AN INFORMATION SYSTEM FOR THE GEO-SCIENCES

W. Kainz
Institute for Image Proc. and Computer Graphics
Wastiangasse 6, A-8010 Graz, Austria
Commission IV, Working Group 8

ABSTRACT

The increase in environmental problems and the demand to have efficient means for capturing, storing, updating, processing, and displaying of spatial and environmental data lead to the development of so-called geographic or land information systems (GIS).

This paper describes how such systems are designed to provide fast and cost effective computer-aided techniques for the required tasks. An implementation at the Institute for Image Processing and Computer Graphics, and results acquired by application of GIS technology are shown.

The process of bringing analog spatial data into a digital computer is called digitizing. This can be done either manually, semi-automatically, or automatically. An interactive digitizing system with powerful editing capabilities is part of every GIS.

For the storing of spatial data, vector as well as raster data structures are applied. The assignment of alpanumeric or symbolic attributes to spatial items leads to the concept of a spatial data base.

Data manipulation and processing are provided by a data retrieval and analysis system. Various kinds of spatial analysis and image processing techniques are used to yield useful results. Graphs, charts, and maps are displayed on appropriate output devices through the aid of computer graphics and computer-assisted cartography.

INTRODUCTION

Growing pollution and the increasing number of environmental hazards make regional and environmental planning, and the monitoring of environmental systems more and more important for local and federal governments. A lot of data from various sources have to be collected, stored and made available for further processing. These data include topographic information, hydrology, geology, meteorology, census and many other geo- and space related data.

Data capturing is mainly done by governmental and private agencies thus producing a great amount of information. Unfortunately these data occur in different formats and are filed in different places, making combination and update of data a difficult task.

So we are aware of three main problems in conservative geo-data processing today:

- incompatibility of data from different sources
- lack of data revision
- lack of efficient procedures for data analysis

These problems can be overcome by application of computers and computer-assisted information systems for the geo-sciences, i.e. automated geographic or land information systems, that share a central data base.

1. DESIGN

The tasks of an automated geographic information system are capturing, storing, updating, analysis and display of spatial data from various sources by application of computers and data processing equipment.

Each GIS has three main parts: an input system, a storage and data management system, and a retrieval, analysis and display system (Fig. 1).

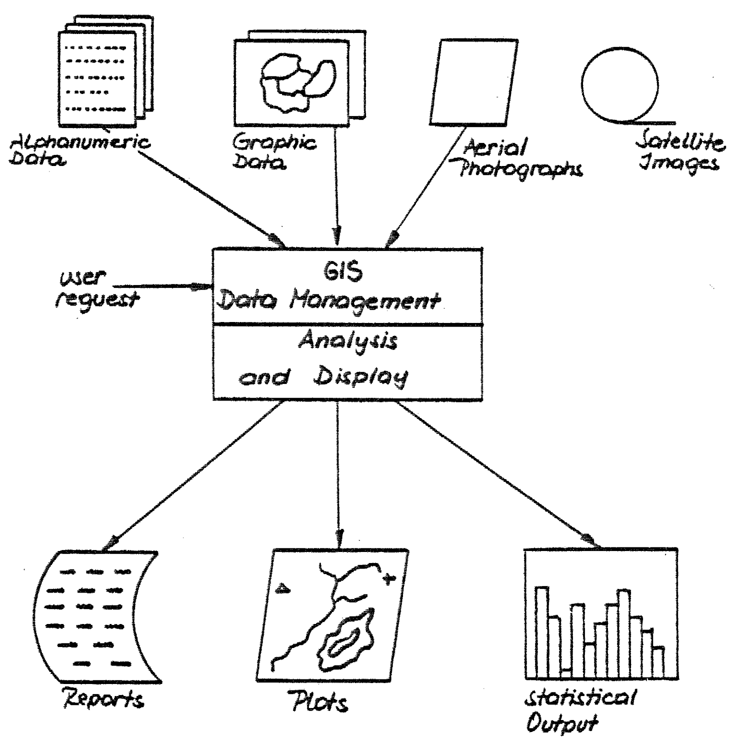


Fig. 1: Main parts of a GIS

1.1 Data Input

Input data for a GIS come from many sources and in different formats. To make these data comparable we define a four dimensional data space. Data are assigned according to their dimensions [13].

