4.2

Spatial data and geographic information systems

Thomas Gumbricht

- **Understanding the spatial context of an agricultural and resource management problem will probably be an important part of solving it, so can not be ignored in your research**
- **Geographical information systems (GIS) allow you to manage and manipulate spatial data**
- **Simple manipulation of data sets that has already been prepared can be learned quickly. However, using data from multiple sources for more complex tasks can be a major undertaking**
- **Many basic spatial data sets are available for Africa but poor Internet connections may limit access to them**
- **Freely available software is now sophisticated enough to be useful in many spatial research projects**

Introduction

The Earth is a sphere with an average distance to the Sun of 150 million km. The Sun radiates energy, which is received by the rotating Earth in diurnal cycles with annual modulation as the Earth completes its annual ellipse. The energy that hence reaches the Earth is mainly dissipated at the Earth's surface. It rotates the hydrological cycle, releases nutrients that feeds the ecosystems, and drives photosynthesis (all which have been largely altered by man since the Industrial Revolution). Thanks to these processes life exists and the Earth's surface has developed a 'natural' logic. In dry areas with poor resources vegetation is sparse, in valleys where water and resources accumulate the vegetation is more luxurious. If there is a trough and enough water a body of water will form. In a similar way the human landscape is also logical, with fields in fertile valleys and dwellings along the ridges. Cities have to be close to large sources of water. These logical landscapes are also evident on a much smaller scale. Most vegetation is bound to specific habitats narrowly defined by conditions of climate, soil and water; that can shift within a scale as small as one metre. At an even finer scale a human thought is also dependent on energy dissipation at interfaces – in a very well described spatial context between the synapses of nerve cells. Image analysis and location information systems are hence very important tools in medicine, sociology, anthropology, biology, ecology, geology, hydrology and many other sciences.

For a particular study the spatial information needed might only be a map – as were the descriptive studies conducted by the first European explorers. In most instances a researcher is probably more interested in extracting more information in order to test a hypothesis. This could be comparing two district-level data sets, perhaps one on poverty and one on incidence of malaria. This is easily done in a geographic information system (GIS), and you still only need a single map, with attributes (databases) on both malaria and poverty. But malaria is a vector-borne disease, and the mosquito carrying the parasite breeds in water, so proximity to water is most probably important. To test that hypothesis an additional data layer of water availability is needed. This step is a major complication that has yet to be fully taken in the case of malaria. Rivers and lakes can easily be found, and their proximity to each population group calculated. But now you ideally also want population and malaria data on village level, not just for districts. Then you realise that mosquitoes can breed in water tanks, small puddles, or even water trapped in an old bucket or boot. Now the comparison becomes almost impossible, and you need to get data on rainfall and temperature in order to calculate the daily water balance. This calculation is possible; the data are there (as you will see below), but the calculation is not a trivial task.

In general, a thematic study will need more refined data, whereas an interdisciplinary study must probably be satisfied with more generalised data. Often this is because detailed data of different origin are seldom compatible in their spatial resolution. However, the use of GIS and spatial data can be very rewarding. The first level, including a map, is almost always welcomed and very simple, it will only take a few days. The second level, comparing ('overlaying' in GIS jargon) attribute data related to the same spatial context is also quite simple, and will take a week to a month. The third level, analysing spatial relations introduces complexity, but can still be done by most standard GIS packages (and some of the freeware packages listed on page 142), but it will take some months to a year. The fourth level of integrating GIS with dynamic (time-resolved) models is quite complicated. This level will demand in-depth knowledge of both GIS and modelling, and most probably of programming as well. It will take longer than a year.

Mapping and modelling with GIS – A game of chess

Capturing the chess board

It is seldom convenient to have a sphere to portray the Earth's surface; so humans have used two-dimensional (2D) maps for at least 5000 years. Most GIS are also static 2D, even if the processes occurring on the Earth's surface are often 3D and dynamic. The most common GIS spatial data model represents space as 'vectors' (points, lines and polygons). This model is suitable for human-created objects and concepts (wells, roads, states, cities, rivers and lakes). Natural phenomena are better represented as continuous fields (elevation, land cover, vegetation density), which in GIS translate to a 'raster' or grid data model (Figure 1). For simplistic reasons, the raster model is often preferred in modelling. It is also the implicit format of satellite images or photographs.

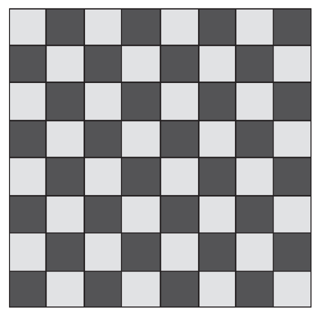

Figure 1. A raster

The raster landscape in Figure 1 is also the landscape where the game of chess takes place. Let us assume that you are unaware of the game of chess, but want to understand it by using GIS. Once you have identified the problem from a GIS perspective you must decide which data model (raster or vector) to use and how to *capture* the data. The chessboard can be captured as primary data (from satellite image or digital photograph) or from an existing analogue (secondary) source (digitising or scanning). Whatever you choose you will, implicitly or explicitly choose a certain grain size (or spatial resolution) when you capture the data (Figure 2). *Meta*-information on capture technique, resolution, and who did it should ideally always follow the GIS data, but is frequently lost on the way to the end-user.

