Journal of Nepal Geological Society, 2000, Vol. 21, pp. 103-108

Engineering applications of Global Earth Information System (GEIS)

N. Gurung[1], Y. Iwao[2], and A. Saitoh[2]

[1] School of Civil Engineering, Queensland University of Technology (QUT),Brisbane, Q 4001, Australia

[2] School of Civil Engineering, Saga University, Honjo -1, Saga 840, Japan

ABSTRACT

Database is a systematic collection and storage of certain base information related in space and time. Any such database that is a systematic collection and storage of certain base information needs further intelligent analysis for the engineering application. Also, the organised listings, sorting, analyses, and instant retrieval of this ground information are highly useful in management. We developed the new information system named Global Earth Information System (GEIS) as a powerful assistant for researchers to query, analyse, and map data in support of the decision-making process. The real world information can be collected of thematic layers that can be analysed and linked together by geography. Secondary data are induced from original data as a representative of area or time.

INTRODUCTION

An information system used to input, store, analyse, manipulate, retrieve and output geographically attributed data so to assist in decision-making process for planning and management of geological environment is defined as Global Earth Information System (GEIS). The concept and character of original data are aimed to treat all of the data in the same rank on multidimensional space. Any kinds of data, such as position data in space, geological qualitative and quantitative data, and time series data can be positioned in multidimensional space. It gives one a power to integrate information, visualise scenarios, create maps, solve complicated problems, produce reasonable ideas and generate clever solutions. Any combination of data that are chosen by user's intention can be displayed as a mutual relation figure or time series chart. Some kinds of data combination are profitable for analysing natural disasters, landslide, debris flow, rainfall, slope stability, land subsidence and earthquake. The systems run on the Windows system of PC. Two examples are given to show applications in Japan.

The organised listings, sorting, analyses, and instant retrieval of this ground information are highly useful in management. With the advancement of the computer technology, a huge volume of information has become easily manageable in a small floppy disk. Key components are computer systems, geospatial data, and users itself. Computer system may consist of hardware and software that store, process, analyse, display, and produce output; Geospatial data may be geological information, maps, tables, satellite imagery, aerial photographs, and last but not least human intelligence itself. The geological information can be stored in a form of digital maps and databases. The digital graphic concept of analyses has born due to this development. Fig. 1 depicts typical key components of the user interface used in the computer for geospatial information. Man has always a quest for knowledge of the territory around him. Maps were prepared even in ancient civilisations. The

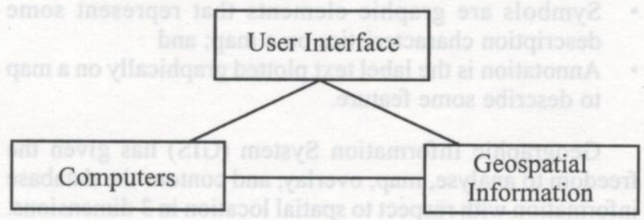

Fig. 1: Key components

survey and mapping works for various topography, soil/ rock type, agricultural land, water resources, forest cover, mineral resources, and urban population have been very common. The process of manual preparation of maps has been costly, time consuming, rigid, and less accurate. Now, computer-aided mapping and design has become cheap, quick, reproducible, and more accurate.

Our system uses space and time coordinates (x,y,z,t). Multivariate statistics as well as regression analyses are possible in these spaces and time dimensions (Fig. 2). It has been programmed in Windows 95 compatible mode and can be operated easily in a personal Pentium computer of 100 MHz speed. Moreover, any sub-routines may be added block by block using a very simple Visual BASIC language.

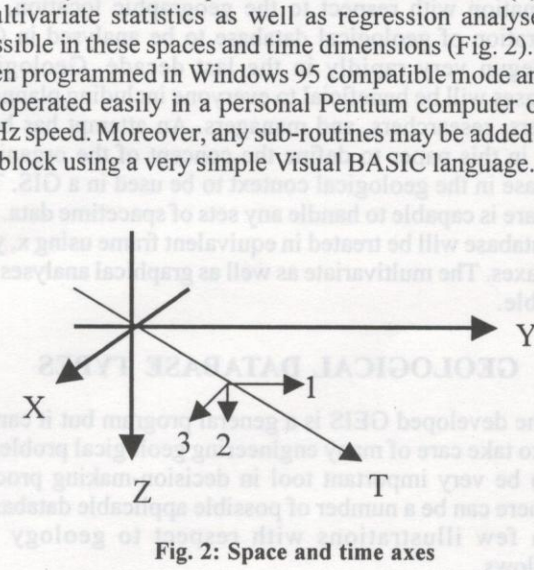

Fig. 2: Space and time axes

GRAPHICAL ELEMENTS

Information is classified in geometric (graphical) and thematic (attributive) data. The graphical data concept use map images in a digital computer useable form of vector or raster to represent a geometry of topology in terms of shape, size and position. It basically uses six types of graphical elements to describe the maps and annotation. These are points, lines, areas, grid cells, pixels, and symbols. The graphical database concepts have basically the following elements:

- Point is a feature with location (*x*, *y*);
- Node is an intersection of lines or strings;
- Line is a chain between 2 points: start and end;
- Area is a bounded, continuous two-dimensional feature represented as polygons;
- Pixel is a two-dimensional smallest indivisible unit element of an image;
- Grid cell is a two-dimensional object showing a single element of a continuous surface;
- Symbols are graphic elements that represent some description characteristics on a map; and
- Annotation is the label text plotted graphically on a map to describe some feature.

Geographic Information System (GIS) has given the freedom to analyse, map, overlay, and contour the database information with respect to spatial location in 3 dimensions. One can easily combine and correlate the database information with respect to the location in a point or line or area or volume in a given space. It will assist studying the cause and effect parameters in addition to mapping and contouring being carried out with the possibility of further regression analyses. But the basic assumption is that the collected information should be accurate, reliable, and worthy to use in an expensive set-up of computer-man-hours.

Because of the great applicability of the geological information with respect to the geographic location, the preparation of geological database to be analysed in GIS has begun very rapidly in the last decade. Geological databases will be beneficial to everyone including planners, builders, researchers, and managers. An attempt has been made in this paper to define the concept of the organised database in the geological context to be used in a GIS. The software is capable to handle any sets of spacetime data. All the database will be treated in equivalent frame using x, y, z, and t axes. The multivariate as well as graphical analyses are possible.

GEOLOGICAL DATABASE TYPES

The developed GEIS is a general program but it can be used to take care of many engineering geological problems. It can be very important tool in decision-making process and there can be a number of possible applicable databases, but a few illustrations with respect to geology are as follows.

Point features

- Boreholes;
- Groundwater wells;
- Mines;
- Test sites;
- Soil/ground property (G, e, N, bearing capacity, groundwater discharge etc.);
- Buildings;
- Historic sites;
- Unique natural areas; and
- Epicentre of earthquakes.

Linear features

- Geological boundary (soil/rock, regional, national and international);
- Joints;
- Folds;
- Faults;
- Dikes;
- Ground unconformity;
- Contour of soil/ground property;
- Earthquake Intensity, climatic factors;
- Lava flow lines;
- Waterways;
- Transportation systems;
- Tunnels;
- Shorelines; and
- Utility systems.

Area features

- Soil/rock types (texture, depth, permeability, agricultural potential, quarry and mine);
- Land cover (urban area, forest, field, wildlife habitat, runoff potential, landslide/flood hazard areas);
- Bedrock geology (mineral resources, oils, unstable slopes, extent of volcanic eruption);
- Land ownership (land use, agricultural field, low land, developed area, land value);
- Hydro-geology (surface/sub-surface) (waterways, wet land, flood plains, aquifer, recharge, flow); and
- Climate (rainfall, temperature, humidity, wind, fog zone, snow area).

Volume features

- Soil deposit;
- Aggregate volume;
- Mineral volume;
- Ground water volume;
- Volume of water in pond, lake, ocean; and
- Volume of atmospheric gases, ozone and green house effect.

MAP PROJECTION

The geological database needs proper collection, testing, and computation. The interpretation of these databases needs correlation and plotting by some composite analyses. For example, the effect of slope, erodibilty, and runoff of a terrain may require composition of soil/rock geology, topography, and watershed maps. The scale of all these three may need a set of adjustments. The frameworks of these maps are fixed in common grid points for so-called overlay analysis. Data encoding formats may be a point, line, or area. The computer program for the storage and analysis may be in grid cell (raster) format or polygon (vector) formatting. The common grid points and conversion of geographic data on to a map require a specific transformation of location. A relative co-ordinate system based on grid for the study area or an absolute co-ordinate system specific to the geographic location of latitude and longitude may be used.

In an absolute co-ordinate system maps can be projected in a planar area by using Universal Transverse Mercator (UTM) system or Lambert Conformal Conic System (LCC) depending upon the shape and orientation of the study region. In LCC Projection, the points on the earth are projected onto a cone that intersects the earth surface at two parallels of latitude. Along these two circles the scale will be exact. Lambert projection is useful for a state that is relatively wide in an east-west direction. Further, in the UTM System, a series of 6 degree of longitude in a north-south strip of 60 zones are covered worldwide. It is useful for a state that is relatively wide in north-south direction.