Dimension one and two (x,y) define spatial elements, thematic elements belong to the third dimension (z), and time elements to the fourth (t).

Spatial elements are divided into point, line, and area elements, which are defined by their planimetric or geographic coordinates (e.g. drill holes, rivers, and lakes).

Thematic elements can be either numeric (height, depth, interval), alphanumeric (river or district name), or symbolic (graphic symbol for map fetaures). They are used as attributes for spatial elements.

Time elements are only needed when certain thematic elements get different values at different times (e.g. temperature during seasons, census data).

To bring all these data into a GIS, they have to be transformed from their analog appearance to digital representation. Thematic elements are typed on a keyboard or digitized (symbols) unless they are already available in machine readable form. The process of transforming analog spatial data into its digital representation is called digitizing and is one of the most important parts of every GIS. Digitizing is done manually or automatically by vector and raster techniques.

Applying manual rastering, a grid of arbitrary cell size is superimposed on a map sheet. Each raster cell is one data item which is assigned a numeric value according to the theme or class it belongs to. If more than one value could be assigned, a predefined procedure selects one (e.g. by choosing the value at the cell center). Manually rastered data are stored in the computer as a matrix or more efficiently "run encoded", which means that a cell-value is stored only when its value is different from its predecessor. Manually rastering is simple and cheap, but rather inaccurate or time consuming.

Automatic rastering (scanning) is achieved by use of scanners. The source map is scanned line by line producing a matrix of grey or binary values. The resoluton reaches 40 points per millimeter. Automatic postprocessing, editing, noise reduction, and vectorizing yield compact vector data.

A straight line between two points, which are defined by their coordinates, is called a "vector". Vector methods for digitizing are manual digitizing and semi-automatic line following. Using the manual method coordinates of points are measured by pressing a button on the cursor of the digitizing device. This is called point mode. In stream digitizing coordinates are sent continuously. Most of the data input systems of a GIS apply manual digitizing. Device resolution is high (0.025 mm), but accuracy depends on operator's skill and experience, and reaches 0.15 mm.

In semi-automatic line following systems a camera is positioned at the beginning of a line. Then it traces along the line, recording coordinates. Operator interventon is required at line crossings or branches to decide which direction to go.

1.2 Storage

Once data are input to a GIS they have to be stored in appropriate data structures to provide fast and effective retrieval for further processing and update. As with input data, there are two main types of data structures, vector and raster structures. Each GIS should provide facilities for storing and processing of both types. Storage techniques for raster data range from simple matrix (space consuming) and "run length encoding" to sophisticated quadtrees [12]. Using quadtrees the study area is divided into four rectangles with equal area, each of them into another four rectangles. This process is repeated for each rectangle until it contains only elements with identical values or no relevant elements at all, or the smallest cell size has been reached.

Vector data are stored in special vector data structures. Point items are stored by simply taking their coordinates (x,y). Lines are defined by a sequence of x-y coordinate pairs, area elements are formed by their bounding edges. To avoid redundant storage of coordinates a so-called edge-node structure is introduced. Each edge is bounded by its start and end node. Line and area structures can be defined and stored as graphs by using edges and nodes. Areas or regions are defined by assigning left and right attributes to each edge viewed in digitizing direction [2,14].

Thematic elements have to be stored in a way which provides fast access and the possibility to establish links to corresponding spatial elements. Storage techniques for alphanumeric data are sufficiently well known as relational or hierarchical data bases. The need to establish relations between alphanumeric and spatial (graphic) data leads to a relational spatial data base. For each thematic element all corresponding graphic elements can be found and vice versa.

A data base management system for a GIS should include software that provides

- simultaneous and fast access of more than one user
- fast retrieval and data revision
- non-redundant data storage
- data security and integrity

1.3 Retrieval, Analysis, and Display

Information systems are usually designed to meet special requirements. There exist systems for geology, forestry, hydrology, land use, etc. But each of them will at least have a set of basic functions for data retrieval, spatial analysis and display of results.

Basic functions for data retrieval include retrieval according to spatial and thematic aspects. Inquiries such as "find all items inside a study area defined by arbitrary polygons", or "report all items which have a certain attribute" will be satisfied.

Spatial analysis functions include simple length and area calculations, basic statistics such as frequency counts, mean and standard deviation. More complicated methods such as map overlay for vector as well as raster data, conversions between data structures (vector to raster and vice versa), inferential statistics, hypothese testing, image processing, and digital terrain models make a system to a poweful planning instrument [8,11].