For most spatial phenomena that are studied, you usually have a conceptual idea about the spatial patterns and dynamic processes that are occurring. If you assume that you already have some existing knowledge of chess, you can decide on a stratified sampling of data. For each square of unit distance, take one sample at a randomised point. You can further assume that there is neither an error in position, nor in the obtained value. A point is the simplest kind of vector data, and you can assign attributes to it – you can form a

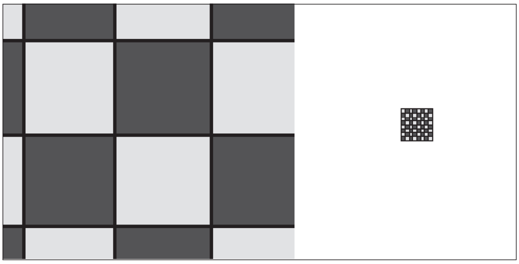

Figure 2. Two raster data sets with different grain size. The ease with which you will be able to understand chess will obviously be dependent on the resolution or grain size used to capture the chessboard

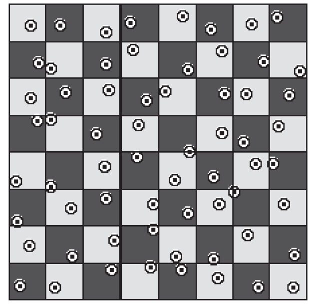

Figure 3. Sample points (vector data) and rasterised pattern

database describing the properties of this point (in this case colour, but it could also be some other capacity such as depth, elevation, or type). Then you can *manipulate* the point data by *rasterising* it to arrive at something that looks like a chessboard. If you honour the value of the measurement in each cell (or *pic*ture *el*ement – pixel) you will arrive at the correct landscape (that you happen to know in this simple case) (Figure 3).

If instead you manipulate the data by using a geostatistical interpolation function, which do not honour the observed value *per se*, you get more or less erroneous results (Figure 4).

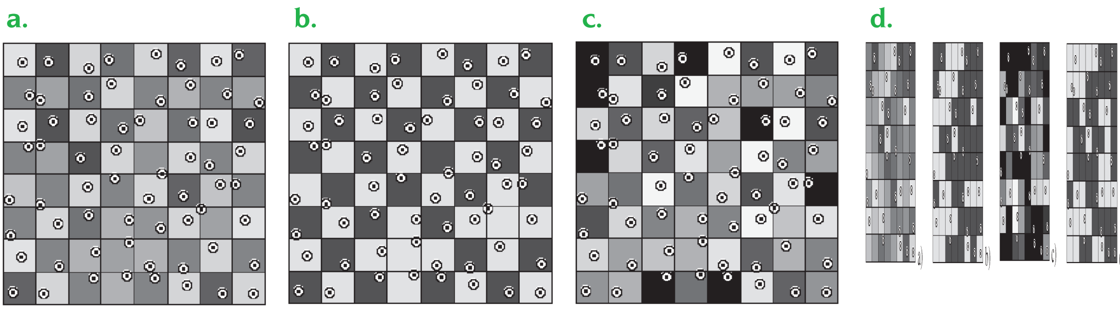

Figure 4. Interpolated 8 x 8 raster image from 64 Boolean sample points, randomly placed in each grid cell: a. Inverse distance weights (IDW) to 8 neighbours, b. Reclassification of a, c. Spline smoothing function to 8 neighbours, d. Reclassification of c. The reclassification is done as a threshold using the value 0.5. Both illustrated interpolation methods can be parameterised to get a true chessboard, that, however, demands iterations and skills, together with knowledge about the pattern of the generated surface

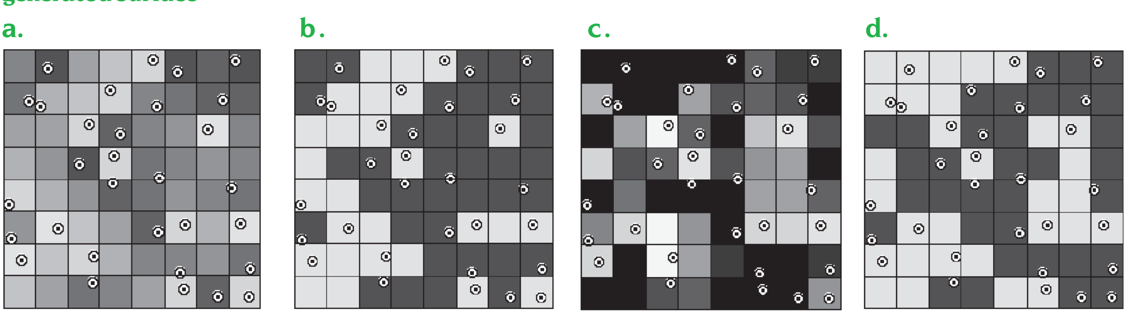

Figure 5. Interpolated 8 x 8 raster image from 31 randomly selected points (see Figure 4) a. IDW to 8 neighbours, b. Reclassification of a, c. Spline smoothing function to 8 neighbours, and d. Reclassification of c. The reclassification is done as a threshold using the value 0.5

The high cost of field and inventory work requires the fullest use of existing data and the application of interpolation methods. Hence, the sampling grid is generally much sparser than the interpolated grid (Figure 5).

