BEST-GIS

ESPRIT/ESSI project no. 21580

Funded by:
The European Commission
DGIII - Industry
ESPRIT Programme

Guidelines for Best Practice in User Interface for GIS

Section 7
"Recommendations for the best use of key GIS UI functions"

7. Recommendations for the best use of key GIS UI functions

7.1 Objectives

The aim of this section is to provide a set of recommendations to guide the end user in exploiting the more common functions of a GIS. Thus it integrates both the GIS operation checklist of section 6, and the checklist for defining user requirements of section 5. This section explicitly considers the following issues:

- end user skill level
- type/class of application
- data structure
- UI technology currently available on the market
- generic UI practice available

Most of the recommendations in this section are aimed at the final user with low or intermediate experience in GIS technology, but with a sound background in his/her own discipline or professional sector (i.e. topography and surveying, civil engineering, hydrology, geology, urban planning, etc.). The key functions discussed here have been selected from among hundreds of possibilities because they are frequently applied or critical in the domains of environmental control and urban planning.

Both vector and raster operations are considered, highlighting the difference in working with these two data structures.

Today a major breakthrough is taking place in the GIS industry. New systems are under development that will have greater capabilities and, more importantly, flexible and easily-customisable user interfaces. However, since most of the GIS available on the market are still bound by current technology, recommendations will mainly refer to the state-of-the-art of commercial products. That is considered a meaningful contribution to address the current practical needs of users.

Since GIS functions, operations and capabilities are illustrated or discussed in a number of good books, reports and manuals (see the selected list of references), attention will concentrate on suggestions and advice that are *complementary* to what the interested reader can already find in literature.

7.2 Recommendations

The recommendations cover the activities that are most common in a project for GIS development and implementation, namely:

- project set-up
- data entry
- data conversion
- data validation
- data visualisation/rendering
- map database management
- attribute data management
- map processing/analysis
- map/report production

Additionally, for ten selected key functions a more detailed description is provided in Tables 7A to 7J.

7.2.1 Project set-up

When the user begins a GIS project, initially and for each daily work session, many commands and parameters need to be initialised or defined. These range from parameters controlling the screen viewing characteristics (i.e. default colours for the background and foreground objects, etc.) to commands that determine the resolution at which operations such as feature editing or generation are interactively or automatically performed. Most of these settings can be registered in some macro or script files; hence the user can easily and consistently start all the subsequent sessions of the work.

The user with limited experience either in GIS techniques or in the application domain should carefully verify the input data spatial resolution (i.e. the minimum distance between two features that can be distinguished). Parameters controlling resolution (i.e. tolerance, weed, rasterisation cell-size, etc.) of data acquisition and editing must be in agreement with the characteristics of the input data and the final objectives of the project.

Usually, such a choice will require a sound project evaluation analysis, following the guidelines described in section 5. The user should be careful not to allow existing data to overly influence the collection/creation parameters of the rest of the dataset. He/she should select these parameters according to genuine project objectives and established quality control measures, not on the possible shortcuts of adjusting everything to match existing data.

7.2.2 Data entry

Data input is frequently the most costly and tedious phase of any GIS development project. Errors in designing the data entry strategy and methodology may eventually result in a complete failure of the GIS project.

Regardless of the strategy used, manual digitising of maps remains an error-prone, long and tedious activity. In spite of the technological evolution, map raster scanning and vectorisation can be successfully applied only to simple maps. When a complex map (i.e. a map displaying many classes of objects) needs to be digitised, it may be convenient or feasible to do the following:

- hand-redraft the map using a set of stable transparency sheets (one for each class of objects);
- scan each sheet using an appropriate resolution (i.e. from 300 to 600 dpi);
- vectorise each sheet using either automatic or semiautomatic algorithms;
- interactively/automatically label each vector feature (point, line, polygon).

In general, users should avoid the direct digitisation of paper maps that invariably change in size and shape with temperature and air moisture.

Map digitising requires a (costly) digitising table whose characteristics and dimension (the most appropriate tables are A0 or A1 in size) should match the project objectives.

7.2.3 Data conversion

Geographical data can be organised into two basic structures or models: *raster* (cellular) or *vector*. These two models are inherently different, the former representing the phenomena within a field (e.g. land use within a group of cells), the latter representing the objects which delimit a phenomenon (e.g. the polygon defining a land parcel, which may have the land use value attached as an attribute). These are often distinguished as *field* versus *object* views of geographical information.

In grid-based GIS, map data entry takes place through digitisation of map features and subsequent vector-to-raster conversion of vector files. In vector-based GIS, scanned maps or digital imagery need to be converted to vector structures. In general, efficient spatial data manipulation and analysis require converting structures from vector to raster or from raster to vector.