Graphic presentation of results is done by application of computer graphics [9]. Maps, charts, and cartograms are available on screens for fast preview, or are designed interactively and then sent to various output devices (plotters, film recorders, printers with graphic capabilities).

2. IMPLEMENTATION

Implementations of geographic information systems vary from cost effective microcomputer applications to large and expensive systems based on mainframes. Many vendors offer turnkey systems including the host and all required peripherals, but it is also possible to get systems tailored for special needs.

Every automated GIS should at least have a graphic input device (digitizer), a graphic screen, and some kind of plotting device (plotter, matrix printer, electrostatic plotter, ink-jet plotter).

The implementation of manual digitizing systems naturally leads to menu driven systems. Because there is a limited number of functions required for digitizing and editing, these can easily be implemented by selecting a certain branch of a screen menu or a field on a so-called software keyboard, i.e. a graphic menu on a digitizing tablet.

Automatic digitizing needs no operator intervention during the scanning process, but very extensive software tools are required for post processing of scanner output. Scanned map data have to be vectorized and map features have to be detected and stored. This is supported by manual intervention and graphic editing.

For data analysis the choice of menu techniques is not so comfortable as for digitizing. The great amount of different analysis functions, options and modifiers suggests the use of a command language. One gets greater flexibility for more training effort, because command languages have to be studied more thoroughly than the use of menus. However, help functions are required for both menus and commands. Very often interfaces to standard software packages (e.g. SPSS) are provided.

Graphic output is, and has been very device dependent. But with the rise of GKS, the Graphic Kernel System [3], device independent applications can be implemented, thus allowing for further upgrade in system hardware or migration from one hardware environment to an other.

3. APPLICATIONS

To show some applications of GIS technology we introduce an automated geographic information system that is currently being developed at the Institute for Image Processing and Computer Graphics in Graz [4,5,6,7,10]. It is called DESBOD (Digitale Erfassung, Speicherung und Bearbeitung ortsbezogener Daten; digitizing, storing, and processing of spatial data) and is used for digitizing suitability maps, and environmental planning.

The system is configured around a DEC VAX 11/750 minicomputer hosting a variety of graphic input and output devices (Fig. 2).

