TEACH YOURSELF GEOGRAPHIC INFORMATION SYSTEMS: THE DESIGN, CREATION AND USE OF DEMONSTRATORS AND TUTORS

N. P. A. Green
D. W. Rhind

Birkbeck College,
University of London,
7-15 Gresse Street,
London, W1P 1PA

ABSTRACT

In the UK - as in many other countries - interest is growing rapidly in the nature and capabilities of Geographic Information Systems (GIS). Given the relatively small number of systems currently in use and, in particular, the very restricted number of individuals who have used GIS for a wide range of applications, a conflict of interests arises. It is desirable that education and training is provided on demand and that ad hoc queries are answered: yet the provision of this advice and teaching falls at present upon the most skilled and expert proponents and diminishes their capability to develop or exploit GIS in other ways. The increase in demand for training can - in principle - be satisfied by the creation and use of a computerised tutor which deals with the fundamental concepts of GIS. ARCDEMO is a demonstrator - a primitive tutor - developed to show the capabilities of the ARC/INFO GIS, and its design and implementation is described. The system contains both text and multi-colour graphics, sections of which can be viewed in sequence or selectively from a menu. Operations covered by the demonstrator include: automatic data validation and correction; projection change; selective retrieval of spatial data; map overlay; and network analysis. A second demonstrator, called ECDEMO, was developed from the basic structure of ARCDEMO but illustrates an environmental data base established for the European Community. We describe the extent of use and successes and failures of these demonstrators and, on the basis of this experience, we set out desiderata for fully-fledged tutors, suggest the contributions which could be made by knowledge-based systems, outline the machine and human resources needed for running such tutors and conjecture the merits of an international collaborative project to set up a GIS tutor with appropriate national data sets.

INTRODUCTION

The interest in Geographic Information Systems (GIS) has grown rapidly over the last few years. Thus far, however, systems are still relatively few and generally difficult to access. Moreover, many users or potential users are still low on the learning curve so far as the advantages and limitations of GIS are concerned. As part of a Natural Environment Research Council (NERC)-funded project to produce a conceptual design for a GIS, a survey was conducted to evaluate NERC user requirements (Green et al., 1985b). To us, one of the most surprising aspects of this survey was the low level of understanding shown by most earth scientists when questioned on the functions and capabilities of a GIS: this was all the more surprising because we have good reason to believe that those individuals are generally more aware of computer-based analysis of spatial data than researchers in other disciplines or many involved in 'production' organisations. At the same time, the small Birkbeck group was being inundated with requests for visits, demonstrations and explanations. It was decided, therefore, that the best way to educate people was to create a small on-line computer demonstrator to illustrate some basic GIS concepts and functions. Computer-based training (CBT) was considered preferable to other means for a variety of reasons, some of the most important of which are given below:

(i) A computer-based demonstrator, especially one available over telecommunication links, could be accessed by a large number of users relatively easily. Although widely distributed, most NERC Institutes are on the Joint university Academic NETwork (or JANET, see Wells, 1985) and would be able to use the system through this link: all universities in Britain are linked to this network.

(ii) Individuals could work through the demonstrator at their own pace and skip sections which they initially might not be able to comprehend or found irrelevant to their needs.

(iii) By using interactive colour graphics in addition to text, the principles behind GIS might be more easily understood than if they were presented in a paper or by seminar; users also seem more inclined to explore systems if colour graphics are used.

(iv) By incorporating a mail facility, users of the demonstrator could provide feed-back on the aspects of GIS with which they found difficulties and the demonstrator could be continuously updated to improve its teaching capabilities.

COMPUTER BASED TRAINING

Computer Based Training or CBT has also grown in importance in recent years, its use is now spread across many fields. Characteristically, it has been successful where large numbers of people need to be trained in routine tasks and it has a high initial cost: Austin-Rover, for example (according to Kaewert 1986), claim that it costs them 35,000 to train 600 people in a particular function using CBT but that the alternative, traditional, means of training might well cost more and be less effective. Moreover, there is some evidence to show that information is learned more rapidly, more predictably and retained longer than with traditional methods: one British credit card company trains its basic grade staff using CBT in one third of the time required by traditional methods.

