26th 2004

Course Manual



Manual prepared by Albrecht Ehrensperger

C E T R A D

Centre for Training and Integrated Research in ASAL Development

P.O.Box 144 - Nanyuki, Kenya
Tel +254 (0)176 31 328
cetrad@africaonline.co.ke

Contents

Figures

CETRAD

1. Foreword

Geographic Information Systems (GIS) are relatively new tools for the generation, manipulation, storage and analysis of spatial information. All the same, GIS is a discipline and a science, which has reached its maturity and significantly contributes to the definition and understanding of problems, the description and analysis of on-going spatial processes, and the formulation of pathways and scenarios for planned development initiatives.

These potentials and prospects are further facilitated by the ever improving accessibility by larger circles of users to increasingly user-friendly GIS software. Some desktop GIS applications are built-up in a way that evokes the same simplicity as the interfaces of conventional office software. And that is where some of the dangers inherent to these recent developments loom: If one knows how to pile up layers in a view document type of ArcView and to print out the result, it does not necessary imply that one has in-depth knowledge about spatial data analysis, that one knows about cartographic rules and conventions, or that one is able to maintain a coherent spatial database. Unfortunately, the consequence of the easy access policy pursued by the producers of most GIS application is the mushrooming of GIS products presenting sometimes very low quality standards.

GIS has become an integral part of the approaches pursued by CETRAD in the frame of most of its activities, which focus on the arid and semi arid lands of Kenya and beyond. In order to set a counterpoint to the above described emerging culture of mediocrity, which, at times, threatens to discredit GIS as a scientific discipline, CETRAD has been paying special attention to the consistence of the quality of its cartographic and other GIS outputs. In the same spirit, the upcoming GIS training course will lay emphasis on the quality of GIS outputs and the necessity for integrative understanding of spatial data analysis, and will therewith enable the participants to distinguish themselves, their institutions and their products from what is being offered in average by other representatives of this discipline.

The concepts, contents and exercises presented in this manual build on previous experiences and consultancies conducted by the author and some of his colleagues of the Centre for Development and Environment (CDE) in Laos, Thailand, Vietnam, Cambodia, Eritrea and Ethiopia. The CDE is a department of the Institute of Geography of the University of Berne in Switzerland and acts as the main counterpart of CETRAD on behalf of the Swiss government.

Albrecht Ehrensperger
GIS expert and technical advisor
CETRAD - Kenya

2. Introduction

2.1. Background

The Centre for Training and Integrated Research in ASAL Development (**CETRAD**) is concerned with the sustainable development in ASAL areas of Eastern Africa in general and Kenya in particular. The objectives of CETRAD are to provide institutional capacity building and human resource development to partner institutions, among others through the transfer of relevant information and knowledge pertaining to enhanced planning and decision making in ASAL areas.

Since several years, Geographic Information Systems (GIS) contribute to fulfil CETRAD's mission which includes the assessment of potentials and utilisation in the ASAL areas, the enhancement of knowledge among planners and development agents, as well as the empowerment of local communities through information dissemination. Hence, GIS plays a key role in a wide range of CETRAD's activities and it is believed that this tool could also contribute in reshaping and enhancing the way other governmental, or private institutions plan and implement their activities. The purpose of the planned *GIS Training Course* is therefore to help such institutions in embracing GIS in an efficient and responsible manner; to show them the potentials and the width of its application possibilities, but also the limitations and the possible pitfalls that go along with its use.

2.2. Aims of the workshop

The *GIS Training Course* will enable the participants to acquire a sound theoretic, but more so practical basis in the use of the most common GIS operations. Particular attention will be given to unequivocal definition of quality standards and to the broadening of each participant's palette of concrete implementation possibilities. The *GIS Training Course* will focus on the following modules:

GIS principles:
Data types and formats; spatial reference and projection; topology issues; overlay analysis; database structure and the need to come up with a meta-database; data safety and user rights.

Pragmatic use of GIS software:
Software palette, strengths and weaknesses; the use of extensions; getting help through internet user forums; software installation procedures and license issues; finding low cost solutions through a pragmatic approach

Data generation:
Digitizing and editing with different GIS software; logical steps and sources of errors; backup procedures and data safety; the use of Global Positioning System (GPS) and the integration of GPS data into a GIS environment

Map layouts:
Cartographic principles; compulsory and additional map elements; dealing with scale; layout extensions in ArcView; excursion into the world of AML

Topographic analysis:
The concepts of Digital Terrain Models (DTM); DTM calculation; calculation of DTM derivatives; excursion into 3-dimensional display possibilities; calculating visible areas; working with topographic classification models

Combining topography and hydrology:
Concepts of sustainable watershed management; hydrologic analysis extensions in ArcView; calculating catchment areas; simulating dam construction; calculating reservoir volumes and estimating water balance

Introduction to satellite image analysis:
The types of sensors and the image formats and resolutions; referencing satellite images; simple image classifications; the NDVI index; calculating forest cover changes on the basis of multitemporal satellite images

2.3. Presentation of the workshop data

A wide range of spatial data from Kenya will be used during the workshop. The course data will be availed to the participants on the first course day and they will be requested to copy this data onto their workstations following a particular data structure.

The data used for the course mainly covers the areas of the Upper Ewaso Ngiro Basin, the municipality of Nakuru, the Mau Escarpment Kenya as a country and Ethiopia. The vector data contained in the course data was generated at the Centre for Training and Integrated Research in ASAL Development (CETRAD), except for the data covering the municipality of Nakuru, which were generated by the Nakuru Local Urban Observatory (LUO) Project and availed to the course for training purposes. The grid data used in the course was entirely generated by CETRAD. The remotely sensed data includes low resolution Landsat Images of different years and different proveniences. The following institutions have availed Landsat Images: CETRAD, the Centre for Development and Environment (CDE) of the University of Berne (Switzerland), the United Nations Environment Programme (UNEP) – Nairobi and the Kenya Forest Working Group (KFWG) – Nairobi. Additionally, a high resolution Quickbird satellite image is used in one of the exercises. This image was availed by the Nakuru LUO Project.

A documentation on the data delivered upon the start of the workshop is included as a text file and a PDF-file on the CD-ROM containing the data (course_data.doc). The data provided follows a coherent naming convention. Participants will be advised to create a 'work' directory on their computer into which they will copy the data layers they are working on. In this work directory naming conventions can be handled individually. However, when archiving a data layer into the established data structure, the naming convention should be scrupulously followed and a comprehensive documentation of the working steps performed should be stored together with the data. Depending on the level of detail, documentation can be stored either in the properties of the theme (short documentation), or in a readme.doc or docu.doc in the same directory as the spatial data set. Please refer to the data_naming.doc file provided on the data CD for more information about both data structuring and naming conventions.

CETRAD

3. Background Information

3.1. What is a GIS?

A geographic information system (GIS) is a computer-based tool for mapping and analyzing things that exist and events that happen on earth. GIS technology integrates common database operations such as query and statistical analysis with the unique visualization and geographic analysis benefits offered by maps. These abilities distinguish GIS from other information systems and make it valuable to a wide range of public and private enterprises for explaining events, predicting outcomes, and planning strategies.

3.2. Components of a GIS

A working GIS integrates five key components: hardware, software, data, people, and methods. Hardware is the computer on which a GIS operates. Today, GIS software runs on a wide range of hardware types, from centralized computer servers to desktop computers used in stand-alone or networked configurations. GIS software provides the functions and tools needed to store, analyze, and display geographic information. Key software components are:

- Tools for the input and manipulation of geographic information
- A database management system (DBMS)
- Tools that support geographic query, analysis, and visualization
- A graphical user interface (GUI) for easy access to tools

Possibly the most important component of a GIS is the data. Geographic data and related tabular data can be collected in-house or purchased from a commercial data provider. A GIS will integrate spatial data with other data resources and can even use a DBMS, used by most organizations to organize and maintain their data, to manage spatial data. GIS technology is of limited value without the people who manage the system and develop plans for applying it to real-world problems. GIS users range from technical specialists, who design and maintain the system to those who use it to help them perform their everyday work. A successful GIS operates according to a well-designed plan and business rules, which are the models and operating practices unique to each organization.

3.3. GIS Tasks

General purpose geographic information systems essentially perform five processes or tasks:

3.3.1. Input

Before geographic data can be used in a GIS, the data must be converted into a suitable digital format. The process of converting data from paper maps into computer files is called digitizing. Modern GIS technology can automate this process fully for large projects using scanning technology; smaller jobs may require some manual digitizing (using a digitizing table, or on-screen).

3.3.2. Management

For small GIS projects it may be sufficient to store geographic information as simple files. However, when data volumes become large and the number of data users becomes more than a few, it is often best to use a database management system (DBMS) to help store, organize, and manage data. A DBMS is nothing more than computer software for managing a database. There are many different designs of DBMS's, but in GIS the relational design has been the most useful. In the relational design, data are stored conceptually as a collection of tables. Common fields in different tables are used to link them together. This surprisingly simple design has been so widely used primarily because of its flexibility and very wide deployment in applications both within and without GIS.

3.3.3. Manipulation

It is likely that data types required for a particular GIS project will need to be transformed or manipulated in some way to make them compatible with your system. For example, geographic information is available at different scales. Before this information can be integrated, it must be transformed to the same scale (degree of detail or accuracy). This could be a temporary transformation for display purposes or a permanent one required for analysis. GIS technology offers many tools for manipulating spatial data and for weeding out unnecessary data.

3.3.4. Query and Analysis

GIS provides both simple point-and-click query capabilities and sophisticated analysis tools to provide timely information to managers and analysts alike. GIS technology really comes into its own when used to analyze geographic data to look for patterns and trends and to undertake "what if" scenarios. Modern GIS's have many powerful analytical tools, but two are especially important:

Proximity Analysis
How many houses lie within 100 m of this water main? To answer such a question, GIS technology uses a process called buffering to determine the proximity relationship between features.

Overlay Analysis
The integration of different data layers involves a process called overlay. At its simplest, this could be a visual operation, but analytical operations require one or more data layers to be joined physically. This overlay, or spatial join, can integrate data on soils, slope, and vegetation, or land ownership with tax assessment.

3.3.5. Visualization

For many types of geographic operation the end result is best visualized as a map or graph. Maps are very efficient at storing and communicating geographic information. While cartographers have created maps for millennia, GIS provides new and exciting tools to extend the art and science of cartography. Map displays can be integrated with reports, three-dimensional views, photographic images, and other output such as multimedia.

3.4. GIS Data Representation

A GIS stores information about the world as a collection of thematic layers that can be linked together by geography. This simple but extremely powerful and versatile concept has proven invaluable for solving many real-world problems.

Spatial data is organized thematically into different layers, or themes. There is one theme for each set of geographic features or phenomena for which information will be recorded. For example, streams, landuse, elevation, and buildings will each be stored as a separate spatial data sources, rather than trying to store them all together in one. This makes it easier to manage and manipulate the data, especially as much of the power of working geographically comes from being able to analyze the spatial relationships between different geographic themes.

Note: The Geodatabase data model of ArcGIS works in a different manner, in that it tries to integrate the different layers of a database in order to allow for permanent relational and spatial analysis between these layers. A short introduction into the Geodatabase model will be given during the workshop.

GIS Theme
A GIS links sets of features and their attributes and manages them together in units called themes. A theme is a collection of geographic features, such as cities, roads, rivers, parcels or soil classes, together with the attributes for those features. GIS work with two fundamentally different types of geographic models: The "vector" model and the "raster" model.

Generally, geographic data sources are **represented** in two ways: vector and raster.

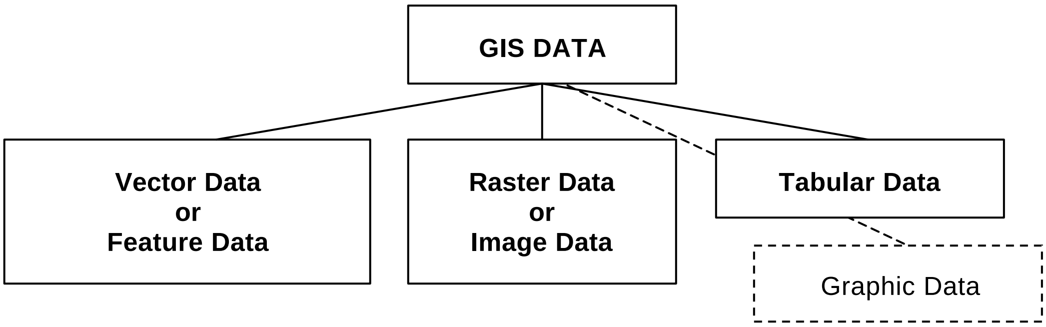


Figure 3.1 *Models of GIS data representation*

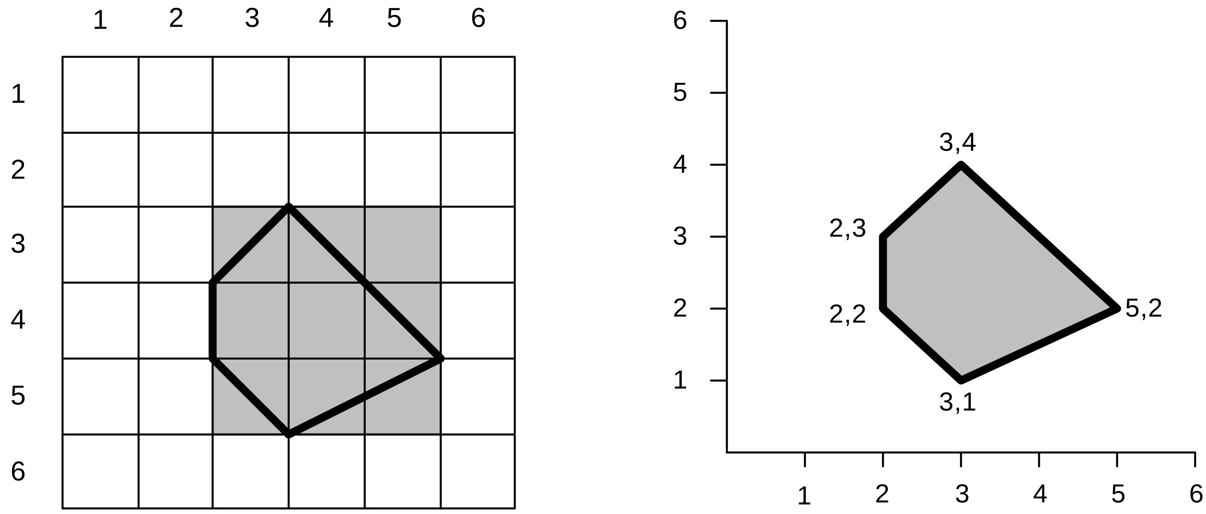
A **vector representation** organizes geographic information using the Cartesian coordinate system. Information about points, lines, and polygons is encoded and stored as a collection of x,y coordinates. Points are stored as single pairs of XY coordinates, whereas lines are stored as a collection of point coordinates. For example, a line representing a road may be described as a series of X and Y coordinate points: one point for the start of the line, one point for the end of the line, and as many points in between as are required to define the line's shape. Polygonal features, such as river catchments, can be stored as a closed loop of coordinates. Vector representations are often used for representing data with exactly known locations, such as streets, light poles, or the legal boundaries of lots.

Vector data is designed to enable specific geographic features and phenomena to be managed, manipulated and analyzed easily and flexibly to meet a wide range of needs. In vector data or feature data there is an **explicit relationship between the geometric and attribute information**, so that both are always available when you work with the data. For example, if you select particular features displayed on a view, ArcView will automatically highlight the records containing the attributes of these features when the attribute table is displayed.

A **raster representation** organizes geographic data using cells arranged in rows and columns. Each cell has a row number and a column number, with the cell in the upper left identified as row 1, column 1. The exact location of each cell is not stored, just the origin, cell size, and number of cells from the origin. Values associated with each cell describe the geographic attributes in the region of space covered by the cell. Each cell stores a numeric value representing a geographic feature. Raster representations are often used for geographic data with less discrete locational boundaries. This is often the case with environmental data, such as soil type polygons of forest boundaries, where the mapped features may not have sharply definable boundaries. The raster model has evolved to model such continuous features. A raster image comprises a collection of grid cells rather like a scanned map or picture. An image is the simplest form of raster; it stores a single value for each location. A grid is a special type of raster where the value stored is a record in a table that stores additional descriptive information for the cells. Raster data (or image data) can come from photographs, remotely sensed data, scanned data, satellite data, and graphics.

Both the **vector and raster** models for storing geographic data have unique advantages and disadvantages. Modern GISs are able to handle both models.

A **tabular representation** organizes geographic data using a table. A tabular representation is mostly used together with a raster or vector representation. Tabular data can include almost any data set, whether or not it contains geographic data.



Raster Data Representation

Vector Data Representation

Figure 3.2 *Vector data representation model and raster data representation model*

3.5. GIS Data Formats usable in ArcView

In addition to representation, the user of geographic data must also be concerned with **format**. For example, TIFF Image File or GRID are different file formats for data represented in a raster form. ArcView supports a variety of data representations and formats. These are discussed below.

The **ArcView Shapefile** format is used for storing vector data. Shapefiles are non topological, which means that limited information is kept in the shapefile regarding the relationships of features to one another, such as what specific line shapes are used to define a specific polygon shape. The lack of topology means that shapefiles are less appropriate for sophisticated spatial analysis than other formats. However, it also provides some advantages, such as improved drawing times. The shapefile specification is openly published; therefore, shapefiles can be created by anyone. The shapefile format consists of at least three elements: shape, index, and attribute. Each of these elements is stored as a separate file on disk; therefore, a shapefile actually consists of three or more disk files (one for each element). Other, optional files can be included by a shapefile, such as the projection file (.prj) or the legend file (.avl).

The **shape element** is the portion of the shapefile describing the geometric shapes that represent the geographic features. These shapes are described by their X and Y coordinate locations. The disk file containing the shape element has a file name extension of .shp. The **index element** of the shapefile provides an optimized means of accessing the geometric shapes described in the shape element. The element contains a sequential index of offsets into the shape data. This spatial indexing, as it is called, provides for faster drawing times and faster queries of geographic features represented by geometric shapes. The disk file containing the index element has a file name extension of .shx. The **attribute element** contains tabular data associated with geographic features. For example a road may have associated information regarding its condition, the year it was constructed, and its width. This information can be stored in the attribute element of the shapefile and associated by the key value with the specific road shape to which it pertains. The attribute element of the shapefile is stored as a standard dBASE file with one record per shape. The disk file containing the attribute element has a file name extension of .dbf. The basic elements of a shapefile are shown in the following illustration.

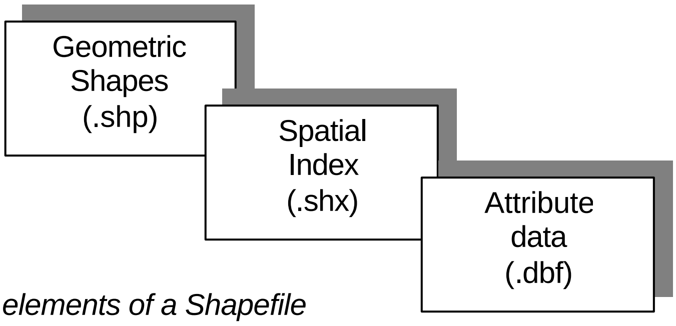


Figure 3.3 *The basic elements of a Shapefile*

<u>**ArcInfo coverages**</u> are another format for storing vector data. There are several differences between coverages and shapefiles: In contrast to shapefiles, coverages are a topological data structure. This means that the format is much more sophisticated in its ability to track the relationship between features, such as what specific line shapes are used to define a specific polygon shape. Because of their more sophisticated data structure and the inclusion of topological information, coverages are better suited for larger data sets and for applications requiring complex spatial analysis. Coverages are represented by subdirectories existing within an ArcInfo workspace. The workspace is a directory that serves as a work area and storage area for the ArcInfo coverages. The coverage subdirectories each represent a single coverage and are named with the coverage names. These subdirectories contain the geographic data stored in file names such as TIC, BND, and ARC. This data can be created and maintained using the ARC portion of ArcInfo.

A separate subdirectory within the workspace contains the attribute data associated with the geographic features of all coverages stored in that workspace. This single subdirectory, named **info**, stores the attributes in database tables that can be accessed using the INFO portion of ArcInfo. The following illustration shows an example of an ArcInfo storage structure.

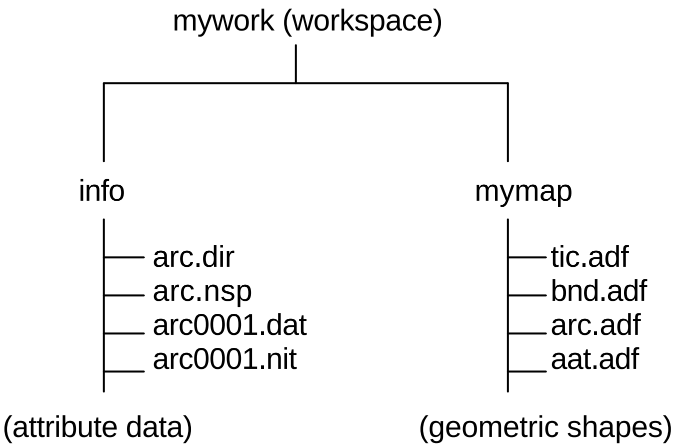


Figure 3.4 *The ArcInfo data storage structure*

Important Note: In a workspace data storage structure you should never move, rename or delete themes with the windows explorer or an other file operating tool. To operate with coverages or grids you should only use **ArcInfo** prompt commands, the **Arc Catalog** or the **Spatial Analyst** extension of ArcView.

CAD drawings are a format for storing vector data. They are produced by computer-aided design applications and can be used in ArcView if ArcView's CAD Reader extension is loaded. ArcView supports MicroStation design files (*dgn* files), and two kinds of AutoCAD drawing files (*.dwg* files), and Drawing Interchange files (*.dxf* files). All this CAD formats can be converted to the ArcView Shapefile format.

ArcInfo Grids are a format for storing raster data. Grids are especially suited to representing geographic phenomena that vary continuously over space, and for performing spatial modelling and analysis of flows, trends, and surfaces such as hydrology. Grid themes use a matrix of cells to represent geographic features or phenomena. The size of the cells used is important to any analysis. The cell size is the smallest unit you are interested in mapping, and defines the limit of your spatial accuracy. The smaller the cell size the more accurate the representation. This is at the expense of larger data sets and slower processing time. Feature and grid themes have some similarities, but they are different in the way they model or represent spatial data. Grid themes are always stored in the ArcInfo data storage structure. That means grid themes are always stored in a workspace, in which you should never move, rename or delete files with the Windows Explorer. An (integer) grid has an associated value attribute table or VAT. In the VAT you can store only one attribute and not more. Never edit the attribute table (VAT) in ArcView.

Grids are stored either as **integer or floating-point** data. An integer grid has an associated value attribute table or VAT. This table stores a single record for each unique value in the grid, as well as the number of cells of that value, and one additional attribute you wish to attach. A floating-point grid has no VAT. Generally speaking, integer grids will be smaller and faster to operate on than floating-point grids.

Every **image data** format is a format for storing raster data. Image data can be used as background for feature based themes. Features that appear on an image can be digitized to create a new vector database. Image data can be organized in a number of ways depending upon the particular image format. Typically, the image data file contains a header record that stores information about the image such as the number of rows and columns in the image, the number of bits per pixel, the color requirements and the georeferencing information. Following the image header is the actual pixel data for the image. The internal organization of the image data is dependent upon the image format. Some formats contain only a single band of data, while others contain multiple bands. ArcView can display and print black and white, greyscale, pseudocolor and true color images. The most important image formats you can use in ArcView are *TIFF, ERDAS, JPEG*.

3.6. Spatial Data

There are major differences between spatial data and data such as vector graphics, and non-georeferenced images such as scanned paper documents like photos: In spatial data there is an explicit **relation between the geometric and attribute information**, so that both are always available when working with the data. Spatial data is **geo-referenced to known locations on the Earth's surface**. To ensure that location is accurately recorded, spatial data always employs a specific coordinate system, unit of measurement and map projection. When spatial data is displayed it has a particular scale just like any paper map. Spatial data is **primarily feature based**. It is designed to enable specific geographic features and phenomena to be managed, manipulated and analyzed easily and flexibly to meet a wide range of needs. Other types of graphic data may be oriented solely towards presentation and display, and may store features such that they can only be accessed in a limited number of ways. Spatial data is **organized thematically** into different layers, or themes. There is one theme for each set of geographic features or phenomena for which information will be recorded.

3.7. Feature Classes and Attributes

Objects represented on maps, whether natural or man-made, are called **map features**, or simply features. Each feature has a location, a representative shape, and a symbol that represents one or more of its characteristics. A GIS stores information about map features in a database and links the information to map features. This information is referred to as **attribute information**, or simply attributes. Information about the features on a map are stored in rows and columns. Each row relates to a single feature, each column contains the values for a single characteristic.

The following feature classes can be differentiated in the ArcInfo data models: **Points** represent features found at discrete locations, such as telephone poles, wells and mountain peaks. **Arcs or lines** represent linear features such as streams, streets and contours. **Polygons** represent areas enclosed by specific boundaries, such as countries, states, land parcels and soil types. **Label points** are points inside polygons that have the same attributes as the polygons. One way to use label points is to symbolize polygons using different point symbols according to their attributes. Unlike coverages, polygons within shapefiles do not have label points. **Nodes** are the endpoints of arcs making up arc features and polygon features. A node occurs at the point where two or more arcs meet and at each end of a dangling (unconnected) arc. **Annotation** is text stored in a coverage to label the features it contains. Annotation stores text strings, text characteristics and shape points used to position each text string.