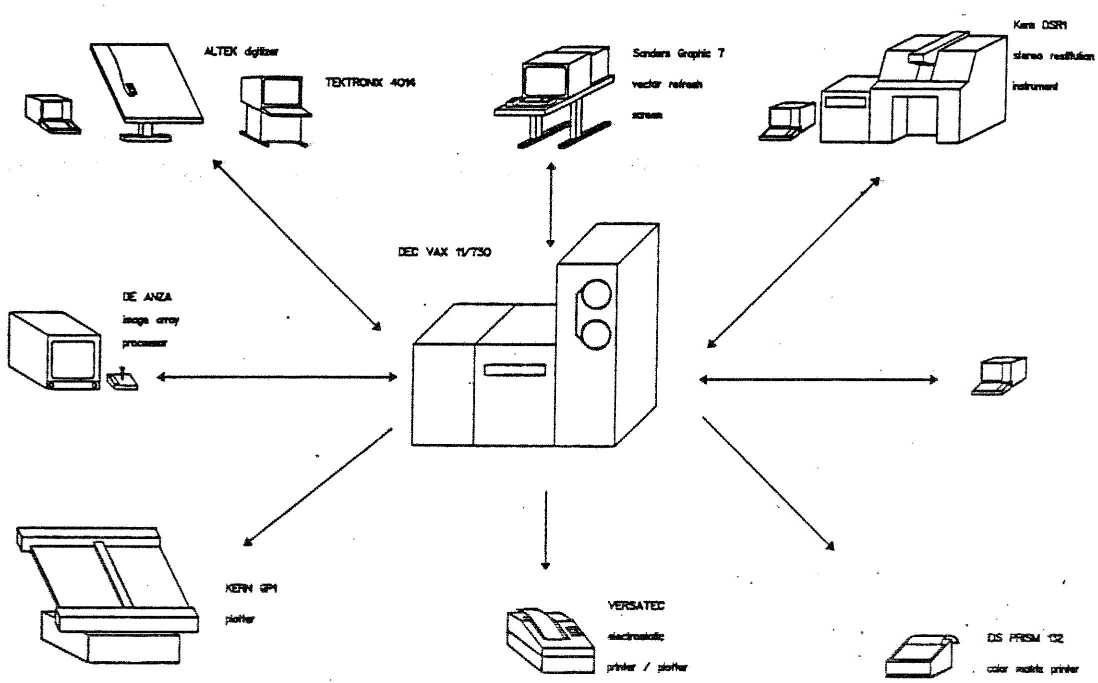


Fig. 2: DESBOD hardware system

Maps and photos are manually digitized, an analytical stereo restitution instrument (DSR1) serves as input device for three-dimensional terrain data derived from a pair of stereo photographs. An interactive editing system provides tools for corrections of digitizing errors. Automatic verification and polygon formation yields topologically correct maps that can be transferred into the data base.

The alphanumeric data base has a hierarchical design with up to 20 levels and an arbitrary number of items in each level. A maximum of 10 (overlapping) data fields for each item are allowed. Graphic data is stored in an edge-node structure with a reference grid for fast spatial retrieval.

Attribute assignment is done interactively during or after the digitizing process by establishing links between thematic and corresponding graphic items. This process is transparent to the user thus increasing operator's comfort.

Analysis is done on a graphic screen and on an image processing system. Output is sent to pen plotters, electrostatic plotters and color matrix printers. Software interfaces allow the exchange of data between DESBOD and other systems (e.g. GTM, a system for the generation of digital terrain models; DIBAG, the institute's digital image processing system). Data flow in DESBOD is shown in Fig. 3.

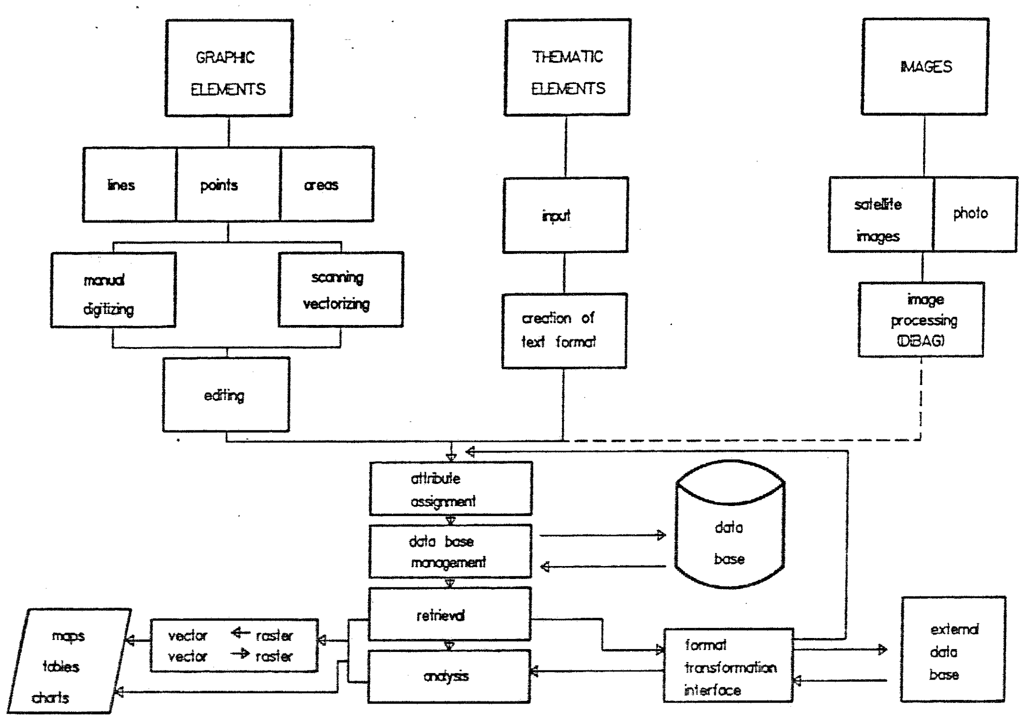


Fig. 3: DESBOD flow diagram

The system is being applied to a district of Styria. The area is covered by four sheets of the 1:50000 topographic base map of Austria. A total of 24 themes are available including geology, morphology, hydrology, soils, vegetation, and reservations. Topographic information (roads, railways, towns, political boundaries, cultural sites) have been digitized and stored for use as situation as well as for analysis purposes. Yet 70 sheets have been digitized representing about 300000 coordinates. Analysis is done using an interactive system for map data processing [1].