Today, virtually all systems are capable of rasterising vector maps through batch commands. Traditionally, the main source of errors in the conversion derived from the selection of a grid cell (or pixel) too large (i.e. 100x100 m or greater) to faithfully resolve spatial details of the input maps. The current tendency to use very small grid-cells (i.e. 10x10 m) greatly reduces the spatial inaccuracy of the operation. On the other hand, to cover even a small region (i.e. 200-500 km^2), an overwhelming number of grid-units is required leading to unmanageable problems of both CPU time and numerical instability when data have to be processed by complex statistical techniques. Hence, before performing any vector-raster conversion, it is necessary to carefully select the most appropriate size of the raster cell.

As outlined in Tab. 7A, raster to vector conversion involves either a single batch process on a raster image file, or a semiautomatic process of line-following. The former offers simplicity at the expense of accuracy (when ambiguities occur and the algorithm “takes a guess”), while the latter offers more user control at the expense of complexity (more labour-intensive).

7.2.4 Data validation

Like data entry, data validation constitutes a long, costly operation. Unfortunately, very few GIS products provide valuable tools which facilitate the operations of checking the quality and integrity of the digitised data and of the related attribute information. This is quite surprising because governmental institutions invest a large amount of money/time in attempting to evaluate the quality of data digitised elsewhere (generally, by some private company). A few GIS do provide this checking mechanism, via user-programmable rules which, unless followed, do not allow the user to continue.

Besides the detection and correction of topological errors (Tab. 7B), most GIS software incorporate a limited number of functions which aid the user in identifying categorical errors in the input data (i.e. a mislabelled polygon of a land-use map). The problem becomes crucial when such data have a large volume. Under these circumstances, by displaying the map along with spatially related terrain features (i.e. ground relief and slope), it may be possible to detect and correct unrealistic or wrong associations of different features (i.e. a rice field located on a steep slope, a landslide laying over a highly stable rock type, etc.).

By superimposing a plot of the digitised map on top of the input map, it is frequently possible to visually detect several errors and discrepancies.

7.2.5 Data visualisation /rendering

One of the remarkable features of modern GIS tools is that map displays are no longer static. Rather, they provide a highly interactive medium for rapidly editing and exploring geographical data. Today, most systems allow for *moving* and *resizing* both the display window and all map (layout) components (frame, title, legend, feature labelling, etc.). Such tasks are achieved using different approaches, among which the most effective appear those derived from the windows environment (Tab. 7C).

Additionally, linear or polygonal features can be coloured according to user-selected colour schemes; while point features can be displayed according to graduated colour or size/shape legends or labelled with any alphanumeric string. More advanced functions allow for displaying objects with symbols or colours that are dependent on the value of an attribute dynamically linked to the object.

Raster (grid) data can be portrayed and explored exploiting a wide family of functions such as image *rotation*, *flipping*, *mirroring*, linear *stretching*, *equalisation*, *filtering*, etc.. Additionally, the ability of many systems to display simultaneously vector data over a raster image constitutes a fairly new functionality, which finds application in a variety of fields (Tab. 7D).

It is worth mentioning that many display operations are tightly linked to analytical operations (distance/cost analysis, buffering etc.); the first and the second are indeed the two faces of the same issue: *understanding geographical phenomena and processes*.

7.2.6 Map database management

A great number of different functions and operations fall under such a heading. Some of them are common to many other graphic tools (i.e. CAD systems), others are specific of GIS environments. To the fist class belong the operations for adding, deleting, moving and rotating features. To the second appertain functions such as projection change, conflation and georeferencing (see section 6). Source maps may come in very different projections (Albers equal area conic projection, UTM eastings and northings, etc.); so they need to be converted to a common coordinate system by means of projection changes. Since the operation may require knowledge of basic topography and cartography, it may be advisable for some users to consult an expert in these matters before embarking on a complex GIS project.

Likewise, to rectify and georeference a satellite image, the input raster must be resampled to a new grid, where the coordinates are spaced according to the new common coordinate system. The task requires a warping of the original image to another through the use of selected ground control points (GCP). In general, GIS incorporate a subsystem to accomplish this operation through a more or less friendly user interface. In some cases, GCP may be copied from one layer to another (i.e. from a cadastral vector layer to a raster layer of land-use cover) within the subsystem menu.

In spite of the functionality of the GIS UI, all the above operations, along with those aimed at matching adjacent maps or images (map joining), are time consuming and frequently frustrating even for users familiar with the matter.

The user should be fully aware that any spatial error introduced into the map database at this stage may adversely affect all subsequent spatial and tabular analyses. Furthermore, error removal at the time of analysis could be very difficult or even impossible.

7.2.7 Attribute database management

In spite of what is frequently stated, most GIS activities require rather simple data base management operations (query and retrieval) on the tabular (attribute) data associated to the geometrical features. However, under some circumstances, users must dedicate considerable time to the (conceptual and physical) design of the database.