Note that the interpolation of the chessboard data are truly 2D, whereas the Earth's surface is a spheroid and interpolation with different geoids and projections render different results. To choose the right projections for a particular purpose is not trivial, but is beyond the scope of this book).

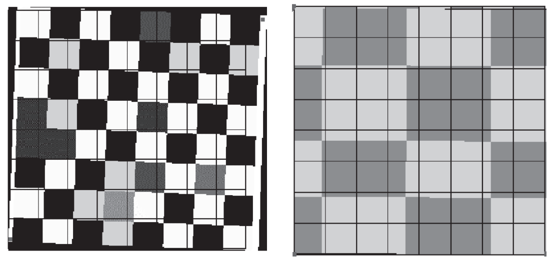

Figure 6. Schematic examples of problems with using remotely sensed data to portray the Earth's surface: a. Georeferencing and spectral properties of the observed phenomena, b. Grain size and geometrical distortions in the sensor

Primary data capture from remotely sensed imagery to GIS is an important part of the integration of GIS and modelling, also in social science for updating or downscaling census data (see below). Remotely sensed data have a definitive grain size and thus resolution. Apart from grain size, problems with sensor quality, spectral properties of the observed phenomena and georeferencing introduce errors when interpreting and classifying remotely sensed data (Figure 6).

Monitoring the dynamic game

Having established the chess-playing arena, a working hypothesis for the processes that are occurring needs to be formulated. For most dynamic phenomena an initial inductive approach is almost inevitable. Only after a set of observations is available is it possible to use coincidental data to formulate a deductive hypothesis.

Observations of natural and human phenomena are often made at regular intervals. Satellite images over an area are usually taken at the same time of day with a given interval (approximately 14 days for Landsat), as are many climate station data, water flow and water quality measurements. Measuring the chess game every morning at 09.30 (cheapest because it is outside the coffee room) always gives the same result (Figure 7).

But if you work late one evening and chance to look at the chessboard, and suddenly see

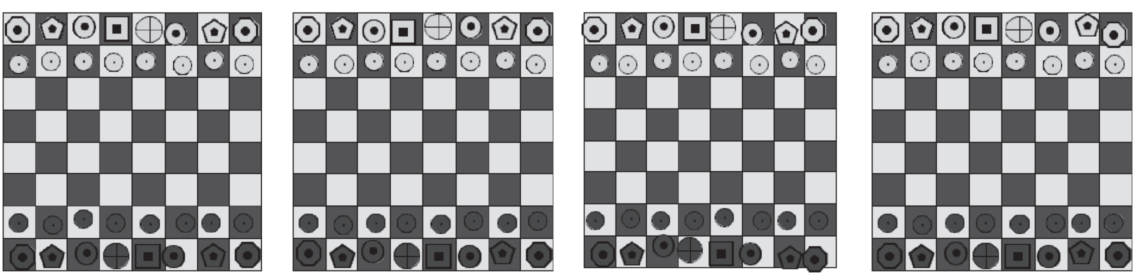

Figure 7. Observations of a chess game on four occasions. At first the game is apparently static. Only with a more detailed scrutiny it is revealed that the players actually are shifted a little between each observation. However, as we have no hypothesis or information of sub-cell pattern or process we neglect this as observation error

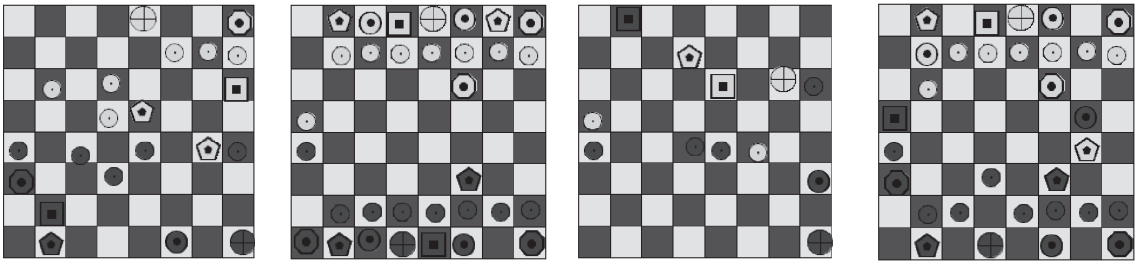

Figure 8. A series of temporally random observations of the chessboard

something has happened, you realise that there is obviously another time scale to the daily one (weekly, monthly or annual). So you start to observe the game regularly when you are working late. A rather erratic series of observations turns up (Figure 8).