Graphical measurement and shape analyses are the main features in vector computations. In case of raster computations, area, perimeter and centroid can be easily by using the following obtained by summing pixels of polygon and averaging coordinates.

Formulae:

Area, $A = \frac{1}{2}\sum_{i=1}^{n}(x_{i+1} - x_i)(y_{i+1} - y)$

Perimeter, $P = \sum_{i=1}^{n}\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$

Centroid, $x_{cg} = \frac{1}{6A}\sum_{i=1}^{n}(y_i - y_{i+1})(x_i^{\,2} + x_i x_{i+1} + x_{i+1}^{\,2})$

$$y_{cg} = \frac{1}{6A}\sum_{i=1}^{n}(x_{i+1} - x_i)(y_i^{\,2} + y_i y_{i+1} + y_{i+1}^{\,2})$$

where, $x_{n+1} = x_i$ and $y_{n+1} = y_i$ are the co-ordinates and A is the area of polygon.

The errors on the map may result from various reasons, e.g. digitising, co-ordinate transformation, and errors in the original map itself. The standard deviation (*S.D.*) may be used to define the accuracy between measured and true values.

$$S.D. = \pm\sqrt{\frac{\sum \Delta x_i^{\,2}}{n}} \qquad \text{for } i = 1, n$$

where, n and x_i are number and measurements, respectively.

COMPUTER APPLIANCES

GIS is a tool invented, developed, and advanced over time to this stage for solving information management problem with respect to the geographic location. Depending upon the storage system, it can be a vector-based or a raster -based program. But any GIS program should be capable to store, manipulate, map, overlay and display the geo-referenced databases. The difference in GIS and other types of information is its spatial analysis functions. The spatial as well as non-spatial attribute data can be analysed in the GIS database system. In simple terms it can be categorised to provide 1. storage and retrieval, 2. constrained query, and 3. modelling functions. We applied the GEIS to analyse and map the Saga Plain (Iwao 1986) of Ariake clayey deposits, geological thickness, groundwater condition, engineering soil properties profiles in horizontal as well as vertical directions, and multivariate regression analyses for various correlations. Like in any GIS system, we need input, storage, and output devices. Some of these general needs and various choices are listed below.

Input devices

- Manual encoding
- Co-ordinate digitizer
- Scanner (video, array, electronic)
- Drum or laser scanner

Storage devices

- Magnetic tape
- Hard disk
- Floppy disk
- CD-RAM, ROM
- Optical disk

Output devices

- Line plotter
- Monochrome /colour printer
- Film recorder
- Other plotters

MULTIVARIATE ANALYSES

The multivariate data analysis (Quantification method) is a tool to investigate the correlation of any sets of interrelated system in a given domain. The quantification method of the multivariate data analysis is a method for the independent variables assigned by item categories to use similarity by regression analysis on dependent variables.

We define all the data sets first in equal space time dimension of x, y, z, and t. Subsequently, the results of analysis are established to predict the outside variable according to the response pattern represented by the item-category of the independent variables. The following two applications are briefly described for this purpose.

ARIAKE CLAY DATABASE

The database of the Saga Plain was analysed (Izumiyama 1998) using this approach. The Ariake Clay is very sensitive and soft. The upper part of this unconsolidated sediment deposit is a product of geological processes of alluvial transgression and regression. It consists of silt and clay and some sand lenses. Table 1 is a summary of the geotechnical database of the Ariake Clay.

Quantitative correlation using multivariate analysis from vast soil test database was done on each geographical location to obtain the following consistency limits.

Plasticity Index PI to liquid limit LL, with correlation coefficient of 0.96:

$PI = 0.73(LL-20)$ ……….(1)

Void ratio e to natural water content W_n with correlation coefficient of 0.98:

$e = 0.142+0.02W_n$ ……….(2)

Unit weight, g_t to void ratio e with correlation coefficient of 0.95:

$\gamma_t = 1.788-0.814 \log e$ ……..(3)

Regression summary on other mechanical characteristics such as compression index (C_c), compression ratio (C_r), unconfined compression strength (q_u), and preconsolidation pressure (P_c), with depth (Z), and soil consistency are given below.