The ideal situation for CBT, then, appears to be where numbers to be trained are insufficient in any one place at any one time to justify heavy investment of teachers, where students vary considerably in their abilities and hence pace of learning, and where individual, anonymous tuition is appreciated (as by many professional, status-conscious staff). All of these are common situations in GIS training.

What is also evident, however, is that CBT is most useful when viewed as a complement to rather than a substitute for an instructor. Human communication skills remain important in encouraging students and in perceiving why a particular student is misunderstanding a topic. Even in the latter situation, however, progress has been claimed in the development of prototype CBT systems which, based upon cognitive modelling concepts, monitor the students progress, develop a fund of knowledge about student reactions and vary the training accordingly.

DESIGN CONSTRAINTS AND OBJECTIVES IN CREATING A GIS DEMONSTRATOR

The design of the demonstrator was bound by a number of constraints and objectives. The most important objective was to construct a system that was easy to use and could be freely accessed by academics and by NERC employees. Other considerations dictated that the system should:

(i) Use minimal disk space and computing resources and run on the departmental VAX 11/750 computer under the VMS operating system.

(ii) Be easy to use. From the outset, it was certain that the system would be used by some people with very little, perhaps no, computing experience. For this reason, instructions would have to be clear, simple and unambiguous.

(iii) Be robust. To test empirically a system's response to every possible request is almost impossible, but the demonstrator had to be as near fool-proof as possible. Fragility of a product designed to introduce a user to GIS would be counter-productive. This ensured that the system had to be rigorously tested, be based on a simple structure and that interrupts and possible error conditions had to be trapped and handled appropriately.

(iv) Be flexible. A structure which allowed people to view examples sequentially, or look at individual sections on request and to review pages of test or graphics was considered desirable.

(v) Provide high resolution colour graphics and text information. With this, those users with high quality displays would be able to gain an accurate impression of what is possible.

(vi) Support as wide a range of graphics and non-graphics terminals as possible. It was clear that, although the system needed to support high resolution colour graphics, it was possible that many people might be accessing the system using low resolution black and white terminals operating at low baud rates. For this reason, the graphics would have to be simple and easy to draw. Indeed, as many users might not have access to graphics terminals, it was decided that both graphics and text-only sections be included in the demonstrator.

(vii) Require minimal human resources to construct, implement and support. This was so critical that, in the early days, we chose the networked facility rather than distributing tapes because of our minimal resources.

THE STRUCTURE AND FORM OF THE DEMONSTRATOR

To provide free access to the demonstrator, the system was installed on a computer account with no access restrictions (no password, etc.) the system being automatically activated after logging on. This ensured that the minimal of instructions had to be given for the system to be accessed but, correspondingly, placed great importance on the prevention of users 'breaking out' into other parts of the VAX computer system. For this reason, all Control-Y and other interrupts are trapped and redirected.

The demonstrator is currently machine-dependant, at least so far as its host is concerned: it consists of a set of VAX files containing DCL commands and text or graphics to be displayed. This decision is not a major constraint: VAX computers are widely (almost ubiquitously) used throughout Britain for GIS work and, furthermore, the remote user logging on to our computer is generally unaware of the type of computer in use. It would, however, be a relatively straightforward matter to port the software onto certain other computer systems.

To comply with point (vi) above, it was decided that the demonstrator should have two main sections. The first consists purely of text information displayable on any 'dumb' terminal. The second section, which could only be used by those having access to a graphics terminal, consists of both text and multi-colour graphics. The graphics produced by the system are designed to support mostly Tektronix and 'Tek compatible' terminals and uses the Tektronix Interactive Graphics Library (IGL) routines. To make the system flexible, it was decided that a hierarchical menu-driven structure would be most appropriate. This was modified, however, to allow the user to 'walk through' the demonstrator in sequence (see figure 1). The following sections are accessible from the main menu:

Introductory Text Section
This section consists of a number of pages of text which give an overview of the ARC/INFO system, how to access the system, and under what conditions it should be used. It does not require a graphics terminal. Here, as in all sections, pages of text are limited to 80 columns by 24 rows i.e. to meet the worst possible case so far as alphanumeric terminals are concerned.

The Graphics Demonstrator
The graphics demonstrator consists of a number of examples depicting some of the main functions of a GIS (Green et al, 1985a). The examples are accessed from a sub-menu and can be viewed in sequence or selected individually. Each example consists of text - never more than two pages in length - describing the function, followed by a number of illustrative graphics 'pages'. Most of the graphics are produced from plot files rather than being created interactively, having the advantage of requiring few system resources and allowing those users who are most able to work more quickly through the package.