At the time of GIS project initiation, some tabular databases may already be available. They need to be properly linked either to the geometric features of the GIS maps or to existing GIS related tables.

These data are often not in the desired format or structure. Tables may need to be joined or split, and new relations between tables formed, according to the desired geographic model. These actions would be executed using several commands within the database management system. Major redefinition of the tabular database may involve such work that it may be advisable to start from scratch which is a decision to be made by a competent database expert and the project coordinator (Tab. 7E).

Lastly, in many circumstances the user can decide to postpone the acquisition of a large and costly DBMS and use the (small) database that generally comes with the GIS software.

7.2.8 Data processing and analysis

Proximity analysis (buffer, distance calculation), spatial joins (overlay), network analysis (optimal routes, allocation of resources), line/polygon generalisation or smoothing, region boundary dissolving, generation of database views and many other fundamental GIS functions (Tab. 7F-7I) fall within this very broad class.

Such functions will differ significantly depending on the data structure (raster or vector) used by the system. For example, network analysis is almost impossible with a grid structure;

while neighbourhood analyses, such as the production of a simple slope map, are difficult or impossible with vector data. In general, grid-based GIS are feasible or optimal where the objects to be manipulated are essentially points or regions which require an intensive use of interpolation, filtering and overlay functions. Vector GIS are more efficient when linear features must be handled and large databases should be linked to the spatial features. Optimal GIS incorporate both vector and grid functionality (this assumes a higher cost).

In Tab. 7G and 7H, we outline a strategy based on the transformation of the spatial data from one structure to another depending on the analysis to be performed. The approach may appear somewhat cumbersome and time consuming. However, in certain circumstances, it is the only feasible way to face problems which are otherwise unmanageable.

Although software reliability is beyond the scope of this document, the user should be aware of the difficulty in detecting inaccurate results from some complex spatial functions.

7.2.9 Production of maps and reports

Because we assume that a map is the best way to communicate georeferenced information, the map product must be good (aesthetically, logically and conceptually) in order to accomplish this task successfully. Most cartographic products of the GIS environment consist of simple maps displaying a few classes of features.

Cartographic products, which are aesthetically similar to traditional, hand-drafted maps are still very uncommon within the realm of GIS. Their production requires sophisticated output devices coupled with advanced software modules. Furthermore, GIS users need to be familiar with the basic rules of information transfer though cartography. In Tab. 7J, basic advice is provided to improve the quality of the GIS map products using current hardcopy technology.

7.3 Exemplar key GIS functions: main characteristics, uses, problems and recommendations (tables 7A to 7J)

Table 7A	**Data Conversion** **RASTER TO VECTOR**
Definition	The conversion of spatial data from raster (cellular) to vector structures (models).
Classical use	Traditionally, the function is used to import *classified* satellite or aerial digital imagery into a vector-based GIS. The result of the operation normally consists of a map displaying polygonal or linear features labelled according to the classes (i.e. forest, pasture, etc.) into which the pixels of the input imagery were classified (i.e. land-use cover) through some algorithm. The operation also permits transferring geodata from a grid-based GIS to a vector GIS. Likewise, it allows for exploiting both data structures within those (few) GIS that support raster and vector representations. More recently, the operation has become part of a process aimed at digitising line or point features from existing paper maps. The process involves: **a)** the *scanning* of the map that produces a (generally huge) file in raster format; **b)** the *thinning* and *vectorising* of rasterised features using semi-automatic or automatic algorithms; and **c)** interactive feature labelling.
Other uses	Fast “quick and dirty” raster analysis of an area where precise linear cartographic analysis and graphic output is not necessary or possible
Frequent problems	In general, conversion from raster maps to a vector polygonal representation consists of a) recognising the borders between differently coded cells; b) reconstructing feature topology; and c) performing a topologically correct and aesthetically acceptable line smoothing. At present, most GIS correctly accomplish steps a) and b), while they either do not perform or fail in step c). Conversion of large raster data sets (i.e. Landsat or SPOT scenes) to a vector dataset is always a computer intensive operation. Usually a batch process, that may require hours to complete, performs it. More importantly, the resulting vector map may consist of an overwhelming number or polygons, many spurious, leading to serious problems at the stage of topology construction. Pixel *thinning* and vectorisation of scanned paper maps that display two or more classes of features (i.e. roads, rivers, etc.) requires so much subsequent interactive editing that the procedure becomes economically unfeasible.
Recommendations	Before importing and converting a satellite/aerial image to the vector map database, the user should carefully classify the image into a few meaningful classes. Likewise, the smoothing/filtering techniques used to eliminate superfluous or meaningless polygons in the vector map, should be selected and used with great care. Before attempting any conversion, the user should evaluate the use of the imagery as read-only background in visual display. To exploit scanner/vectorising technology on complex paper maps, the user should follow the strategy outlined in section 7.2.2. In some cases the correct conversion of complex raster databases to the vector representation may require the advice of an expert.

