A GEOGRAPHIC INFORMATION SYSTEM FOR USE IN THE SUGARCANE INDUSTRY

By G. G. PLATFORD

South African Sugar Association Experiment Station, Mount Edgecombe

Abstract

There is a need to map graphic information of farms together with topographic and crop variables. A geographic information system (GIS) which co-ordinates all the information is described. The requirements needed to operate the system are listed. The accumulated data can be used to answer queries at the industry, mill or farm levels and to produce maps at different scales. The capture of data for maps for legal issues, survey documents, quota or for agricultural management purposes is detailed. The establishment of a digital terrain model (DTM) is explained. Data base files of crop performance from Field Record System (FRS), soils details from Fertilizer Advisory Service (FAS), pathology, entomology, meteorology and hydrology, are attached to specific ground co-ordinates and totally integrated analyses achieved. A trial in which map and database files from 43 farms and 3 miller sections used in the Mid-Illovo area shows that there is great accuracy and repeatability in map production. Queries on conservation planning, crop management and modelling, distribution of resources and ground details were answered. The management of a very large data bank for the application of this system to the whole of the sugar industry is described. The major cost in setting up the operation is in the capture of data.

Introduction

Individual growers and managers know the condition of crops in their fields. They also have maps and aerial photographs with farm and field boundaries to help them associate their visual impressions of crop, soils and slopes with positions on the farm. Also available are different maps (rainfall, temperature etc.) relating to growth factors, which help in reaching management decisions. Many varied factors contribute to the production of the crop. They include all the static physical attributes at the specified location, such as slope, aspect, soil and elevation, as well as the variable inputs of fertilizer, herbicide, rainfall, sunshine, humidity and numerous others. To compare growth results from farms at different locations is difficult if all the input factors are unknown. Often comparisons are made using too few of the growth controlling factors. The use of a system which can allocate the parameters to a specific site can improve analyses. However to achieve a reasonable analysis vast volumes of data are required. At present most of the data are in map or photo form, and for use in a GIS they need to be converted into a digital format. The advantage of a GIS is that it can be used for all the conventional mapping tasks, and effectively analyze spatial or locational data.

Objectives

In the sugar industry many different map types and scales are used. Queries about the crop and ground conditions are also frequently posed together with enquiries about meteorology, soils and slopes. To co-ordinate and control all mapping and to use the data in database files and on the maps, requires an industry GIS. A list of objectives to meet these requirements is set out below.

1. To provide an integrated mapping service for the sugar industry
2. To provide cadastral and survey maps with legally accepted accuracy and to have updating and reporting facilities
3. To provide the means to analyze spacial and alpha-numeric data for specific times with minimal effort
4. To include facilities to interpret data on:
   - 4.1 An industry scale, (i.e. Natal and Eastern Transvaal)
   - 4.2 Mill group scale, (i.e. Umzimkulu, Sezela, etc)
   - 4.3 Ward scale, (i.e number of growers or miller planter sections)
   - 4.4 Farm and field scale, (i.e. 500 ha to 1 ha)
5. To provide planning facilities using topographic, climatic, geographic, agronomic and other data. Farm, irrigation, agronomic and management planning with comparative and resource evaluations are required. Included in this is the requirement to plot grade lines
   - 5.1 To provide decision makers with the answers to particular queries
6. To be based on PC technology either on a single facility or in a network. A high degree of compatibility with other