Because of the strange observation angle (from above or 'nadir' in remote-sensing jargon) the visualisation of the players is poor, and it is difficult to distinguish the actors. However, a few hypotheses on their roles can be put forward:

1. One species (Bishops) seems to be bound to a certain feature type, or habitat, (namely black or white) in the playing ground.
2. The smallest and most common species (Pawns) seems only to be able to move in one direction like water downhill.
3. The species in the corners (Rooks) seem to be the most home-bound.

After some months of random observations hypotheses 1 and partially 2 are corroborated whereas 3 is falsified. After several years of fund-seeking the observations can be transformed into intense evening campaigns. With observations down to 10-minute intervals some of the players rules crystallise themselves, however the role of the knights escapes a robust formulation. Finally, a sensor connected to a real-time observation can capture the full sequence of activities, and the role of each player can be formulated.

Modelling the full game

The identified role of each player leads to a surge in modelling the game, mostly by using a rule-based (rather than statistical) approach where the roles of each player can be unambiguously defined. The formulation of initial (setting at start) and boundary conditions (edges of the playing arena) are straightforward. The application of an object-oriented approach for each player is favourable; a certain actor can only do a certain action, which cannot be done by another actor.

However, even though the game is spatially defined, it is not possible to use the toolbox of any commercial GIS to play the game. And, only a few softwares have architecture open enough to allow the GIS game to be programmed to them, but with great difficulties. With a customised GIS it is possible to create a graphical user interface (GUI) that can help to set up initial and boundary conditions, and even to allow the set-up of the players' positions in the middle of a game, and the use of that as an initial condition. This leads to the development of an intermediate coupling of the chessboard and the game simulator through their sharing a common file format. It is a bit cumbersome to use and never reaches widespread use to improve the social awareness of the game.

For the game itself, the combination of such advanced machine-learning as artificial neural networks and faster computers, mean more alternative game outcomes can be foreseen after each activity (draw). Finally, one computer (Deep Blue) succeeds in winning the game. This is now more esoteric interest among the chess community, but the general public, policy- and decision-makers are unaware of this development.

Implications

A game of chess always aims at checkmate – which is unambiguously defined, as is the role of each player. The rules of the game show no evolution, neither in space, nor over time. If you change the extent of the arena, the role of the players or the outcome for checkmate to an unknown event, the computer would have little chance of winning. In a transient social or natural environment that is how the evolutionary game is played. In the simple case of chess there are only two scales that are of importance, that of a cell and the whole board. Furthermore, the game as such has no influence on the arena. In a landscape all discretised scales are arbitrarily chosen, the real landscape is a continuous nested hierarchy: but some scales have dominance-generating spatial architectures and temporal cycles, entrapped by key stone species and related processes. This also leads to the conclusion that the processes are forming the patterns rather than the other way around – and that the systems has feedback loops at various scales. All those aspects can be disregarded in the special (and simple) case of the chess game.

The general conclusion that can be drawn is that modelling in GIS is hampered by several shortcomings, that care must be exercised when using distributed data for modelling, and that the quality of many GIS integrated models is poor. They are also poor because they have poor GUIs, fail to visualise the results, and hence do not reach the intended user community. In order to secure high-quality GIS-integrated models the following issues need to be considered:

- Close co-operation between GIS model researchers in general, and particular among
  - researchers studying the same phenomena but adopting different methods and/or scales
  - researchers, planners and decision-makers
- Up- and down-scaling, and nesting models of different resolution
- Spatial and temporal domain, grain size and sampling intensity when integrating data from various sources
- Strategies for sampling spatial phenomena to get representative data
- Selection of spatial interpolation methods and spatially correlated error tracking and tagging
- Methods for evaluating the influence of error and error propagation on model performance, and error visualisation for communication information on uncertainty
- Integration of remote sensing into GIS models
- Integration of temporal processes into GIS (3D- and 4D-GIS)
- Integrated systems that support a complete digital data flow from data collection with mobile field GIS (Global Positioning Systems, GPS) to visualise and exchange results via networks
- Formulation of versatile criteria for evaluating the prediction power of GIS-related environmental models
- Compilation of high quality, accessible (shared) databases to be used as back-drops to evaluate the predictive power of different GIS-related environmental models
- Establishing baseline and framework data
- Development of guiding GUIs that can lead the user to select the best method for the formulated problem and the available data

- Development of friendly interfaces that promote the dissemination of GIS and integrated models to domain experts, planners and managers.

Using GIS in Africa

Studies involving spatial dependence and GIS in Africa are hampered by lack of data and computer resources, and poor knowledge and communications infrastructure. However, with the growth of geoinformatics over the Internet, global and continental-scale data are becoming increasingly available. Together with more powerful free GIS and remote-sensing software, there is a good chance that the data and software needed for many studies are available, either directly or via map algebraic modelling and other manipulations applied to available GIS data in combination with satellite imagery. The global trend in adopting remote-sensing data for spatial studies is strong in traditionally data-poor regions. Free high-resolution satellite images [Landsat Thematic Mapper (TM) and Enhanced TM (ETM)] are now available for the whole African continent. Access to this data in Africa, however, is often illusive due to poor Internet connections. The global data sets derived from satellite data (including land cover) are seldom adjusted for continental needs, leading to semantic discrepancies and interoperability problems when merging data sets. Local knowledge is mostly disregarded. Further, studies employing global data in Africa are often esoteric, and seldom used for policy or management inside Africa.