Table 7B	**Data validation** **IDENTIFICATION/CORRECTION OF TOPOLOGICAL ERRORS**
Definition	Locating and correcting the digitised data to reduce or eliminate spatial and topological errors. Making spatial data usable for certain types of analyses (i.e. network tracers) requires topological construction, which makes the relationships explicit between features (line connectivity, area definition and contiguity, etc.). Correction of errors means adding missing data, and removing or replacing inaccuracies with correct data. The basic steps to make sure that data in a digital map is topologically correct are: **a) Construct topology**: procedure for explicitly defining spatial relationships between features in a digital map. It stores data more efficiently and allows for geographical analyses, **b**) **Identify digitising errors**: procedures for identification and marking of potential errors, **c) Correct errors**: process of adding missing data, and removing and replacing inaccuracies with correct data, and **d) Reconstruct topology**: necessary process of reestablishing the spatial relationships between features after the correction of topological errors.
Classical use	Topological construction is the necessary process to create *intelligent* maps from *spaghetti* data (i.e. a simple set of xy coordinates). It recognises and establishes the relationships between map features, thus allowing for subsequent analysis operations where it is mandatory to know spatial connectivity and contiguity. The function is also the basic step in order to create the feature attribute table linked to the cartographic database. The operation adds a primary ID to each feature and prepares links to other external databases.
Other uses	The function is also used to perform iterative quality control within and between different map layers; it allows regeneration of original layers form overlaid-derived layer. Some GIS software allows *ad hoc topology,* very convenient for some applications.
Frequent problems	The most common errors that users face when constructing topology consists of arc that do not connect (i.e. undershooting, overshooting, etc.), polygons that are not perfectly closed, *sliver* polygons, polygons not labelled or mislabelled. Some GIS do not support vector arc-node topology. In many GIS the procedures for identifying topological errors are not well automated, creating a labour-intensive environment and leaving many editing decisions to the end-user, who may not be prepared for this type of work. Identification and correction of topological errors may become a major issue when the volume of the map database is large or very large (say over 50 Mb).
Recommendations	In order to reduce the number of digitising errors, the user should invest more time both in properly designing the workflow for data entry, and in accurately executing the digitising operation. When the input map exhibits very many digitising errors, the user should carefully evaluate two following alternatives: correcting such errors or restarting the digitisation process from scratch. Under many circumstances, the second alternative may prove to be the most feasible. Since error propagation through overlaying operations is a frequent problem, the user should carefully check the quality of each base map before starting any analytical operation.

Table 7C	**Data visualisation/rendering** **ZOOM and PAN**
Definition	Zooming is the process of magnifying or reducing the scale of a map or image displayed on the monitor. Panning is the process of changing the position at which the view is displayed, without modifying the scale.
Classical Use	Zooming may take place in such a way that: ***(a)*** the new rendition is centred at the same location as the previous one (*zoom in/out centre*); ***(b)*** the new zooming centre is selected by the user; ***(c)*** the user can magnify a sub-region by selecting the opposite corners of the rectangle encompassing the area of interest; ***(d)*** the user can assign the *exact, real* scale to the map visual display by typing the scale denominator. The ability to redisplay the map before the last zooming ("*previous*") is a fairly common option. All the above techniques can be accompanied by the capability of creating, on user request, a second ancillary window where the whole map is displayed for reference (bird's eye view). Panning can take place by discrete movements of the view point using hot keys, continuously through traditional scroll bars or, more intuitively, by means of cursor dragging (the icon often assumes the shape of an open hand).
Other Uses	Zooming of type (**d**) proved to be particularly tailored when the user needs to exactly know how the map will look when plotted at that given scale. In a few systems, during the execution of an editing command (i.e. screen line digitising), the mouse pointer either remains constantly at the centre of the display window (dynamic panning) or is magnified within an ancillary window (dynamic zooming).
Frequent problems	Some GIS do not allow for multiple zooming options. Few options are still available for map panning. Some raster-based GIS still lack such functionality. Panning and zooming may become very slow operations when map data volume increases and, thus, must be read (paged) from disk to memory.
Recommendations	When the user is mainly concerned with the exploration of relatively small data volumes (say < 2 Mb), the recommended zoom and pan functions should fulfil the following requirements: a) they should be *versatile*, that is, the operation may be accomplished using different approaches, such as those outlined at the points ***a***, ***b***, **c** and ***d***; b) each approach should be *intuitive; and c)* each approach should allow for *repeating* the operation many times without excessive stress on the user. At present, several small and large GIS tools meet the above requirements. When the user has to deal with large data volumes, data structure and software optimisation become of ultimate importance on the GUI. Under such circumstances, the user must select a GIS that fulfils the requirements of both an efficient GUI, and optimised software architecture. At present, few GIS are available on the market that meet the above requirements. In order to evaluate software efficiency (algorithm speed), the potential customer should test zoom–pan functions on a large (say over 20 Mb) data set (map) obtained from the GIS vendor or elsewhere. Always consider that (re)drawing speed is affected by the system memory (more is better) and by the frequency of disk reads.