Exercise 3.1 – Getting familiar with ArcView

Steps	Data	Remarks	Time
Loading extensions	Spatial.avx	- Go to the EXT32 folder and make sure that the spatial analyst extension is available in this folder (spatial.avx). - Start ArcView and load the Spatial Analyst extension	10 min
Loading data	Keewlco1 Keewriv1 Keewtow1	- Open a view document type - Find the layers in the data structure and load them into the view (use data_naming.doc for orientation). - Be aware that data can be of various types!	15 min
Getting familiar with the view		- Try out the zooming, panning, full extent and active layer extent tools to move around in the view. - Get additional information on the towns by using the identify tool (the layer has to be active!)	15 min
Changing symbology of layers	Keewlco1 Keewriv1 Keewtow1 Land_code.doc	- By double-clicking on the legend editor of each layer change their symbology in a way that suits you. For the landcover grid use the land_code.doc file for orientation.	20 min
Using a brightness theme	Keewhil1	- Load the hillshade grid into the view. - Enter the hillshade grid as a brightness theme in the advanced option of the legend editor of the landcover grid - Observe the colour shifts that may have appeared as a result of using a brightness theme. - Try out different minimum and maximum cell brightness options and see what happens	20 min
Naming the view		- In the View – Properties menu change the name of the view to “Ewaso” - Observe the changes on the view window header - Find the “ewaso” view in the project window - Observe the other document types listed in the project window and learn about their functions	10 min
Saving the project	Exercise3-1.apr	- Save the project in auxil\apr as “exercise3-1.apr” - Close ArcView and look for the APR file with the explorer, double click on it and observe what happens - Draw conclusions as to how data is saved in ArcView	20 min
		Total time:	110 min

Exercise 3.2 – Querying information from layers

Steps	Data	Remarks	Time
Loading data	Nk_schools Nk_dtm1 Nk_hil1	- Start ArcView - Load the data into a view and display it - Change the legend of the DTM to be the beige to brown colour ramp and use the hillshade as a brightness theme with 20 – 100 for minimum and maximum brightness - Check what happens when you invert the colour ramp	10 min
Querying information on raster data	Nk_dtm1 Nk_hil1	- Using the query builder find out the areas which are above 2000 meters above sea level, then find the ones that are below 1800 meters above sea level. Finally make a request that will identify areas between 1800 and 2000 meters above sea level. -	10 min
Querying information on vector data	Nk_schools	- Change the legend of the school layer in a way that schools funded by the MCN appear in one colour and all other schools in another colour (use the “unique value” legend type to do this) - Query the layer to find primary schools funded by the MCN (disregard those which are mixed primary with other levels)	15 min
Viewing selected features in the attribute table	Nk_schools	- Open the attribute table and promote the selected features to the top of the table. - Find out how many schools match the query criteria entered in the previous step. - In the table launch a new query to find which ones among the selected schools have boarding facilities. - Switch back to the view to find out the locations of the two schools matching all the query criteria	10 min
Measuring distance	Nk_schools Nk_roads	- In the view – properties menu set the map and distance units to meters - Find out the approximate distance in meters between the two selected schools (in a straight line) - Load the road layer and display it - Measure the approximate distance between both schools when following the shortest path along the roads	10 min
Labelling features	Nk_schools	- Zoom in such that the two selected schools lie in opposite corners of the view - In the Window – show symbol window menu set the font to Arial 10. - Label the features with the school names by using the theme – auto label menu - Check which features have been labelled - Remove the labels with the theme – remove labels menu - Unselect all features and run the labelling operation again - Check which features have been labelled.	15 min
Saving project		- Save the ArcView project in the auxil\apr folder under ewaso and call it exercise3-2.apr	
		Total time:	70 min

4. Spatial analysis concepts

4.1. Compilation of spatial analysis concepts

This section offers a recapitulation for further use of the main spatial analysis tools of ArcView. Generally speaking, there are six main types of spatial analysis, each requiring different tools or a combination of tools. These six types of analysis are presented within the sections 4.1.1 to 4.1.6.

4.1.1. Polygon – Point Analysis

This type of analysis is frequently used, e.g. to find out the number of villages in each catchment of a specific area, or in each administrative unit of a province. Such questions are solved in ArcView by using the **Spatial Join** option of the Geo-Processing Wizard. This tool allows three analysis options: 'nearest', 'inside' and 'part of'. In ArcGIS you will find the spatial join in the layer context menu .. Joins and Relates .. Join. In this dialog use the “Join data from another layer based on spatial location”-option in the first selection box.

Nearest
If you're assigning data from a point theme to another point theme, or if you are assigning data from a point theme to a line theme, a 'Distance' field is automatically added to the theme you are assigning data to, along with any other data in that theme. This 'Distance' field contains the distance to the nearest feature. For example, if you assign data from a theme containing locations of bus stations in Nairobi to a theme containing locations of schools in the same town, a ‘distance’ field would be added to the theme containing the location of schools. This field would contain the distance of the nearest bus station to each school. If you perform the reverse operation, the ‘distance’ field will be added to the bus station theme and will contain the distance of the nearest school to each bus station.

Inside
If you are assigning data from a polygon theme to a point, line, or polygon theme, the data will be joined to the point, line, or polygon that is contained by each of the polygons. For example, the Ministry of Public Health could join a polygon theme containing demographic data (including health parameters, age distribution, etc.) to a point theme showing the location of hospitals. The demographic data could thus be used for population profiling, in order to adapt the facilities and services of each hospital.

Part of
If you are assigning data from a line theme to another line theme, data will be assigned from lines that are 'part of' (a sub-set of) the lines you're assigning data to. For example, a transportation engineer might have a theme of streets under construction and another theme of all the city streets. To help prioritise which streets to fix first, the engineer might want to join the theme of streets under construction to the theme of all city streets. Assign data by location will join the construction data for the streets that were part of the city streets theme. Any streets under construction that are not part of the city street network will not be joined.

4.1.2. Polygon – Grid Analysis

Grids are a raster data format specific to the ArcGIS software but equally usable in ArcView. In order to carry out analysis on Grid data, the ArcView extension ‘Spatial Analyst’ needs to be installed and made active. Once the Spatial Analyst extension is loaded, a new menu, called Analysis, is added to the ArcView toolbar (in ArcGIS the Spatial Analyst toolbar is activated and the Spatial Analyst Tools become available in ArcToolbox). Polygon – Grid Analysis is implemented with the Tabulate Area option in the Analysis menu. In ArcGIS the Tabulate Area tool is located in the “Zonal” toolset of the Spatial Analyst Toolbox.

Tabulate Area performs a cross tabulation of the zones between two input themes. The zones in the **row theme** create the rows of the resulting table and the zones in the **column theme** produce the columns. The values in the resulting table identify the area of each zone in the column theme encompassed within each zone in the row theme. You can use it, for example, to summarise the area of each land use type within each district.

In the **dialogue box**, the row theme defines the rows in the output table. There will be a record for each unique zone in the row theme. Each record contains the area of each zone in the column theme within that zone. The row theme can be a point, line, polygon, or grid theme. The **row field** is the field used to define zones in the row theme. The column theme defines the columns in the output table. There will be a field for each unique zone in the column theme. Each field will hold the area for that zone within the zones of the row theme. The **column field** is the field used to define the zones in the column theme.

Notes: Very often before carrying out the tabulate area analysis, the grid data has to be re-classified. The analysis of unclassified grids representing a continuous surface (e.g. DTM, slope.) does not make a lot of sense. Reclassify will always create an integer grid theme.

Equal Area:	Reclassifies so there is an equal number of cells in each output class.
Equal Interval:	Divides the range of cell values into equal sized sub-ranges.
Natural Breaks:	Identifies breakpoints between classes using a statistical formula (Jenk's optimisation), which minimises the sum of the variance within each of the classes. Natural Breaks finds groupings and patterns inherent in your data.
Quantile:	Reclassifies so there is an equal number of features in each of the output classes. For a grid theme this is exactly the same as Equal Area.
Standard Deviation:	Finds the mean cell value and then places class breaks above and below the mean at intervals of either 1/4, 1/2, or 1 standard deviations until all the cell values are contained within the classes. Any cell values that are beyond three standard deviations from the mean are put into two classes, greater than three standard deviations above the mean ("> 3 Std Dev.") and less than three standard deviations below the mean ("< -3 Std. Dev.").

Only the Equal Interval and Standard Deviation methods are available for floating point grids.

4.1.3. Polygon – Line Analysis

This type of analysis is less frequently used, but in some cases it may be very useful. It is used, for example, to find out how many kilometres of which road class there is in each district of Kenya. To carry out such a calculation, use the **Intersect** tool of the Geo-Processing Wizard (the Analysis Toolbox in ArcGIS). The Intersect process computes the geometric intersection of two themes and adds it to the view as a new theme named *itsct.shp* (by default). The input theme can be a line or polygon theme. The overlay theme must be a polygon theme. The overlay theme's features will split the input theme. Any features in the input theme that are not overlaid by features in the over-lay theme will not be added to the new theme. The output shapefile's features will be of the same type as the input theme features. When intersecting features of the input theme with the polygons of the overlay theme, the attribute tables are also updated. The attribute table for the output shape-file includes the attributes from the input and overlay theme.

4.1.4. Polygon – Polygon Analysis

Polygon – polygon analysis is best conducted using the **Union or Intersect** options of the Geo-Processing Wizard (Analysis Toolbox in ArcGIS). If one of the polygon layers has a smaller extent than the other one and you do not need the surrounding information contained by one of the layers only, choose intersect as the processing time will be shorter. Use the **Union** process when you want to produce a new theme containing the features and attributes of two polygon themes, for example, to find out how crops are distributed on different soil types. The analysis of the spatial distribution of agricultural crops in relation to different soil types might give indication on optimisation possibilities regarding best possible use of available soil fertility.

Attention: Before using Union or Intersect you should:

- Add a new field to your polygon layers which is clearly defining as the area of your respective polygon shapefiles. For example, if you wish to analyse the soil distribution per district add a field area-di to the district layer and a field area-so to the soil layer.
- In each table highlight the newly added area item and enter the following Avenue request in the calculate window:

  [shape].returnarea

- In order to calculate the area **in ArcGIS**, check the Advanced option in the field calculator and add the following statements to the Pre-logic VBA-script text box:

  Dim dblArea as double
  Dim pArea as Iarea
  Set pArea = [shape]
  dblArea = pArea.area

  In the text box directly underneath type **dblArea.**

The aim of this manipulation is that after performing union or intersect there will be two items called Area, and you may get confused regarding which area belongs to which original layer. ArcView has a **major bug** concerning the union and intersect tool: It does not automatically calculate the area of the new polygons resulting out of union or intersect. However the Area item from the original polygon layers (e.g. soil types and crops) are being added to the attribute table of the intersected theme, which could lead to **serious misinterpretation**. Hence, the area of the newly intersected polygons has to be calculated manually immediately after intersecting, or unioning. Use the same Avenue or VBA-script syntax as above to calculate the area of the new polygons.

Buffering

Very often one is interested in areas that are within a certain distance of an object or a feature. For example, it is of primary interest to identify areas within a certain distance from rivers, whenever planning water distribution, or irrigation schemes. When buffering a feature (point or line) the output will always be a polygon layer covering the area within the specified buffer distance.

Notes:
- If the **Create Buffers** option is disabled in the View menu, it is most probably because the map and distance units have not been set in View - Properties.
- Buffering large or complex shapes may require more system resources than is minimally recommended for running ArcView. If you find that your system runs out of memory, try increasing the amount of virtual memory, or divide the buffering job into smaller pieces by selecting subsets of features to be buffered. The merge option in the GeoProcessing Wizard will help you to reassemble the various buffers into a single result.

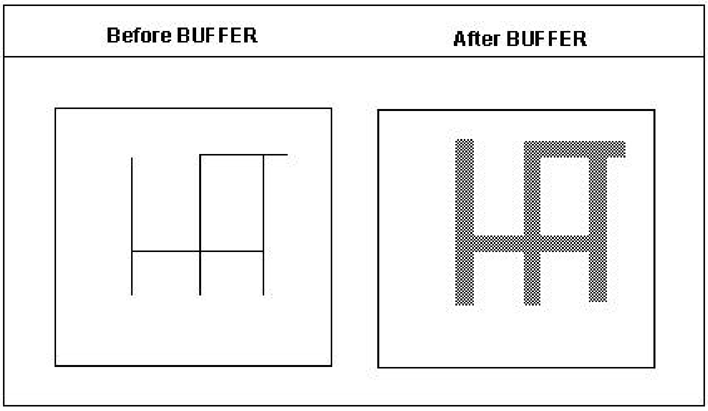


Figure 4.1: *buffering with a fixed distance*

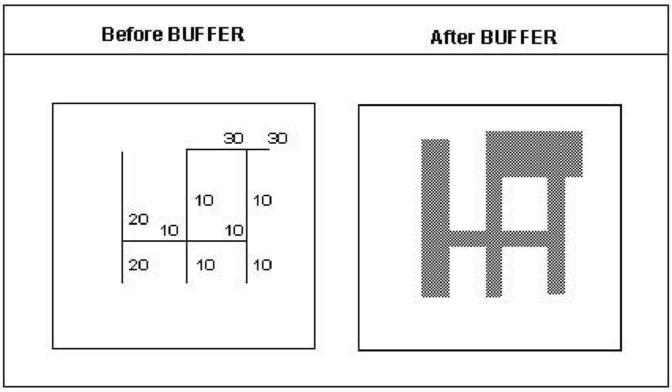


Figure 4.2: buffering with a distance defined by an attribute field

4.1.5. Point – Grid Analysis

The point - grid analysis is used if you want to get information joined from a grid to a point layer, for example, if you want to know the elevation of all towns in Kenya. The default tools of ArcView can not manage such a task and thus, an additional extension, which can be downloaded from the internet at www.esri.com and **is called getgridvalue.avx** will be necessary. In ArcGIS use the "Extract values to points"-tool in the Extraction toolset of the Spatial Analyst toolbox.

Note: To be able to activate your extension copy the getgridvalue.avx to your **EXT32 directory**. To make extensions active in an ArcView project select, in the **File** menu, the **Extension** field and then tick the extensions to be activated (in this case "getgridvalue"). If you want these extensions to be active upon each opening of an existing project or starting a new project tick the "**make default**" field.

Use the following steps to carry out point – grid analysis:

1. Load your point and grid data into a view.
2. Loading the getgridvalue.avx extension adds a new button to the Graphic User Interface :
3. After pressing this button you will be asked to enter the point and grid layers and to specify the name and definition of the item in the point layer into which the information of the grid shall be written.
4. Check in the attribute table of the point layer whether the desired information was correctly added.

4.1.6. Grid – Grid Analysis

Grid is a **raster- or cell-based geo-processing system** integrated with ArcInfo Workstation / ArcGIS. Like all modules of ArcInfo Workstation, Grid uses a geo-relational model for geographic data. The model used in Grid, however, differs from the other modules because it is based on a combined raster-based (grid-cell) spatial model and a relational attribute model. In Grid, the inherent power of the grid-modelling structure is coupled with the capabilities of a relational database that manages all attributes associated with the cell values.

If two or more input grids in an expression are at different resolutions, GRID automatically resamples them (using the nearest-neighbour resampling technique) to the coarsest resolution unless specified otherwise before processing. Processing grids at different resolutions causes no reduction in speed.

The two greatest advantages of a grid-cell-based structure for representing geographic space over other GIS structures are:

- The ability to represent continuous surfaces
- The ability to store points, lines, polygons and surfaces uniformly

A continuous surface is best represented by a cell-based system since the attributes of a location in the cell-based system are a function of where the location is relative to a phenomenon that is progressively changing as it moves across space. The problems that the locational structure and processing mechanism resolve best can be divided into four main types:

- The attributes of a location are influenced by the attributes of the locations that surround it, e.g. the pH of a particular location can vary depending on the pH of the surrounding areas.
- The attributes of a location are influenced by where its position falls within another feature, e.g. the visual preference for a particular location can be influenced by the diversity of positive or negative visual effects within the area.
- Cell-based systems can be used for continuous surface-data modelling.
- Optimal allocation and surface determinations that can only be calculated when a surface is divided into discrete units are well-suited for cell-based modelling

Exercise 4.1 – Spatial Analysis on Vector Basis (carry out exercise 5.1. prior to this exercise)

Steps	Data	Remarks	Time
Deleting layer	Keewwsh1	- In exercise 5.1 we saw that transformation with the ArcView Projection Utility gave slightly different results. For this reason, we delete the layer keewwsh1 using the File – manage data source menu. Make sure that this layer is not currently opened in ArcView before going ahead.	10 min
Geoprocessing wizard	Keewwsh2 Kenards1north Kenards1south	- Load the three layers and display them - The roads of Kenya have been availed in two separate layers for areas north and south of the Equator. Our area of interest (Ewaso basin) is located at the junction of both areas. Use the View – Geoprocessing menu and the **merge** option to bring both shapefiles together. Save the result as Kenya\gis\infra\kenards1 - As we are interested in the roads within the basin area, we will now clip the merged shapefile with the watershed boundary. Do this with the Geoprocessing wizard and the **clip** option. Save the result as ewaso\gis\infra\keewrds1.	30 min
Buffering rivers and roads	Keewrds1 Keewriv1	- In order to carry out a field survey on river water, we are interested in finding out those areas located at the same time within less than one km from any river and within less than 1 km from any road. - Load the buffer wizard extension - **Buffer** both roads and rivers with one km distance and save the results as keewriv1buf and keewrds1buf in their respective directories. Make sure you have set the units of the view in the view – properties menu before carrying out the buffering with the theme – create buffer menu. - **Intersect** both results in order to get the areas which are at the same time within 1 km from any river and within 1 km from any road. Save the result as ewaso\gis\mapping\keewriv-rds	30 min
Dissolving and unioning	Keewlus1 Keewriv-rds	- We wish to find out how many square kilometres of the areas identified in the intersect step above are located within large-scale ranches. - Load the land use layer and dissolve the polygons according to the lcode (land use code). Dissolving is done with use of the geoprocessing wizard. Save the resulting layer as ewaso\gis\land\keewlus2 - Union the dissolved land use with the keewriv-rds. Save the result as ewaso\gis\mapping\keewunion1 - Carry out a query to find out the areas within the buffer zones (bufferdis = 1000) and located in large-scale farms (use the land_use_code.doc) file to help you finding out which code corresponds to this type of land use). - Go to the attribute table and promote the selected polygons to the top of the table. - Start editing the table (Table – start editing) and add a new field (edit – add field) which you call area (width = 8, decimals = 2). Click on the header of the new field in order to select it and go to the field – calculate menu. In the window write the syntax [shape].returnarea/1000000. This will return the area in km^2 for the selected polygons. - Go to field – statistics and under SUM check at the number of square km^2 matching the query.	30 min
		Total time:	100 min

CETRAD

5. Mapping, map projections and map layouts

5.1. Short introduction on maps

5.1.1. Map attributes

Basically, maps have three attributes: Scale, projection and symbolisation. (Monmonier, 1996).

Scale: provides indication about the ratio between a distance on the paper and the same distance in reality. A scale of 1:50,000 indicates that 1 unit on the map is equal to 50,000 units in reality, e.g. one centimetre on the map is equal to 50,000 centimetres, or 500 metres in reality.

Projection transforms the curved, three-dimensional surface of our planet into a flat, two-dimensional plane. The appearance of a large scale map (e.g. 1:10,000 or 1:50,000), which covers a very small portion of the globe, as for example a small catchment, a district, or a village area, will not be heavily affected by the type of projection chosen to map it. Visible differences arise on small scale maps (e.g. 1:1,000,000) representing larger areas, as for example countries, continent, or the world. Refer to the next section for more information on map projections and on how to adapt them in ArcView.

The **symbolisation**, finally complements the map scale and projection by making visible the features, places, and other locational information represented on the map (Monmonier, 1996). Mostly the symbolisation underlies national or international convention. One of the conventions used in cartography is for example to set the azimuth of the sun to 315 degrees, i.e. to north-western direction. The azimuth determines which slopes will be exposed to the light of the virtual sun and which ones will be lying in the shadow. With an azimuth of 315 degrees, all slopes oriented into north-western direction will be exposed to the sun light and all slopes facing south-east will be in the shadow. The visual effect of relief is created, in a hillshade map, by shading the surface of the map according to the exposure to direct sun light. Our eyes have become so much used to this convention, that when a map is held upside-down, one has a tendency to perceive valleys as being mountains and vice-versa.

5.1.2. Functions of maps

Maps are used for various purposes. “The most important function of maps is probably the function of orientation or navigation. In any case, most of the maps the general public comes across, with the exception of weather charts, are produced as an aid to orientation and navigation. (...) Physical planning maps are maps that inventorise the present situation, maps that define development processes, and maps that contain propositions for a future situation, e.g. future land use. (...) Maps used for management tasks or monitoring purposes are generally large-scale maps that are manufactured bearing in mind the management and maintenance of objects, e.g. roads, railways, forests, dikes, canals and airports. (...) For educational objectives, special purpose map material has been produced since around 1750; school atlases, wall maps and work books provide the pupils with a spatial frame of reference in order to be able to understand national and world-wide developments. Another map function is codification, e.g. showing the legal situation as regards property rights.” (Kraak, M.J. and F.J. Ormeling, 1998, 53).

5.1.3. Some mapping rules

Disregarding the function of the map you are producing, whether it is orientation, planning, education, or codification, some map elements should not be absent from your map layout. Hereafter is a (non-exhaustive) list of these elements:

Map field: This is the area of your map layout, in which the actual map is positioned. Depending on the theme and on the function of the map, different elements will be included in the map field. When working with ArcView, these elements are prepared in the view document type and added to the view frame in the layout document type. Pay attention to choose clearly readable symbols, colours, etc. for the different elements. Some conventions are widely used and it is advisable utilising them as well (see paragraph on map symbolisation above).

Map frame: The frame around the map field is commonly used to indicate at regular intervals the co-ordinates of the portion of space represented in the map field. The interval between displayed co-ordinates should be chosen depending on the scale of the map.

Map margin: The map margin contains all additional and relevant information concerning the map. Among this additional information the following should not be left out ("**compulsory elements**"):

1. **Map title** clearly summarising the contents and the intention of the map
2. **Map legend** giving unequivocal information on all punctual, linear and spatial map elements. It is advisable to separate the key for linear and punctual information from the key for spatial information showing hatches or shades.
3. **Scale bar and map scale** in numbers. The scale of the map could actually be derived from the co-ordinate grid, but the indications provided with the scale bar can be considered as compulsory. If you intend to photocopy, or scan your map to use it at a different scale, there is no point in indicating a scale (e.g. 1:50,000) as this one will not be correct any longer, the indications provided by the scale bar, however, will still be valid.
4. **North arrow**. Though the overwhelming majority of maps are north oriented, the indication of the cardinal points is still a required element of the map layout.
5. **Projection information**: All parameters of the map projection used to display the spatial data layers should be clearly indicated. Read the next section on map projections before writing down these indications
6. **Map source**: The map elements portrayed on your map layout may originate from your own base data (e.g. GPS measurements in the field), or from external sources (e.g. satellite imagery purchased from an official provider, vector information bought from a topographic institute, etc.). Wherever your data comes from, the source(s) of this data have to be indicated on the map. The map source information should also contain indication about the age of the primary information used.

Next to these "compulsory" elements, some **other elements** may be added to the map margin:

7. **Project information**: The map you are producing is maybe prepared for a specific project. In a text frame, you can give some summarised information on this project (e.g. implementing agency, donor agency, runtime of the project, main aims of the project, etc.).
8. **Second map field:** A second, smaller map field can be used to locate the geographic extract shown in the main map field, a second map field can also be used to show the same area as in the first map field, but at a different period in time (e.g. land use in a district in 1980 in the first map field and in 1990 in the second map field).
9. **Charts**: Some of the thematic contents of the map could be worth being portrayed as a statistical chart (e.g. statistical distribution of slope classes in a catchment, statistical distribution of soil types in a village territory, etc.)
10. **Pictures**: Some graphs or photographs portraying the area displayed in the map field can help designing an attractive map layout. The use of such graphic elements is left up to the map author. It is advisable, not to overload the map layout with too many elements.

5.2. Map projection

5.2.1. Definition of projections, spheroid and datum

Projections are flat representations of the Earth drawn on paper or displayed on a computer screen. In other words, projections express a three-dimensional surface in two dimensions. Mathematical formulas are used to transform spherical geographical co-ordinates to the two dimensions of a plane. There are different ways in which such a projection can be conducted (see Figure 5.1). The transformation from the three-dimensional ellipsoid to the two-dimensional plane is not possible without some form of distortion. The distortion affects shapes, distances and directions. Each of the many formulae available will result in different distortions. This determines whether each map projection will be suitable of unsuitable for a certain purpose (Kraak, M.J. and F.J. Ormeling, 1998):

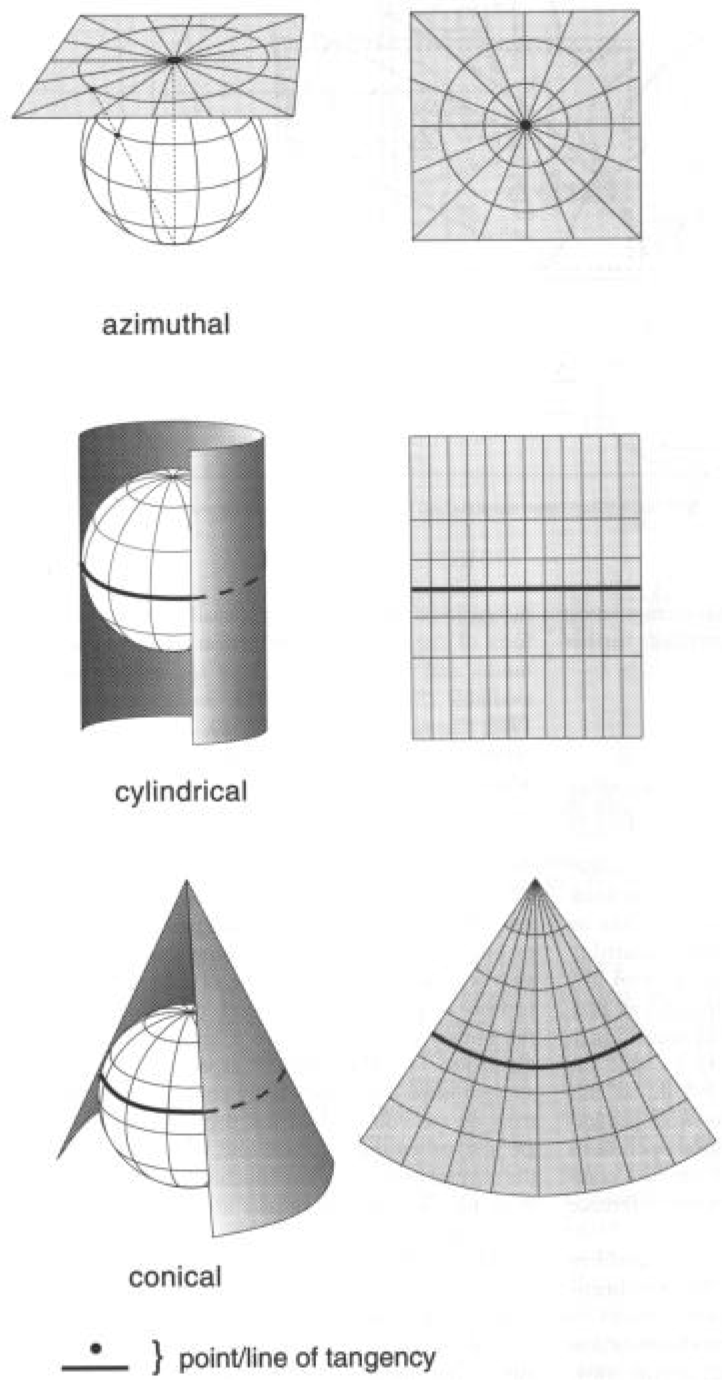


Figure. 5.1: Projection methods
Adapted from Kraak, M.J. & F.J. Ormeling, 1996, p. 82.

Equidistant projections (e.g. Sinusoidal projection) represent distance scaled correctly, but only in one direction (usually north-south). Equidistant projections are often more aesthetically pleasing for representing large portions of the earth's surface than other projections.

Conformal projections (e.g. Universal Transverse Mercator projection) represent angles correctly. The Mercator projection is one example of a conformal projection and it has a primary use in navigation where correct representation of bearings and angular measurement are important.

Equal-area or equivalent projections (e.g. Robinson projection), finally, portray areas on the earth's surface in their true proportion. Such representations have application where the cartographer wishes to show distributions on the earth's surface which can be compared and contrasted.