GIS and remotely sensed (RS) databases for Africa

In this chapter spatial databases have been divided into framework databases and field databases. **Framework databases** are base maps holding mostly information on anthropogenic-derived features – e.g., political boundaries and infrastructure, but they sometimes also have more object-oriented physical themes like elevation contours and hydrography. These databases are typically object-oriented and in vector format. They can be used to create simple thematic maps. Framework databases available for Africa typically contain data at district level, and hence simple descriptive statistical analyses (population density, travel distances, etc.) can be done at a level based on this data. Framework data can seldom be used directly for advanced analyses and modelling (environmental studies). Environmental studies demand field data, usually in raster format for such parameters as population density, soil classes, drainage, elevation, temperature, and precipitation.

Framework databases for Africa

The foremost baseline framework database for Africa (and other parts of the world) is the Digital Chart of the World (DCW). DCW is a 1:1 million scale thematic map developed by the Defense Mapping Agency (DMA) and compiled by Environmental Systems Research Institute, Inc. (ESRI). For large parts of Africa these base maps are the largest scale maps available, either due to lack of other data or to the larger-scale maps being classified. Themes in DCW include political boundaries, populated places, roads/railroads and other infrastructure, hypsometry, hydrographical data, and rudimentary land coverage.

Based on the DCW, ESRI has assembled a more easily accessible database and has also developed a more field-oriented World Thematic Database. Several other GIS software producers have also established databases based on DCW (and additional sources mentioned below) for bundled delivery. The most comprehensive probably being the Mud-Springs Geographers – AWhere Almanac Characterisation Tool (ACT). This and other software tools (listed on **page 142**) are a very good way to learn GIS using data over Africa. AWhere-ACT is especially powerful for analysing climate data (supplied with the software) for agriculture and natural resource management applications. In many cases the software and bundled data are free for use in Africa by non-profit organisations.

Several recent efforts in creating more-detailed (large-scale) regional framework databases for Africa have been made. The most comprehensive is probably the Africover project by the Food and Agricultural Organization of the United Nations (FAO). This database also includes detailed land cover derived from combinations of Landsat ETM data and topographic maps. The agencies of the UN have also initiated an attempt to create a common depot for their GIS data – which has led to the Data Exchange Platform for the Horn of Africa (DEPHA) (see **page 140** for a more complete list of framework data sources available).

Field databases for Africa

Elevation

For Africa the elevation data in DCW (contour lines and spot elevation data) together with generalised 3-arcsecond digital terrain elevation data form the primary source for the global 30-arcsecond (approximately 1 km) GTOPO30 elevation database released by the United States Geological Survey (USGS) in 1996. The data in GTOPO30 have been hydrographically corrected and resampled to a 1-km grid, to create the HYDRO1k database. From the hydrologically corrected HYDRO1k Digital Elevation Model (DEM) seven derivative themes have been extracted: flow directions, flow accumulations, slope, aspect, compounded wetness indices, stream-lines and basin areas. Several individual countries have better elevation databases. The next elevation data set covering the whole of Africa will be the Shuttle Radar Topography Mission (SRTM) database (90 m resolution), expected to be released during 2004.

Land use/cover

Two global land cover data sets covering Africa in 1-km resolution are presently available. The latest is derived from TERRA-MODIS (Moderate Resolution Imaging Spectroradiometer) data (2000/2001) and was created by the University of Boston. MODIS has also been used to create a global tree cover database in 500-m resolution available from the University of Maryland. The older land cover is produced by the USGS from NOAA–AVHRR (Advanced Very High Resolution Radiometer) data (1992/1993). It exists in several versions useful for different applications and also includes monthly vegetation data from April 1992 to March 1993. The Africover database mentioned above is superior to these global databases but does not yet cover the whole continent.

Climate and vegetation

The United States Agency for International Development (USAID), as part of the Famine Early Warning System (FEWS), continuously provide 10-day composites of vegetation density (Normalised Difference Vegetation Index – NDVI) derived from NOAA–AVHRR in 8-km resolution covering the whole African continent. The data set goes back to 1981 and is archived and disseminated by the USGS. It can be retrieved from the African Data Dissemination Service (ADDS). Thermal Meteosat images together with 760 ground precipitation stations are used to estimate precipitation over Africa as part of USAID FEWS. Processing is based on 30-minute image intervals for cloud top temperature combined with the ground data and derived fields of humidity, winds and DEM. The data extends from 1995 and are archived and disseminated by USGS (ADDS webpage) as 10-day composites. More coarse resolution databases that cover climate together with scenarios of climate conditions under various assumptions of human impacts on the climate are available from the Climate Research Unit (CRU), University of East Anglia, UK, either directly via the Internet, or from the Intergovernmental Panel on Climate Change (IPCC) as a CD.