Table 7D	**Data visualisation/rendering** **MANAGEMENT OF BACKGROUND IMAGES**
Definition	The process of displaying and managing on the same view vector objects (point, lines and polygons) in the foreground and georeferenced raster imagery (i.e., satellite imagery) in the background.
Classical Use	Interactively update spatial objects that change in time (i.e. road network, built-up areas, waste deposits areas, coastal lines, river courses, etc.), using updated, properly georeferenced imagery (i.e. a satellite imagery, a scanned aerial photograph or topographic sheet). Interactively add new spatial objects (i.e. buildings, power lines, etc.) to an existing collection using background imagery. Visually investigate and analyse the spatial relations between some objects in vector format (i.e. landslide deposits, recreational areas, etc.) and some landscape or environmental features acquired as images (ground relief, ground slope, soil composition, etc.). The above operations are by far less expensive than the vectorisation of the imagery and the subsequent use to update the existing map database.
Other Uses	Handling of ortho-photomaps.
Frequent problems	When the original source document (i.e. topographic, soil, geological maps) is scanned at high resolution, the resulting background raster file may be very large (say 20-40 Mb) and cannot be readily managed by means of a small computer platform. Additionally, when attempting to send to a (inkjet) plotter such huge files, unmanageable problems may arise. Usually the background image needs to be carefully georeferenced. The task may be rather cumbersome. Ground control points of a topographic reference sheet cannot be easily identified in some type of imagery (i.e. SAR satellite imagery). Transformation of the image into real world coordinates could require either a rotation or an affine transformation; many GIS tools are not able to perform such operations. If the projection of the coordinates of the vector objects is different from that of the imagery coordinates, a projection transformation is needed. Many GIS products lack such a capability; additionally, some projection change will require some knowledge in topography. Scanned aerial photographs cannot be exactly georeferenced because of their perspective distortion; they can be used for qualitative work. Some systems do not allow for modifying the colour classes of the background image (from colour to grayscale palette, linear stretching, etc.)
Recommendations	In scanning the documents that are the source for the background image, the user should carefully select the most appropriate resolution, balancing raster file size and “readability” of the image. If the available GIS system lacks image analysis capabilities, the scanned file could be first pre-processed using one of the shareware image modules that can be found on the Net; then transferred to the GIS for visual exploration and analysis. In many circumstances, the use of a grayscale palette (colourmap) for the background image yields the best results in terms of clarity and interpretability (i.e. a shade relief image on top of which the vector stream network is placed). If the raster file is large, plot it at the lowest resolution (i.e. 150 dpi). Keep in mind that the monitor (screen) resolution is about 70-90 dpi.

Table 7E	**Attribute database management** **ESTABLISHING MORE COMPLEX RELATIONS**
Definition	Implementation of graphic/alphanumerical connections between graphical and alphanumerical databases. This connection is based on the use of a GIS internal table as linkage table to other tables in external databases. The set or collection of data that describes the characteristics of real world entities or conditions is frequently too large to be stored in a single table associated to the graphic elements. This data are usually managed by a relational database management system (RDBMS). The usual procedures are based in the connection of each graphical element to a line of a column of the alphanumerical table containing its attributes (record). This action can be performed automatically or not depending on the GIS software used.
Classical use	Connection of a cartographic base map, as for example administrative boundaries with large thematic databases allowing the production of thematic maps from views resulting of SQL queries to the external database. The use of SQL queries performed over the external alphanumerical tables allows an extensive list of possibilities on combining information and defining conditions that result in the generation of highly comprehensive thematic maps.
Other uses	Use of large external databases, not designed for geographical analysis purposes, but useful in a specific context. Connection of cartographic base maps to external databases containing data captured from sensors on a real time basis (i. e: taxation information organised in a GIS; dynamic GIS applied to traffic control).
Frequent problems	Consistency of more complex relations like graphical alphanumerical connection can be affected by changes on the elements of the system. These changes affect the topological definition of the graphical elements that is the basis for the graphical-alphanumerical connection (i. e: creation of new polygons, arcs or nodes; creation of new tables and/or columns in the external database). Other problems are related to the nature of geographical information, namely, its description by alphanumerical attributes can be difficult as alphanumerical databases usually deal with fixed record sizes and graphical elements topological description is often too long.
Recommendations	The graphical alphanumerical connection with external databases must be implemented when graphical and alphanumerical elements are perfectly defined considering the possibility of establishing relational links. Control of the topological definition of systemic elements and systematic verification on the consistency.