Mail and News Facilities
The provision of a mail facility, exploiting the VAX MAIL utility, enables users to make comments and suggestions. Some modifications have been made to the demonstrator in the light of these replies: these are discussed in a later section. First-time users of the system are encouraged to leave their name and address, which are then read into INFO and used to produce a mailing list to keep people informed on new up-dates, conferences, etc. A frequently up-dated bulletin board is also provided to allow people to keep track of forthcoming events and news.

THE COMPOSITION OF THE GRAPHICS DEMONSTRATOR

When called from the main menu, the graphics demonstrator first displays a table listing the types of terminals supported and prompts the user. After the terminal type has been specified, a sub-menu is displayed which lists the examples held in the demonstrator. Any example can be selected from the menu, or all can be gone through in sequence. It is also possible to 'step back' to a previously displayed page of text or graphics, or to skip further stages in an example and continue with the next example or to return to the sub-menu. Currently, the graphics demonstrator is composed of eight sections, each one of which is summarised below.

Data Structuring and Validation
The section illustrates the process of 'cleaning' an unstructured 'spaghetti' digitised file. The first frame shows digitised lines (arcs in ARC/INFO terminology), control points (tics), and the flagging of possible topological errors at unconnected line intersects nodes. Subsequent frames show how software can be used to remove automatically the 'overshoots', to 'snap' together undershoots, to geometrically transform and then to generalise the mapped data.

Data Enhancement
The generation of corridors or 'buffer zones' around features is used in this example to show how a GIS can enhance existing data. The example involves the creation and display of a corridor where width varies around roads of different class (figure 2).

Data Integration
Using the buffer zones generated in the previous example, this section shows the intersection of these areas with others depicting land use (figure 3). Further frames in this section show the intersection of postcode sector boundaries for the same area with land use zones and the subsequent selection of the portions of land use zones falling within a single postcode sector.

Data Retrieval
The section originally initiated an interactive dialogue session with INFO, a relational data base management system coupled to ARC; the former is used to generate a report summarising the results of the integration of postcode sectors and land use given in the previous example. This example also serves to demonstrate how INFO handles thematic data associated with a map and introduces the concept of relational file structures.

Data Manipulation
The ability to manipulate map data is one of the most useful aspects of a GIS. Comparing maps draw at different scales and using different map projections is the most basic of map manipulation requirements. Taking the example of map projection change, this example shows the transformation of a simple map of the continents of the world through a series of azimuthal, polar and Mercator projections (figure 4).

Network Analysis
Network analysis is becoming increasingly important in GIS for the modelling of flows, for planning the distribution of services and the allocation of resources and for route finding. Taking the example of service centre allocation, the frames show how different parts of the road network in the centre of London can be assigned to specified centres based on a travelling time criterion.

Data Structure Conversion
With the advent of an increasing number of mapping and information systems, the ability to transfer data between systems is becoming more important. The ARC/INFO system has a number of two-way interfaces for data conversion to a number of different formats and has been further enhanced by some of its users. The interfaces available on the Birkbeck version include DIME, DLG, DMC, GIMMS, and a GRID structure. The ability to convert to and from a grid structure is one of the most useful facilites. The demonstrator shows the effect of converting a map to grid format (i.e. rasterizing the vector data) and then converting it back to vector format (figure 5). Parameters were chosen so as to ensure some difference is discernable between the initial and final products!

This section also includes a plot illustrating the creation of Thiessen polygons (Dirichlet tessellations) abstracted from point data i.e. space is partitioned so as to allocate all locations in space to the nearest data points (figure 6). The resulting polygons are then available within ARC/INFO and can be treated as if they had been digitised.

Graphical Symbolism
The last section in the demonstrator illustrates the various types of graphical symbolism which can be generated automatically by the system.

Examples of point, line, area and text symbolism are displayed in a selection of bright colours, chosen so as to map to shades of grey on a monochrome terminal.