Population

The best and latest population figures are the 1-km resolution Landscan project data for 2000, 2001 and 2002 from the Oak Ridge National Laboratory, USA. These figures are created from census data and downscaled using intelligent interpolation (using relations such as light at night, slope, or elevation, which correlates strongly with population density). The Center of International Earth Science Information Network (CIESIN) hosted by University of California, has compiled global population data for 1990 and 1995. The data has an original resolution of 5 arc-minutes (approximately 10 km), but for Africa the data mostly represent averages for larger regions. United Nations' African population figures for selected countries covering the second half of the 20th century are available from Central African Regional Program for the Environment (CARPE) (see **page 140**).

Soil map

FAO has produced a Digital Soil Map of the World (DSMW) in 1:5 million scale. Soil classes are given as polygons, with derived characteristics attributed. The soil map is only available as a CD. For some regions FAO also has a 1:1 million scale soil map.

Satellite imagery

Remote sensing (RS) data are increasingly important for creating and updating both physical/biological and socio-economic databases. Access to RS data is constantly improving thanks to: lowered prices, declassification of historical high-resolution data, a new generation of multi-sensor satellites (TERRA and ENVISAT) that are now operating, improved computing power and better software–user interfaces.

For national to continental studies NOAA-AVHRR and TERRA-MODIS data and their derivatives are the most easily accessible. Other data of similar resolution that can be easily accessed include the European Space Agency (ESA) ERS-2 satellite and its ATSR 7-band sensor (which can be downloaded from the Internet in near real time), and the SeaWiFS 6-band sensor.

Full coverage, high-resolution Landsat TM and ETM data are now also freely available for the whole of sub-Saharan Africa via the University of Maryland. Landsat E(TM) composites in Mr-SID compressed formats of the whole globe are more easy to download and available from NASA. To find all available Landsat MSS, TM and ETM scenes, and other satellite data sources use the NASA Earth Observing System Data Gateway.

The original TERRA–MODIS and NOAA-AVHRR scenes that were used for the land use/cover classifications (see above) are all freely available as composites from University of Maryland (TERRA–MODIS) and USGS (NOAA–AVHRR). The Africa NOAA–AVHRR tiles for vegetation are also available from the International Centre for Insect Physiology and Ecology (ICIPE). Additional, raw, NOAA–AVHRR data are available via the NOAA Satellite Active Archive on the Internet, or from USGS at the cost of reproduction.

Georeferenced time series point data for Africa

Time series point data on climate (weather station data) and hydrology are available via a variety of web pages. As this is outside the main scope of this chapter we recommend the ICIPE data server as a source of archived weather station data from around Africa.

Data accuracy and merging

Most of the older global and regional databases are not quality labelled, neither for positional error nor attribute accuracy, thus potentially leading to large problems in interpretations and applications. Even if most data that can be downloaded are georeferenced, their accuracy often does not allow mapping at higher resolution than 1:1 million. The dataset with highest spatial accuracy is the NASA geocover (downoadable as MrSID images – see **page 141**), which is within 50-m and can hence be used for geocorrecting other data sets. Another problem is that the semantics used, for instance, for the land use/cover maps are not coherent with those used in different parts of Africa. This is also due to a poorly developed unanimous semantic cover of natural geography in Africa. Semantic inconsistencies lead to information loss and prevent sound conclusions being drawn.

For most spatial studies, it is necessary to merge data. Most satellite images must be georeferenced to a projection that fits the geographic location (and the framework data) before they can be used for analysis and further studies. This is not a trivial task. Field data sets continuously vary over time and space on different scales. A satellite image is already an aggregation of the land surface over the pixel size. Field data, including socio-economic data are often up- or downscaled, or aggregated. The quality difference between data sets of the same origin but presented at different scales is seldom reflected in the metadata. It is extremely important to know the timing of acquisition, grain size and scaling of field data when analysing, interpreting, and applying such data. For most of the global data sets available this is seldom a problem. However, most local users are ignorant of the problem and secondary data sets derived from such sources often lack meta-data.

Points to remember

- Increased data availability and the ease with which distributed data layers are created from point and line data, and remote sensing, have led to a widespread coupling of GIS and remote sensing to existing (non-topological) cause-effect models in, for example, hydrology and erosion studies, and to updating and downscaling land cover and population density maps
- Data availability for Africa has now reached a point where it is possible to do such studies, often with freely available data
- The major problem for the individual researcher is in accessing the data, and in acquiring the skills of GIS and RS needed to 'massage' the data into a coherent database
- The free GIS software programmes available today are powerful enough for you to learn GIS, and to create basic databases
- The bottleneck for using GIS for research in Africa is poor Internet access and poor GIS skills
- If you want to use GIS you should download the necessary data or order it via CD/DVD (usually possible for a small fee), and you should learn GIS by using one of the listed free software programmes
- As most of the software programmes have very similar interfaces, learning one means that learning a second becomes an order of magnitude easier. So going from DIVA-GIS to ArcView is very simple (also because they share data formats).