Table 7F	**Data processing/analysis** **CREATE, SAVE DATABASE VIEWS**
Definition	Database views may be created through SQL queries based on logical operators or similar *select* options. Logical operations deal directly with the database (alphanumeric information) and allow the user to identify and select features by a specific set of criteria. Generally, features are identified and selected according to a combination of several conditions. In a typical application, a specific item in the database is employed to differentiate features that satisfy different sets of selection criteria. Graphical selections, using the mouse to define an area of interest, may also form part of a view.
Classical use	Database views may be created and reused at a later time, by matching the defined criteria selected features and saving them in a new file for further manipulation, classification, or analysis. These views are often the building blocks of complex analyses, whereby the GIS does not allow complex operations in a single pass. For example, the user selects all clients (points); then reselects only those clients worth some amount; then determines which of that subset live within 2 km of a certain highway. That final subset of elements (points) would be a new view, and could be saved in a file called "clients_delivery_route_1".
Other uses	Summary statistics can often be derived for the selected features. Basic summary statistics include mean, variance, maximum, minimum, range and frequency of every attribute selected by the user. Additionally, the same process of creating database views serves for any purpose whereby only a certain subset of the dataset should be selected, highlighted, edited, deleted, etc. It therefore provides added control to potentially dangerous situations where the entire database should not be at risk.
Frequent problems	Database design, if not carefully thought out at the beginning of the project, can frequently leave the user with dead-ends at the time of complex query formulation. Some GIS vendors are steering toward implementations whereby all data, attribute as well as spatial, are stored in tabular format in RDBMS. This may have some implications at the time of GIS acquisition and implementation.
Recommendations	Many novice GIS users see the tool as purely a cartographic one, whereas seasoned users quickly find that, once the cartography has been automated, many of the difficult issues in GIS are database-related. Therefore, training in the optimal use of the RDBMS in question is advised, beyond the few pointers given in the typical GIS training course. Database queries, or views, may be constructed previous to analysis, to be ready for their eventual use. A geographic database with (perhaps) millions of elements may require minutes or even hours to generate certain complex views. Therefore, in many cases these may be created ahead of time and stored (cached). This capability differs according to RDBMS manufacturer.

Table 7G	**Data processing/analysis** **BUFFERING (dilation)**
Definition	The process of generating a polygon that encloses an area within a specified distance from one or more point, line or polygon features.
Classical Use	Perhaps, buffering is the most popular spatial operation appertaining to the wide family of proximity (neighbouring) functions (i.e. interpolation, contouring, trend, spanning, etc.). Buffering can be applied to point, line and polygon features. It produces proximity corridors or buffer zones surrounding the selected (point, line area) features. It can be carried out with either vector or raster data structures. Usually the purpose of buffering consists in creating a map which displays proximity to a selected feature. In the domain of environmental control, the function can be used to identify all features (i.e. dwellings) located within a given distance from a dangerous locality (i.e. a landslide deposit). In urban planning, by creating a corridor along the road network, all the features (i.e. buildings) affected by the acoustic noise produced by the traffic are readily identified. Often, maps showing proximity to selected features are subsequently used for further spatial analyses and investigations, such as cost analysis and optimal pathways analysis.
Other Uses	Buffering can be used to convert point or linear features to polygonal ones, as it might be required by some spatial analysis or cartographic print out. Special kinds of dilation operations are used in mathematical morphology for the analysis and enhancement of images (feature extraction).
Frequent problems	Today most GIS incorporate some kind of buffering function. However, if the features to be buffered are very many (say over 10,000), and the analysis is performed on vector data, the operation can become time consuming and computationally very demanding. This is a serious obstacle to the use of this function as an exploratory tool that should be executed many times varying corridor width and corridor confining conditions. Several vector GIS do not allow for controlling the way the buffer corridor is calculated depending on the values assumed by one or more attributes of the table linked to the feature to be dilated (i.e. roads classified for traffic intensity will be buffered with a corridor width function of the traffic). Owing to the limitations of the UI, dilation operations, which depend on the characteristics both of the feature to be buffered and of the properties of the surrounding area, are rather cumbersome when performed in most GIS environments.
Recommendations	To test a model or simulate a process, very many buffering operations could be necessary, varying both corridor width and constraining conditions. Under such circumstances, dilation can be carried out much more efficiently in a raster environment. Hence, the user could consider following the strategy outlined in Tab. 7.H. In order to obtain meaningful maps displaying proximity of given features, it is of paramount importance to evaluate carefully the physical meaning of the operation. Too frequently, users exploit this powerful function without understanding that the surrounding or the buffered feature is neither homogeneous nor isotropic.