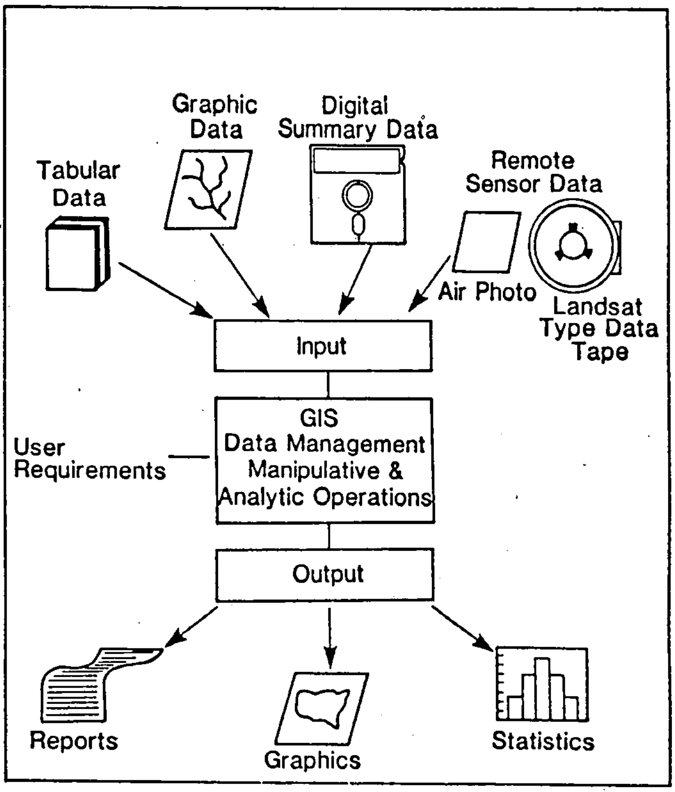

FIGURE 1 General functions of a GIS.

languages and systems is essential. It must be able to accept data from all other sources

7. To be user friendly and fast.

GIS and CAD.

The difference between a true GIS and many advanced Computer Aided Draughting (CAD) packages is that a GIS includes CAD package for the draughting functions, but also has attached database files (see Figure 1) and data management software which give the data some artificial intelligence. The GIS software allows drawing entities to know where line segments are shared. The data files are distributed to specific co-ordinates and can be questioned, and the answers used to make management decisions. The software is built up of different modules contained in the UNIGIS package, as supplied by a commercial company.

Data capture

In any GIS the major and generally the most expensive task is to capture the data (Dinkele[1]). Data consist of alphanumeric or attributive data and graphic or positional elements. The former can be captured as conventional database files, while the positional data must be captured by specialized GIS data capture software.

Alpha-numeric data capture

Data sets in most formats can be entered into the GIS, where they are attached to specific locations. For example, a grower's field records on crop performance can be extracted from the main frame Field Record System (FRS) and read directly into the GIS. Similarly any fertilizer, crop pathology or any other data file can be used.

Positional data capture

Cadastral data may be captured by a CAD system with integrated surveying and town planning software, and other graphic manipulation tools. All the functions of the CAD system (such as drawing parallel lines) are available to the operator at the time of data capture. Text is added using the CAD features. The available co-ordinates are typed into a file and read into the drawing directly. Two paths are now available to the user. If legal data are required, the graphics data are captured by entering all the distances and directions on the legal document (diagram or general plan). Checking is done by conventional survey techniques.

If the legal data are not required, the graphics data can be captured by powerful CAD features such as drawing multiple parallel lines, extending lines to reference lines, merging lines and so on. At this stage, the data are merely a set of lines and text in a CAD drawing, with no topological structure or intelligence. Where necessary, lines and text are differentiated into levels to separate the data into identifiable homogeneous groups. This information is then "massaged" into topological data structures, by batch mode line splitting and polygon processing software, where topological relations and data "intelligence" are added. It is these data structures that allow the GIS interrogation software to extract information such as adjacency, inclusion, intersection, and route analysis. The data can now be said to be object orientated, the objects being polygons. These relate, on one level, to their topological components or entities, and on the other to any number of attributes stored in relational tables, in a relational database.

Topographic data capture

Topographic data are an essential element of GIS information. Two basic categories exist: relief data (the third dimension) and planimetric data.

Relief data have traditionally been represented as contours which are an "analogue" representation of height. They are easily interpreted by the human eye and brain combination, quick to generate by manual techniques, and easy to use on a paper map. However, in the digital environment, they are not only expensive in terms of space, but are clumsy for interpretive analysis. A truly digital representation of relief is the Digital Terrain Model (DTM), where a grid or triangulated network of spot heights is generated, from which all other relief information such as contours, ground heights, ground profiles, and slope analysis are interpolated.

DTMs can be generated from ground surveyed spot heights, from contours digitised from existing maps, from spot heights captured from stereo-photographs in a photogrammetric stereo plotter, or from contours captured digitally from stereo photographs in a stereo plotter. They are often stored in raster format for ease of data retrieval and to conserve the space that vector coordinates would use. Contours are then not needed in the system at all, and need only be interpolated at the map production phase.

Planimetric data can be captured in several ways.