Any projection can take only one of the above three forms. It is important noticing, that it is often impossible to determine the true type of projection used without a statement by the cartographer on the map. Therefore, a map should always contain information about the projection used.

A projection always refers to a mathematical approximation of the shape of the earth. This approximation is called **spheroid, or ellipsoid**. The parameters of such ellipsoids are optimised for local conditions and thus different ellipsoids are used in varying parts of the world. Because of gravitational variations and variations in surface features, the Earth is neither a perfect sphere nor a perfect spheroid. Satellite technology has revealed several elliptical deviations; for example, the South Pole is closer to the Equator than the North Pole. It should be noted that satellite-determined spheroids are starting to replace the older ground-measured spheroids for reference calculations. A factor that must be taken into account before changing spheroids of reference is that it will change all previously measured values. Because of the complexity of changing spheroids, ground-measured spheroids are currently still in use and are still valid.

A **datum**, finally, is a set of parameters defining a co-ordinate system, and a set of control points whose geometric relationships are known, either through measurement or calculation (Dewhurst 1990). A datum is defined by a spheroid, which approximates the shape of the Earth, and the spheroid's position relative to the centre of the Earth. There are many spheroids representing the shape of the Earth, and many more datums based upon them.

A **horizontal datum** provides a frame of reference for measuring locations on the surface of the Earth. It defines the origin and orientation of latitude and longitude lines. A local datum aligns its spheroid to closely fit the Earth's surface in a particular area and its "origin point" is located on the surface of the Earth. The co-ordinates of the "origin point" are fixed and all other points are calculated from this control point. The co-ordinate system origin of a local datum is not at the centre of the Earth. In the last fifteen years, satellite data has provided geodesists with new measurements to define the best Earth-fitting ellipsoid, which relates co-ordinates to the Earth's centre of mass. An Earth-centred, or geocentric, datum does not have an initial point of origin like a local datum. The Earth's centre of mass is, in a sense, the origin. The most recently developed and widely used datum is the World Geodetic System of 1984 (WGS84). It serves as the framework for supporting locational measurement world-wide. GPS measurements are based upon the WGS84 datum.

(*the above description are extracted from Dorling D. and D. Fairbairn, 1997, p. 30ff. as well as from the ESRI – Arc/Info help programme*)

5.2.2. The Projection Utility wizard

Instant change of the projection of spatial data is possible since a few years only, with the appearance of efficient GIS programmes. With version 3.2 ESRI have included a projection utility wizard to ArcView. Previously, projecting in ArcView was possible only when the input data was available in a spherical co-ordinate system in decimal degrees (longitude and latitude). The projection wizard now allows changing from one projection into another one. However, only shapefiles can be projected. Images and grids, which can be projected in Arc/Info, can still not be projected in ArcView. The only work-around would be to manipulate these images/grids with an image-warp tool as it can be found in the Image Analysis extension (see chapter 11).

Notes:

1. It is important to **clearly document any changes of projection** that were conducted on a shapefile, or any other data set. The parameters defining projection, datum and spheroid are many and the use of a wrong definition will result in possibly significant shifts. For the same reason, it is also advisable, if possible, not to mix projection tools when one data set has to be projected several times, or when several data sets need to be projected in order to spatially match. If you use the Projection Utility Wizard to project your data from UTM to Geographic and then transform your data set into an Arc/Info coverage and project the latter back again into UTM, you might end up with an input and an output layer not correctly overlaying, due to slightly different definition of one, or several parameters used in the projection file. Hence, not only should the different parameters used for the projection be documented properly, but also the projection tool that was used to carry out the operation.

2. If available, projection should preferably be carried out in ArcGIS (in the Arc Toolbox or at the Arc prompt with use of a projection file written in an editor). Reliability of the transformations carried out by the ArcView projection wizard seems to be, at times, questionable.

3. Remember, that the projection utility located in the view – properties menu defines projection parameters of the current view in which you are loading, displaying, or editing themes. These definitions are not stored with the shapefile. The Projection Utility Wizard, on the other hand, defines (optionally) the projection parameters for individual shapefiles. The projection definition of the shapefiles is stored in a *.PRJ meta-file having the same name as the shapefile. This file can be deleted any time and the projection of the shapefile redefined.

4. The projection commonly used in Kenya is:

Projection	Universal Transverse Mercator (UTM)
Zone:	UTM Zone 37 North
Central Meridian:	39 degrees east of Greenwich
Spheroid:	Clarke 1880
Vertical Datum:	Arc 1960
Horizontal Datum:	Approximate Sea Level
Units:	Meters
X-shift:	500,000 meters
Y-shift	10,000,000 meters

In order to avoid the appearing of negative values in areas located south of the Equator, an artificial shift– referred to as Y-Shift – is operated towards the north. This shift is equal to 10,000,000 meters.

Exercise 5.1 – Projecting Shapefiles

Steps	Data	Remarks	Time
Loading data	Keewwsh_geo Keewriv1 Keewadm1	- Start ArcView - Load the three layers into a view document type and display them	5 min
Changing view extent	Keewwsh_geo Keewriv1	- Activate the river layer and move to it's extent - Move the mouse in the view and check on the coordinates that appear on the upper-right side of the window - Activate the watershed boundary layer and move to it's extent. Check on the coordinates again. - What is the difference between the reference systems? - Zoom to the full extent and see what happens	15 min
Projecting shapefile in ArcView	Keewwsh_geo	- One layer has to be projected so they match - Load the Projection Utility Wizard and start it - Use the wizard to project the boundary from "geographic coordinates" to Arc 1960 of UTM as described below: - Current projection: geographic (name: GCS_Arc_1960) - New projection: In the "Name" tab change the type to Projected and the name to World_Mercator. In the "Parameters" tab change the system to GCS_Arc_1960, the false easting to 500,000, the false northing to 10,000,000 and the central meridian to 39. In the "Datum" tab change the transformation to WGS84_to_Arc1960 - Save the new shapefile as keewwsh1	20 min
Projecting in Arc Toolbox	Keewsh_geo	- Carry out the same projection with the projection utility of Arc Toolbox. Use the project tool in the Data Management Toolbox (in the Projections and Transformations .. Feature toolset). Pay attention to use exactly the same settings as above. Save the new shapefile as keewwsh2	20 min

		- Compare both results and draw relevant conclusions	
		Total time:	60 min

5.3. Improved layouts

ArcView enables you to quickly design an acceptable map layout. Various wizards assist you in constructing legends, measured grids (co-ordinates), etc. These tools are briefly presented in the next sections. The layout data type of ArcView also presents some **shortcomings**, which are presented beforehand, so that participants using the sections below do not run into problems while working on a map layout:

Graphic elements: ArcView considers all elements in the layout document type to be graphic type elements. This means, that any element in a layout (view frame, scale bar, legend, etc.) can be selected, deleted, or moved like in a graphic programme. The positive aspect of this concept is that map designing is very fast and easy, as the elements of a map can be positioned on the available page with drag and drop movements of the mouse. The problematic aspect is that there is no more geo-reference in a layout document type. For example, when creating a graticule, or measured grid for a view frame, the latter can be moved again and therewith the reference between the view frame, containing the map elements, and the graticule, containing the co-ordinate values, is lost. For this reason, graphic elements that should not be shifted relatively to one another should immediately be grouped after they have been generated (see the section on graticules and measured grids below).

Display weakness: The display characteristics in the layout document type are sometimes unsatisfactory. For example when changing an element in the view document type to which the layout is related, the display of the layout is not refreshed properly and the layout needs to be iconised, or even closed and opened again for the graphic elements to be correctly refreshed.

Mapped extract: When the spatial data in the view includes a larger area than what should be displayed in the layout, the panning in the view until the correct extract is found in the layout can be trying. When at a later stage, the extract in the view is shifted, for example because some editing needed to be carried out on one of the data sets included in the view, the extract mapped in the layout shifts as well and needs to be re-positioned a second time. We have found no tool specifying exactly the extract that should be portrayed in the layout. A work-around to this problem is to create a new polygon shapefile in the view and to draw a rectangle box over the desired area. After editing, panning, etc, one just needs to activate the shapefile in the theme column and to click the 'Zoom to extent of active themes' button, to get back to the desired extend.

Templates: The definition of layout templates and their later use with other data sets is not all that straightforward as suggested in the ArcView documentation. In some cases, layout elements are lost when the template is re-used. In other cases, layout elements have to be eliminated before creating a template in order to be operational with other data sets, etc. As a rule, the map elements in the view frame should be deactivated before creating a layout template, otherwise the latter will not be properly operational.

Printing: Simple and small size layouts can easily be printed with the printing device to be found in the layout document type. But, when handling and printing large size layouts, the default printing options might be unable to handle the data and the only way to produce hard-copies would be to purchase and use the ArcPress extension. ArcPress for ArcView, however, provides less satisfactory than its homologue for Arc/Info. Problems concern mainly the output quality, as well as the processing of transparent, or translucid map elements (brightness themes, or transparent 'NoData' areas), which can cause the entire layout to collapse.

5.3.1. Legend tool

The legend tool is one of the tools that allow improving the layout's design. When starting a layout document type, ArcView automatically adds a default legend to the layout. This legend is a single graphic element and can not be adapted according to one's own wishes. In most of the cases, the default legend is not suitable for a proper map layout. Hence, a custom legend needs to be prepared. This can be done with the custom legend extension. When loading this extension, a new icon is added to the GUI (). By pressing this icon, and pointing the round cursor at the lower left corner of the desired legend location one starts the custom legend wizard, which leads through the legend creation process. This process is fairly self-explanatory. Just make sure to select the correct themes (from the correct view) to be included to the legend. Once the custom legend wizard has been completed, the legend can be selected and the different elements it is composed of can be addressed individually after they have been ungrouped (use the ungroup option in the graphics menu repeatedly, until all elements of the custom legend have been ungrouped). After ungrouping, the elements can be moved, resized and altered in terms of colour, font type and size, etc. After adapting the different elements, they can be re-aligned using the align option in the graphics menu. When you are satisfied with the way your legend looks, select all elements belonging to it and group them to a single graphic element which you can now position on your map layout.

5.3.2. Graticules and measured grids

The graticules and measured grids extension is a tool used to add a co-ordinate grid, or tic lines to a layout's view frame. Upon activating the extension, a new icon is added to the GUI (). The extension is self-explanatory and the current appearance of the co-ordinate grid can be viewed and re-adapted at each step of the extension. The extension is very convenient and versatile. One thing to remember is that the co-ordinate grid which is being drawn after completion of the extension's steps is a graphic element which matches the CURRENT SCALE AND POSITION of the elements in the view frame. This means that, when you change the scale, move the elements in the view, or move the view frame without moving the co-ordinate grid at the same time, the geo-reference of the latter is irremediably lost and a new co-ordinate grid needs to be created. In order to prevent at least the last of these eventualities, the view frame and the co-ordinate grid can be grouped (select both of them and then select the group option in the graphics menu). In any case, the creation of a co-ordinate grid should be – together with the creation of a custom legend – one of the last steps in layout designing, when one is sure not to have to change scale, or position of the view frame.

5.3.3. Layout templates

Layouts can be stored as templates, to be used to a later date, with other data. The idea of templates is that recurrent elements of a map layout, as for example page size, location of the view frame, north arrow, scale bar, text elements describing the projection parameters, or the project set-up in the frame of which the map was created, or the map source, etc. can be stored in the template and retrieved and display in a new layout with different contents (e.g. when making a map series with identical thematic content but adjacent geographic extents).

5.4. The Layout View in ArcGIS

In ArcGIS the Layout is designed in the Layout View (as opposed to the Data View). Legends, scale bars, north arrows, and graticules are adapted when the map extent or the map content is changed in the Layout View. These elements can be converted to graphics in order to change them to the user's needs. However, the resulting graphics are not associated with the map extent any more and do not change / scale appropriately when changing the mapped extent. Some of the restrictions of ArcView 3.x have been dealt with, however, printing maps with transparent rasters has still some restrictions in ArcGIS 9, especially when containing white elements in the symbology. Layout templates in ArcGIS are stored in map template files (*.mxt). There is a number of pre-defined templates to chose from, or custom-made layouts can also be stored as template files and reused at a later occasion. There is a limitation that one map file can only contain one layout, which is planned to be solved at version 9.1.

CETRAD

6. Digitizing in ArcView

6.1. Preliminary remarks to digitising in ArcView

6.1.1. Digitising functions and concepts in ArcView

One of the main problems when it comes to digitising and editing in ArcView is the inability of this programme to create a topology. It is important to note that **shapefiles do not inherently have a topology** (see box below) like Arc/Info coverages (after the latter were "built", or "cleaned").

The lack of topology sometimes leads to **unclear results**, especially when splitting an arc, or when snapping a new arc to an existing one. Furthermore, polygons are not "adjacent" in the normal sense. Two polygons within a shape file can actually overlap each other. And, if two polygons are side by side and appear to have a common arc they actually do not. Each polygon has it's own boundary which is not automatically "shared" with any other feature.

This fact can easily be verified by opening a polygon shapefile, putting it into editing mode and selecting with the 'select feature' tool one polygon which is adjacent to one, or several other polygons. Once it is selected, the polygon can be dragged into a new location. The line that was located between this polygon and its former adjacent neighbour(s) is maintained on both the moved polygon and on the former neighbours, which means that the contact line is actually a double line, which is an absurdity in terms of topology, as the common border should logically be unique and shared and the adjacent polygons should not be "separable" the way they are in ArcView.

These remarks should always be kept in mind when manipulating the geometry of shapefiles, as the particularities of this geometry forces one to take special care in some situations.

What is a topology?

"Topology is a mathematical procedure for explicitly defining spatial relationships. The principle in practice is simple. Topology expresses different types of spatial relationships as lists of features (e.g., as area is defined by the arcs comprising its border). The ability to create and store topological relationships has a number of advantages. Topology stores data more efficiently. This allows processing of larger data sets and faster processing. When topological relationships exist, you can perform analyses such as modelling the flow through connecting lines in a network, combining adjacent polygons that have similar characteristics, and overlaying geographic features. The three major topological concepts of Arc/Info are:

1. Arcs connect to each other at nodes (connectivity)
2. Arcs that connect to surround an area define a polygon (area definition)
3. Arcs have direction and left and right sides (contiguity)."

ESRI, 1997, p. 2-10

6.1.2. Topologies

Points are used to digitise elements which are too small (in relation to the scale of the spatial data set) to be depicted as lines or areas (e.g. wells, villages, houses, trees, etc.).

Lines are used to digitise elements which are too narrow (in relation to the scale of the spatial data set) to be depicted as areas (e.g. roads, rivers, electric lines, etc.).

Polygons are used to digitise elements which are too big (in relation to the scale of the spatial data set) to be depicted either as points or lines (e.g. lakes, countries, fields, towns, etc.).

6.1.3. Before digitising

What should you do before starting to digitise a new theme?

1. Check if the spatial data is not already **available as spatial data set** either as a shapefile, or in another format which can be read by, or imported into ArcView. Very often the same information is being digitised several times by different organisations or governmental institutions because of lack of communication, or because of reluctance to share data among these organisations or institutions. However, great amount of effort in terms of time and manpower could be saved, if the possibilities of data exchange are fully implemented.

2. Check if the data is not already **available in tabular form** and could thus be easily entered into an existing GIS data layer instead of being tediously digitised. Often statistic information bears great potential for inclusion into a GIS. These possibilities should be carefully checked.

3. Check out the option of **scanning an existing hardcopy** (topographic map, cadastral plan, road map, etc.) and to subsequently vectorise the resulting bitmap. Such an option might, in certain circumstances be much faster than digitising linear elements by hand. Various vectorising software provide interesting automating devices and perform quite satisfactory results

There are two main possibilities to digitize data: on-screen, with help of a digital image as a background (scanned map, aerial photograph, satellite image, etc.), or with a digitizing tablet. On-screen digitizing is easier and less infrastructure intensive, but some situations might require the use of a digitizing tablet, even if this method tends to be replaced by scanning of originals and subsequent on-screen digitizing. In the frame of the present course, we are concentrating on the on-screen digitizing option and will have an additional look at the generation of vector information through scanning and vectorising, as well as GPS generated vector data.

6.2. Digitising on screen

On-screen digitizing requires the availability of digital geo-referenced **image, or grid data** on the basis of which the vector information is going to be digitised. The image data (for example a scanned land registration plan, a satellite scene, or an aerial photograph) is used as a background and the linear or punctual information required is read from this background.

The necessary steps for digitizing features in ArcView are listed below. In ArcGIS, the section "An overview of editing in ArcMap" from the ArcGIS Desktop Help can be used as guideline.

6.2.1. Digitising points

AV-help

1. Open a view or create a new view
2. Add the image, or grid data on the basis of which you wish to digitise points
3. From the View menu, choose New Theme. In the dialogue that appears, select Point as the feature type from the dropdown list, and press OK.
4. In the dialog that appears next, specify the name and location of the new shapefile in which ArcView will store the data you add to your theme. Press OK. A new empty theme is added to the view

5. Click the Drawing tool palette and in the list of tools, click the Point tool . Add points to the theme by clicking the left mouse button in the view.
6. When you are finished adding new points, choose Stop Editing from the Theme menu. Choose Yes when you are asked whether you want to save your changes.

AV-help

6.2.2. *Digitising lines*

1. Open a view or create a new view
2. Add and display the image, or grid data on the basis of which you wish to digitise lines
3. From the View menu, choose New Theme. In the dialogue that appears, select Line as the feature type from the dropdown list, and press OK.
4. In the dialog that appears next, specify the name and location of the new shapefile in which ArcView will store the data you add to your theme. Press OK. A new empty theme is added to the view.
5. In the view, hold down the right mouse button and choose Enable General Snapping from this menu
6. Click the general snap tool . In the view click and drag out a circle to represent the tolerance distance (alternatively, use the interactive snap tolerance – see ArcView help).
7. Click the Drawing tool palette and in the list of tools that pops down, click the Line tool .
8. Click where you want the line to start, click each vertex along the line, then double-click the final vertex.
9. When you are finished adding new lines, choose Stop Editing from the Theme menu. Choose Yes when you are asked whether you want to save your changes.

Important note: When a line which was digitised with the 'draw line' tool snaps into an existing line, the latter is not split at the point where it was met by the new line. This has implications on the implementation of some tasks especially on route calculations with the Network Analyst extensions. To avoid these problems use the 'draw line to split feature' tool and slightly overshoot the line to be snapped. Then delete the dangle of the overshoot. With this method, the initial line is split at the spot where it was crossed by the new line and a network analysis can be conducted normally.

AV-help

6.2.3. *Digitising polygons*

1. Open a view or create a new view
2. Add and display the image, or grid data on the basis of which you wish to digitise polygons
3. From the View menu, choose New Theme. In the dialogue that appears, select Polygon as the feature type from the dropdown list, and press OK.
4. In the dialog that appears next, specify name and location of the new shapefile in which ArcView will store the data you add to your theme. Press OK. A new empty theme is added to the view.
5. In the view, hold down the right mouse button and choose Enable General Snapping from this menu
6. Click the general snap tool . In the view click and drag out a circle to represent the tolerance distance (alternatively, use the interactive snap tolerance – see ArcView help).
7. To create **irregularly shaped polygons** click the Drawing tool palette and in the list of tools that pops down, click the Polygon tool . Click where you want the polygon to start. Double-click the final vertex.
8. To create **circles** click the Drawing tool palette and in the list of tools that pops down, click the Circle tools . Position the cursor where you want the centre of

the circle to be, hold down the left mouse button and drag the mouse until the circle has the size you want and release the mouse button.

9. To create a **box** click the Rectangle tool . Position the cursor where you want one corner of the box to be, hold down the left mouse button and drag the mouse until the box has the size you want and release the mouse button.
10. Use the Auto Complete tool to **create a new polygon if it shares part of its boundary with (an) existing polygon**(s). It is not necessary re-drawing the polygon segment already represented in the theme. Just cross the existing line where the new polygon border starts and re-cross the same or another existing line where the new polygon border ends.
11. Use the Polygon Split tool to divide an existing polygon in order to get two or more polygons. The line drawn with the Polygon Split tool needs to cross at least two existing lines.
12. When you are finished adding new polygons, choose Stop Editing from the Theme menu. Choose Yes when you are asked whether you want to save your changes.

AV-help

> Chapter 7 contains additional information on how to edit and transform existing points, lines and polygons. Please refer to that chapter for more details on the manipulation of vector data.

In the exercise below, the image to be used as a basis for digitising new elements is a satellite image of Nakuru. The image was taken by the Quickbird satellite in February 2003 and has a resolution of 2 meters (resampled from the initial 0.6 meters). It represents the latest generation of commercially available remote sensing data.

Exercise 6.1 – Digitising on screen

Steps	Data	Remarks	Time
Digitising River Njoro	Nak_qb_2m.tif	- Njoro River can be distinguished on the satellite image. - Create a new line theme and digitise the river. - Save the edits, add a “Name” field in the theme’s table and label the field accordingly - Name the theme correctly and save it in the correct folder (create a folder if it does not yet exist)	35 min
Digitising roads of block 12	Nak_qb_2m.tif Nk_blocks	- The town is divided into town-planning blocks. Load and display the layer to located block 12 - Create a new line theme and digitise the roads in the block, including the ones running along its borders and pay attention to have correctly networked roads (snap). - Save the edits, name the new theme correctly and save it in the correct folder.	40 min
Digitising plantation forest blocks on Menengai	Nak_qb_2m.tif	- Various blocks of Eucalyptus plantations can be identified on Menengai. The blocks are usually separated by roads or paths. - Create a new polygon theme - Digitize the forest plantation blocks - Use the auto-complete tool to digitize polygons adjacent to others and make sure that the logic of the geometry is intact - Save the edits - Open the attribute table and add a field called “Area”. Calculate the values with the “returnarea” request. Add another field called status. Fill in approximate values from your visual interpretation of the image. - Name the layer correctly and save it in the correct folder.	45 min
		Total time:	120 min

6.3. Scanning and vectorising

6.3.1. Scanning and referencing

One of the alternative ways to transform analogue spatial information into digital spatial information is to first scan the analogue data set and then, if necessary, to vectorise the image resulting from the scanning.

As a general rule, it will be easier for automatic vectorising programmes to process separate layers (e.g. maps containing only land use, or only elevation contours, or only rivers) than to process multi-layer maps (e.g. topographic maps containing elevation contours and rivers and co-ordinate grids and roads, etc.). If separates are available, black and white scanning is advisable, but if multi-layer maps have to be used, coloured scanning is advisable, as **colour separation** conducted on the scanned bitmap might be possible by using a specialised graphics programme. Colour separation allows to filter out some colours (e.g. roads – black, contours – brown) for the vectorising process, such that only the elements one is interested in (e.g. rivers – blue) remain in the bitmap.

After the scanning, the image needs to be geo-referenced. Referencing is done by identifying at least four control points the co-ordinates of which are known. If you do not have a referenced data layer on which you can identify ground control points (GCP) which are also visible on the scanned image (e.g. road crossings, houses, river junctions, etc.), make sure that the image you are scanning contains at least 4 co-ordinate points, or a co-ordinate grid. You will find additional instructions on how to **reference images** with the Image Analysis Extension in section 11.2.

Note: If you are interested in the vector information to be extracted from the scanned image and if you intend to conduct automatic vectorising with a vectorising programme as for example Corel OCR Trace (see below), there is no point in geo-referencing the scanned bitmap with the image analyst, as the geo-referencing information is lost during the tracing operation. In such a case, geo-referencing has to be conducted on the vector data gained from the tracing operation. Section 6.4.3 gives some hints on how to conduct the geo-referencing of vector data.

6.3.2. Vectorising

In some cases, digital data contained on a scanned image can easily be transformed into vector information. Hereafter is an **example**:

A field mapping of the land use has been carried out in a specific location, using a base map and a transparent sheet onto which the land use polygons have been drawn by hand. The transparent sheet, which also contains sufficient ground control points, has then been scanned. In order to be able to address the various land use polygons and to join attributes describing area, land use type, etc. into an attribute table, the land use polygons represented on the bitmap need to be transformed into vectors. This operation can be done either manually, by using the bitmap as a background and editing a new polygon theme, or automatically with a special vectorising, or tracing software. Below is a description of the automatic vectorising (tracing) operation. The manual vectorising is described above in the section dealing with on-screen digitising.

Various software exist which can perform automatic tracing plus additional functionalities, as for example semi-automated contour labelling, etc. Some **graphic programmes**, as for example Corel Draw have tracing modules (Corel OCR-Trace) which can be used to automatically vectorise bitmaps. When using such software, there is no point in referencing the bitmap before vectorising it, as graphic programmes are not equipped for handling geo-referenced information (see above). Other **specialised raster to vector conversion programmes** like R2V produced by Able Software Corp. (www.ablesw.com) or the ArcScan-extension for ArcGIS are able to handle geo-referenced information. However the purchase of specialised tracing software can be quite expensive. Therefore, a basic solution using a common graphic programme (Corel OCR-Trace / Corel Draw) is given as an example hereunder.

After having scanned your image, open Corel OCR-Trace, and load the image. If your image is not north-oriented, rotate it accordingly (Image – Rotate). If your image is not in Black and White format, convert it into this format (Image – Convert To). Apply a Threshold which allows to remove small impurities without interrupting the lines to be traced. Perform the tracing (OCR-Trace – Perform Trace – By Centre Line) and save the traced lines (Save – Vector) as AutoCad file (DXF).

Note: You might have to try various settings until the vectorising results are satisfying. Yet there is no guarantee that the transformation into DXF format will be matching your expectations. Often lines are not joined, such that additional editing will be necessary. Possibly other vectorising software will provide better results, unless it is the import into ArcView which presents some weaknesses as well.

6.3.3. Geo-referencing shapefiles

Once the vectorising procedure is completed, you can open the DXF file in ArcView (make sure that you have activated the "CAD Reader" extension first) and save it as a shapefile. You will note that the vector data is not geo-referenced by looking at the co-ordinate values appearing on the upper-right corner of the ArcView GUI as you move the cursor in the view. This means, that the vectorised data can not be overlaid to other data layers covering the same geographic area. Hence, the vectorised lines need to be referenced first.

The basic functions of ArcView do not include the shapefile to geo-referencing, which requires the use of **additional extensions**, some of which can be downloaded for free from the internet. The extension used here is called **ShapeWarp.avx** and was created by Kenneth R. McVay. It is available from the ESRI homepage (under ArcView 3.X – ArcScripts). The explanations given hereafter are mainly extracted from the document delivered by the author together with the extension. Read the explanations carefully.

Copy the extension **ShapeWarp.avx** to the EXT32 folder. Before beginning a ShapeWarp session make sure you save all your work and then start a new project. ShapeWarp removes all existing documents from the project, thus if you start the session with existing work present you will loose these documents. To begin a **ShapeWarp** session close all views and press the **RedDiamond** button.

You will be asked if you would like to **set the projection** for the TO view. The TO view is where the themes are located that you want to warp to. The FROM view is where the themes that you would like to warp are located. **If your TO themes are in decimal degrees** and you would like to warp the FROM themes to something other than decimal degrees then you should specify the projection that you would like to warp to. **If your data is in something other then decimal degrees** then you should not set a projection. If you find that you set a projection and you decide against it then you can hit cancel in the projection dialog box. You will be asked to confirm setting the projection.