Fig. 4 shows an overlay of two maps. The intersection is computed and shown in Fig. 5. Areas and percentages can be computed and printed. Choropleth maps are produced in high quality (Fig. 6).

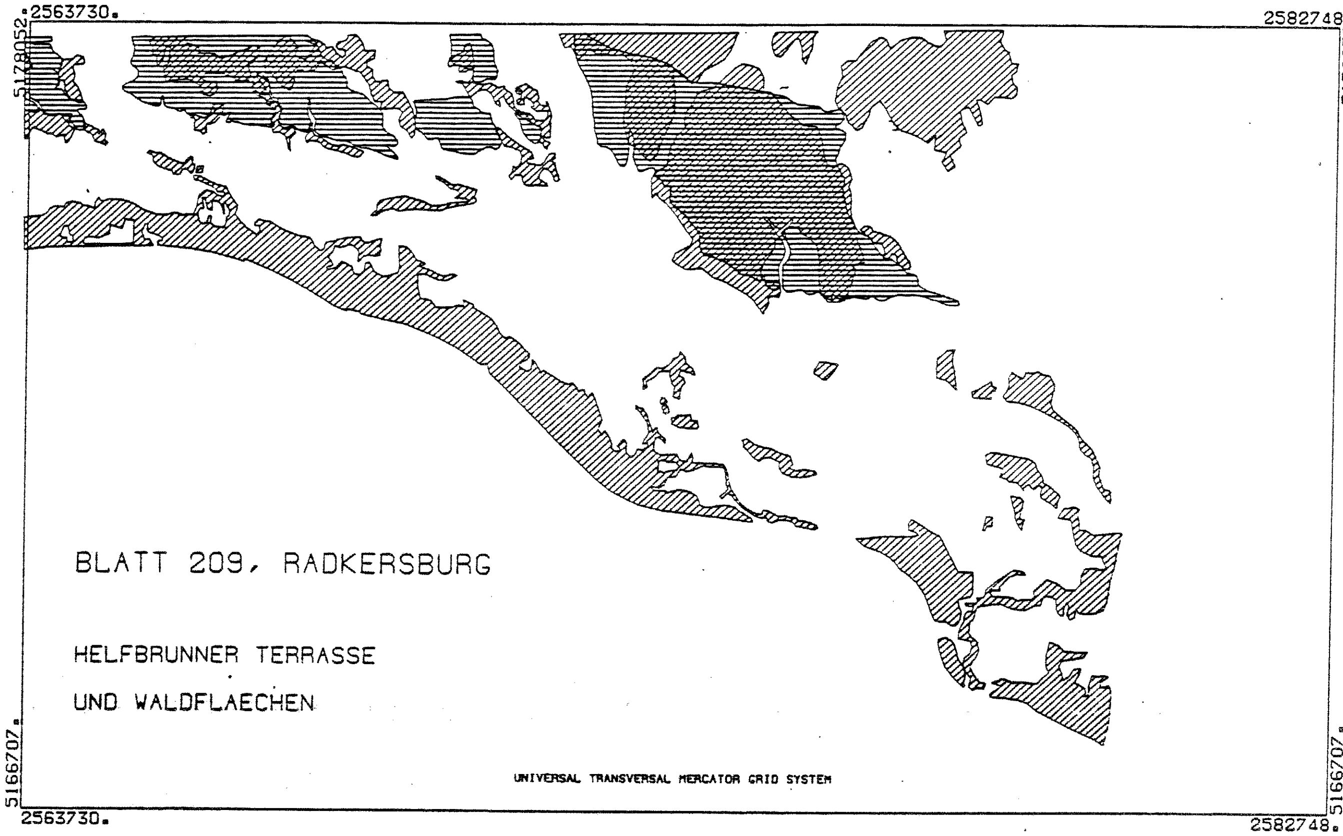


Fig. 4: Map overlay

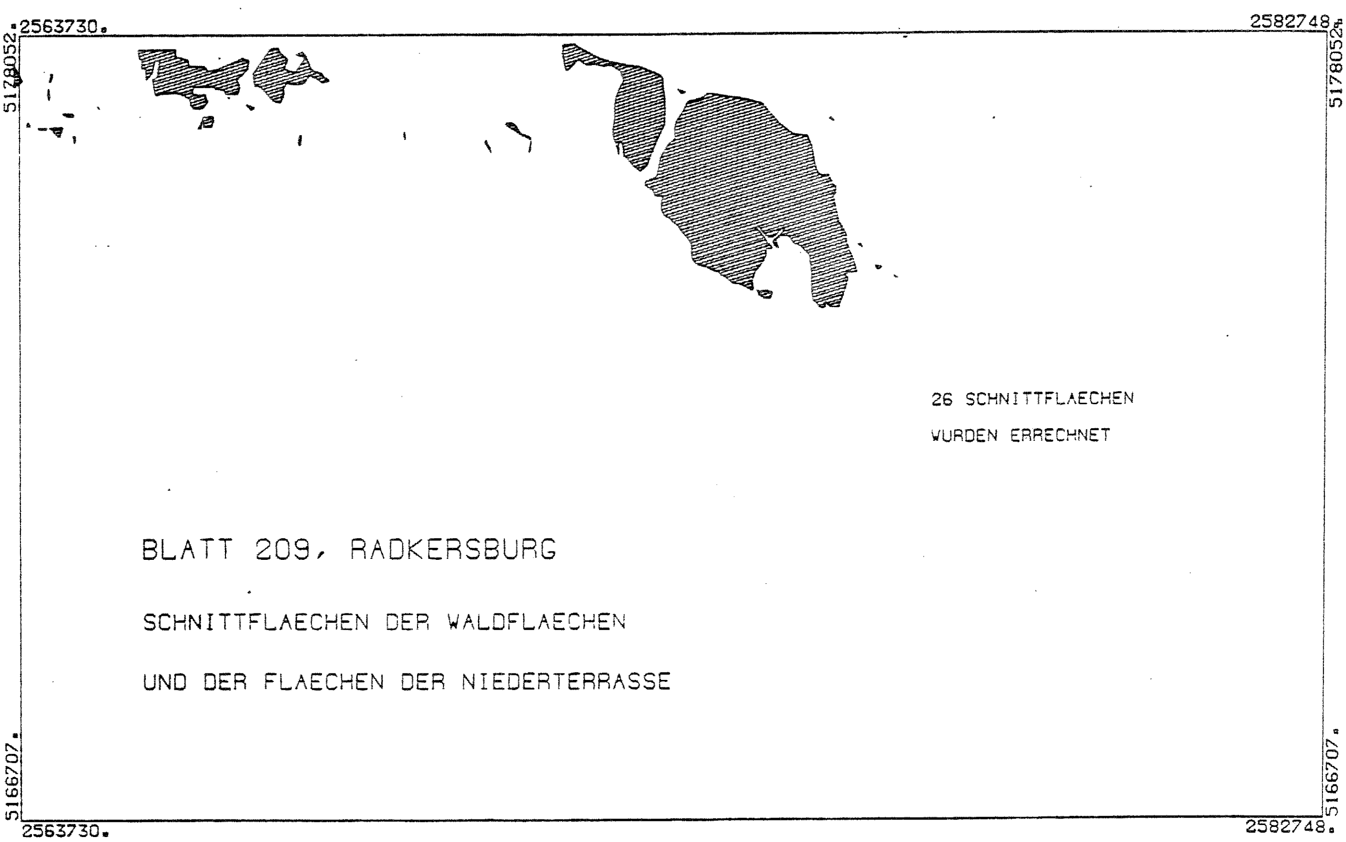
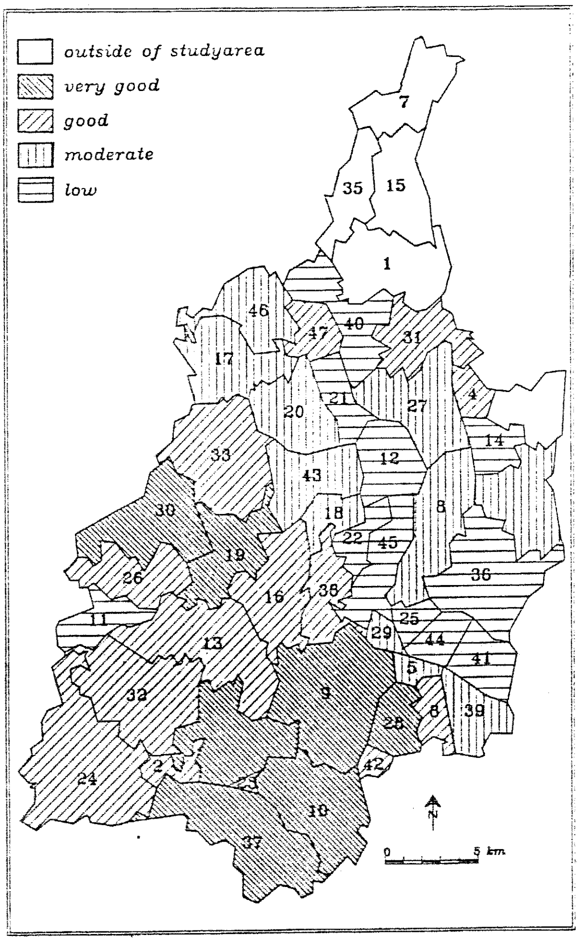