THE OPERATION OF ARCDEMO

Originally installed in September 1985, ARCDEMO became available over JANET in November 1985. All Geography, Geology, Planning and Computer Science departments in Britain were circulated with the JANET address and login details, heads of departments being asked to pass these to individuals who might be interested in such a facility. Between then and end-March 1986 there were over 250 logins; over 150 people have left their names and addresses. Originally designed as a semi-interactive system, ARCDEMO was modified in February 1986 to replace sections requiring processing by ARC/INFO with plot files and with static text. These modifications were made to allow the system to run faster, to prevent file access conflicts under INFO version 9.2 with many people using the system at once, and to enable the system to be used as a stand-alone package. The demonstrator now characteristically uses no more than 2 to 3 % of the cpu on our VAX 11/750. In addition to being located at Birkbeck College, ARCDEMO has now been installed on VAX computers at the universities of Reading, Leicester and Sussex where it is used as a teaching aid; it has also been run on a micro-VAX at DEXPO by DORIC, the company marketing ARC/INFO in the UK.

Due to the success of ARCDEMO, another system, ECDEMO, was produced to be used as part of a major project to create an Information System on the State of the European Environment (Rhind et al 1985, Wiggins et al forthcoming). This system was designed to show how ARC/INFO was being used to validate, edit, integrate and map a variety of European data sets including soils, topology and climatic data. The system has the same structure as ARCDEMO and has been used by other institutions involved in the project and by the project coordinators in Europe over international packet switched computer networks. It differs from ARCDEMO in a number of ways, including the need to offer the user text in any one of the different official languages now in routine use in the European Community.

LESSONS GAINED FROM ARCDEMO AND ECDEMO

These include:

(i) What is obvious to the system creators is frequently arcane to the user. Hence even early system testing must be done by a sample of users, rather than the system designer/creator.

(ii) Many of the difficulties relate to trivial matters which should be answered locally - such as clearing the graphics screen. This implies a need for a hierarchical release of the system to local experts, then a more general release.

(iii) Boredom rapidly results from too much text appearing on the screen.

(iv) Each example must be clearly focussed and should not extend beyond say five 'pages' or screenfuls.

(v) The response rate is important: users are frequently unhappy if it takes more than about 45 seconds to fill a screen with a graphic. It is highly desirable to give reassurance (e.g. sending messages such as '... sorry for the delay I'm still calculating the results') when nothing obvious is happening on the screen.

(vi) Simplicity is essential, if an instruction can be misinterpreted it will be misinterpreted. Only thorough testing, however, can detect some more subtle misconceptions.

(vii) Despite the problems listed above, it has proved entirely possible to run a demonstrator of this type at 2400 baud; running at 300 baud is acceptable for text but not for graphics.

(viii) The happyness of the user with the system appears to be related to the quality of the terminal being used: colour and a resolution of at least 600 x 400 seem desirable.

(ix) The operational reliability of existing networks makes use of a demonstrator hosted in a centre of expertise a technical possibility.

DESIDERATA FOR A GIS TUTOR

Based upon our experience with ARCDEMO and ECDEMO, we believe that a genuine GIS tutor should:

(i) be hierarchical in structure and at least have a comprehensive overview of GIS functionality. This may be provided by an ARCDEMO-type superstructure. It may also usefully include sections on GIS-related concepts e.g. on data structures and data volumes.

(ii) interact with the user by seeking responses to simple questions (e.g. 'do you require further explanations of this ?').

(iii) provide locally relevant examples, with all dialogue in the local language. Thus the design must be highly modular, with local examples slotted in as appropriate. Despite this, cross-references must be made e.g. to the need for 'clean' vector data before overlaying polygons.

(iv) response in manner appropriate to the facilities locally e.g. less complex graphics should be transmitted to a monochrome terminal attached to a 300 baud line than to a multi-colour terminal running at 19.2 kb.

(v) trap user errors and log these, together with a note of the most frequently used sections of the tutor, to build up knowledge of the users. This provides some scope for the use of Intelligent Knowledge Based System approaches.

(vi) provide 'help' facilities at any point, together with a glossary of terms used and a keyword-searchable bibliography.

(vii) be readily capable of expansion to include local details e.g. of the particular tasks carried out in the 'production shop' running the tutor.

One difficulty which we anticipate is that of generalising the tutor: it is easy enough to create one which relates directly to one specific system such as ARC/INFO. At this stage, however - and especially since no known system provides a full range of vector- and raster-based operations - the concepts and approaches embedded in different systems would have to be included in any ubiquitously usable tutor. In the absence of a GIS command language analogous to UNIX, this is likely to cause difficulties.