Then, you will be asked to **enter the theme you would like to warp**. This is the theme that you have just vectorised using OCR Trace and it will go into the FROM view. Next you will be asked to **pick the TO theme(s) which are the reference themes you would like to warp to**. These themes will go into the TO and TO ROAM views. They can be any valid feature, image, or grid theme. The theme(s) in the TO view should have locations which are identifiable on the FROM theme. Hence, before starting the geo-referencing, special attention should be paid to the identification of ground control points (GCP) available both on the referenced (TO) and on the non-referenced (FROM) themes (see also special hints below).

Next you will be asked if you would like to use an **existing table**. You may use any valid table that is created with this extension as long as the data in the table match the TO and FROM themes. If you have an existing table you would like to use select YES from the message box and then select the existing table from the file dialog. If you do not have an existing GCP table then select NO from the message box and then using the file dialog select the directory and name for the new table.

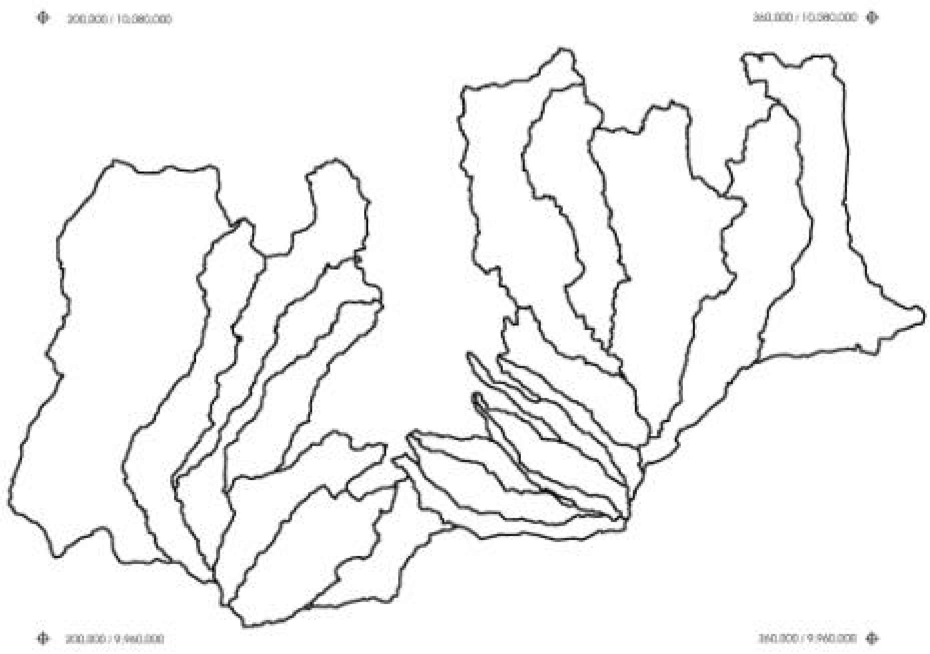


***Figure 6.1**: FROM – theme: Scanned image showing sub-catchment boundaries within the Upper Ewaso Ngiro Basin. This image will be vectorised using Corel OCR Trace and then geo-referenced using the ShapeWarp.avx extension. Four co-ordinate points have been added to the image as GCP to facilitate the geo-referencing procedure.*

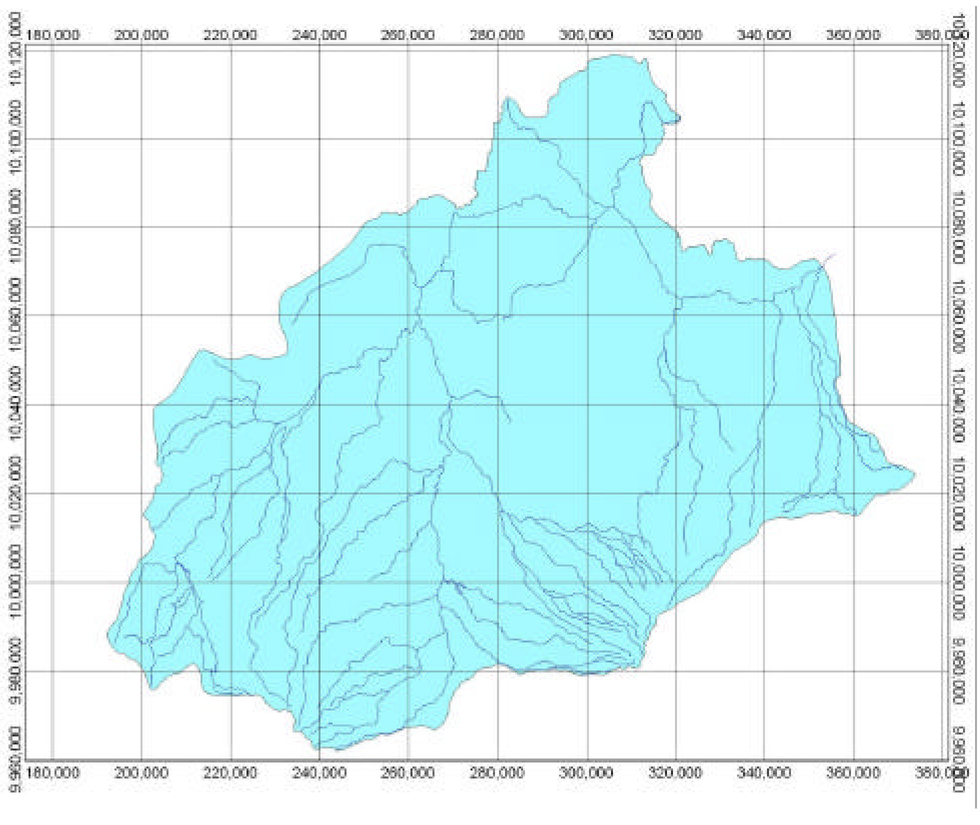


***Figure 6.2**: TO - themes covering the same area as the FROM theme shown in Figure 6.1. The watershed boundary and the rivers can be opened in a view and the co-ordinate grid can be created in this view, using the Coordinate Grid Maker extension. In principle, only the co-ordinate grid needs to be added to the TO view, but other layers can be added as well for more comfortable handling and easier identification of the GCPs.*

Below is a table with the **main tools** of the ShapeWarp extension. The tools which are not included to the table have the same function as the default ArcView tools.

	GCP selection. With it you drag a box around a GCP which selects the point. You can then perform other operations on the selected GCP (move, delete, or turn it off).
	Move GCP. Once a GCP in the TO or FROM view is selected, using this tool you can click on the new position for the point and it will be moved to the newly selected point.
	GCP pick. You use this tool to pick the GCP's in the TO and FROM views. You must always pick a GCP in the TO view first followed by picking a GCP in the FROM view.
	Feature select. Works the same as it normally does. Great for selecting a feature in the TO ROAM view and then choosing the zoom to selection button. The zoom will take place in the TO view.
	Toggle table of contents. The views are not showing the TOC. Pressing this button will toggle them open so that you may change colours etc. on the themes. Pressing it again will close the TOC.
	Toggle View. This button will toggle the views from the main views to the GCP table and back.
	Calculate RMS. This button will calculate the required polynomial and RMS for the chosen GCP's. For a first order fit you need at least 3 GCP's, 6 GCP's for a second order, and 10 GCP's for a third order.
	Delete multiple GCP's. Choosing this button you will be displayed a dialog that allows you to pick one or more GCP's and delete them.
	Toggle GCP's ON/OFF. Sometimes you may not want to delete a GCP you may just want to turn it off. This button allows you to do that by picking the GCP's you want to turn on or off. GCP's that are ON are red, those which are OFF are blue.
	Delete a single GCP. With this button, after using the Select GCP tool to select a GCP, then you press this button to delete the single GCP.
From	Compute From GCP. After having selected the fourth GCP in the TO view, you can then use this button to calculate and add the point in the FROM View. The point will be added and the From view will be centred on the new point. The point will only be close you will still need to make adjustments to its position by moving it.
Go	Warp themes. After you have chosen all your GCP's and calculated the RMS, you can use this button to go ahead an create the new warped theme. Once you press this button you will be asked for the name for the new theme. For every active theme in the FROM view you will be asked to input the name for the new output warped theme. The theme will be created and you can then add it to the TO ROAM View to look at the fit.
Quit	This button allows you to quit the ShapeWarp session
Snap	Snap Dialog.

Special hint: Sometimes, it is difficult identifying GCP among the thematic features on both the FROM and TO files. In such cases it is advisable adding sufficient co-ordinate grid points to the theme to be warped already before scanning the image (see Figure 6.1). These co-ordinate points need to be available on the TO theme(s) as well (see Figure 6.2). One way to proceed, is to open one of the TO themes in a view before starting **ShapeWarp** and to create a co-ordinate grid in this view by using the **Coordinate Grid Maker** extension and saving the co-ordinate grid as a shapefile. The latter can then be included as a geographic reference to the TO view, when running the **ShapeWarp** extension.

Note: When leaving the ShapeWarp extension in the EXT32 folder, its presence can, at times provoke an error message upon starting ArcView. If this happens, you will have to delete the shapewarp.avx file from the EXT32 folder and store it somewhere else until needed again. The reason for this error is not known.

Exercise 6.2 – Automatic tracing and geo-referencing

Steps	Data	Remarks	Time
Scanning a paper copy	Paper copy of sub-catchment boundaries of Ewaso Ngiro Basin	- In groups, the participants will scan the paper copy. Particular attention will be paid to correct resolution and color format. - The scanned image will be named correctly and saved in an adequate folder - Depending on scanner availability, this part of the exercise will be conducted as demonstration	30 min
Tracing the image with OCR Trace	Scanned image	- The participants will prepare the scanned image for automated tracing in Corel OCR trace, or another available tracing programme. - The vector data will saved in a format readable by ArcView (e.g. AutoCAD format - DXF) in the ewaso\gis\hydro folder.	30 min
Geo-referencing the vector data with ShapeWarp.avx	Vectorised data Keewwsh1utm Keewriv1	- The participants will create a co-ordinate grid for the region of the Ewaso Ngiro basin using the overall catchment boundary, the rivers and the Coordinate Grid Maker extension (copy the extension from the extensions folder in the course data and place it under the same name in the EXT32 folder of ArcView - The participants will geo-reference the vector data using the ShapeWarp extension (this extension is also available in the extension folder of the course data). - The resulting shapefile will be stored adequately and compared with other layers for the same area in order to check the precision of the geo-referencing. If significant shifts appear, the warping has to be repeated.	60 min
Total time			120 min

6.4. Importing GPS and tabular data

In the sections above we have seen how to generate new data either by digitising, or by scanning existing or mapped data and vectorising it for subsequent use as shapefiles in ArcView. Another possibility is to use tabular data containing X and Y co-ordinates for selected points in space as well as attribute data describing these points. The most likely situation is the integration of **GPS data** into your GIS. In the third week of the course, a field exercise will be conducted during which participants will learn to handle GPS receivers and to download data for integration into a GIS environment. The theory and practical explanations concerning this part of the course are not included in this manual.

6.4.1. Creating a new point shapefile on the basis of tabular data

Let us assume that you have received by e-mail a table with the co-ordinates of boreholes in Nakuru municipality, which have been measured in the field by a project partner with a GPS receiver or with another measurement device (the table can be found in Nakuru\tables in the course data). The table is shown below and contains the X and Y co-ordinates of the boreholes in decimal degrees as well as a code number, which will help to identify the boreholes when combining the co-ordinate information with other information stored in separate tables. As your GIS project is in UTM zone 37N, datum Arc 1960 and spheroid Clarke 1880, with false northing = 10,000,000 and false easting = 500,000, you will need to import the table and then to project it in order to match your other data layers.

In the present case you will create a new point theme containing the boreholes, the co-ordinates of which are shown below. It is also possible to add tabular data into already existing GIS data sets. Both cases are presented in the frame of the two exercises below.

CODE	X	Y	CODE	X	Y
1	36.059272	-0.285358	20	36.140003	-0.300453
2	36.057381	-0.288524	21	36.144425	-0.305203
3	36.056897	-0.270606	22	36.141489	-0.311132
4	36.055024	-0.271529	23	36.059181	-0.292197
5	36.056858	-0.273952	24	36.046878	-0.281202
6	36.083628	-0.275235	25	36.051304	-0.278901
7	36.090956	-0.279821	26	36.049865	-0.275951
8	36.096407	-0.279622	27	36.045749	-0.278949
9	36.100642	-0.279415	28	36.133336	-0.260719
10	36.118653	-0.285374	29	36.133626	-0.261032
11	36.120901	-0.282308	30	36.133506	-0.260875
12	36.121747	-0.279404	31	36.133619	-0.260321
13	36.134667	-0.291525	32	36.133273	-0.260534
14	36.123973	-0.285522	33	36.071742	-0.288997
15	36.137223	-0.286382	34	36.101802	-0.270741
16	36.137695	-0.285963	35	36.124134	-0.303697
17	36.138434	-0.287227	36	36.123532	-0.304242
18	36.154994	-0.284208	37	36.124133	-0.304364
19	36.132815	-0.296421	38	36.123532	-0.303636

Exercise 6.3 – Creating a new point theme with co-ordinate points

Steps	Data	Remarks	Time
Reformating table as a DBase file	Nk_boreholes_ coord_geo.xls	- The co-ordinates of 38 boreholes are stored in an Excel table (see table above). This table need to be saved as a DBF table (DBase format) in order to be imported into AV - Store the DBF table in Nakuru\gis\water	30 min
Adding tabular data as event theme	DBF table created above	- Open the DBF table from the AV project window - Open a new view and add a new event theme based on the tabular information of the DBF file. Pay attention to select the correct columns for the X and Y co-ordinate information - Save the event theme as a shapefile in …\water	30 min
Project the theme into UTM-37		- Open the Arc Toolbox - Project the point theme from its current Geographic projection to the projection specified above. - Control the position of the boreholes based on expert knowledge available among the participants. - Save the projected shapefile in …\water	30 min.
Total time			90 min.

6.4.2. Joining tabular data to an existing shapefile's attribute table

Before you can begin with the joining of tabular data with your vector layer's table of attributes, you have to convert this tabular data to a DBase file (in excel and access you can simply export your file to a dbase4 file format). This step has been carried out in the exercise 6.3 above. Take into account that DBase does not support long file and item names. If you want to join a file that has Information that has more than 8 letters, you have to save it as comma delimited text, and join this to the shapefile.

Two conditions have to be fulfilled for you to be able to join tabular information:

1. The spatial data to which you want to join the Information has to be a shapefile (in ArcGIS any type of Feature Class can be used).
2. The spatial data and the tabular data you want to join to it have to have a common field.

You can join a table to the active table based on the values of a common field found in both tables. In case of the borehole data shown above, "Code" will be used as the common field. Join establishes a one-to-one or many-to-one relationship between the destination table (the active table) and the source table (the table you are joining into the active table). Typically, the source table contains descriptive attributes of features that you wish to join into a theme's table so that you can symbolise, label, query and analyse the features in the theme using the data from your source table.

To join two tables:

1 Open the tables you wish to join.
2 Make the source table active.
3 Click the name of the common field you wish to use.
4 Make the destination table active.
5 Click the name of the common field you wish to use.
6 From the Table menu choose Join, or click the Join button.

The table that is active when you choose Join is the destination table. The last table that was active is the source table.

In ArcGIS joins, or, in the case of a one-to-many relationship relates, are available from the layer context menu. The two options "join by attributes" (as used in the current example) and "join by location" (described in section 4.1) are available.

As mentioned above, the precondition to adding tabular information into an existing spatial data set is that at least one common information field (item) is available in the tabular data to be imported and in the spatial data's attribute table (e.g. names of provinces, district codes, etc.). This common information field will enable the software to allocate the new information to the correct spatial element. It is very important, that the entries in the "join items" of both tables have exactly the same spelling. If for example the names of cities are used for joining two tables, as it is shown below and if the names of these cities are not written in exactly the same way, the joining of the tables will not be carried out correctly.

Table to be joined

Towns	Popul
Archers Post	*xxxxxxx*
Nairobi	*yyyyyyy*
Nakuru	*zzzzzzz*
....	....

Joining

Attribute table of the shapefile

Towns	Schools	District
Archer's Post	*xxxx*	*aaaaa*
Nairobi	*yyyy*	*bbbbb*
Nakuru	*zzzz*	*ccccc*
....	....	....

In the example above, a table containing the population statistics per towns is joined to the attribute table of the point shapefile describing the location and different characteristics of the towns in Kenya. The join item is "towns". Archers Post (Archer's Post) will not be joined correctly in this example, as the change in orthography will hinder the programme from recognising the entries as being one and the same thing.

Note: Joining tabular data is a reversible process. When you save a project containing a join, ArcView saves the definition of the join rather than saving the joined data itself. The next time you open the project, ArcView rejoins any joined tables by reading their files from disk. In this way, any changes to the source or destination tables that have taken place since they were last joined are automatically included in the project, and reflected in any views, tables, charts or layouts based on the joined data. Because joining tables has no effect on their physical files on disk, you don't need to have write permission to files to be able to join them in ArcView. For example, if you are working with themes representing geographic data accessed across your network from a remote, read-only GIS database, you can still join your own data tables to these themes and work with your data spatially.

With “undo all joins” (in ArcGIS: “Remove all Joins” in the layer context menu) the information that was added to the attribute table of the shapefile will disappear again. There are three possibilities for storing the information permanently in the attribute table. The first one is to choose Export from the Table menu (in ArcGIS: Data .. Export Data from the layer context menu). This is done, when the tabular data is to be used in another application. The next possibility is to convert the shapfefile to a new shapefile (not available in ArcGIS). This will store the joined information permanently into the new shapefile’s attribute table. The other option is to start editing the table, creating a new field for all joined fields and conducting “calculate field” with the option new field = ‘one of the join fields’ for each newly created fields. After the information has been copied from the joined fields to the newly created fields, the join can be undone.

In the exercise below you will learn how to join information from an DBF table containing information on the boreholes in Nakuru municipality that were created as an event theme in Exercise 6.3.

Exercise 6.4 – Joining tabular information

Steps	**Data**	**Remarks**	**Time**
Preparing the tabular data	Nk_boreholes_ attributes.dbf (in \tables)	- Open the DBF table in Excel and control the entries and the correct disposition of the data in the table. Also control that none of the field headers exceeds 8 characters in length - Save the table as a DBF file again. While doing so, make sure that you have either selected none or all of the fields that need to be included in the finale DBF document.	10 min.
Opening the DBF in ArcView	DBF table created in step 1 and output from exercise 6.3	- Open the DBF table in ArcView - Open the attribute table of the borehole shapefile (output from exercise 6.3) - Compare both tables and check if a join could be carried out (there needs to be a common item)	10 min.
Joining the tables	id.	- Join both tables. Pay attention to selecting the correct join items and make sure the attribute table of the shapefile is activated when you carry out the join command.	10 min.
Make join permanent	attribute table of borehole shapefile	- Joins are non permanent links between tables. The “Remove all joins” option in the Table menu allows you to remove the joined data from the table. - To keep the data permanently in the table either add new fields and copy the information from the joined fields into these new fields, or carry out a “convert to shapefile” command on your shapefile	10 min.
Total time			**40 min.**

7. Editing with ArcView

7.1. What is editing?

Editing includes all manipulations altering the geometry of a spatial data set, i.e. adding shapes (e.g. adding a newly built road to an older road coverage), deleting shapes (e.g. removing a mine field from a mine field coverage after that particular area was cleared from unexploded ordinance) and reshaping shapes (e.g. refining a river coverage after more precise base data like aerial photographs, or high resolution satellite data became available). Furthermore, attribute modification, i.e. data manipulation in the attribute tables of shapefiles, is also part of the editing possibilities of a GIS.

While attribute modification tasks are quite user friendly in ArcView, the editing of geometric features (points, lines and polygons) bears some problems. Basically the same remarks as the ones formulated at the beginning of the previous chapter, on the inability of ArcView to build a topology can be made concerning the editing options presented below. An ArcView user wrote in an e-mail:

"*The problem with ArcView is that it's not an editing tool. ESRI designed it originally to be a desktop GIS system mainly for the cartographic output and as a result the editing capabilities are extremely basic. Arc/Info is the intended editing environment of the ESRI products.*"

In the meantime, the own editing options in ArcView have improved and numerous editing extensions can be bought, or in some cases downloaded for free from the internet. One of these extensions is the **EditTools** programmed by Ianko Tchoukanski (ianko@yebo.co.za / http://www.ianko.com/) and distributed by Shareit (http://www.shareit.com / info@shareit.com). This extension, which cost 110 US$ facilitates numerous editing tasks including digitising, snap tolerance setting, extending lines, splitting lines, intersecting shapes, creating centre-lines, removing redundant nodes, etc. This extension could be very useful especially in combination with routing tasks (see chapter 10 on the Network Analyst) which require topologically correct shapefiles in order e.g. to calculate a route. See section 7.3 for further comments on this extension.

For a comprehensive overview of the editing functionality in ArcGIS, see "An overview of editing in ArcMap" in the "Editing in ArcMap" Chapter of the ArcGIS Desktop Help.

7.2. Default editing options in ArcView

7.2.1. Vertex editing

The most common task when editing an existing spatial data set is to correct, refine and complete existing information on the basis of new indications or data. The completing and to some extent the refining tasks can be assimilated to digitising tasks and have been describe in chapter 3. The correcting of existing information mainly consists in reshaping lines and polygons and in moving features. Reshaping becomes necessary when the original digitising operation did not fulfil the precision standards required, or when the digitised features have changed in reality (e.g. increase of built-up areas, new train tracks, etc.).

The above described reshaping manipulations are conducted with the vertex editing tool. Vertexes are points along an arc (a line) which divide this line into a number of straight line segments.

On a straight line few vertexes are required to draw the shape adequately. However in curves, smaller spacing between individual vertexes is necessary in order to smoothly conduct the line (see figures below).

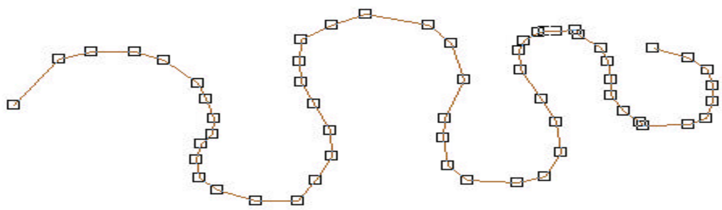


Figure 7.1: Distribution of vertexes on curvy lines

On a curvy line (e.g. road in the mountains, river, etc.) many vertexes are required to draw the line adequately

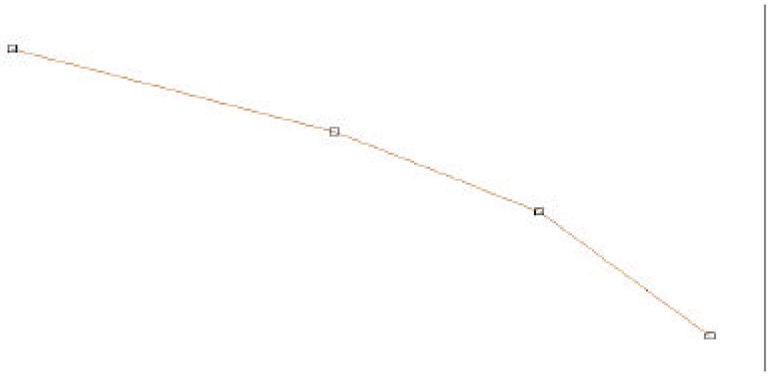


Figure 7.2: Distribution of vertexes on straight lines

On a straighter line (e.g. highway, communication lines, etc.), few vertexes are required to draw the line adequately.

The vertex edit tool allows to move single vertexes in order to reshape the line. In the example below, the second vertex from the left of the above line is moved downwards using the vertex editing tool. The line follows a new route.

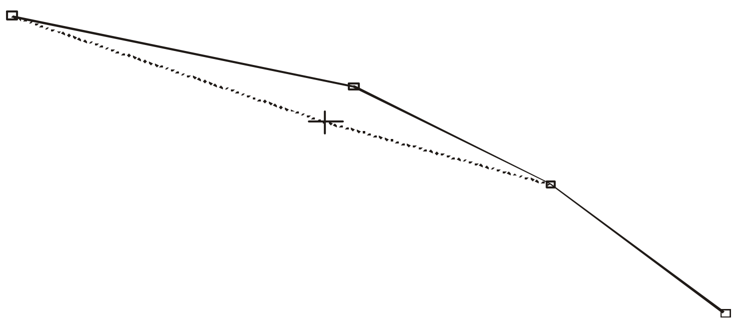


Figure 7.3: Moving vertexes with the vertex edit tool of ArcView

7.2.2. Moving features

Sometimes the shapes of the various features of a spatial data set are precise enough, but the location of these features in space does not satisfy requirements. Such cases happen frequently when using data sets based on different referencing parameters, or originating from different projections. Projection and precise referencing of spatial data are complex issues. Furthermore, in various areas of the world, no, or not enough precise benchmarks with which such spatial data can be referenced have been defined. Because of these reasons, shifts from one data set to the other can often be observed, as it is the case in the example below. In this example, the river beds shown on the satellite image are shifted by approx. 180 meters to the north when compared to the vector lines representing the rivers. The shift is constant over the entire extent displayed.

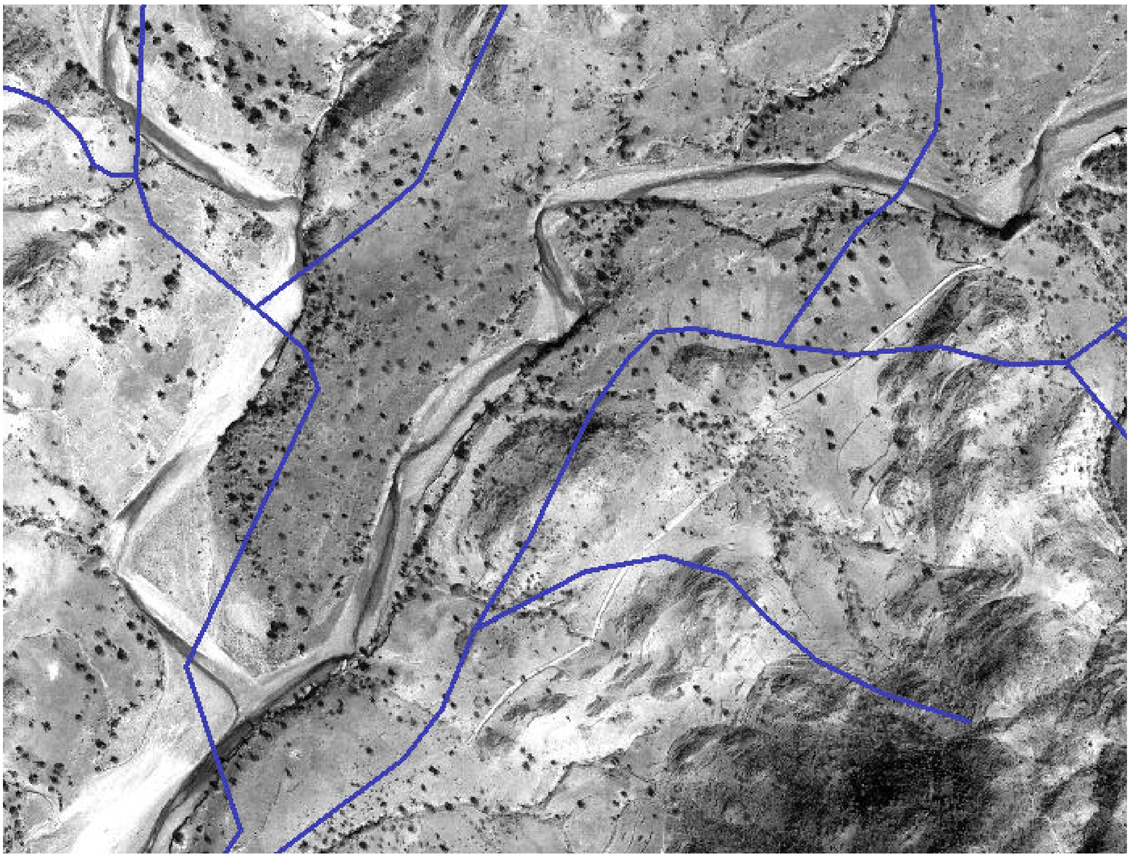


Figure 7.4: Shift between an Ikonos satellite image and corresponding linear features of a shapefile.