Firstly they can be digitized from existing maps. Secondly, the features can be captured directly from stereo photographs in a photogrametric stereo plotter. This involves orienting the photographs in the plotter and then digitizing the co-ordinates from the resultant stereo model. A third method is to digitize co-ordinates from a single photograph and subject them to a transformation, which involves a perspective transformation onto a non-plane surface represented by a DTM.

Hardware

The use of high speed personal computers with large amounts of storage space together with ancillary equipment enables a GIS to be run effectively from a small office. Data banks are established and stored on disk or tape. The hardware requirements for this particular system are:

A PC/AT (80386) with 1 Megabyte extended memory

40 Megabyte hard disk.

2 serial ports, 1 parallel port and a maths co-pressor

Tape streamer or Compact disc for mass storage with AO digitizer tablet and an A1 multi-pen plotter.

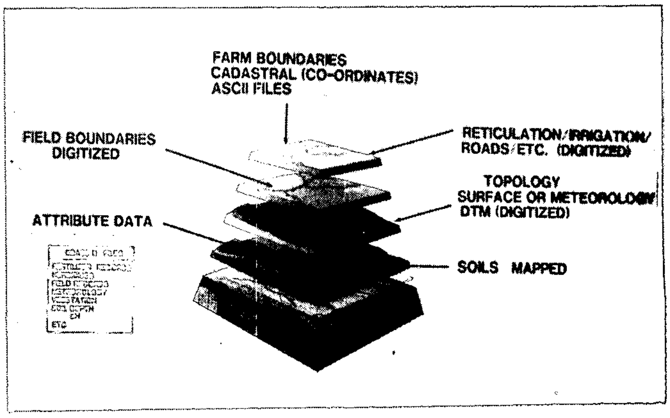

FIGURE 2 Schematic of a geographic information system data capture scheme.

Design of the software.

The GIS consists of a number of individual software packages which are all designed to interact (see Figure 2). The modules included are:

PROCAD - This uses a "head up" approach for ease of operation with minimal use of a keyboard. Instantaneous pop-up menus and simple graphic icon selection are used. All the usual CAD functions such as construction sets, symbol and text insertion, overlays and measuring facilities are included.

PROMAP - The module has been specifically designed for cadastral and related mapping requirements and extends the PROCAD capability to cover topography.

Data capture modules.

PADS - This module (Photo Analytic Digitizing System) controls digitizing from single aerial photographs. Various checking and squaring functions are carried out to correct anomalies in the photograph.

DIGIPAC - For copying maps and plans the module can put in to the computer digital values for contours, spot heights and map features while correcting for paper deformations. The data stored can be used directly by the terrain modelling programme.

Terrain modelling.

SURFACE - The module is capable of modelling any terrain with options to generate contours, calculate bulk earthworks or view a site in perspective (3D). It can be used for slope and aspect analyses.

PROFILE - In conjunction with the surface DTM profiles and sections along roads or pipelines can be generated by this module.

GRADE - Defined gradients can be drawn between certain points on the DTM. This module is designed to allow the siting of diagonal extraction roads, terrace banks or any other line which needs a gradient.

Geographic data management

The data about geographic features are either in raster or vector formats. Data on attributes such as field records, fertilizer records and soil physical properties are also stored in DBASE or other language files. Various linkages are built to allow the whole database to interact. During queries the accumulated data are integrated via the software so that specific extractions and examinations can be achieved and questions answered. Specialized software has been developed to allow all the classes of data to be used for specific queries. An example of this aspect is the design of terrace bank spacings which involve the topography, soils and DBASE files or look-up tables of the nomograph where grower options are allowed.

Accuracy and repetition

Doubts are often expressed about the accuracy and repeatability of mapping carried out with electronic techniques. Much of the concern is centred around the digitizing of existing maps to put the data into a central database. There is no doubt that when cadastral legal data are stored they are exact. Thereafter the accuracy depends to a large extent on the scale at which the data are captured. With existing digitizers, which have a resolution of 100 lines per inch (4 per mm), a map at 1:6000 scale, where one millimeter equals six meters, resolution to less than two meters can be expected.

Special development

It is expected that as the system develops, specific requirements for the sugar industry may well need to be expanded or included in the package. One reason why this particular UNIGIS software was selected, was that the local company had the necessary expertise to set up and develop the model which best suits the sugar industry's needs. While several successful operating GISs are on the market, most are of external origin. It was decided to accept a local company who have an excellent record in mapping and photogrametric procedures so that they can solve any problems.