Table 7H	**Data processing/ analysis** **MAP OVERLAY (Spatial Joins)**
Definition	The process of superimposing two maps such that the resulting map contains spatial and attribute information from both input maps.
Classical Use	Map overlay is one of the most fundamental and versatile spatial functions of a GIS, aimed at detecting and highlighting the spatial relations between two or more phenomena or processes. In the vector structure, the overlay of two polygon maps (i.e. soil and land-use maps) leads to the generation of a new set of polygons (and attributes) that explain the relations existing between the two input spatial data (i.e. vegetation type and soil nature). Likewise, by overlaying point (i.e. seismic epicentres) or line (i.e. streamlines) features over a polygon map (population density), populated areas at seismic or flood hazard can be readily delineated. In vector and raster structures, map overlay can follow Boolean algebra (AND/OR/NOT/XOR) applied to feature attributes. Depending on the spatial model to be adopted in the different cases (i.e. union vs. intersection), the resulting map (and tables) will greatly differ in form and content.
Other Uses	By sequentially overlying pairs of raster or vector maps, each displaying a certain land characteristic (soil, geology, land-use, slope classes, etc.) a final map showing the subdivision of the land *into unique-condition-units* is obtained. Such a map and associated table constitutes the input for developing many spatial models.
Frequent problems	When overlaying two maps, each having polygon features with convoluted boundaries, the resulting map may have very many small and seemingly unimportant polygons (*slivers*) that are generated when lines on the two maps were almost but not quite identical. If the two input maps have very many polygons (say over 10,000 each), *sliver* polygons can become enormous (say 100,000 or over) demanding, among others, very heavy computation for building the topology. In order to rebuild a clean map, slivers must be filtered away or merged with adjacent polygons. Although most GIS have the capability of eliminating such unwanted polygons, few systems allow for controlling how filtering/merging are accomplished. Thus, often considerable manual, very time-consuming, intervention is required.
Recommendations	In order to obtain a meaningful overlay of two maps, it is of paramount importance to carefully classify or reclassify the value of the entity displayed in each input map. If very many overlay operations are needed on a large set of vector geographical data, it is worth evaluating the following strategy: **a)** reclassify all maps into a limited number of meaningful classes, **b)** convert each map into a grid structure with a grid-cell size close to the resolution of the input data; **c)** perform all spatial analyses in such new format; **d)** reconvert the results of the analysis to the original vector structure. As already discussed, the last step may prove to be rather cumbersome.

Table 7I	**Data processing/analysis Data analysis** **DIGITAL TERRAIN MODELLING**
Definition	The process of generating and handling a digital terrain model (DTM), a numerical representation of the earth surface based on a set of x, y and z coordinates. The z value may also represent any other spatially continuous attribute such as sea depth, air pollution or population density.
Classical Use	DTM and its derivatives (slope, aspect, local slope curvature, etc.) are used in land-use planning (i.e. agriculture), engineering works (i.e. cut and fill problems), military activities, and analysis of cross-country visibility. DTMs permit the production of perspective views of the relief that may be associated with landscape features, which are spatially related with ground elevation (map *draping*). DTM products such as shaded-relief maps are typical background images for exploring several landscape or man-made phenomena and processes.
Other Uses	Support for simulation of natural phenomena and in urban planning.
Frequent problems	Errors or uncertainty in the elevation source data (i.e. incorrectly labelled contour lines or spot heights; inadequate number of heights over very steep or very flat regions), can seriously affect the generated DTM. Grid formats and triangulated irregular networks (TIN), which are the fundamental DTM structures implemented in GIS, can be almost equally affected by errors. In addition, GIS systems provide little information on the limitations of the implemented algorithms. Grid-based DTMs suffer from large errors derived from inefficient algorithms that attempt to interpolate the original contour lines without exploiting the simple topological proprieties of contours. Likewise, most TIN generators are unable to generate triangles in areas of high contour curvature (ridges or stream courses) or in flat regions. Many users find it difficult to assess how faithfully the generated DTM reflects the original ground surface as expressed by the input data (digitised contour lines, spot heights, profiles, etc.). Additionally, the selection of the most appropriate resolution of the final DTM (grid or TIN density) is not an easy task.
Recommendations	Depending on the type of investigation and landscape characteristics, either grid-based or TIN DTMs provide the best information on the relief and its derivatives of the region under investigation. In general, DTM structure and data resolution should allow for preserving in the 3D model the morphological information pertaining to the original input data (contour lines, spot height, etc.). If a grid DTM has to be generated from digitised contour lines, the following strategy is highly recommended: **a)** rasterise contours with a grid-cell size less than one third the input contour interval; **b)** resample (by pixel *thinning*) the resulting, usually too dense, raster file to the desired grid spacing. When a TIN DTM has to be produced, the user should find out if the algorithm is able to generate automatically ancillary points for creating meaningful triangles in ridge or valley areas. Alternatively, such points have to be added manually. In some cases, the best way to produce a faithful DTM consists in calculating elevations directly from the aerial photographs through digital photogrammetric techniques.