Expert systems

GIS and RS (or geoinformatics) have developed from being tools for data storage and presentation to also include analyses and modelling. Overlaying two or more thematic maps (see Figure 9) is a simple but often illustrative means of identifying relations in spatial

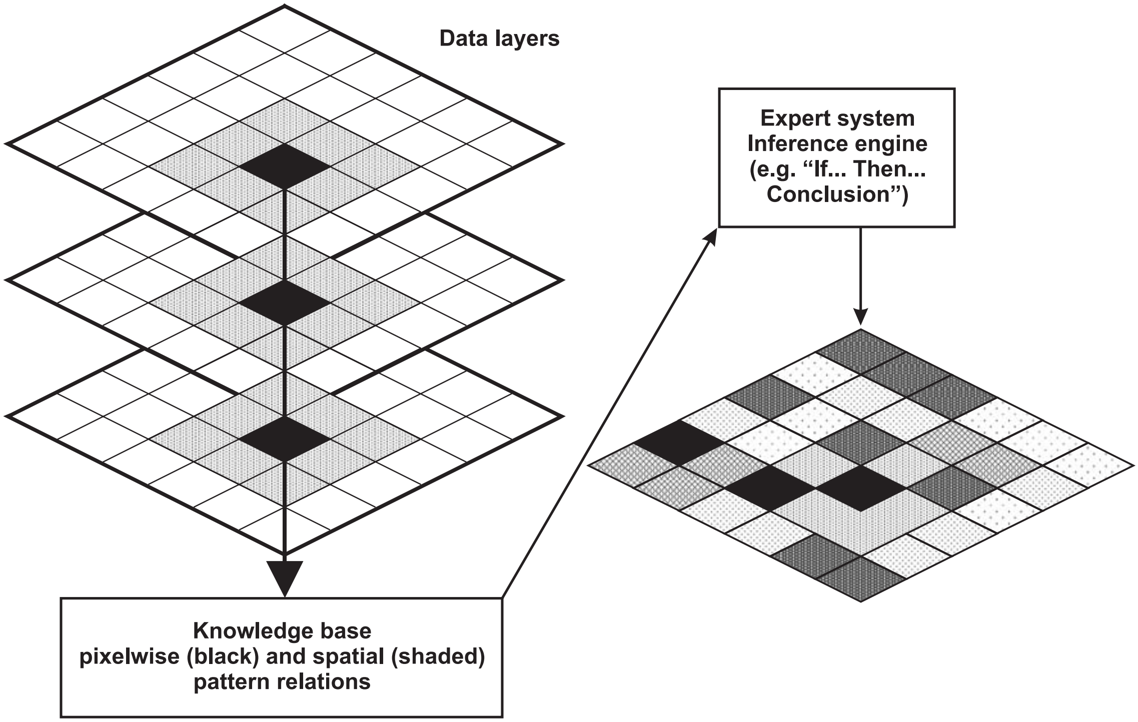

Figure 9. Schematic structure of an expert system approach for spatial data analysis

patterns. More advanced analyses include using map algebraic formulae combining several thematic layers. Such 'expert system' approaches are widely used to rank vulnerability of natural resources, food security, or water availability. One example is the DRASTIC method (Depth to groundwater, Recharge, Aquifer media, Soil, Topography, Impact of rootzone, Conductivity) for groundwater vulnerability analysis, where each of the seven factors has a physical related value. Development is towards more advanced expert systems including object-oriented methods, and considering ancillary and multi-temporal data, and spatial relations (Figure 9). Expert systems are like the game of chess – unambiguously defined with a set of strict rules. Expert systems are thus said to be data- or forward-driven. However, GIS is also becoming a decision-support system (DSS), e.g., for ill-structured (localisation) problems. Used as a DSS GIS becomes more of a tool for discussions and illustration of decision alternatives. Formal methods have been developed to involve various stakeholders in such discussions, including multi-criteria evaluation (MCE). In contrast to expert systems based on predefined rules and weights of physical parameters, DSS are related to different stakeholders perceptions, and as the aim is to reach a solution (for allocation of land use/development, water or nature protection), the method is said to be goal-driven.

Whether studying natural or social science, GIS can be very useful, and there is a plethora of methods, models and techniques that you can apply to analyse or present data that deals with spatial relations. But it is critical that you formulate a sound hypothesis and use adequate data of sufficient quality. To avoid mistakes a parsimonious approach, and rigourous meta-data description is essential. This will make it easy to update and eventually publish your data and results.

Resource material and references

Burrough, P. and McDonnell, R.A. 1998. *Principles of Geographical Information Systems.* Oxford University Press, Oxford, UK.

Fotheringham, S. and Wegener, M. 2000. *Spatial Models and* GIS. Taylor and Francis, London, UK.

Goodchild, M.F., Parks, B.O. and Steyart, L.T. 1993. *Environmental modeling with* GIS. Oxford University Press, New York, USA.

Goodchild, M.F., Steyart, L.T., Parks, B.O., Johnston, C., Maidment, D., Crane, M. and Glendinning, S. 1996. GIS *and environmental modeling: Progress and research issues.* GIS World books, Fort Collins, Colorado, USA.