FUTURE DEVELOPMENTS

In-house Developments

Under an agreement with NERC, the structure of the ARCDEMO demonstrator is soon to be extended and modified to include further GIS functions, covering sections on edge-matching, sliver polygon removal, and map library concepts. Major new sections on remote sensing, topographic and thematic data handling are also to be included, along with an interactive bibliographical system, an index of active individuals in GIS and a glossary of terms. This development will make it more general, allowing the input of sections derived from other systems. In addition, the demonstrator is to be modified to allow closer monitoring of how people use the system. This should give an indication as to which of the various aspects of GIS specifically interests users. In particular, we intend to produce different versions of the demonstrator to monitor users responses to different types of system interface.

Possible International Developments

In addition to the within-UK developments above, we believe there is much scope for international collaboration in creating GIS tutors. The obvious vehicle to achieve this is the International Geographical Union (IGU) Commission on Geographical Data Processing and Sensing, which has been active in the area of GIS for twenty years and has established a world-wide net of contacts. Our suggestion is that this should take place in two phases, the first creating and distributing of a prototype and, if successful, the second proceeding to a full-scale operational tutor. Our experience indicates that a prototype tutor could be produced to meet the less ambitious of the desiderata set out in an earlier section well within an 18 month timespan, particularly if it built upon experience thus far. Clearly, such a development must be proceeded by agreement on the overall objectives to be met and the structure and design of the tutor and not all the work could be achieved on a voluntary basis: the proposal therefore has some funding implications. Nonetheless, these may be relatively small, despite the need for international collaboration - we now routinely make use of computer messaging facilities to the USA and to mainland Europe and such communication is very inexpensive. In short we can now use computer facilities to expedite international collaboration to teach others about computerised facilities.

An informal proposal to this effect has been forwarded to the Chairman of the IGU Commission, Prof. Duane Marble, and proposed at the Commission-sponsored meeting in Seattle in June 1986. Details of the status of this proposal will be given orally at Auto Carto London.

CONCLUSIONS

To the best of our knowledge, ARCDEMO is the first freely available, networked computer demonstrator using computer graphics in the UK. Clearly, as a teaching tool, it is somewhat primitive; however, we believe that Computer Aided Learning and Computer Based Training have an important roles to play in the development of GIS and automated cartography, especially if combined with expert systems. We claim no great prescience for this insight: Fraser Taylor, for example, in his book on Education and Training in Contemporary Cartography stated,

> "... we must look for complementary methods of educating the current map user ... the use of computer-aided learning may be useful, although again few such packages exist." (Taylor, 1985, p. 21)

Our own experience with large numbers of users suggests that it is worthwhile to devote considerable resources to creating and making readily available tutors and demonstrators and we expect to see many other developments in this area.

ACKNOWLEDGEMENTS

Thanks are due to the Economic and Social Research Council for the funding of the purchase of the ARC/INFO software. Part of ARCDEMO was written in the course of a project financed by the Natural Environment Research Council. Thanks to Jo Elliot for assistance with the figures. We also thank Jack Dangermond and his ESRI colleagues for much assistance over the last three years. Finally, we thank all those who sent us helpful suggestions on ARCDEMO by use of the mail facility.

REFERENCES

Green, N. P., S. Finch, J. Wiggins (1985a) The 'State of the art' in Geographic Information Systems. Area 17.4, 295-301.

Green, N. P., D. W. Rhind, S. Finch, K. E. Anderson (1985b) Conceptual Design of a Geographic Information System for the Natural Environment Research Council: Report 1, User Needs and Design Constraints. (NB this report was not generally available at the time of writing).

Rhind, D. W., B. W. wyatt, D. Briggs, J. Wiggins (1985) The Creation of an Environmental Information system for the European Community. Proc of the Int. Conf. of the Remote Sensing Society and CERMA. Int. Conf. on Advanced Technology for the Monitoring and Processing of Global Environmental Data (supplementary papers), London, Sept. 1985, 11-21.

Kaewert J. (1986) How CBT could shape up your teaching programme. DEC User April 1986, 45-48.