Table 7J	**Map/report production** **GENERATE COMPLEX GRAPHICAL PRODUCTS (MAP COMPOSITION)**
Definition	The combination of map layers to compile a graphical representation of data, both geographic features (points, lines, polygons) and cartographic annotation (titles, text, legend, scale bars, etc.). The resulting product can also incorporate other graphical elements such as histograms, scatterplots, variograms, and the like.
Classical use	Most of the maps generated by GIS are experimental, and not designed to be of professional cartographic quality. In general such maps are generated in order to try particular data combinations, to test hypotheses and to browse through spatial data in search for patterns and meaning. As a result, most of maps produced during the development of a GIS project are simple and aesthetically poor. In order to facilitate the transfer and dissemination of the final results of a GIS project, high-quality map products may also be generated through the use of appropriate subsystems permitting to select among very many symbols, shadings, colour palettes and annotations.
Other uses	The operation is also frequent for the production of simple draft prints on transparency paper that allow for checking digitising errors. Recently, GIS providers are developing subsystems for publishing digital maps on the Internet.
Frequent problems	Many users are totally unfamiliar with the most basic rules for producing a map which effectively conveys information to the reader. Frequently, the colours in the printed copy of the map do not match the colours that appear on the monitor screen. Sometimes, such a mismatch also affects some kind of symbol patterns. When using inkjet plotters, large maps (i.e. maps plotted on A0 paper size) imply tens of megabytes of RAM and possible problems with map output.
Recommendations	It is important to always think about the map reader needs and focus on production of outputs that meet reader requirements, namely to have a final product that is easy to read and interpret. Before producing a complex, high quality map, the user should: **a)** determine the purpose of the map; **b)** identify the map reader; **c)** design the components of the map; **d)** determine the map scale; and **e)** select the most appropriate symbols, colours and annotations. The production of an aesthetically valuable map composition will be greatly facilitated by properly organised map and attribute databases. Before plotting a map, the user should view and inspect relevant details of the maps (mainly annotations) on the monitor screen using a zoom scale equal to that of the final plot. To avoid printing problems with large size (A0) inkjet plotters, the user should add maximum (say more than 40 Mb) RAM to the output device.

References

Aronoff, S. (1989). Geographic Information Systems: a management perspective (294 pp.). Ottawa: WDL Publications.

Bonham-Carter, G.F. (1994). Geographic Information Systems for geoscientists. Computer Methods in Geoscience v. 13, Pergamon, Ottawa, 394 pp.

Bugayevskiy L. M., & Snyder J. P. (1995). Map Projections, a reference manual (328 pp.). London: Taylor & Francis.

Burrough, P.A. (1986). Principles of geographical information systems for land resources assessment (194 pp.). Oxford: Clarendon Press.

Carrara, A., & Guzzetti, F. (eds.) (1995). Geographical Information Systems in Assessing. Natural Hazards. Dordrecht: Kluwer Acad. Publ.

Green R., & Rix D. (eds) (1995). The 1995 AGI yearbook for Geographic Information Systems (392 pp.). London: yearly published by the English Association for Geographic Information.

Kovar, K., & Nachtnebel, H.P. (eds.) (1994). Application of Geographic Information Systems in hydrology and water resources management. Proceed. HydroGIS 93, Vienna, April 19-22 , IAHS Publ. No. 211., 694 pp.

Laurini, R., & Thomposon, D. (1992). Fundamentals of spatial information systems (680 pp.). San Diego: Academic Press.

Maguire, D. J. (1989). Computers in Geography. London: Longman.

Maguire D. J., Goodchild M. F., & Rhind D. W. (eds). (1992). Geographical Information Systems. Harlow-Essex: Longman Scientifical and Technical, 2 volumes.

Peuquet, D. J., & Marble D. F. (1990). Introductory Readings in Geographic Information Systems. London: Taylor & Francis.

Raper, J., (ed.) (1989). Three dimensional applications in Geographical Information Systems. (189 pp.). London: Taylor & Francis.

Star, J., & Estes, J. E. (1990). Geographic Information Systems an introduction. Englewood Cliffs, New Jersey: Prentice Hall.

van Westen, C.J. (1993). Application of Geographic Information System to landslide hazard zonation. ITC-Publication No. 15, ITC, Enschede, 245 pp.

As a complement to sections 6 and 7 and related also to the tables of section 7.3, some pictures are added, to further illustrate the outlined functions, and/ or to show how these functions can be used and joined in the flow of the GIS process.