Usage

1. Map production.

   The production of legally acceptable maps of different types and scales, and the alteration of previously mapped internal field boundaries or road positions are major tasks needed from the system. As growers develop their farms according to prepared plans or as new photography becomes available, changes will be easily accomplished. As there is a scale change facility built into the system, the need to change scales, often by expensive optical methods, will fall away.

2. Map or spatial data interrogation.

   The type of questions which could be answered by the GIS are illustrated later when the trial area is discussed.

   2.1 Land Use Plan production.

   Much time is spent on converting cadastral details and drawing terrace bank roads, drainage lines and specific gradient haulage roads onto topological maps. The software developed in UNIGIS can draw on the map, semi-automatically, many of the required roads at the locations specified, and can do so according to previously linked data criteria, such as slope, soil type and management practice. The surface area net function, together with the other topology query functions, will also be a useful tool to identify wetland positions or drainage area problems. The estimates of areas is also possible by the software, and this can be produced in report or map form.

   2.2 Soil Parent Material Maps.

   The conversion of existing soil parent material maps to more meaningful soil "form and series" maps, using the DTM to locate slope position or aspect and elevation, will allow catina effects in soil development to be used. The import of data from external sources is possible and FAS or soil profile information can be used.

   2.3 Grid and Surface Survey Queries.

   Such surveys can be completed quickly and much more efficiently than at present if the module is used for surface topology. Maps to desired scales can be generated and the DTM can also be used when slope analyses are required with aspect.

   2.4 Disease, variety and pests queries.

   If sufficient information is available on the above subjects, then the data can be used in the GIS and allocated to its specific location on the map. Once this has been completed the data can be queried together with any of the other relational databases and geographic information stored in the system. Direct access to the GIS is possible using Pest and Disease Committee files and surveys. The advantage over the existing procedures would be that the system could act as a distributed spatial model, and links could be set up between all the different P&D Committees.

2.5 Meteorological data queries.

All existing meteorological data can be incorporated into this GIS. It is a simple operation to enter the data and also to use it in the form of a DTM. As each site is located in space, the generation of isolines of the different recorded values for rain, temperature, evaporation, sunshine and other factors can be easily generated. These can be integrated with any of the other relational database values to answer specific or general questions. Access to the Computing Centre for Water Research data banks is also possible, and either Ascii or other pre-generated map files can be imported.

2.6 Trial area

Forty-three private farms and three sections (26 000 ha) belonging to the C.G.Smith, Illovo group in the Mid-Illovo district in the Natal midlands, were selected as a trial area (Fig 3). For these farms all the available maps and photographs were assembled. These included quota maps, soils maps, ortho-photos, form line maps, land use plans and various aerial photographs. Cadastral information from the Surveyor General's office was used to set up the frame work for the selected farms. The data banks of crop-related details were transferred from the main frame computer on to DBASE files on floppy discs. These consisted of FAS and FRS records and could also have included any pest and disease survey or varietal records. Dummy locations were established for meteorology sites, so that a network of met data could be generated over the area of the model. Real long term met data were attached to the dummy met sites

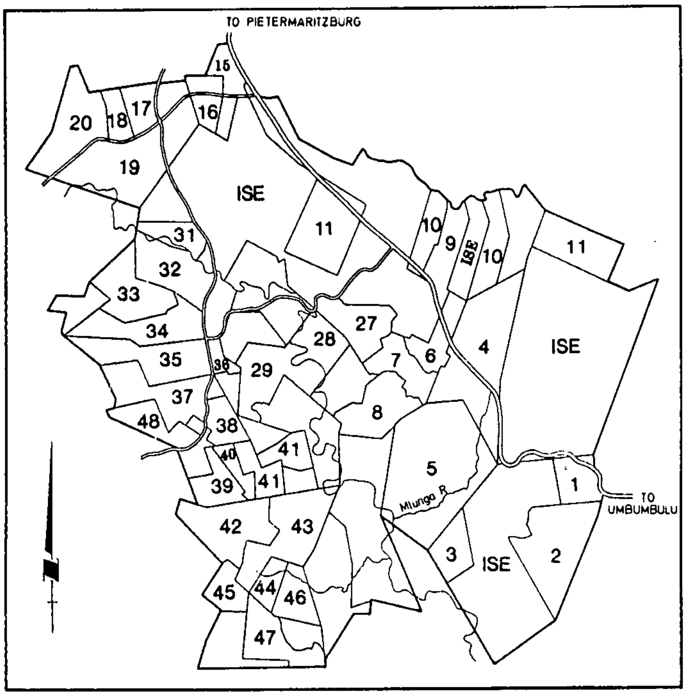

FIGURE 3 Midlands South area used to test the GIS. Shaded areas are farms with FRS or MCP records.