Wells M. (1984) The JANET Project. University Computing, 6, 56-62

Wiggins, J. C., R. J. Whittaker, R. J. Hartley (forthcoming) Computing Aspects of a large Geographic Information System for the European Community. Paper in Proc. Auto Carto London.

Figure 1 .

PAD>CALL A5163000

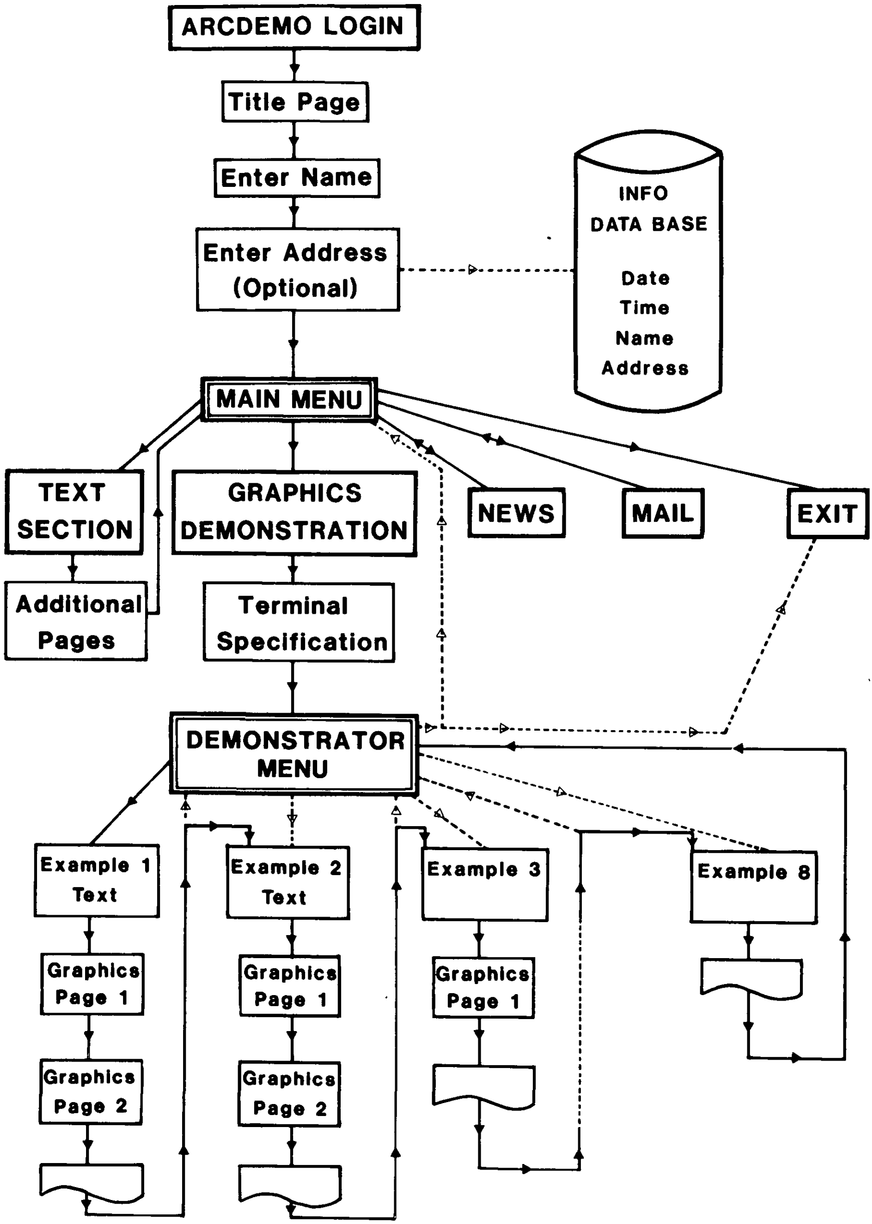


Figure 2.