Fig. 5: Map intersection



RECREATIONAL SUITABILITY

suitability value		municipality
very good		
52	6	Eichberg
45	10	Glanz
42	37	Schlossberg
41	28	Ratsch
41	9	Gamlitz
41	30	St.Andrae
40	19	Kitzeck
good		
39	42	Sulztal
38	2	Arnfels
38	13	Grossklein
37	3	Berghausen
36	32	St.Johann i.S.
36	47	Wildon
35	23	Leutschach
35	24	Oberhaag
35	26	Pisdorf
33		St.Nikolai i.S.
31		Seggauberg
31		Breitenfeld
30		

Fig. 6: Choropleth map

CONCLUSION

The automated approach to planning processes has proved to become more and more important for planning agencies and governmental agencies. Experiences have shown a great need for automated information systems in all domains of administration. Quick availability of data and cheap graphic presentation considerably ease decision making processes.

REFERENCES

1. Autometric, Inc. "MOSS User's Manual", Autometric, Inc., Falls Church, Virginia, 1982

2. Corbett J.P. "Topological Principles in Cartography", U.S. Department of Commerce, Bureau of the Census, technical paper 48, 1979

3. Enderle G., Kansy K., Pfaff G., Prester F.-J. "Die Funktionen des Graphischen Kernsystems", Informatik Spektrum 6, pp. 55-75, 1983

4. Kainz W., Ranzinger M. "DESBOD Jahresbericht 1981/82", DIBAG Bericht Nr. 5, Institut für digitale Bildverarbeitung und Graphik, Forschungszentrum Graz, April, 1982

5. Kainz W., Ranzinger M. "DESBOD, A Geographic Information System", Proceedings of CAMP'83, Computer Graphics Applications in Management and Productivity, Berlin, 1983

6. Kainz W., Ranzinger M. "DESBOD Jahresbericht 1982/83", DIBAG Bericht Nr. 9, Institut für digitale Bildverarbeitung und Graphik, Forschungszentrum Graz, April, 1983

7. Kainz W., Ranzinger M. "DESBOD Jahresbericht 1983/84", Institut für digitale Bildverarbeitung und Graphik, Forschungszentrum Graz, April, 1984

8. Monmonier M.S. "Computer Assisted Cartography", Prentice Hall, 1982

9. Newman W.M., Sproull R.F. "Principles of Interactive Computer Graphics", Mc Graw Hill, New York, 1979

10. Ranzinger M., Kainz W., Leberl F. "DESBOD Systemspezifikation", DIBAG Bericht Nr. 2, Institut für digitale Bildverarbeitung und Graphik, Forschungszentrum Graz, November, 1981

11. Ripley B.D. "Spatial Statistics", Wiley, New York, 1981

12. Samet H. "An Algorithm for Converting Rasters to Quadtrees", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 3, No. 1, 1981

13. Weber W. "Geographische Informationssysteme - ein Überblick und Gedanken zur weiteren Entwicklung", Nachrichten aus dem Karten- und Vermessungswesen, Heft 75. Verlag des Instituts für Angewandte Geodäsie, Frankfurt am Main, 1978

14. White M. "A Survey of the Mathematics of Maps", Proceedings of AUTOCARTO IV, Vol I. Published by American Congress on Surveying and Mapping, American Society of Photogrammetry, 1980