to simulate the conditions which would exist in the true GIS. The legal cadastral data of farm boundaries were entered from files to set up an accurate framework for the model. The remainder of the information from the maps was captured in digital format using the software procedures described.

Of the farms which were included in the trial area only nine are currently on the FRS. However, as long as individuals keep field records there is no limit to the input of data in more informal modes. Similarly for records of other parameters the same procedures could be followed.

Exercises

The first exercise carried out on the data was to check the mapping accuracy. The original cronoflex quota maps for three quota holders were selected. The cadastral information in legal format was entered and the planimetric data was digitised using the AO digitizer. Text and other data bases were added. Measurements were made of the areas digitized using the software. The resulting data were stored on a 4,5" floppy disc, and also plotted out using a standard A1 multipen plotter. The transparent cronoflex original was placed on top of the plotted paper copy to check for any errors. No significant differences could be seen between the original and the computer generated drawing.

The same map was then produced at a scale of 1:10000 and 1:5000

Some examples of queries used in the trial area were:

(a) Show and measure the area concerned which:
is in quota 7531053
between slopes of 15-23%
has a NE aspect
has erodible soils less than 500 mm deep
consists of variety N12
has produced more than 5,6 t/c/ha/m
has an elevation of less than 350 m
has received more than 600 mm rain during past six months

(b) Show and plot the area which:
is in the Illovo mill group area (7531)
above 700 m elevation
has TMS (mistbelt) soils
which has had more than 500 mm rain in the past nine months
where the mean monthly evaporation has been greater than 150 mm

(c) Draw, measure and report the areas:
in the Illovo mill group area
with slopes less than 2%
less than 500 m from a drainage line
with or without cane.

(d) Show and then plot:
quota number 7531-0213
all fields cut between dates 10-04-1988 and 12-12-1988
variety NCo376
Dwyka soils deeper than 350 mm
slopes less than 15%
aspect NW to NE

(e) Using the PADS module a coloured, small format, aerial photograph of one of the farms was digitised onto the cadastral framework. The software made the necessary changes in scale required by the distortion in the photograph.

(f) A series of extraction exercises were carried out using the accumulated data. Drainage areas along streams and drainage lines were extracted using the DTM with the SLOPE module and then this was queried for area planted to sugarcane. Soils were then added to these areas.

(g) A single farm was used to site extraction roads from selected points.

(h) DTMs of the meteorology data were established and isolines for rainfall were calculated. Areas on north-west to north-east slopes between 10 and 20% slopes, soils deeper than 500 mm which had received more than 800 mm of rain were isolated.

Discussion

A great deal of data have to be captured before any meaningful analysis can be completed. The exercise carried out in the Midlands South area showed conclusively that once the data-base has been established it can be effectively used. It is a flexible system, which is one of its real strengths. The limits to the capabilities of the system are more related to the amount of data which can be accumulated and handled. With the advent of faster machines and more advanced data storage and retrieval techniques, it is likely that the system will be able to cope with even more complex queries than it can at present.

While the GIS is powerful as it stands it does need competent and skilled personnel to operate the system. Management queries must be addressed to skilled operators who can then interrogate the GIS. The input of data, by digitizing, must also initially be supervised by trained staff, but as junior staff become familiar with the routines only periodic cross checks will be necessary.

Already the mapping facilities and data queries which are possible using the GIS, justify the initial outlay. As technological advances are made the system will become even more powerful. The ability to communicate with other CAD and GIS packages and to advance within the system, make UNIGIS an extremely useful management and mapping tool.

REFERENCE

1. Dinkele JM, (1989) Cadastral and topographic data for Geographic Information Systems. *Nagis News No. 1* June 1989. p 6