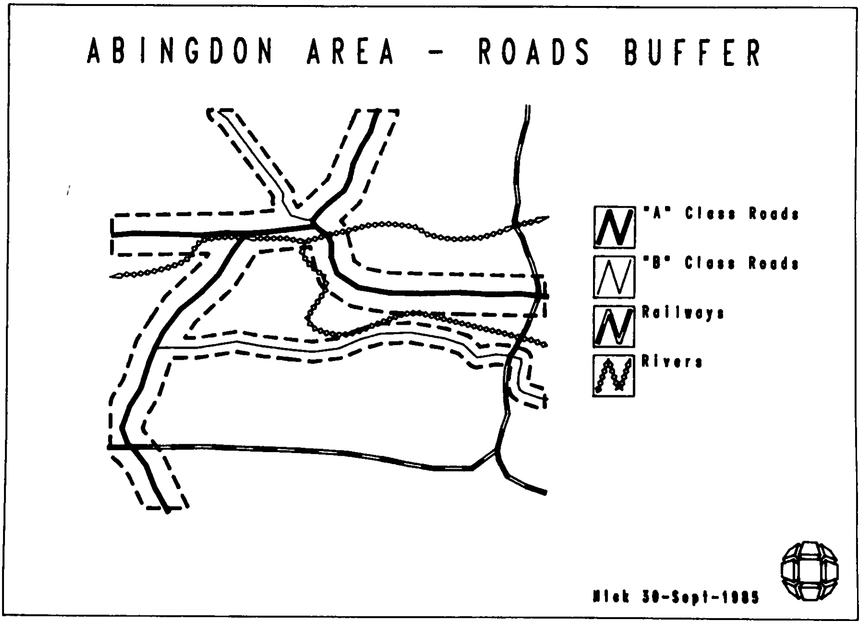


Figure 3.

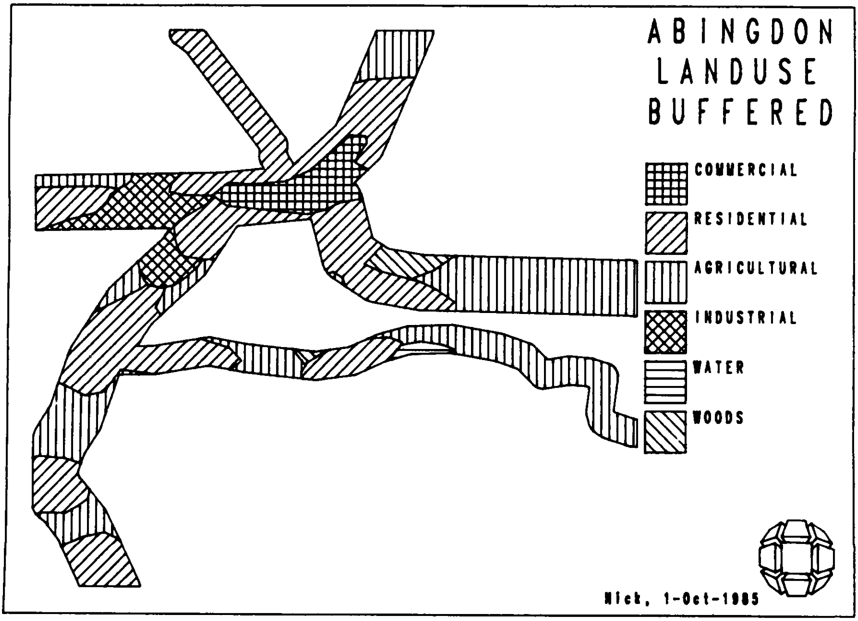


Figure 4.