In such cases, presenting a constant shift over the entire data set, one possibility to correct the incoherence is to move all the features of the concerned shapefile an approximate 180 meters to the north. Alternatively, the satellite can be warped so that it matches the features of the shapefile. In the event of such manipulations it is extremely important to always **keep a clear track** of the shifts and warps undertaken on the various data sets. Furthermore, it is primordial to **choose one and only one set of references which will then be considered to be "correct"** and which will serve as a basis for the correction / adaptation of the other data sets.

Features of a shapefile are easily moved by putting the shapefile into editing mode, selecting the features to be moved (most commonly all of them) and moving the set of features with the mouse, until they are positioned correctly. This technique has the disadvantage, that it is almost impossible to move several shapefiles exactly the same distance and direction when using the mouse. The default tools of ArcView do not allow to enter values in map units and then to move the selected features using these values. However, there is one extension, called **transform2d.avx**, programmed by William Huber of QuantDec and available on the ESRI homepage (provided with the course data). This extension allows to shift all the features of a theme in X or Y directions, to rotate the theme, or to change its scale.

In other cases, the shift might not be uniform and thus the differences between two or more data sets might rather be related to varying projection parameters. Sometimes, one or several parameters defining the exact projection of a spatial data set, as for example the datum, or the spheroid are not precisely defined, or the benchmarks used for the definition of these parameters are not accessible anymore because of a war, a conflict, or other reasons.

The **lesson to be learned** from the above remarks are:

1. When building up a spatial data base, a consensus on the projection, including all parameters of the geo-reference of the spatial data sets to be entered into this data base should be reached and rigorously applied to all the data.

2. Before integrating a new data set into an existing spatial database, all projection parameters of this data set should be known and recorded in a safe place

3. Standards for the projection of new data sets should be defined (which software to use, which parameters to enter into the projection file / projection software, etc.)

4. Each projection / transformation manipulation should be recorded precisely.

5. A copy of each data set in their original projection should be kept in a safe place.

7.2.3. Flipping lines

Another editing manipulation which is used especially for the preparation for further manipulation of hydrologic features (rivers, streams, canals, etc.) and themes to be used in the frame of a network analysis is the line flipping tool. As it is shown in the figure below, a line always has a beginning node, and end node, vertexes and a direction. Vectors are oriented.

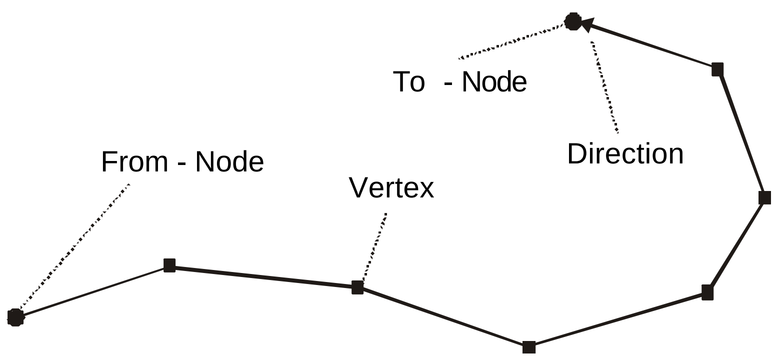


Figure 7.5: Anatomy of a vector

Some calculations (some DTM algorithms, some steps in hydrologic modelling, etc.) require a correct orientation of all the lines (rivers) of the spatial data used for the calculation. The problems that can arise, when a DTM is calculated with an incorrect hydrologic entry are shown in the comparative figure below. The algorithm interprets the rivers as being the lowest point of a valley and the direction of the vector indicates the direction of the slope. In the case below, calculated with an option which gives priority to the indications provided by the hydrologic data set, the algorithm tries to make the river flow by all means through the valley, only in the wrong way!

Several extensions provide a flipping option. Among them the above mentioned EditTools extension, which is presented in the next section and a small extension called flip lines (flip.avx) only fulfilling the flipping option. Both extensions are provided with the course data.

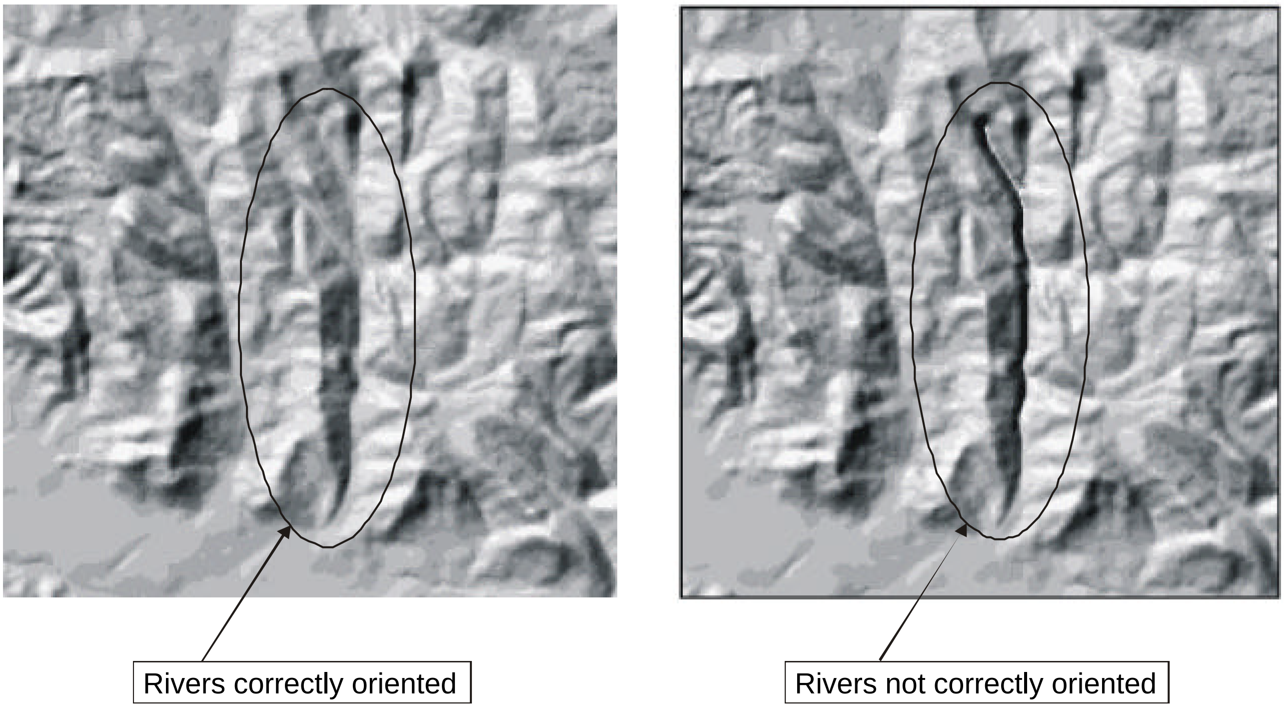


Figure 7.6: The effect of wrongly oriented river lines on the calculation of a DTM using the Topo-grid Algorithm of Arc/Info.

7.3. Editing with the Edit Tools extension

The Edit Tools extension has been shortly presented above. Hereunder is a short list of the most useful features of this extension, the way they are found in the version 2.3. As already mentioned, the extension is not available for free (it costs 110 US$), but the buyer is entitled to free delivery of all upgrades by e-mail. The current version is 3.0.

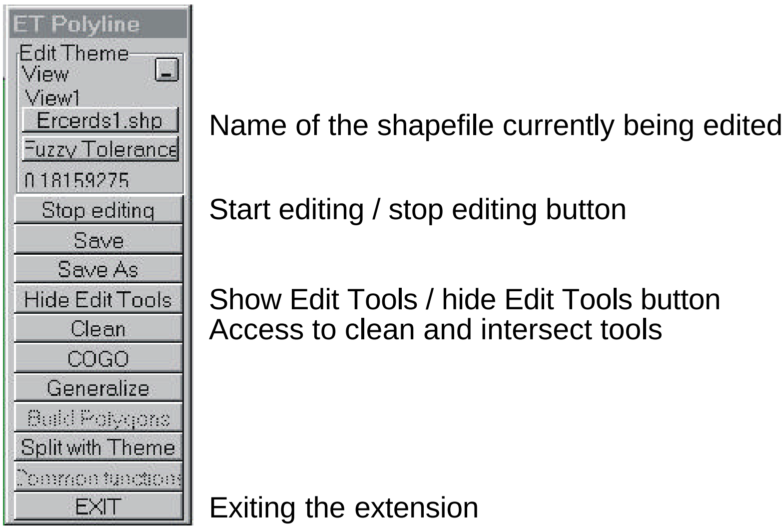


Figure 7.7: The menu selection window of the Edit Tools extension

Before starting the Edit Tools extension, the view's map units and distance units need to be set in the view – properties menu. When starting the extension Edit Tools asks for the shapefile to be edited and then a menu selection window is opened. In the menu selection window above, a line theme was opened. The shapefile type is indicated in the window title line (polyline). The editing tools of the extension are started with the 'start editing / stop editing' button. All topology building tools are started with the 'access to clean and intersect tools' button.

The menu selection window appearing upon editing a polygon is slightly different. In addition to the selection possibilities shown above, it also features an **eliminate** button, with which it is possible to eliminate certain polygons from a shapefile based on a logical request, e.g. to eliminate all polygons with an area smaller than X. Furthermore, the selection window displayed when editing a polygon shapefile features a clean gap function which allows to get rid of 'sliver polygons'. Finally, two very useful tools offered for polygons are the **Create Label Points** and the **Attributes from Points** tools. With these tools, attribute values can be taken over from point shapefiles. For example you have freshly digitised a polygon shapefile with the districts of a province. This shapefile has no attribute information yet. Additionally, you have a point shapefile with the administrative centre of each district. This shapefile contains valuable information which is relevant for the district area as well. The Attributes from Points tool enables you to take over all attribute information from the point shapefile into the polygon shapefile. The Create Label Points tool enables you to create a point shapefile containing the attribute information of a polygon shapefile. Each point in the point shapefile will be located in the centre of a polygon.

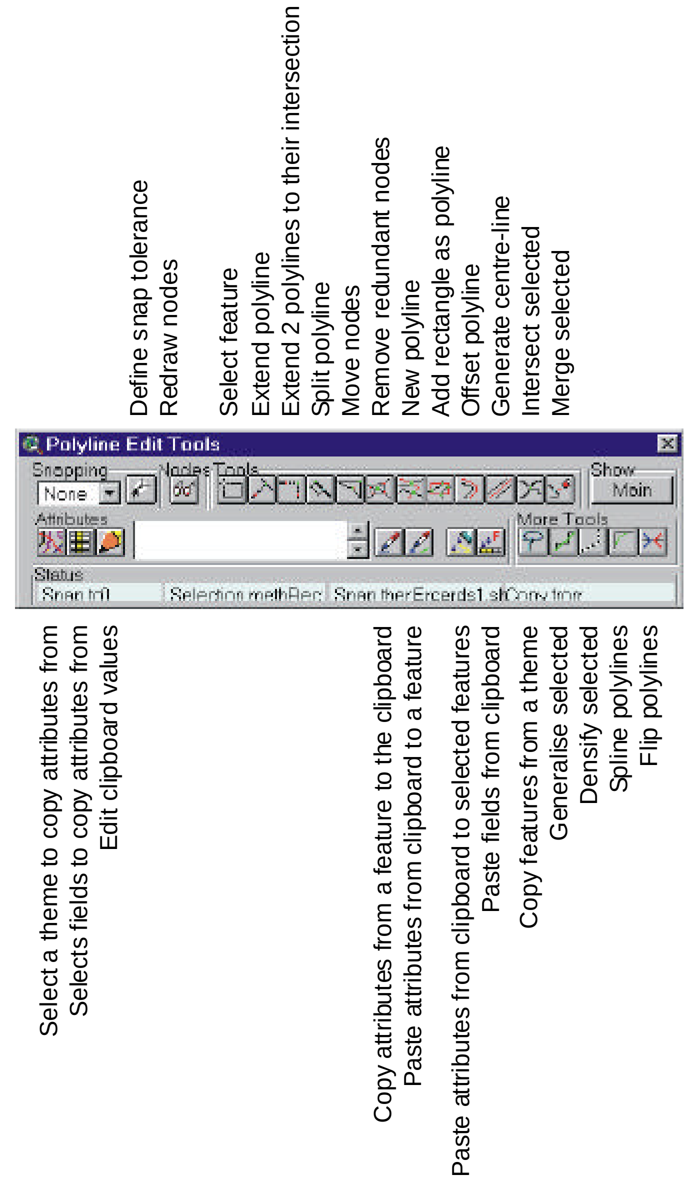


Figure 7.8: Editing tools of the Edit Tools 2.3 extension used for the manipulation of polylines

Editing lines

The **edit tools menu for polyline shapefiles** are shown in Figure 7.8 above. They include a variety of tools which will for sure speed up editing jobs in ArcView. The tools on the upper line of the Edit Tools window contain shortcuts to functions which can also be found in the ArcView's GUI, as well as new tools the most useful of which are certainly the line manipulation tools allowing to intersect, split, extend and merge polylines and therewith help building up a coherent data layer. Please note, that some manipulations are not refreshed automatically in the view. To refresh the view use the redraw nodes tool.

The lower line of the Edit Tools window for polylines contains attribute manipulation tools. Attributes can be copied from one theme to another and from one feature to another and can be edited in the process. These tools can be found on the Edit Tools window for polygons (see Figure 7.9). On the right side of the same line are three functions which are taken over from Arc/Info and which help to adapt the line geometry to the specificity of the data set. Generalise, densify and spline are all functions steering the vertex spacing on a polyline as well as the geometric shape of this line. The flip function allows inverting the direction of selected polylines, which is an important tool when orienting lines in a road traffic data set, in an electric network map, or in a river network (sees explanations above).

Editing polygons

The editing tools menu for polygon shapefiles is shown in Figure 7.9. Worth mentioning is the reshape polygon tool which allows to quickly change the shape of a polygon without having to select and drag, or even add new vertexes as it is being done with the vertex edit tool of the ArcView GUI. Furthermore, the **fill holes** and **remove gaps** tools are useful to "repair" a polygon shapefile the "topology" of which was affected by some faulty manipulation, or by non precise vectorising of an input image data. The remove gaps tool identifies areas between polygons where the latter are not exactly adjacent. There is no shortcut in the default ArcView GUI to cleaning up sliver polygons. Without the Edit Tools extension you would have to go vertex by vertex and snap them all together.

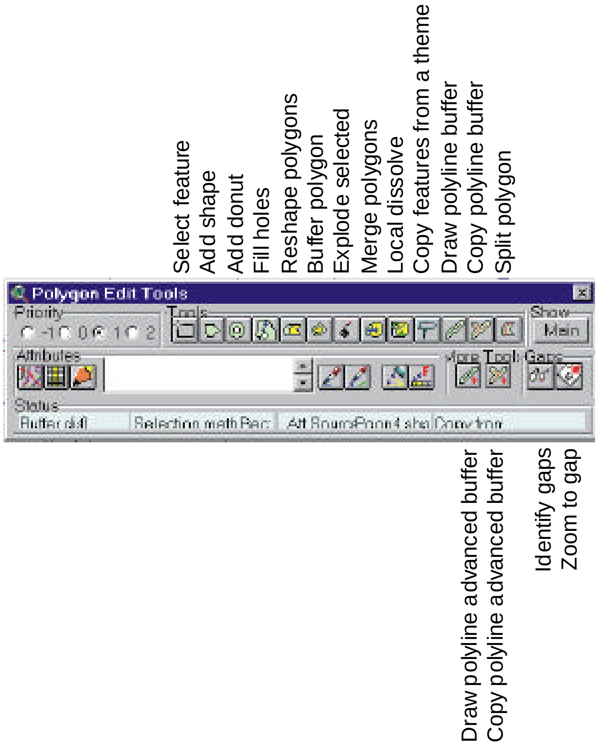


Figure 7.9: Editing tools of the Edit Tools 2.3 extension used for the manipulation of polygons

Some of the tools provided in the extension can also be found in the geoprocessing wizard extension (view – geoprocessing menu). One of these tools is the **dissolve polygons** tool, with which adjacent polygons can be dissolved on the basis of selected attributes. The merge function can also be found in the geoprocessing wizard.

Another useful function, probably taken over from Arc/Info, is the **eliminate polygons** function. This function allows eliminating polygons of a shapefile based on a logical request. For example a classified slope grid is transformed into a polygon shapefile, Depending on the area portrayed, this shapefile might be almost impossible to manipulate due to the huge amount of features. As a matter of fact, areas belonging to the same slope class and covering only a few cells of the input grid will also be converted to polygons. In such a case, the smallest polygons, which have almost no incidence on the overall thematic statement can be eliminated with a request of the style: "eliminate area smaller than 500 square meters".

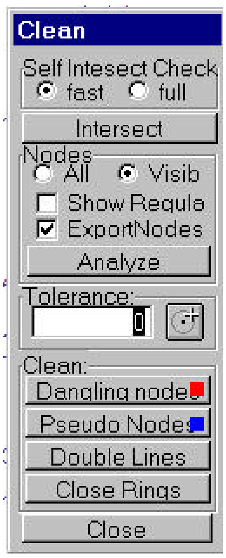


Figure 7.10: Cleaning and intersection menu of the Edit Tools 2.3 extension

Finally, the Edit Tools extension provides different "topology maintenance" tools shown in Figure 7.10. These tools are similar to equally named commands of Arc/Info. Though, the tools suggest the building up of an Arc/Info-like topology, ArcView's geometric concept and the lack of topology inherent to the data structure of shapefiles remains valid also when using the present extension. Nevertheless, the functionalities of the clean menu are very convenient when digitising and editing feature themes, especially thanks to their visualisation, analysis and cleaning of pseudo and dangling nodes. The intersect tool builds nodes on line intersection and is therefore an invaluable help for the preparation of a shapefile to be used in the frame of a network analysis (see chapter 10).

Exercise 7.1 – Improving shapefile with EditTools

Steps	Data	Remarks	Time
Loading sub-catchment layer	Output of exercise 6.2	- The shapefile generated in exercise 6.2 presents an important number of geometric inaccuracies which occurred during the transformation of the information into a DXF layer. In the present exercise, the participants will try eliminating some of these inaccuracies with the Edit Tools extension. - Start ArcView, load and display the sub-catchment shapefile generated in exercise 6.2.	10 min
Setting view properties and loading extension	---	- Set the map and distance units in the view – properties menu to the correct values. - Make sure the ET36.avx extension has been copied to the EXT32 folder of ArcView - Load the extension through the file – extension menu.	10 min
Starting Extension and cleaning dangling nodes	Output of exercise 6.2	- Start EditTools and specify the layer to be edited - Start editing the layer and delete the lines that were created while vectorising the GCPs and the coordinate numbers. - Move to the clean menu and set the tolerance interactively with the mouse. Then analyze dangling and pseudo nodes and then clean them. You might have to run this operation severally with different tolerance values in case the initial trials do not give satisfactory results. After cleaning analyze again and zoom to the remaining dangles, adjust the tolerance and clean again until the shapefile has a coherent geometry.	20 min
Transforming lines into polygons	Cleaned Shapefile	- Once you are satisfied with the result save the edits and leave the EditTools extension. - Re-start the extension and after specifying the layer to be edited move to the “build polygons” menu. - Enter a folder and a name for the temporary polygon layer to be generated and do the same for the final shapefile to be generated. - The new polygon shapefile is automatically displayed in the view. Change the symbology to unique value in order to check whether polygons are actually delimited from each other or not.	20 min
Eliminating polygons	Polygon shapefile	- Unless manual editing was conducted in exercise 6.2 it is probable that some sliver polygons remain between some of the “real” catchment polygons. In order to remove them start the EditTools extension and specify the polygon shapefile as the layer to be edited. - Move to the “eliminate” menu and start a selection with the general logic [area < x]. You should see a significant gap in size between the sliver polygons and the other ones. The selection will show you, whether you entry was correct. - Eliminate the sliver polygons	20 min
Total time			80 min

CETRAD

8. Advanced spatial modelling with ArcView

8.1. Handling Grid data

An introduction to the concept of raster data in general and the Grid data format was given in chapter 3. The present chapter will focus, in section 8.2, on spatial interpolation, specifically on the generation of digital terrain models (DTM). Before approaching this subject, let us however become familiar with some of the most important operations for handling Grid data. These operations are summarized in the frame of exercise 8.1 on the next page.

Most of the steps conducted in the exercise 8.1 rely on the functionalities provided by the Spatial Tools extension, which is delivered with ArcView. The Spatial Analyst extension does not cover all the Grid handling functions one requires.

8.2. Creating a digital terrain model (DTM) in ArcView

8.2.1. DTM basics

A DTM, or Digital Terrain Model, is a map containing one piece of elevation information for each and every point on its own surface. This particularity distinguishes the DTM from a topographic map containing elevation information along linear, 2-dimensional elevation contours only (on topographic maps, the elevation between 2 lines has to be approximated by the user).

Several **characteristics** describe a DTM:

1. A DTM is a **computer-stored** representation. It is a type of spatial information processed and digitally stored as a file, or a set of files in a Geographic Information System (GIS).
2. A DTM attributes **elevation information** to each and every point of the area it covers.
3. The elevation information in a DTM is not measured in reality, but calculated from a preliminary data (topographic map, elevation contours, spot heights).
4. As the DTM is digitally stored, topographic information (elevation, landform, shape, slope, etc.), can be retrieved in subsequent steps.
5. In combination with other data sources, the DTM provides an ideal basis for numerous simulations and interpretations.
6. A DTM **represents reality** (in this case the topography of an area) as accurately as possible, while being as simple and handy as possible.

The **generation of a DTM** involves 5 main steps:

1. **Gathering and, if necessary isolating elevation information** from the preliminary data (e.g. topographic maps). Sometimes, the elevation information is available as a separate layer, either in analogue form (on paper, or transparent foil), or in digital form (as bitmap, or as vector data). The quality and density of the preliminary elevation information will determine not only the quality of the DTM but also the minimal cell size (resolution) with which the DTM can be calculated.

Exercise 8.1 – Handling Grid Data

Steps	Data	Remarks	Time
Loading Spatial Tools	---	- Make sure the Spatial Tools extension has been copied to the EXT32 folder of ArcView - Activate the extension in the File – extensions menu	10 min
Merging grids	Kemadtm_n Kemadtm_s	- We are interested in conducting analysis on the Njoro River. However, the DTM that is available for this area comes in two parts (north and south). - Merge, or mosaic both parts in order to come up with one grid	30 min
Calculating an integer Grid	Result from step two	- In order to reduce the size of your input Grid, calculate an Integer of the same using the map calculator in the Analysis menu. The request for the calculation of an integer is .INT and is positioned after the name of the input grid.	
Clipping a Grid	Result from step three Kenjwsh1.shp	- The resulting grid has an extent which is too big as compared with the area of interest (Njoro watershed). Therefore, create a new theme and digitize a rectangle, which will be used to cut out the portion of the DTM we are interested in. Save the shapefile in \mapping. - Open a new script and load the thm2gphx.ave script located in the \samples\scripts folder of ArcView. Run the script on the shapefile you have just digitized. - Set the Analysis extent to be equal to the rectangle that was digitized, select the grid to be clipped and carry out the clipping with the Spatial Tools.	30 min
Derive slope and hillshade	Result from step four	- Use the result from the previous step in order to calculate a hillshade and a slope map with the Spatial Analyst extension.	15 min.
Map calculator	Results from steps four and five	- For your field work you are interested in areas located above 2800 meters. Use the map calculator in the analysis menu to identify these areas. - You are also interested in identifying the areas located above 2800 meters and at the same time presenting a slope higher than 10 degrees. Carry out a second map calculation in order to identify these areas	30 min
Grid to polygon transformation	Results from step six	- For technical reasons you need to transform the grid showing the areas above 2800 meters into a polygon shapefile. Carry out the transformation with the Theme – convert to shapefile menu. - Some polygons are of insignificant size in the final product. Open the shapefile's attribute table, start editing and create a new attribute called "Area". Calculate its values with the request: [shape].returnarea. - Set the units in the View - properties menu and use the EditTools extension in order to eliminate the polygons with a surface too small to be of further use.	30 min
Total time			**145 min.**

2. **If the preliminary data is available in analogue form only, it has to be scanned**, either directly, or after having separated manually the elevation information from other, unwanted, information displayed on the map (land use, hydrology, infrastructure, etc.). The result of this step will be a digital bitmap. A bitmap is a raster data format consisting of cells arranged in rows and columns, according to the characteristics of the area described. Each cell of a bitmap has a value. The smaller the cells the more precise the information of the entire data set. When scanning a paper map showing elevation information, the resulting bitmap will most probably be **binary**, whereas the value 1 will be assigned to the foreground (the lines representing the elevation contours) and the value 0 will be assigned to the background.

3. **The bitmap** generated in step 2 **has to be vectorised** and errors inherent to the vectorising process have to be corrected as partly described in the previous chapter. Such errors especially show up in areas of the map, where elevation information is very dense and thus, lines are sometimes touching each other. Other sources of errors are mainly linked to additional information present on the preliminary data and which could not be properly separated from the elevation information. In order to avoid scanning and vectorising errors, the elevation information should not be too dense. However, it should also not be too distant, in order to avoid interpolation errors in the fifth step (see below).

4. **The elevation information has to be assigned to each vector**. Each vector of a spatial data set is an individual which can be addressed by the software and which is related to an attribute table containing information about all the vectors found in the data set. The elevation information will be stored as one information item in the attribute table of the data set.

The steps 2, 3 and 4 are shown in Figure 8.1 (step 5 follows below the figure). A picture is scanned and the resulting product is a raster file. Subsequently, the raster file is vectorised. The point information is transformed into vectorial (linear) information.