Framework databases for Africa

Geography network
http://www.geographynetwork.com/

ESRI downloadable data
http://www.esri.com/data/download/index.html

Data Exchange Platform for the Horn of Africa: UN organisations Geo data depot.
www.depha.org

Digital Chart of the World (DCW): Basemaps for all the countries of the world.
http://www.maproom.psu.edu/dcw/

Food and Agriculture Organization of the United Nations (FAO)
Very good land cover maps over East and Central Africa
www.africover.org

Global GIS database – Digital Atlas of Africa
http://webgis.wr.usgs.gov/globalgis/

Field databases for Africa

GTOPO30 global topographic data
http://edcdaac.usgs.gov/gtopo30/gtopo30.html or
http://www.ngdc.noaa.gov/seg/topo/globe.shtml

Hydro1K HYDRO1k Elevation Derivative Database
http://edcdaac.usgs.gov/gtopo30/hydro/index.html

Landscan population data Oak Ridge National Laboratory
http://web.ornl.gov/sci/gist//landscan/index.html

Global landcover from NOAA-AVHRR (1992–1993 data)
http://edcdaac.usgs.gov/glcc/glcc.html

MODIS land cover from Boston University (2000–2001 data)
http://duckwater.bu.edu/lc/mod12q1.html

MODIS Global Vegetation Continuous Fields from 500m MODIS data 2000–2001
http://modis.umiacs.umd.edu/vcfdistribution.htm

CIESIN (Center for International Earth Science Information Network) Colombia University
Including climate data and global gridded population data from 1990 and 1995
www.ciesin.org

CARPE (Central African Regional Program for the Environment)
http://carpe.umd.edu/products/

Satellite imagery and related data

University of Maryland Global Land Cover Facility
A very good source of free Remote Sensing (Landsat (ETM) and TERRA MODIS) scenes
http://glcf.umiacs.umd.edu

USGS Land Processes Distributed Active Archive Center
http://edcdaac.usgs.gov/main.html

USGS Earth Explorer
http://earthexplorer.usgs.gov

USGS NOAA-AVHRR used to create global landcover – 93 original scenes (1992 to 1996)
http://edcdaac.usgs.gov/1KM/1kmhomepage.html

Africa Data Dissemination Service (United States Geological Survey – USGS)
http://edcw2ks21.cr.usgs.gov/adds/data.php

NASA Earth Observing System (EOS) Data and Information System
http://edc.usgs.gov/

NASA Earth Observing System Data Gateway.
http://edcimswww.cr.usgs.gov/pub/imswelcome/ or
http://redhook.gsfc.nasa.gov/~imswww/pub/imswelcome/

NASA global Hydrology and Climate Centre (Weather satellite data)
http://wwwghcc.msfc.nasa.gov/GOES

Goddard Institute for Space Studies
http://www.giss.nasa.gov/data/ and http://xtreme.gsfc.nasa.gov/

National Geophysical Data Center
http://www.ngdc.noaa.gov/

MrSID images (excellent geocorrected – can be used for georeferencing other spatial data)
https://zulu.ssc.nasa.gov/mrsid/

ICIPE (International Centre for Insect Physiology and Ecology) Africa Data Bank
(Including remote sensing data and weather station data Over Africa)
http://informatics.icipe.org/databank/

Japan Aerospace Exploration Agency
(Free JERS-1 radar images over most of Africa can be ordered on CD)
http://www.eorc.jaxa.jp/eorctop.htm

SRTM (Shuttle Radar Topography Mission)
http://www.jpl.nasa.gov/srtm/

Visible Earth – NASA site with preprocessed satellite images of Earth
http://visibleearth.nasa.gov/

Microsofts image database (terraserver)
http://terraserver.com/ or http://terraserver-usa.com/

Digital Globe (very high resolution data sets over selected cities)
http://archive.digitalglobe.com/

Geocommunity spatial news (incl Landsat viewer)
http://spatialnews.geocomm.com/

Free software sources

Dynamic Maps (A free ware GIS with many predefined functions for natural resource management)
http://www.skeinc.com/ or via www.africover.org

DIVA-GIS (A fully functional GIS developed by the International Potato Center)
http://diva-gis.org/

Arc-Explorer (Light-weight GIS by ESRI that also produced Arc-Info, Arc-View and Arc-GIS)
http://www.esri.com/software/arcexplorer/index.html

ERViewer (viewer for many image formats)
www.ermapper.com

WINDISP (A simple freeware for image processing from FAO)
http://www.fao.org/WAICENT/faoinfo/economic/giews/english/windisp/dl.htm

Mud Springs Geographers (A fully functional GIS bundled with free GIS data for Africa)
http://www.mudsprings.com/home.aspx

Mapmaker basic (Light-weight GIS freeware)
www.mapmaker.com

SILVICS (Satellite image processing for forests)
http://eurolandscape.jrc.it/forest/silvics/

GRASS (Advanced GIS and image analysis for UNIX or Linux)
http://grass.itc.it/

Microdem
http://www.usna.edu/Users/oceano/pguth/website/microdemdown.htm