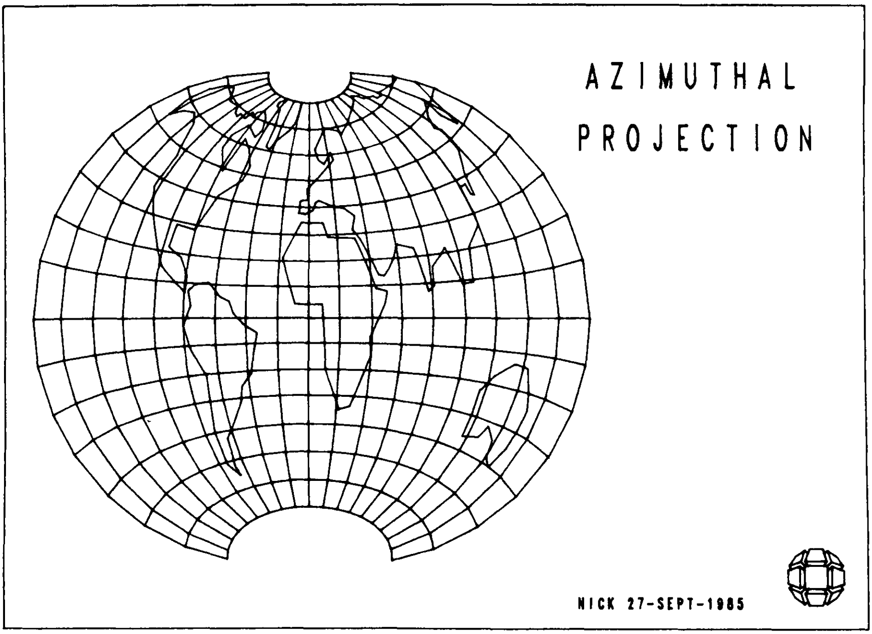


Figure 5

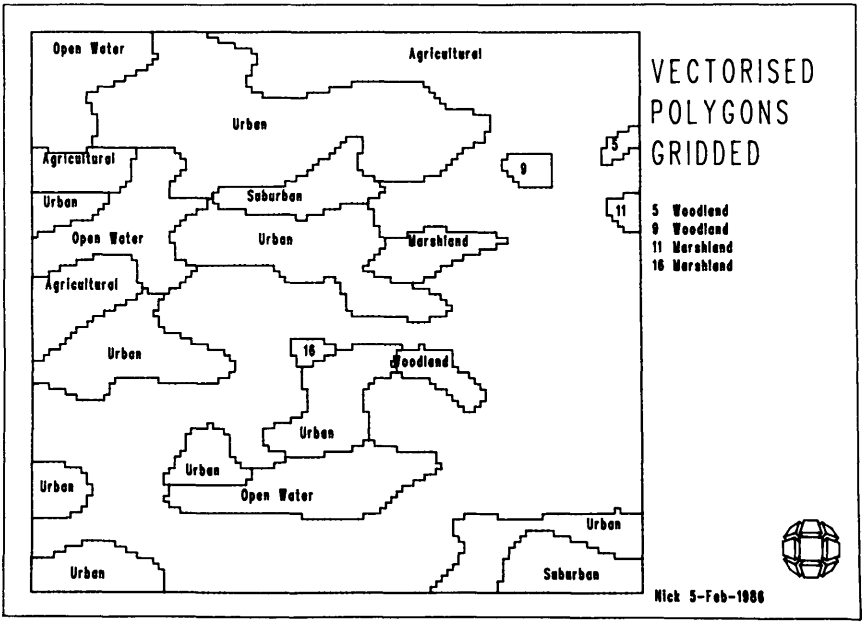
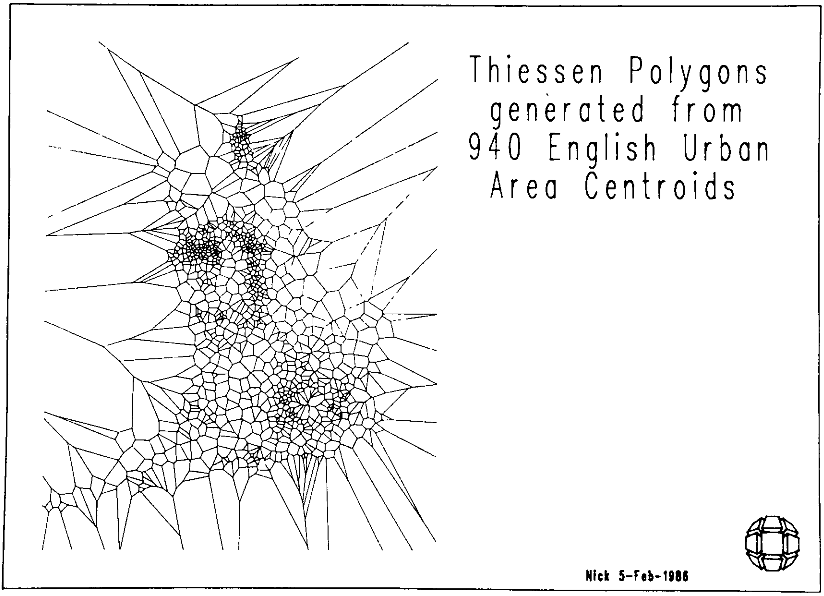


Figure 6.


Thiessen Polygons generated from 940 English Urban Area Centroids
Nick 5-Feb-1986