5. **Interpolating the elevation information** (= DTM generation). The DTM is generated by interpolation, using a GIS. Linear information (the elevation contours of the vectorised files described above), or punctual information are transformed into spatial information (DTM). In other words, the information gaps between the elevation contours are filled up with calculated (not measured) elevation information. When looking at a contour map, **a trained human eye** intuitively recognises the landscape 'underneath' the map. The landscape on the contour map is usually interpreted as a continuous surface; the eye immediately 'computes' elevation and slope for any spot on the map. The basic challenge when generating a DTM out of vectorised contour lines is to make the computer recognise the landscape underneath the contours; to make it calculate the elevation of any given spot on a particular surface. This is done by interpolation.

   The interpolation is done on a mathematical basis, through so-called algorithms. There are several algorithms which can be used all of them having their specificities, strengths and weaknesses. Some are more complex, taking into consideration the hydrologic information of an area in order to calculate the relief, such as the topogrid algorithm of ArcInfo (in ArcGIS: Topo to Raster tool in Spatial Analyst Toolbox).

8.2.2. The spline algorithm

ArcView offers **two options** for general spatial interpolation and more specifically for DTM calculation: The Inverse Distance Weighted (IDW) interpolator and the Spline interpolator:

The **IDW interpolator** assumes that each input point has a local influence that diminishes with distance. It weights the points closer to the processing cell greater than those farther away. A specified number of points, or optionally all points within a specified radius, can be used to determine the output value for each location.

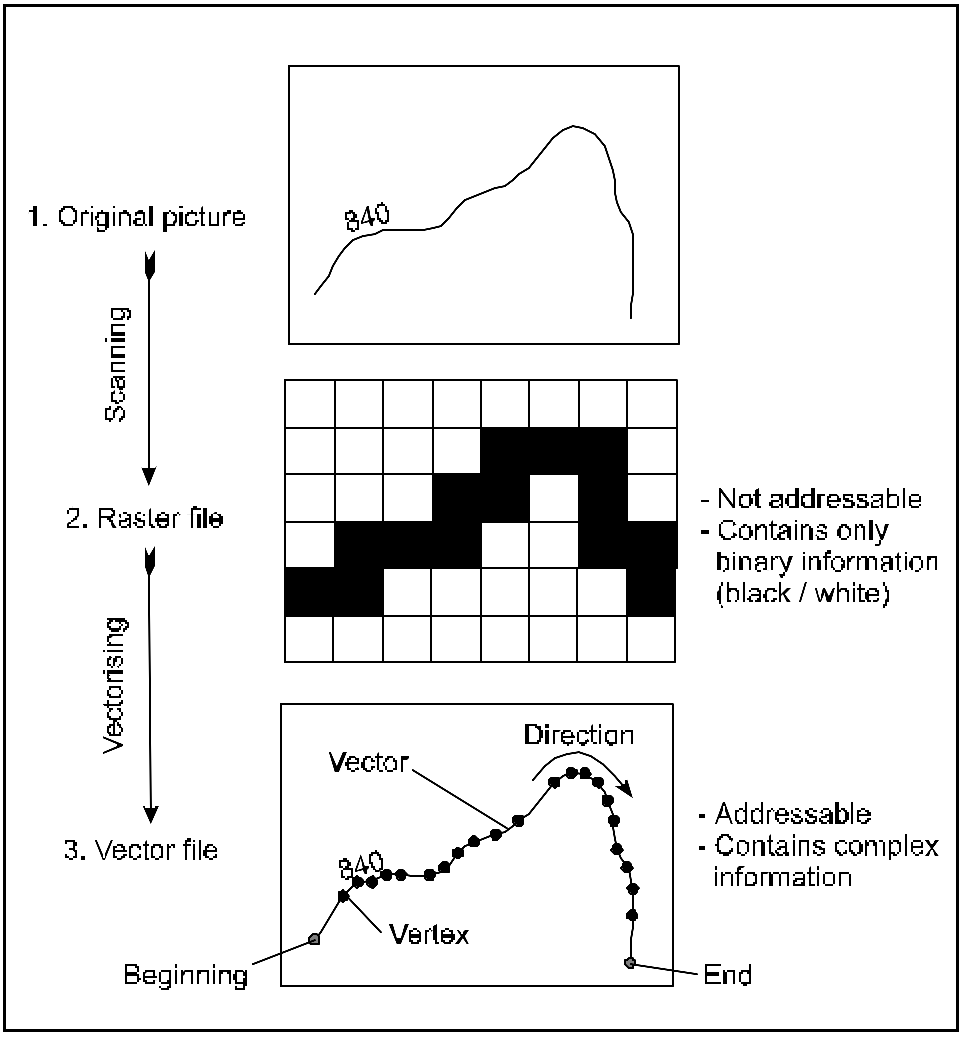


Figure 8.1: Scanning and vectorising input data for DTM generation.

The **Spline interpolator** is a general purpose interpolation method that fits a minimum-curvature surface through the input points. Conceptually, it is like bending a sheet of rubber to pass through the points, while minimizing the total curvature of the surface. It fits a mathematical function to a specified number of nearest input points, while passing through the sample points. According to ESRI's interpretation, "this method is best for gently varying surfaces such as elevation, water table heights, or pollution concentrations. It is not appropriate if there are large changes in the surface within a short horizontal distance, because it can overshoot estimated values". At CDE (University of Berne), Spline was used to calculate a huge DTM in Southeast Asia. The interpolator was found to give best results in areas with high relief energy. Mountainous regions offer a high density of elevation information, which facilitates the interpolation with Spline. The most problematic regions are those where areas with high information density (mountains) are located next to areas with low information density (plains). This problem can be solved by the addition of further elevation points in the areas containing only low information density, prior to calculating the DTM.

All in all Spline provides **rather satisfactory elevation and slope models** - if parameters are set appropriately - and it can be adapted to the particular situation in the project area. However, special attention should be given to areas where the information density changes abruptly (see above). The fact that Spline – as well as IDW - operates with **points only**, is a bit problematic when using ArcView. As a matter of fact, the default functions of ArcView do not allow to transform lines (elevation contours) into points (elevation points). Special scripts or extensions are necessary to carry out this operation (see next section). Some of these scripts yield unsatisfactory results, others are not easy to use. Another weak point is that spline sometimes is unable to fit a surface through the input points and therefore bails out. There is almost no way to find out why and where this error occurs, and thus, one is sometimes forced to divide the input data into several subsets, to run spline on each subset and therewith to try identifying where the problematic area is located.

8.2.3. Using Spline: Line to point conversion

Several scripts and extensions aiming at transforming lines into points have been tested at CDE. The most frequent weakness of these programmes is the definition of the spacing between points in the output theme. Some programmes do not allow any manipulation of this characteristic, others only offer limited options. The right spacing between the points is crucial for a satisfactory DTM calculation, as too narrow spaces will result in a "stairway" effect, whereas too broad spaces will lead to insufficient information density. As a rule of thumb, the distance between two points on the same elevation contour should be approximately equivalent to the distance between two adjacent elevation contours.

The most convenient among all extensions and scripts for line to point transformation tested at CDE is an extension called poly2pts.avx, which was programmed by William Huber and which can be downloaded from the ESRI homepage. The extension allows specifying a fixed distance between two points, or, when entering 0, the programme will set a point at the location of each vertex.

8.2.4. Using Spline: DTM calculation

The actual calculation of the DTM in ArcView does not require lots of manipulations. After having activated the relevant point theme, and made sure that the Spatial Analyst Extension is loaded, the Interpolate Surface option in the Surface menu can be selected. There, the interpolation method (spline or IDW) can be selected. When choosing Spline, the dialogue asks for the item containing the elevation information. The interpolation type has two options: The ***regularized*** method yields a smooth surface. The ***tension*** method tunes the stiffness of the surface according to the character of the modelled phenomenon. When you choose *regularized*, the weight parameter defines the weight of the third derivatives of the surface in the curvature minimisation expression. If you choose *tension*, the weight parameter defines the weight of tension.

Exercise 8.2– Mount Kenya DTM calculation

Steps	Data	Remarks	Time
Transforming elevation contours into elevation points	Kemkcnt1	- Load the poly to point conversion extension - Load the shapefile containing the elevation contours of Mt Kenya area - Transform the elevation contours into elevation points (decide on an adequate distance between elevation points).	30 min.
Calculating first DTM	Point theme calculated in the first step	- Make sure the spatial analyst is loaded - Calculate a 50 m resolution DTM using the correct elevation item, the regularized interpolation type, the default weight and 20 points.	30 min.
Calculating second DTM	Point theme calculated in the first step	- Calculate a 50 m resolution DTM using the correct elevation item, the tension interpolation type, a weight of 5 and 10 points.	30 min.
Calculate hillshades and make a first visual analysis	Both calculated DTMs	- Using the surface menu, calculate a hillshade for both DTMs, using the default settings. - Display the hillshades using a grey monochromatic ramp and 255 "equal interval" classes. Make a rough visual interpretation of both calculations. What are the main differences? What are the strengths and weaknesses of both methods? Which method would you choose and refine with more accurate tuning of weight and number of points?	30 min
Total time			**120 min.**

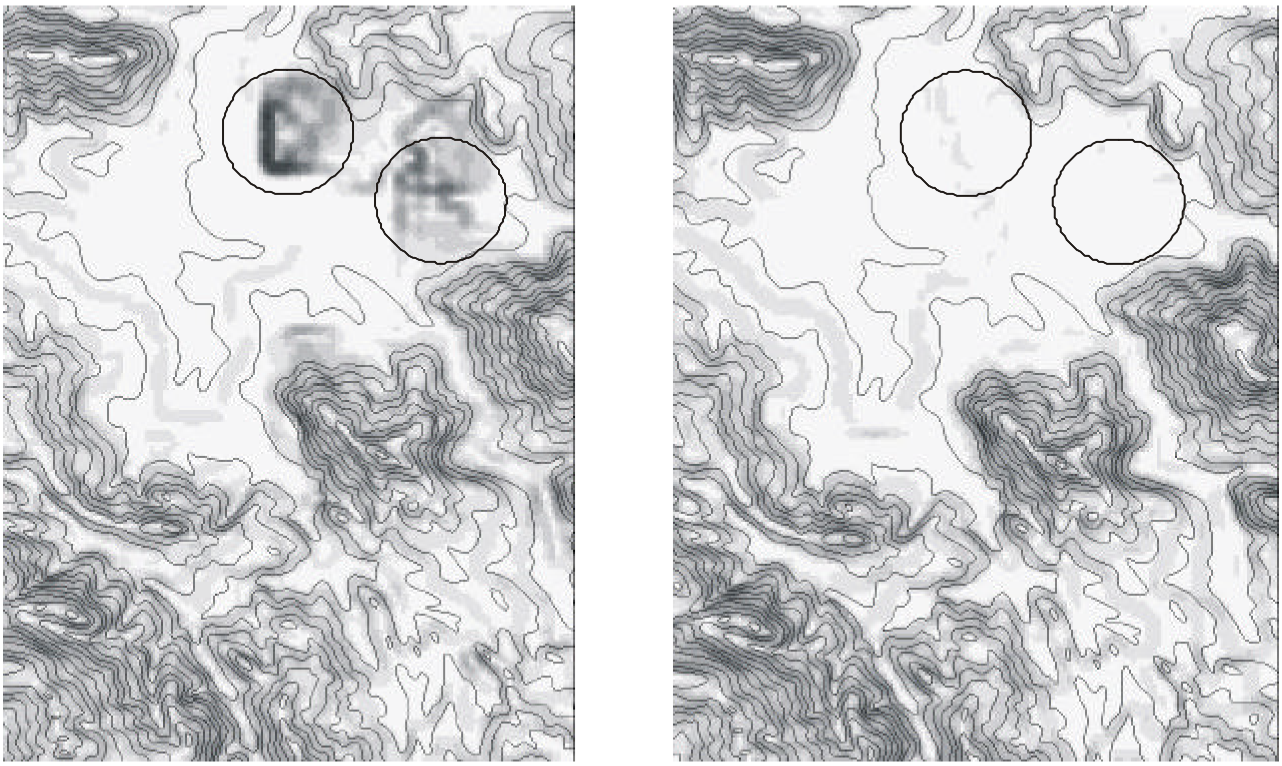
8.3. Comparing DTMs

8.3.1. Checking DTM quality

Digital Terrain Models are usually visualised by computing a **hillshade**, which provides a realistic view of the relief of the area covered by the DTM. The hillshade allows to localise rough errors, like wrongly labelled contours provoking unrealistic relief features (gullies or ridges), or wrongly oriented rivers, affecting the DTM calculation (see the figure in section 7.2.3). However, the hillshade does not provide sufficient control over smaller inaccuracies resulting from the interpolation method used, the overall quality of the input data, or the fine-tuning of the tolerances during interpolation.

A second control possibility is to check the **DTM's value range,** on the one hand in the legend and, on the other hand, by making a histogram of the DTM, with the histogram tool (in ArcGIS go to the Symbology tab in the layer properties, switch to "Classified" and press the "Classify"-button). Check if the values of the DTM are within the range of what can realistically be expected for the mapped area. Some interpolation overshoots or undershoots, resulting from abrupt changes in the information density might produce unrealistic elevation values. A **map query** from the analysis menu (in ArcGIS use the Raster Calculator from the drop-down menu of the Spatial Analyst's toolbar) displaying the areas above and below the realistically expected elevations allows to localise the problematic areas quickly and efficiently. Supposing that you do not expect any location in your study area to be higher than 2300 meters above sea level, or lower than 300 meters above sea level, you can formulate the following map query to find out locations that do not fulfil this condition:

([DTM] > 2300) or ([DTM] < 300)



Slope map from DTM calculated with:	Slope map from DTM calculated with:
Spline	Spline
Weight = 0.1	Wieght = 5
Points = 20	Points = 10
Type = Regularized	Type = Tension

Figure 8.2: Comparison of two DTMs - through their slope maps -, both calculated with the Spline algorithm, but with different calculation parameters.

Yet another way to control the DTM quality is to **derive slope values** from the DTM and to view the resulting slope map in a view. The slope map shows quite well whether the DTM interpolation led to a stairway effect or not, as the latter would lead to alternating steep and less steep stripes which could easily be identified optically on the slope map. Unrealistic depressions in plain areas as well as areas with extreme roughness values can also be identified, as for example in Figure 8.2: The regularized option on the left side led to two undershoots in the middle of a plain area (circles). These undershoots are due to the low information density in this area in combination with the “elastic” interpolation method of the regularized option, as opposed to the more “rigid” interpolation method of the tension option (on the right).

Slope values in degrees are easily calculated from the DTM, using the Surface menu of the Spatial Analyst extension. If, for some reasons, slope values have to be calculated in percents, the detour via map calculator (analysis menu) is necessary. Enter the following syntax into the map calculator dialogue window:

([DTM].slope (nil,true))

In ArcGIS: Slope([DTM])

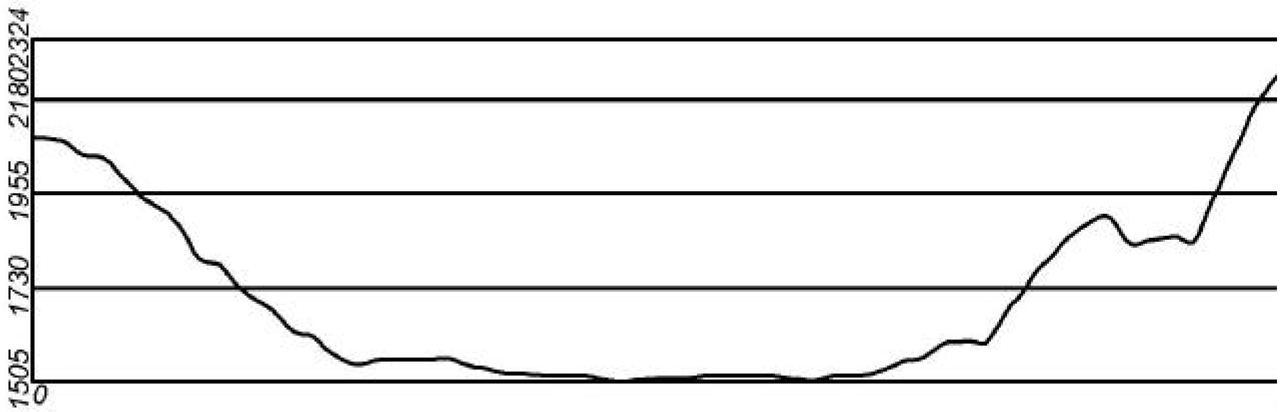


Figure 8.3: Profile through DTM calculated with the tension option

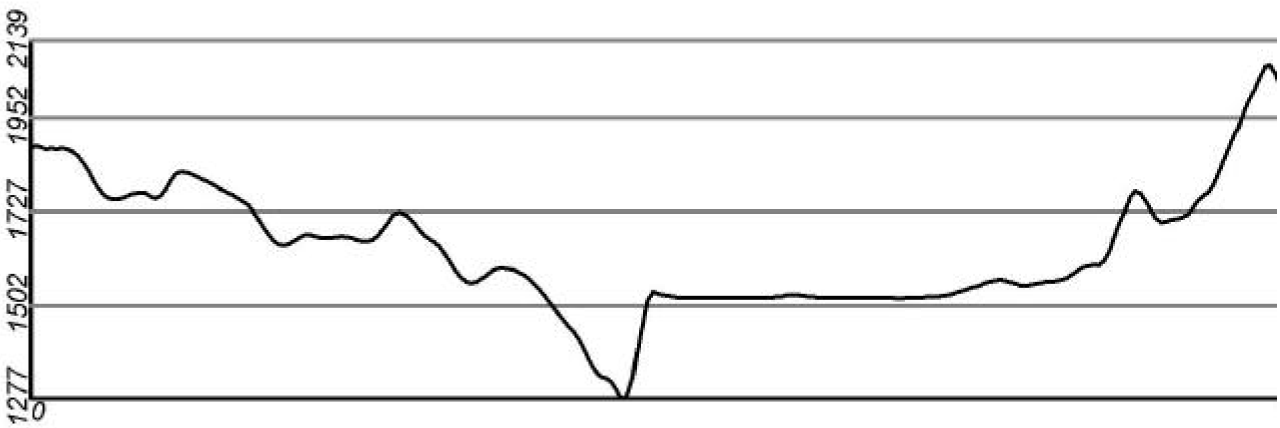


Figure 8.4: Profile through DTM calculated with the regularized option

Finally, a **profile** through selected parts of a DTM will help visualising the interpolation style and adapting it to the situation found in reality.

In order to **create a profile**, the 3D-Analyst extension must be activated and the DTM(s) from which the profile line(s) shall be calculated must be activated as well and displayed. Then, interpolate a graphic line with the interpolate line tool . Make sure the relevant DTM is activated when you draw this line. If you want to compare two DTMs, draw the line again, while the other DTM is activated. The interpolated line can be straight (two points, one at the beginning and one at the end), but it can also contain numerous vertexes and follow a specific path, a road, a boundary etc. Open a new layout and draw the box into which the first profile will be drawn by using the profile graph tool . Follow the instructions in the up-coming dialogue. Repeat the profiling, while the other interpolated line is being selected, either in the same layout, or in a different one. The figures 8.3 and 8.4 show two profiles through the plain shown in the slope map above. The profile in figure 8.3 reflects the result from the DTM calculated with the tension option and the profile in Figure 8.4 reflects the result from the DTM calculated with the regularized option. The abrupt depression in the middle of the plain is evident, as well as differences in the relief of the surrounding mountains. In these mountainous areas, the regularized option seems to offer a more diversified picture of the reality than the tension option.

Exercise 8.3 – Profile calculation

Steps	Data	Remarks	Time
Calculate a profile through the first DTM calculated in Exercise 8.2.	First DTM calculated in the Exercise 8.2	- Load the 3D-Analyst extension - Activate and display the first DTM - Draw a line with the interpolate line tool - Open a new layout and set the paper format - Draw a profile with the profile graph tool	20 min.
Calculate a profile through the second DTM calculated in Exercise 8.2.	Second DTM calculated in the Exercise 8.2	- Activate and display the second DTM - Draw a line with the interpolate line tool along the same track as the first line - Make sure the second line is selected - Open a new layout and set the paper format - Draw a profile with the profile graph tool, making sure to use the same vertical exaggeration factor as for the first profile	20 min.
Compare both profiles	Profile 1 and Profile 2	- Display both profiles next to one another and compare them - Draw conclusions on the way the spline algorithm interpolated the DTM with the different options	20 min.
Total time		-	**60 min.**

8.3.2. Comparing DTMs arithmetically

Another way to evaluate the characteristics and the quality of a spatial interpolation - whether it is a topographic interpolation, or a different topic (see section 8.4 below), as for example rainfall, population density, etc. – is to arithmetically combine the data layers to be compared.

Figure 8.5 was obtained by subtracting the regularized DTM from the tension DTM in the map calculator, as follows:

[DTM_tension] – [DTM_regularized]

The differences between both DTMs are displayed as standard deviations from the mean value. The red tones indicate positive standard deviations, i.e. areas where DTM_tension is higher (in meters above sea level) than DTM_regularized and blue tones indicate negative standard deviations, i.e. areas where DTM_tension is lower (in meters above sea level) than DTM_regularized. By observing the spatial distribution of the positive and negative standard deviation, relevant conclusions can be drawn concerning the interpolation characteristics of both DTMs.

It appears that DTM_tension is systematically higher than DTM_regularized in valleys and in some parts of plains. The opposite is true on ridges and also in some parts of plain areas. Figure 8.6

shows these differences as a profile. The regularized option allows for much more "swinging" of the interpolation than the tension option which interpolates in a much more linear way. The latter is definitely better suited for flat areas with low information density. However, its tendency to flatten hilltops could be a serious impediment for applications in mountainous areas.

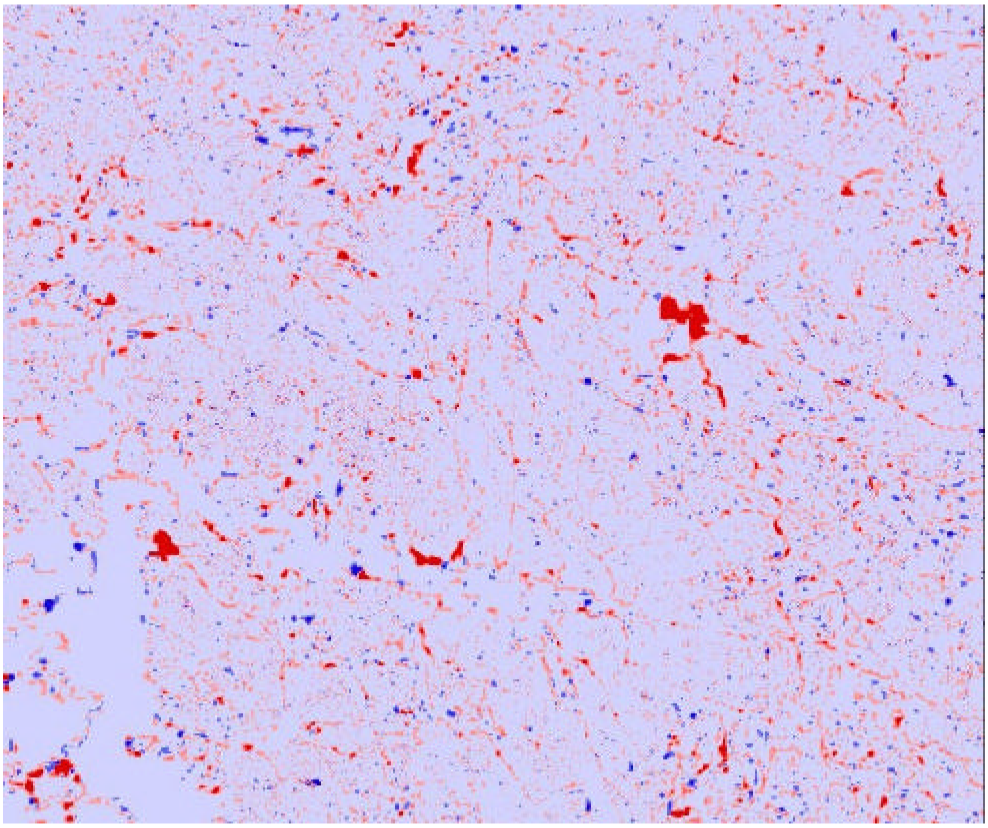


Figure 8.5: Arithmetic comparison between the regularized and tension DTMs.

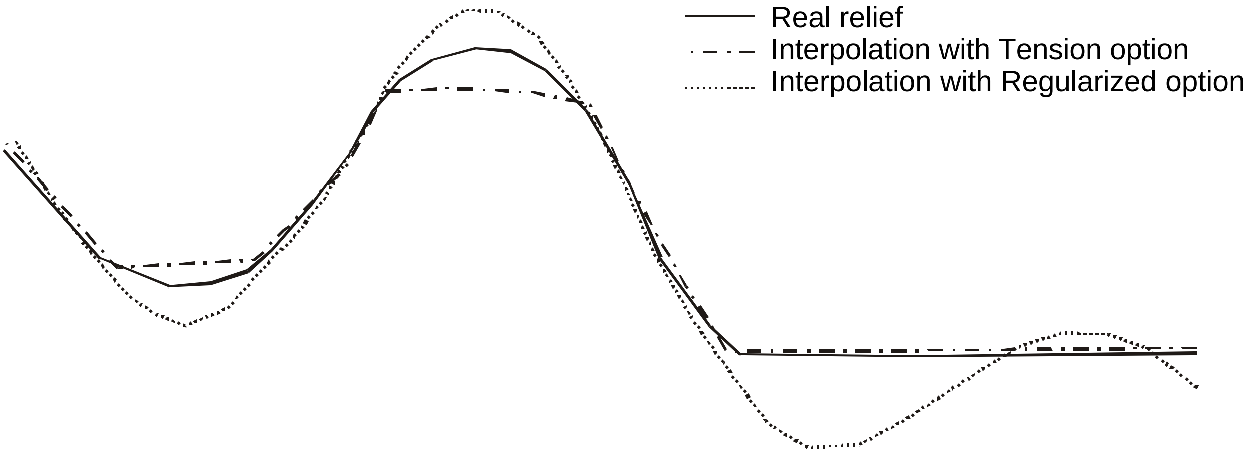


Figure 8.6: Idealised profile comparison between reality, tension and regularized algorithms.

From the above analysis, it is clearly noticeable, that both the regularized and the tension method have strengths and weaknesses. Depending on the use to be made from the DTM, one method will be used rather than the other one. A thorough investigation of the aims of the modelling, the expectations towards the DTM and the main area of activity (e.g. mountains rather than plains) will help in choosing the most satisfactory interpolation method.

8.4. Interpolating non topographic information

It has been mentioned earlier, that interpolation can be used in a much wider scope than just for the modelling of the topography. Interpolation can be conducted to gain exhaustive spatial information for any gradually changing themes having a spatial dimension and for which only punctual or linear information is available. In such a way, a rainfall map can be modelled from punctual rainfall measurements at meteorological stations and a pollution map can be interpolated from punctual pollution measurements. The quality of the interpolation will, of course, entirely depend on the quality of the measurements at the measurement stations and on the spatial density of such stations. Basically it can be said that the higher the measurement density, the better the interpolation will be.

In both cases, additional spatial parameters can be integrated into the model to refine it and improve its quality. A DTM for example would bring significant improvement to both the rainfall and the pollution models, as rainfall distribution and the spreading of pollutants highly depends on the relief. Winds, as well as many other data layers would also help improving such models. The question is always, whether this data is available or not.

The exercise below is carried out on the basis of a rainfall data set for Ethiopia. Due to the vast geographic extent to be interpolated the resolution at which interpolation can be carried out is 1000 meters. The point data used in the exercise 8.3 contains rainfall data from 226 meteo stations. The attribute table contains monthly and average annual data.

Exercise 8.4 – Interpolating rainfall data

Steps	Data	Remarks	Time
First interpolation of the point data	Etnarnf3	- The point coverage contains rainfall data for Ethiopia. The information density is very low. - Conduct an IDW interpolation on the point theme, using the total yearly rainfall, a cell size of 1000 m, 12 neighbours and a power of 2	20 min.
Second interpolation of the point data	Etnarnf3	- Conduct a second IDW interpolation on the point theme, with same parameters as above except power = 1.	20 min.
Comparing the results of the interpolation	Both rainfall grids calculated in the steps above	- Compare the rainfall grids and draw relevant conclusions concerning their quality and accuracy. What is the influence of the “power” option? Which version should be kept in such a case? - Which additional data could be added to the interpolation model to improve its accuracy?	20 min.
Displaying rainfall map with hillshade	rainfall grids calculated in the steps 2 and 3 etnahil1	- Add the hillshade of Ethiopia to the view - Find an adequate colour ramp for the rainfall grid and add the hillshade as a brightness grid.	10 min
Total time			**70 min**

CETRAD

9. Hydrologic modelling in ArcView

9.1. What are watersheds?

The practical definition of a watershed is a term used to describe the land area that is drained by a particular river, stream, or creek. When hydrologists use the term they refer to a drainage basin or area of land that discharges its surface waters through a single outlet or stream. A large stream like the Nile River can drain a huge land area and encompass a watershed of thousands of square kilometres. Most large watersheds are made up of many smaller watersheds associated with tributary rivers and feeder streams that contribute flow from areas known as sub-basins. Sub-basins are separated by drainage divides, topographically high places like ridge tops, where water will tend to run in a particular direction depending on which side of the divide the water is falling on. The **picture** hereunder shows a small watershed of a tributary to a larger river. The watershed boundary is indicated as a purple line.

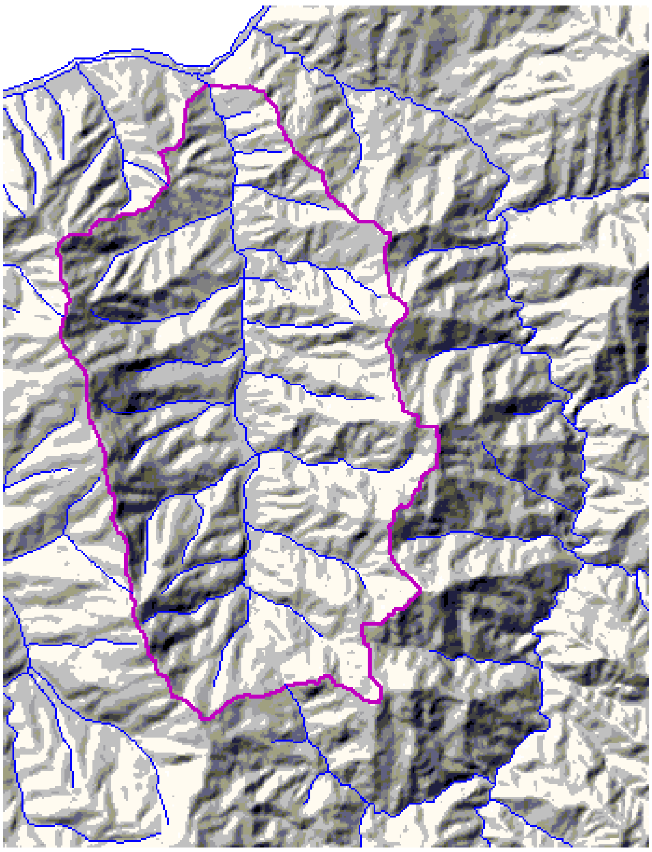


Figure 9.1: A watershed (or catchment) of a perennial tributary to a larger stream

9.2. Watershed management

In recent years watersheds have heavily gained importance as entities for planning activities and natural resource management. Watershed management has become a priority for many development projects.

According to the US American Environmental Protection Agency, watershed management, or, as they formulate it, the watershed approach "is a co-ordinating framework for environmental management that focuses public and private sector efforts to address the highest priority problems within hydrologically defined geographic areas, taking into consideration both ground and surface water flow". The goals of watershed management are oriented on the concerns expressed with a watershed focus. Depending on the concrete situation, these goals are therefore concentrating on aspects like increase of low water flow, reduction of high water flow, reduction of sediment transport and accumulation, or improvement of water quality. In more abstract terms, one could say that the goal of watershed management is to reduce risks and to increase potentials that are caused by water-related ecological processes.

Water resource management is a crucial issue for Kenya. Especially the conservation and retention of water during the rainy season for continued use during the dry season. Hence, the understanding of hydrologic processes and the availability of tools allowing conducting of various simulation and modelling of hydrologic parameters is of prime importance. In the present chapter we are going to see some of the possibilities offered by ArcView and to carry out some watershed calculation and dam simulation.

9.3. Hydrologic modelling extensions

There are at least two freely available extensions for ArcView, with which basic catchment modelling can be conducted: A **hydro.avx** extension and a **hydrov11.avx** extension. Apart from these two, some more powerful hydrologic modelling programmes are available on the market. Some of them, like for example the **Mike Basin** software, can be integrated into an ArcView environment and therefore take advantage of the full GIS capabilities of this platform. A short demonstration of the Mike Basin software will be done during the course. For practical exercise purposes, however, the course relies on freely available extensions, accessible to all ArcView users.

The following description of the hydrov11.avx extension is extracted from the **hydrov11.hdr** information file, which comes along with the extension.

This extension provides functionality to create watersheds and stream networks from a DTM, calculate physical and geometric properties of the watersheds, and aggregate these properties into a single attribute table that can be attached to a grid or shapefile. The extension has two ways of approaching this functionality. If you simply want to create watersheds, or stream networks, or a particular property of them, work directly with each of these functionality choices under the Hydro pulldown. If you want to create watersheds and calculate many additional attributes for them, use the Hydrologic Modelling choice under the Hydro pulldown.

Hydro pulldown menu:

Hydrologic Modeling *- launches a custom dialog for the creation of watersheds and their characteristics (see the "Hydrologic Modeling dialog" section below).*
Flow Direction *- computes the direction of flow for each cell in a DTM.*
Identify Sinks *- creates a grid showing the location of sinks or areas of internal drainage.*
Fill Sinks *- fills the sinks in a DTM, creating a new DTM.*
Flow Accumulation *- calculates the accumulated flow based on a flow direction grid.*
Watershed *- creates watersheds based upon a user specified flow accumulation threshold.*
Area *- calculates the area of each watershed in a watershed grid.*
Perimeter *- calculates the perimeter of each watershed in a watershed grid.*
Length *- calculates the straight-line distance from the pour point to the furthest perimeter point for each watershed in a watershed grid.*

Flow Length *- calculates the length of flow path for each cell to the pour point for each watershed if you choose Yes, or the length of flow path for each cell to the furthest perimeter point of each watershed if you choose No.*
Flow Length by Watershed *- calculates the maximum distance along the flow path within each watershed.*
Shape Factor by Watershed *- calculates a shape factor for each watershed. Shape factor is calculated as watershed length squared, divided by watershed area.*
Stream Network As Line Shape *- creates a vector stream network from a flow accumulation grid based on a user specified threshold.*
Centroid as Point Shape *- creates a point shape file of watershed centroids.*
Pour Points as Point Shape *- creates a point shape file of watershed pour points.*
Mean Elevation *- calculates the mean elevation within each watershed.*
Mean Slope *- calculates the mean slope within each watershed.*
Mean Precipitation *- calculates the mean precipitation in each watershed.*
Mean Curve Number *- calculates the mean curve number for each watershed.*

Hydrologic Modelling dialog:

The Hydrologic Modelling dialog is designed to be a quick 1-step method for calculating and then aggregating a set of watershed attributes to a single file. To use this dialog, you can start with an existing DTM or a shapefile of watershed boundaries.

To create watersheds from the DTM*, choose Delineate from DTM and select an elevation surface. You will be prompted to fill sinks. Though some* ***sinks*** *are valid, most are data errors and should be filled. You will then be prompted to define the* ***minimum number of cells*** *to define a watershed. This will create watersheds based on the number of cells, or up-slope area you have defined as being the smallest watershed you want. After adding the new watersheds as a theme, you will be prompted to use this for further analysis. If you select No, you will be asked to specify a new number of cells for defining basins. If you select Yes and have the 'Output as shapefile' option checked on, you will be prompted for an output shapefile name and the attributes you have checked will be calculated.*

To use an existing set of watershed boundaries*, choose Use Existing Watershed, select a corresponding elevation surface and watershed data source. You will be prompted for a key field for conversion, which will be the common attribute you will want in both data sets (most commonly, the Id field is chosen).*

The watershed attributes available in the Hydrologic Modeling dialog are the same as those described above, except for the addition of Stream Flow Length. Stream Flow Length is the longest flow distance on the stream network defined by the flow accumulation threshold used to define the minimum watershed size. By default, Stream Flow Length, Watershed Length, and Shape Factor are not checked on, because they are slow to calculate, especially for a larger number of basins.

This extension requires the Spatial Analyst v1.1 extension and Dialog Designer Extension. This extension will automatically load the Spatial Analyst if it is not already loaded.

The hydrov11.avx extension lacks one function which can be found in the hydro.avx extension and which allows to delineate a single watershed based on the location of a pour point (outlet, or point where all water drained through surface runoff passes through). In the following exercises we are going to work mainly with this extension.

In nearly all cases, you will also need to load and to activate a Digital Terrain Model in order to carry out hydrologic modelling.

9.4. Defining watersheds in ArcView

The hydrologic modelling exercises will be conducted on the watershed of the Siakuu river, which is a tributary to the Ewaso Ngiro River and is located partly in Laikipia and partly in Isiolo Districts.

9.4.1. Filling sinks in a DTM

In order to be able to correctly calculate watersheds, the DTM used to carry out the calculations needs to be free of sinks. Sinks are areas of internal drainage. In this sense, parts of the Rift Valley, e.g. the Lake Nakuru Basin are sinks. The sinks to be eliminated in our case however, are small areas of internal drainage caused by inaccuracies, or by misinterpretation of the elevation information through the algorithm during DTM calculation. Such sinks can be just a few grid-cells in size, but can also cover wider areas. Some algorithms (e.g. Topogrid of ArcInfo) have an option to identify such sinks during DTM generation and to write the identified sinks into a separate data layer.

9.4.2. Calculating a flow direction grid

The next step on the way to watershed calculation is to produce a flow direction grid from the filled DTM. A flow direction grid is a grid that indicates the direction into which surface runoff will flow on a cell by cell basis.

9.4.3. Calculating a flow accumulation grid

Finally, the hydrologic modelling extension also needs a flow accumulation grid in order to carry out proper watershed delineation. A flow accumulation grid shows for each cell the potential number of up-slope cells that will drain into this cell.

9.4.4. Delineating a watershed

All intermediate products required for watershed delineation are now available. The only thing that remains to be done is to instruct the extension about the data layers to be considered: In the first hydro menu select Properties and enter the names of the flow direction and flow accumulation grids in the dialogue window.

The delineation of a single watershed is done by activating the watershed delineation icon (W) and pointing in the view the location of the outlet (pour point).

Exercise 9.1 – Defining watershed of Siakuu River

Steps	Data	Remarks	Time
Preparing input data	Keewdtm1 Clip_siakuu	- The DTM at our disposition for the areas of interest covers the entire Ewaso Ngiro basin. If used as it is, some calculation steps will be quite slow. Therefore we clip the DTM using the clip_siakuu shapefile and the Spatial Tools extension. Don’t forget to set the Analysis extent before proceeding. Derive a hillshade from the clipped DTM.	20 min.
Filling sinks	Clipped DTM of Siakuu area	- Load both hydrological modelling extensions, run the fill sinks function, name and save the output correctly	15 min
Calculating a flow direction grid	Filled DTM of Siakuu area	- View the filled DTM of Siakuu area and try to identify changes from the original DTM. - Run the flow direction utility on the filled DTM - Name and save the output correctly	15 min
Calculating a flow accumulation grid	Flow direction grid of Siakuu area	- Run the flow accumulation utility on the flow direction grid - Name and save the output correctly	15 min
Delineating the watershed of Siakuu river	Flow accumulation grid Flow direction grid	- Delineate the watershed of the Siakuu river. Place the outlet on the co-ordinate 304,130 – 10,067,100	15 min
Total time		-	**80 min**.

9.5. Simulating a water reservoir

GIS is a powerful visualisation tool. In the field of hydrologic modelling, the visualisation of **scenarios for water reservoir construction** is one of these visualisation possibilities. In the present section, the participants will simulate the construction of a water dam on Siakuu river. It is important to know before starting this exercise, that **the simulation is entirely fictive and is not supported by any feasibility study**, neither from the point of view of sufficient water drainage in the catchment of Siakuu river (as the river originates from the lower slopes of Mount Kenya, in a semi-arid area, the water potential is most probably insufficient for any project of that type), nor from an engineering point of view, nor from the point of view of soil and geologic suitability, nor from the point of view of potential uses. The exercise only aims at showing the application possibilities in the field of scenario visualisation.

9.5.1. Calculating a DTM including the lake surface

The first operation to be carried out is to locate the ideal location for the dam. In the present case, it will be where Siakuu River crosses the coordinates 306,180 / 10,044,000. Next, the catchment area located upstream of this point has to be calculated in the same way as was explained in exercise 9.1. Then, the maximum height of the dam, or rather the elevation of the water surface after completion of the dam should be defined by identifying the elevation of surrounding relief elements (the water surface can not be higher than the lowest ridge, or pass in the reservoir area. In our case this elevation has been identified as being 1600 meters above sea level. Next, a DTM (hereafter referred to as DTM2) with a uniform elevation equivalent to the projected water surface elevation (1600 meters) has to be calculated from the sub-catchment theme lying upstream of the dam site. As that theme is in grid format, a conditional statement in the map calculator will be necessary to obtain a grid with a value of 1600 for all cells located within the sub-catchment and a very low value for all other cells (e.g. 10). By specifying a very low value for the area outside the catchment, one ensures, that in this area, the values of DTM1 (see figure 9.1) will be used, even if there are areas lower than the projected water surface elevation (which is the case for all the areas close to the bottom of the valley and downstream of the dam site). Finally, the lake surface is integrated into the area's DTM (DTM1) with a conditional statement in the map calculator, which specifies that:

“If DTM2 is higher than DTM1, then return the values of DTM2 and
otherwise return the values of DTM1”

After finishing the map calculation, a hillshade of the newly created DTM, as well as a reclassified elevation grid can be calculated for display purposes. Furthermore, a simple map query with the statement

new_DTM = “elevation of lake surface”

will return a value of 1 for the lake area and a value of 0 for the rest of the data set. By choosing a blue colour for value 1, the lake can be displayed realistically, both in the view and in a 3D scene.

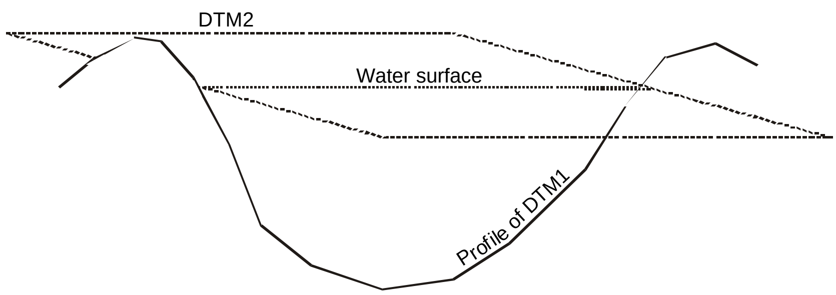


Figure 9.2: Graphic explanation for the simulation of a water reservoir on the basis of a Digital Terrain Model (DTM).

Exercise 9.2 – Simulating a water reservoir

Steps	Data	Remarks	Time
Calculating sub-catchment up-stream of dam site	Flow direction and flow accumulation grids calculated in previous exercise	- The input grids for watershed delineation were calculated in the previous exercise. Use them to calculate a watershed the pour point of which is at 306,180 / 10,044,000. - Save the output as a grid.	30 min.
Creating horizontal DTM	Sub-catchment calcu-lated in previous step	- The sub-catchment grid most probably has a value of 0 for the area located within the sub-catchment and a value of "No Data" for all other parts of the grid. - Use the following conditional statement in the map calculator to change the values to 1600 (inside the catchment) and to 10 (out-side the catchment): ([Subcatch].isnull).con (10.AsGrid, 1600.AsGrid). Whereas Subcatch is the grid calculated in the first step. - Save the result of the map calculation as lake_dtm	30 min.
Calculating new DTM including lake surface	Filled DTM of Siakuu area Lake_DTM	- Start the map calculator and enter the follow-ing conditional statement to return the values of the DTM if the latter is higher than the lake surface and to return the value of the lake surface if the DTM is lower than the lake surface: ([DTM_Siakuu] > [Lake_DTM]).con ([DTM_Siakuu], [Lake_DTM]) - Save the result as Siakuu_dtm_lk	20 min.
Calculating hill-shade	Siakuu_DTM_lk	- Create a hillshade with the newly created DTM and control the result - Save the hillshade as Siakuu_hil_lk	10 min.
Map query	Siakuu_DTM_lk	- Conduct a map query to return the areas exactly located at the elevation of the lake surface	5 min.
3D representation	Siakuu_DTM_lk Siakuu_hil_lk	- Create a 3D scene with Siakuu_DTM_lk and Siakuu_hil_lk. Use a z-factor of 2 for the dis-play. Navigate in the 3D view until you have a nice overview. - Enhance the resolution of the 3D View with the 3drasres.ave script to be found in the samples folder. - Save the 3D view as a bitmap and print it out	45 min.
Total time		-	**140 min**.

9.5.2. Calculating water volume

The visualisation of the water reservoir alone does not allow exhaustive quantitative description. The elevation of the lake surface, the approximate length and width of the lake can be described from the view by using the information tool and the measuring tool. The total area of the surface can be calculated as follows:

1. With the map query option of the analysis menu, query all areas of the newly created DTM (the one containing the reservoir volume) which are exactly at the elevation of the water surface (see above)
2. Transform the resulting grid into a polygon theme
3. Open the theme table, start editing and add a new field called "area"

4. Calculate the new field with the following statement: [area] = [shape].returnarea

The result of this calculation is given in squared map units. Alternatively, and maybe more easily, you can also calculate the surface of the lake by querying as in step 1., then by opening the Grid table and promoting the selection to the top of the table. The number of cells selected and the size of each cell can be used to arithmetically calculating the surface of the lake in square meters or kilometres.

In order to calculate the volume of the reservoir, two ways can be followed:

The manual way:

1. Clip the original DTM (without reservoir) with a polygon shapefile representing the lake surface.
2. Do the same with the flat DTM equivalent to the elevation of the water surface.
3. In the map calculator calculate the difference between lake surface and relief underneath.
4. Make sure the resulting grid is in integer format (no decimals). If this is not the case, calculate an integer grid in the map calculator using the following syntax: [InGrid].int (in ArcGIS: Int(InGrid)).
5. Add a new field to the attribute table of the integer grid and calculate the new field = [value] * [count]. Then get the field statistics and write down the sum (In ArcGIS go to layer properties of the grid layer. The statistics are available on the source tab). If the map units are meters and if the grid resolution is 50 meters, the volume of the reservoir can be calculated as follows:

[sum] * 50 * 50

Using CutFill:

1. Clip the original DTM (without reservoir) with the polygon shapefile representing the lake surface.
2. Do the same with the flat DTM equivalent to the elevation of the water surface.
3. Load the 3D-Analyst extension, make both clipped grids active and select CutFill in the surface menu (in ArcGIS: 3D Analyst Toolbox).
4. Chose the flat DTM (lake surface) as before grid
5. Open the attribute table of the newly created cut-fill grid, select the volume field and conduct statistics on this field (field – statistics). The sum in the statistic report is equivalent to the lake volume in map units.

Exercise 9.3 – Volume calculation

Steps	Data	Remarks	Time
Volume calculation 1	Original DTM (without lake) and flat DTM (lake surface) Map query (lake area)	- Calculate the lake volume with the manual method, using the instructions above	30 min.
Volume calculation 2	id.	- Calculate the lake volume with the CutFill method, using the instruction above.	30 min.
Comparing volume calculation	Results from steps 1 and 2	- Compare the results of both calculations above. Do the results match?	10 min.
Total time			**70 min.**

CETRAD

10. The Network Analyst

10.1. What is the Network Analyst?

AV-help

The ArcView Network Analyst can solve common network problems on any theme containing lines that connect. This theme can be a **shapefile**, an **Arc/Info coverage**, or a **CAD drawing**. Before solving a problem, you can model networks precisely, including setting up average travel times, one-way streets, prohibited turns, overpasses and underpasses, and closed streets. The main functions of the Network Analyst are:

a) Find efficient travel routes
Find the best way to get from one location to another, or the best way to visit several locations. You can specify the locations by pointing to places on a line theme, by entering addresses, or by using a point theme. You can decide the order they are visited, or you can let the Network Analyst find the best visiting sequence for you.

b) Determine which facility or vehicle is closest
Find the closest facility to any location on a network. The Network Analyst identifies the closest facilities and displays the best way to get to or from them. To get these results, you need only specify a location on a line theme and the name of the point theme representing the facilities.

c) Generate travel directions
You can generate sophisticated, easy-to-use travel directions for any route you find with the Network Analyst, whether it's a route between two locations, one that visits several locations, or a route to the closest facility. You can decide how to report travel times and distances, how to report streets along the way, and whether to incorporate landmarks. Once generated, directions can be displayed, edited, printed, and saved for future reference.

d) Find a service area around a site
The Network Analyst provides 2 tools that allow you to learn what is near a particular site, service networks and service areas. Service networks identify the accessible streets within a specified travel time or distance via the road network. Service areas identify the region that encompasses the accessible streets. Once you have a service network or service area, you can use ArcView's theme on theme selection capabilities to evaluate the accessibility of the site.

e) Customize your work
All of the problems you solve with the user interface can also be solved by writing and running Avenue scripts. When you install the Network Analyst, you get several new Avenue classes and requests for solving network problems. These classes and requests can be used to automate tasks, add new capabilities, and build applications.

10.2. Input data requirements

When working with the Network Analyst extension, special attention needs to be given to the **topology** of the data layer used. When working on an **Arc/Info line coverage** the geometry of which matches the specific requirements of the network task to be performed, no specific problems should be encountered. It is however important to insist that the provider of the Arc/Info coverage runs the BUILD LINE command followed by the RENODE command on that line coverage and carefully checks that all lines that need to be connected (snapped) according to the logic of the network are really connected. Otherwise it can happen, that the network analyst can not perform a best route calculation on the coverage, unless it is being transformed into a shapefile.

If the network analysis needs to be performed on a **line shapefile** which has to be digitised in ArcView, the following possible pitfall needs to be considered during digitising: When a line which was digitised with the 'draw line' tool snaps into an existing line, the latter is not split at the point where it was met by the new line. In such cases, the Network Analyst extension is unable to perform a route calculation. When trying to calculate a best route, passing through such a point, the dialog will prompt 'not reached' and no best route will be indicated. There is a snap.avx extension at the ESRI ArcScripts website that is useful for editing in ArcView. However, this still does not resolve the network routing problem. Although the vertex is snapped to another arc, network analyst does not recognise this connection and a route does not get built. To avoid this problem use the 'draw line to split feature' tool instead of the 'draw line' tool and slightly overshoot the line to be snapped. Then delete the dangle of the overshot (see Figure 8.2.1). With this method, the initial line is split at the spot where it was crossed by the new line and network analysis can be conducted normally. Another possibility, if you have acquired the Edit Tools, is to carry out a "Self Intersect Check" in the clean option of this Extension (note: the shapefile needs to be in editing mode before carrying out this option). The intersection check will split the arcs at intersections and thus, allow the network analyst to find a route.

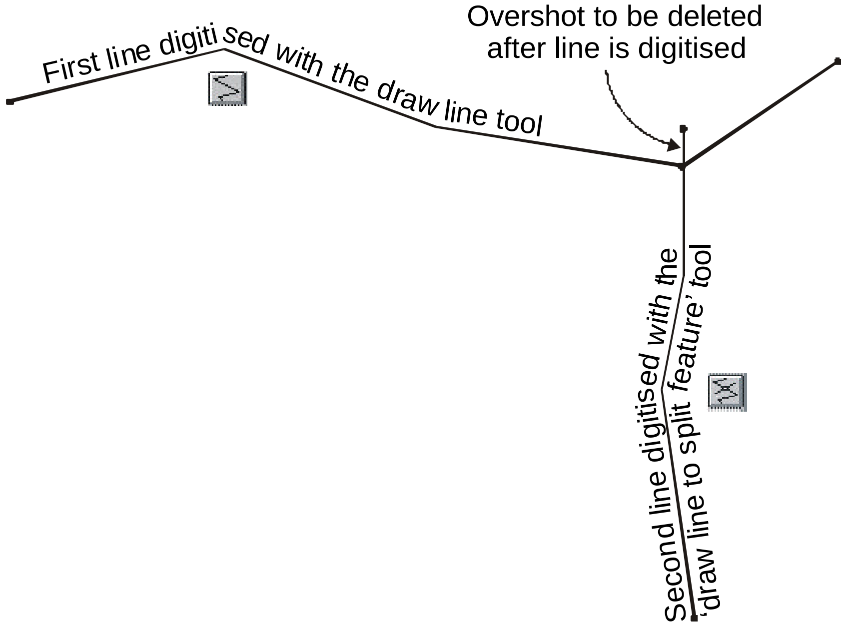


Figure 10.1: Digitising lines with the 'draw line to split feature tool' in order to enable correct network calculations.

10.3. Calculating a best route

10.3.1. Defining travel costs

To find the shortest route between two points is a question which can be relevant in many fields. For the planning of a public transport system, for the organisation of an electrical distribution network, for faster emergency intervention in urban areas, etc. Distance in this context can be understood as simple metric distance, or as a combination of metric distance, time and financial effort. The distance from A to B might be shorter via road 'a' than via road 'b', but road 'a' might be unpaved and therewith allowing a travel speed only half as fast as on road 'b'. Thus the way via 'a' might end up being longer in terms of time, despite being shorter in terms of kilometres. Similarly, travelling on highway 'd' might be shorter than travelling on either 'a', or 'b', both in terms of time and kilometres, however, the highway toll fee might cause the journey via 'd' to be more expensive than the two other routes.

The above example shows, that there are different ways of perceiving transportation efficiency, depending on the priority given to distance, time and costs. Other rules might influence the selection, or planning of a route as well:

AV-help

One-way streets:	Streets that can be travelled in one direction only.
Prohibited turns:	Turns that are not allowed. For example, it may be prohibited to make a U-turn at one intersection and to make a left turn at another.
Overpasses and underpasses:	A street that passes over or under another street. You cannot make a left or right turn when on an overpass or underpass.
Closed streets and other streets to avoid:	Streets currently closed to traffic or certain types of streets to avoid. For example, you might want to avoid residential streets when routing a truck through a city.

All these parameters can be simulated in ArcView, simply by defining them in the attribute table of the shapefile used to conduct the network analysis. Working with network analysis problems requires you to have a **numeric value representing the cost of traversing each line feature**. Cost values allow the Network Analyst to find the best route through the network. This cost can be the **length** of the line feature, the **time** it takes to traverse it, or **any other system of measurement**. Fields containing costs must be named as in the tables below, otherwise the Network Analyst won't recognise them.

When setting up travel cost in your line theme, you first need to decide whether to set the **travel cost to be different for each direction** of travel on a street. In many cities, particularly large congested ones, it often takes longer to travel in one direction along some streets than it does to travel in the opposite direction. This is typical during rush hour traffic. However, at other times in these same cities, or at all times in smaller cities, there is not enough congestion to affect travel times. When this is the case, it takes approximately the same amount of time to travel in one direction as it does in the other.

If it takes approximately the same amount of time to travel in one direction as it does in the other, you can either use the default or use a specially named field that's in your line theme feature table or in a table joined to it. Use the default if you want to use the length of each line feature as the travel cost. The Network Analyst always calculates the length of each line feature using the theme's co-ordinates, and stores this information in a special directory it accesses when solving a problem. If you're going to use the default, you should set the view's map units if you haven't done so already. This will enable you to work with the interface in meaningful distance units, such as kilometres or miles.

AV-help

If you want to represent travel cost in other units, or in units of time, add a numeric field to your line theme feature table or to a joined table. Use the field to store the travel cost for travel along the line in both directions.

To store travel cost in these units	Name the fields
seconds	SECONDS
minutes	MINUTES or DRIVETIME or IMPEDANCE or TRAVELTIME
hours	HOURS
millimetres	MILLIMETERS
centimetres	CENTIMETERS
meters	METERS
kilometres	KILOMETERS
inches	INCHES
yards	YARDS
feet	FEET
miles	MILES
nautical miles	NAUTICALMILES
non-time/non-distance units	COST or UNITS

If it takes a different amount of time to travel along some streets in one direction as it does in the opposite direction, add two numeric fields to your line theme feature table or to a joined table. These fields are known as directional cost fields because each of the fields represent travel cost in one direction only. Use one field to store the cost of travelling from the start of the line to the end of the line, which is the same direction the line was digitised. The names of these fields must be prefixed with FT_ or FT-. Use the other field to store the cost of travelling from the end of the line to the start of the line, the opposite direction the line was digitised. Always prefix the names of these fields with TF_ or TF-.

AV-help

To store travel cost in these units	Name the fields
seconds	FT_SECONDS and TF_SECONDS
minutes	FT_MINUTES and TF_MINUTES, or FT_DRIVETIME and TF_DRIVETIME, or FT_IMPEDANCE and TF_IMPEDANCE, or FT_TRAVELTIME and TF_TRAVELTIME
hours	FT_HOURS and TF_HOURS
millimetres	FT_MILLIMETERS and TF_MILLIMETERS
centimetres	FT_CENTIMETERS and TF_CENTIMETERS
meters	FT_METERS and TF_METERS
kilometres	FT_KILOMETERS and TF_KILOMETERS
inches	FT_INCHES and TF_INCHES
yards	FT_FEET and TF_FEET
feet	FT_YARDS and TF_YARDS
miles	FT_MILES and TF_MILES
nautical miles	FT_NAUTICALMILES and TF_NAUTICALMILES
non-time/non-distance units	FT_COST and TF_COST, or FT_UNITS and TF_UNITS

Notes: A hyphen (-) may be used in place of an underscore (_) in all field names.

Use a cost value that would for sure cover all routes on your actual vector data and calculate the service area. Like that you'll find out the lines not connected to the rest of the network.

You can also use non-standard cost field names by setting aliases for them in the Table Properties dialog. For example, if you have fields named FT_TIME and TF_TIME, you could use them as cost fields if you set alias names of FT_MINUTES and TF_MINUTES for them.

AV-help

Exercise 10.1 – Calculating travel costs and travel time for Nakuru roads

Steps	Data	Remarks	Time
Preliminary remarks		The road layer of Nakuru was digitised on the basis of a Quickbird satellite image and is differentiated into five road types. Travel speed on these roads is not uniform. The following average travelling speed for each road type has been defined: **Road type** **Road code** **Speed** Main road: A 80 Secundary road: B 80 Small streets: C 60 Unpaved streets: D & T 50	
Calculating travel costs	Nk_roads	- Set the units in the view properties - Start the Edit Tools extension, start editing the shapefile, go to the clean menu, tick the select regular box define a tolerance and analyze the layer. Make sure that there is a regular node at each intersection and that there are no unwanted dangling nodes. - Open the theme's attribute table - Add a field called speed - Query the entries type by type and enter the speed for each type - Add a field called cost. This field is the one that will be used for the best route calculations - Select all entries and calculate the cost field as follows: [cost] = [length] * (1 / [speed] * 100) - Save the table edits	30 min.
Calculate travel time for road segments	Nk_roads	- Open the theme's attribute table - Add a new field called hours (see table above). Make sure to allow 3 decimal places - Calculate the new field with the following syntax: [hours] = ([length] / 1000) / [speed] The new field should now show the time necessary to travel on the various road segments considering the pre-defined average travel speed	30 min.
Total time			**60 min**.

10.3.2. Calculating best routes and service areas

The cost and travel time calculation per road segment which was conducted in the frame of exercise 10.1 allow us now to conduct best routes and service area calculations on the shapefile containing the roads of Nakuru. These calculations will be carried out in the frame of exercise 10.2. Before starting the exercise however, let use have a short look at the tools provided by the network analysis extension.

To start with, make sure the extension has been copied to your EXT32 folder and make it active in the file – extensions menu. A new menu option called “network”, as well as two new icons (the ‘solve network problem tool’ and the ‘add location tool’) appear on the GUI. The Network menu offers three options:

1. Find best route
2. Find closest facility
3. Find service area

When choosing the **find best route** option, a new theme is automatically opened and given the name Route#, whereas # stands for a number depending on how many routes have already been calculated and are still stored in the working directory of ArcView (usually c:\temp).

When choosing the **find closest facility** option, a new theme is automatically opened and given the name Fac#, whereas # stands for a number depending on how many closest facilities have already been calculated and are still stored in the working directory of ArcView (usually c:\temp).

When choosing the **find service area** option, two new themes are automatically opened and given the names Sarea# (service area) and Snet# (service net), whereas # stands for a number depending on how many closest facilities have already been calculated and are still stored in the working directory of ArcView (usually c:\temp).

The handling of all three options is similar, therefore, only the first option is described in details below. The exercise that will follow will, however involve all three options.

To find the best route on a line shapefile, open the shapefile in a view and make sure the topology of this shapefile is suitable to the implementation of a network analysis. You can control the topology of the shapefile by using the clean and intersect options of the EditTools2.3 extension. After having activated the network analysis extension click on the network menu and select the ‘Find best route’ option. A window with the name of the route to be calculated (e.g. route1) is opened. Next, click on the add location tool () and select two locations on the line shapefile you would like to be linked by a best (shortest route). The locations are written down in the route window and labelled graphic pick 1 and graphic pick 2. When through, check the properties menu, you’ll see that the <line length> is being used as a cost item to calculate the shortest route. Click on the ‘solve network problem tool’ () in the route window. The shortest route (shortest regarding distance only) is calculated and the route cost (again as a metric distance) is indicated in the ‘cost’ column of the route window. If your shapefile contains a cost field different from the line length and you wish to calculate the best route on the basis of this field, select the latter in the properties menu before running the solve network problem tool.

Exercise 10.2 – Calculating best routes and service areas according to different cost criteria

Steps	Data	Remarks	Time
Finding shortest distances	Nk_roads Nak_qb_2m.tif	- Find the shortest way (line_length) between Pipelines (Elementaita Road) and Prisons. - Find the shortest way between the New Sewage Treatment Plant and Flamingo Bottlers Ltd.	20 min.
Finding fastest way	Nk_roads Nak_qb_2m.tif	- Repeat both route calculations by using the hours field instead of the line_length field. - Check if the same route was calculated or not. - Calculate again the best routes by using the cost field. Check whether they match the routes calculated with the hours field.	20 min.
Finding service area	Nk_roads Nak_qb_2m.tif	- Check which part of town can be reached from the central police station, by car, within 5 minutes, assuming, there are no one-ways and no traffic jam. - Considering that average walking speed is 3 km per hour, calculate which part of town can be reached from the Post office within half an hour.	30 min.
Total time			**70 min**.

CETRAD

11. The Image Analysis Extension

11.1. Introduction to the Image Analysis extension

AV-help

Imagery of the earth's surface, both aerial and satellite, has become an integral part of today's GIS and desktop mapping systems. The Image Analysis extension allows ArcView GIS users to utilise the digital image data that is available today, as well as the data that is planned for in the near future.

With the Image Analysis extension, you can perform many tasks:

- **Import and incorporate** raster imagery into ArcView.
- **Categorise** an image into a number of classes corresponding to land cover types, like vegetation.
- **Evaluate** images at different time periods to identify areas of change.
- **Identify** and automatically map a land cover type with a single click.
- Find areas of **dense** and thriving **vegetation** in an image.
- **Enhance the appearance** of an image by adjusting contrast and brightness or by applying histogram stretches.
- **Align** an image to a map co-ordinate system for precise area location.

What you can do with the Image Analysis Extension:

The software is designed **primarily for natural resource and infrastructure management** applications. These applications include **forestry and agriculture, environmental assessments, engineering and infrastructure management** projects like facility sitting and corridor monitoring, and general GIS map database update and maintenance.

With the Image Analysis, you can perform tasks that range from simply displaying images to performing detailed **spectral analyses** and detecting **temporal change**.

The **tools** available in Image Analysis are designed to support a complete imaging workflow in ArcView GIS, and include:

- Access to industry-standard image data types and formats with simple-to-use importers and direct read/write utilities.
- Fast image display and manipulation for productive use of large image files within an ArcView GIS project.
- Three levels of image enhancement capabilities created to optimise image visualisation and interpretability for any given project. These functions let you enhance the contrast and brightness of your images, apply standard histogram stretches, or apply custom stretches to enhance specific data or features of interest.
- Spatial enhancement filters designed to optimise interpretability and output of both continuous and thematic imagery. These functions will let you sharpen or smooth imagery to suit your interpretation and output needs.

- Image-to-map rectification tools for co-registering aerial and satellite imagery to a project map base. This will also let you perform image shifts and local registrations.
- Analysis capabilities for multi-spectral categorisation, vegetation mapping, automated feature mapping and temporal change mapping with both continuous and thematic imagery. These functions let you easily incorporate imagery into any project, extract usable information, and integrate the results into a project.
- ArcView GIS integration utilities of input and output of GRID files for use in Spatial Analyst, integration of shapefile information, and integration of Image Analysis results into the ArcView GIS database.
- Customisation with Avenue for special application development."

AV-help

11.1.1. What are images

As you work with geographic applications, you will probably encounter many kinds of images, from pictures of buildings to satellite images of the earth. All of these can be grouped into two categories: **images associated with features in a geographic space and images that occupy the geographic space**. A picture of a building would fit in the first category, as the building would be represented as a point feature in a coverage. An aerial photograph, on the other hand, fits into the second category, as it covers a recognisable area. Some images can actually fit into either category depending on the scale. An architectural drawing would fit into the first category on a map of a city. On this map the building would only be a point. If you were working on a map showing the detail of a city block or building site then the architectural drawing fits in category two and could be used as a background to the other features on the map.

Images from both categories are typically produced by optical or electronic devices, and thus stored as raster data. That is, the data about an image is laid out in a grid structure, where each cell or pixel is referenced by a particular row and column value. Features on a map, however, are stored as vector data, discrete x, y co-ordinate pairs.

For images in the first category, actual spatial extent is not relevant. The other types of images, however, since they are directly related to a geographic space, must first be geo-referenced to real-world co-ordinates before they can be used with any other spatially referenced data.

11.1.2. How to enhance an image with band combinations?

Much of the data you work with using the Image Analysis extension is made up of multiple bands of data. While the Image Analysis extension brings images into the view with a default band setting, you have the option of changing that setting at any time. You can see the band and display colour combinations by looking below the theme's title in the table of contents.

To change the band combinations of an image, you need to access the Image Analysis Legend Editor. There, you can not only change the bands assigned to the red, green and blue output colors of the display, you can also check to select and deselect bands from being displayed in the view.

11.1.3. How to change band combinations?

1 Double click the theme's title in the table of contents to open the Image Analysis Legend Editor.
2 In the right hand portion of the dialog, use the dropdown lists for Red Band, Green Band and Blue Band to select the band you want to display in conjunction with the display color.
3 Click Apply in the Image Analysis Legend Editor to apply the change to the image in the view.

AV-help

11.1.4. How to keep bands from displaying in the view?

You can also choose not to display bands and display colours in the view.

1 Double click the theme's title in the table of contents to open the Image Analysis Legend Editor.
2 In the right hand portion of the dialog, use your mouse to click the boxes next to Red Band, Green Band and Blue Band to select or deselect them from displaying in the view.
3 Click Apply in the Image Analysis Legend Editor to apply the change to the image in the view.

Note: If a checkbox is checked, then that colour is displayed in the view; if a checkbox is not checked, then that colour is not displayed in the view.

If you are working with a single band image, you cannot change the band and display colour settings in the Image Analysis Legend Editor. You can however, change the brightness and contrast settings of a single band image. You can also adjust the histogram of a single band image.

11.1.5. What is an Image Analysis Theme?

When you load the Image Analysis extension into ArcView, you can work with a new type of theme, the Image Analysis theme. Image Analysis themes provide new functionality and capability in the ArcView environment.

An Image Analysis theme can be either continuous or thematic. Both can come from many different sources. Depending on the source, some data must be brought into the Image Analysis extension environment through importers. Other data can be loaded directly into a view.

A continuous data Image Analysis theme is made of raster data that are arranged in a system of columns and rows of pixels. The columns and rows of pixels, when combined, form an image. Continuous data is quantitative in nature. It is used to represent measured characteristics. Most of the continuous data you'll be working with in the Image Analysis extension environment is measured based on pixel values that represent reflectance values of land cover. Continuous data can be either single or multiband. In multiband continuous data, each pixel has multiple values — one for each band that makes up the image.

A thematic data Image Analysis theme deals with categories of information in an image. A good example of thematic data is an image made up of classes such as different soil types or land cover types. Thematic data is usually the result of an assignment or categorisation of pixels that make up the image. When you categorise an image, the pixels that make up the image are grouped together based on their spectral similarities. The new groups or classes of pixels can then be assigned a class name that corresponds to what they represent, like land cover. Unlike continuous data in which a pixel can have multiple values based on the band it is located in, thematic data has only one value stored for each pixel. Therefore, you will only work with single layers of thematic data.

Like other themes in ArcView, you may want to access the Theme Properties to control how an Image Analysis theme functions in a view before you progress to image analysis.

AV-help

11.1.6. Pyramid Layers

When you add an image as an Image Analysis theme to the view, you have the option of creating pyramid layers if the image does not already have them. Images created using the Image Analysis extension, however, automatically have pyramid layers of data. You may also know pyramid layers as reduced resolution data sets.

Pyramid layers are copies of the original bands of data where the spatial resolution has been consecutively reduced by a power of 2. Then, the data file values are placed in a new smaller grid. The Image Analysis extension takes the original image and extracts every other pixel to create the first layer.

For example, a 4,000 x 4,000 (4K x 4K) pixel file is reduced to a 2,000 x 2,000 (2K x 2K) pixel layer. (Note that the Image Analysis extension does not save this first pyramid layer. It accounts for one-fourth of the data.) Then, it extracts every 4th pixel from the original file to create the next layer. Thus, the 2K x 2K pixel layer is reduced to a 1K x 1K pixel layer. Then, it extracts every 8th pixel from the original file to create another layer. So, the 1K x 1K pixel layer has become a 512 x 512 pixel layer. The process continues until the original file is re-sampled down to a 64 x 64 pixel layer. In this way, depending on the zoom factor, the software can go straight to the pyramid layer that represents the scale.

11.1.7. Import File Dialog

To import data, follow the steps below.

1 Open a view in the project.
2 Insert the diskette or CD that contains the data you wish to import into the appropriate drive.
3 From the File menu, choose Import Image.
4 In the Import File dialog box, use your mouse to navigate to the drive and directory.
5 Click the List Files of Type dropdown list and choose the data type.
6 Click to select the file you want to import from the list of File Names. Click OK.

Depending on the size of the file you are importing, the process may be lengthy. You can see the progress of the import in the ArcView GIS status window. The image is imported as an Image Analysis theme with the title Imported Image.

1 Click the Theme menu, and choose Save Image As.
2 Name and save the image in the directory of your choice.

Note: If you do not name and save the file, you must complete the import process each time you want to use it.

11.2. Referencing images with the image analysis extension

Images are stored as raster data, where each cell in the image has a row and column number. Shapefiles and Arc/Info coverages are stored in real-world co-ordinates. In order to display images with coverages or shapefiles, it is necessary to establish an image-to-world transformation that converts the image co-ordinates to real-world co-ordinates. This transformation information is typically stored with the image.

Some image formats, such as ERDAS, IMAGINE, BSQ, BIL, BIP, GeoTIFF, and grids, store the geo-referencing information in the header of the image file. ArcView uses this information if it is present. However, other image formats store this information in a separate ASCII file. This file is generally referred to as the world file, since it contains the real-world transformation information used by the image. World files can be created with any editor. They can also be created using Arc/Info's REGISTER command. In ArcGIS georeferencing of images is done using the Georeferencing toolbar. For guidance read "About geo-referencing" in the ArcMap - Working with rasters chapter of the ArcGIS Desktop help.

AV-help

11.2.1. The align tool

With the Image Analysis extension, you have access to the **Align tool** to help you perform image rectification. The tool is located on the ArcView tool bar and works in conjunction with the Alignment Panel. **Typically**, you use the Align tool when you have loaded two themes of overlapping geographic extent such as an Image Analysis theme of an image and a feature theme of a map that do not display together in the view. This usually happens when one of the themes is in a map co-ordinate system and the other is not. The Align tool can help you correct this situation. However, you can also use the tool in these other applications:

- You can use the Align tool with two Image Analysis themes, one of which contains a map co-ordinate system. In this image-to-image rectification example, the Image Analysis theme without the map co-ordinate system aligns to match the one that does.
- You can use the Align tool to perform image-to-image rectification where neither Image Analysis theme has a map co-ordinate system. In this example, neither theme has a map co-ordinate system, yet they can still be aligned.
- You can also use the Align tool to rectify an Image Analysis theme to co-ordinates collected from fieldwork. In this example, you could have a set of GPS co-ordinates that corresponds to a point on your Image Analysis theme. You can enter those points directly into a dialog, or import the co-ordinates to a shapefile.

11.2.2. To start rectification

To start, click the Align tool to display two themes together in the view. If you know about what portion of the map the image covers, then zoom to that area first. Collect control points in the two themes that correspond to the same geographic location. The point you select in the theme without the map co-ordinate system is called the **From point**; the point you select in the theme with the map co-ordinate system is called the **To point**.

As you continue to collect control point pairs, the theme without the map co-ordinate system is rotated and adjusted to agree with the theme that does have the co-ordinate system. Once you have chosen four control point pairs, the **RMS error** is reported in the ArcView GIS status area. The lower the total RMS error, the better the conversion of the image from a pixel co-ordinate system to a map co-ordinate system. You can see this by the way the image and the feature theme fit together in the view. Once you have taken enough control points and achieved a low RMS error, you can choose to save the theme in its new rectified position (calibration), or you can create a new theme with the co-ordinate system (rectification) and preserve the original theme.

11.2.3. Rectification steps

1. Display both the Image Analysis theme and the feature theme in the view.
2. Optionally, if you know approximately where the image should go, zoom to that location.
3. Make sure the image you want to align is active.
4. From the ArcView GIS tool bar, select the Align tool. The two themes display in the same approximate area in the view.
5. Click a point in the image. A rubber-banding line appears indicating you are to click the corresponding point in the reference theme, which may be any type of theme.

6 Click the corresponding point in the reference theme. (You may need to use some of the right mouse button tools to assist you.)
7 Continue to select control point pairs in this manner. The pairs should be distributed throughout the image. You should select at least four control point pairs if you want the RMS error measure.
8 Once your RMS error is acceptably low, you have two options. You can click the Theme menu and choose Save Image to calibrate the image, or you can choose Save Image As to resample the image.

AV-help

Exercise 11.1 – Geo-referencing a Landsat image on the basis of another one

Steps	Data	Remarks	Time
Image rectification	16960_03p.tif 16960_1990.tif	- Two Landsat images are available for the Mau Escarpment area. One for the year 2003 is referenced according to UTM Zone 37. The other image lacks proper referencing. - Load the image analysis extension - Add the TIFF images as image analysis themes - Rectify the 1990 image with the align tool by defining landmarks identifiable on both the 2003 and the 1990 image. Use the "image to front" and "image to back" options to facilitate referencing - Think about activating the RMS error display (menu upon clicking the right mouse button)	30 min.
Saving image	Referenced image	- Save the image (Theme – Save image as) and the ground control points as well in the RS folder of mau.	5 min.
		-	
Total time			**35 min.**

11.3. Unsupervised automatic classification

ArcView Image Analysis allows you to conduct unsupervised classification on multi-band satellite images. This is a fast method to try to gain an overview over a complex land cover, but it gives only very limited possibilities to influence the outcome.

Exercise 11.2– Unsupervised classification

Steps	Data	Remarks	Time
Classify	16960-03p.tif	- Conduct an unsupervised classification on the 2003 image using 5 classes	15 min
		- Display the classes with colours you think fit the land cover that is represented by the class - Draw conclusions on the potential of such a classification method. What are the limitations? Where do the biggest errors occur?	10 min
Total time			**25 min.**

11.4. Normalized Differencial Vegetation Index

Spectral enhancement functions transform the values of each pixel on a multiband basis. One approach to spectral enhancement of images is the application of specific band ratios or indices. In numerous remote sensing applications, **vegetation indices** have proved their suitability to enhance important object characteristics for crop and vegetation monitoring. The Normalised Difference Vegetation Index (NDVI) has been a standard means in many remote sensing applications, to estimate the biomass, the leaf area index and the amount of water and chlorophyll within the leaves. The NDVI is calculated with the formula: NDVI = NIR – red / NIR + red.

Exercise 11.3– NDVI

Steps	Data	Remarks	Time
NDVI 1990	Rectified 1990 image	- Run the NDVI calculation (Image Analysis – Vegetation Indexes) on the rectified 1990 image. Make sure that you specify the correct bands for the IR and the R channels before running the calculation	30 min.
NDVI 2003	16960_03p.tif	- Repeat the exercise with the 2003 image.	5 min.
Classification	Both NDVI layers calculated in step 1	- NDVI layers can have values between -1 (no biomass at all) and 1 (maximum biomass). In order to classify forested areas and differentiate them from others, one has to set a threshold value beyond which it is assumed that land cover is forest. - Examine the NDVI layers with the identify tool and with simultaneous comparison with the original image and decide on a threshold value. - Transform the layer into a grid and run a map calculation on it which will give an output value of 1 to all input values greater or equal to the threshold you have set and an output value of 0 to all other input values.	
Comparing with “find like areas” classification	Classified NDVI layers Kemaforest03.shp	- A forest classification has been carried out in the same area with the “find like areas” tool in ArcView. The result is provided in the data folder. - Compare the classification originating from the NDVI and the one from the find like areas tool and the original picture - Draw important conclusions.	
Total time			**35 min**.

11.5. Identifying similar areas

This tool allows identifying areas having identical or similar cell values around a selected cell, or on the basis of training sites identified through graphic elements or with polygon features in a shapefile. The kemaforest03.shp layer used in the above exercise was generated in this way. In the following exercise, the participants will try conducting their own supervised classification by identifying training sites containing different patches of forest and conducting the classification on this basis.

Exercise 11.3 – Identifying similar areas

Steps	Data	Remarks	Time
Identifying forests	16960-03p.tif	- Add a new polygon theme to the view, name and save it adequately - Start editing and draw polygons (rectangles or circles) containing forest patches. Pay attention to identifying different types of forests, both in shadowy and in sunny areas. Identify at least 15 sites. - Make sure that the training sites cover forest areas in the 1990 image as well. If not delete the concerned features. If after this verification you are left with less than 10 sites, digitize new ones. - Save your edits - Run a classification on both the 1990 and the 2003 images using the same training sites and using the Image analysis – find like areas menu.	30 min.
Observations		- Observe the change in forest cover between 1990 and 2003. What are the main conclusions to be drawn from this change?	20 min
Layout		- Think of additional layers one should add to the forest information in order to come up with a significant map layout. - If these layers are available in the CETRAD database, add them to the view and adapt their symbology. - Create a layout with the spatial elements in use and print it out.	60 min
Total time			**110 min.**

CETRAD

12. Bibliography

Dorling D. and D. Fairbairn, 1997: “Mapping. Ways of representing the world”. Insights into human geography. Series Editors: Paul Knox and Susan Smith. Longman Editions. ISBN: 0 582 28972 6.

Ehrensperger A., 2000: “Fieldwork with WSC data and GIS. A Methodological Test in the Mountains of Muang Sing District – Lao PDR”.

ESRI, 1996: “ArcView GIS. The Geographic Information System for Everyone”. Publication of the Environmental Systems Research Institute (ESRI) Inc., Redlands, California, USA

ESRI, 1997: “Understanding GIS. The Arc/Info Method”. Publication of the Environmental Systems Research Institute (ESRI) Inc., Redlands, California, USA.

ESRI, 1998: “ArcView Image Analysis. Enhanced Image Integration, Display and Analysis”. Publication of the Environmental Systems Research Institute (ESRI) Inc., Redlands, California, USA. ISBN: 1-86242-033-5

Hutchinson S. and L. Daniel, 1997: “Inside ArcView GIS – Second Edition”. OnWorld Press, Santa Fe, USA. ISBN: 1-56690-116-2

Kraak, M.J. and F.J. Ormeling, 1998: “Cartography. Visualization of spatial data.” Addison Wesley Longman Limited. Essex, England. ISBN 0-582-25953-3

MRCS, 1997: “Mekong Watershed Classification. The WSC Map Users Guide”. Field Manual prepared by the Centre for Development and Environment (CDE), Institute of Geography, University of Berne, for the Mekong River Commission Secretariat (MRCS), Phnom Penh, Cambodia.

Monmonier. M, 1996: “How to Lie with Maps”. The University of Chicago Press. ISBN 0-226-53421-9