Table 1: Summary of Ariake clay properties

Soil characteristics	Value
Physical properties	
Specific gravity, Gs	2.26-2.82
Clay content, (%)	10-81.5
Void ratio, e	0.4-4.53
Liquid limit, LL (%)	32-150
Plasticity Index, PI (%)	7-95
Liquidity Index, LI	0.04-4.64
Natural water content, Wn	12-173
Total unit weight, γ, (gm/cm^3)	1.2-2.11
Mechanical properties	
Unconf.Comp. Strength, q_u (kg/cm^2)	0.04-1.98
Compression index, C_c	0.19-2.81
Pre-Consolidation press, P_c (kg/cm^2)	0.12-2
Standard Penetration value, N	0-5
Sensitivity, S	> 8-16

$Cc = 0.612\,e - 0.409$

$Cc = 0.015\,W_n - 0.342$

$Cc = 0.011\,LL + 0.128$

$Cc = 1.265 - 0.02\,Z$

$Cc = 0.658e-0.002LL-0.34$

$Cc = 0.016\,W_n - 0.002LL-0.288$ ……..(4)

$Cr = 0.096\,e + 0.071$

$Cr = 0.00212\,W_n + 0.081$

$Cr = 0.002\,LL + 0.159$

$Cr = 0.004\,Z + 0.106\,e + 0.01$

$Cr = 0.004\,Z + 0.003\,W_n + 0.019$ …….(5)

$q_u = 0.042\,Z + 0.061$

$q_u = 1.571\gamma_t - 1.859$

$q_u = 0.989-0.211\,e$

$q_u = 0.992-0.006\,W_n$

$q_u = 1.221\gamma_t - 0.056\,e - 1.201$

$q_u = 0.036\,Z + 0.526\gamma_t - 0.653$

$q_u = 0.037\,Z - 0.002\,W_n + 0.27$

$q_u = 0.037\,Z + 0.674\gamma_t + 0.025e - 0.936$

$q_u = 0.036\,Z + 0.631\gamma_t + 0.086e - 0.002\,W_n - 0.85$

$q_u = 0.999-0.061\,e-0.004\,W_n$

$q_u = 0.989-0.211e$ ……..(6)

$P_c = 0.068Z+0.117$

$P_c = 1.39-0.282\,e$

$P_c = 1.368-0.007\,W_n$

$P_c = 0.023Z+0.0835\,q_u+0.127$

$P_c = 0.063\,Z-0.07e+0.338$

$P_c = 0.064\,Z-0.001\,W_n+0.3$

$P_c = 1.134\,q_u - 0.068\,e + 0.363$

$P_c = 1.139\,q_u - 0.002\,W_n + 0.337$

$P_c = 1.39 - 0.003\,W_n - 0.174\,e$

$P_c = 0.02\,Z + 0.84\,q - 0.047e + 0.272$

$P_c = 0.02\,Z + 0.839\,q_u - 0.001\,W_n + 0.243$

$P_c = 1.137\,q_u + 0.001\,W_n - 0.104e + 0.36$

$P_c = 1.195\,q_u - 1.859$ ....(7)

OHITA LANDSLIDE DATABASE

The other example (Iwao et al. 1986) of the system by using the multivariate analysis in space and time axes (x,y,z,t) is the Ohita landslide database. It was used to reveal the fundamental factor for landslide and collapse in the Ohita region of Kyushu, Japan. The topographical map of the Ohita area was divided into 5 minute x 5 minute meshes based on the standard mesh system of numerical system of Japan. Table 2 illustrates the five parameters selected for this analysis.

The category used were the same as in Table 3. The drainage, mean elevation, and slope angle can be easily taken from the topographical map of 1:25,000 scale. Land use and surface geology can be taken from other map layers. The quality and quantity aspects of these items were classified in several categories. The analysed results are shown in Table 3.

Table 3 is useful to judge the range value (difference between maximum and minimum) to landslide susceptibility. In this regard, dense drainage density, mean elevation >140 m, unused area, consolidated sediments, and slope inclination (25-30 degrees) are the most effective factors for landslides and collapse during the heavy rainfall. The accuracy of the multivariate analysis was about 81.4 %.

Table 3: Category and score for various attributes

Item	Category	Score	Range
Drainage density	1	0.000	0.286
	2	0.005	
	3	-0.007	
	4	-0.006	
	5	-0.150	
	6	0.136	
Mean elevation (metre)	1-20	0.000	0.335
	20-50	0.086	
	50-80	-0.019	
	80-110	-0.070	
	110-140	-0.227	
	>140	0.108	
Land use	Water	0.000	0.225
	Urban area	-0.050	
	Unused area	0.175	
	Forest	-0.031	
	Rice field	-0.040	
	Vegetable field	-0.008	
Surface geology	Unconsolidated sediments	0.000	0.249
	Consolidated sediments	0.128	
	Pyroclastic flows	0.038	
	Tuff	0.015	
	Water	-0.121	
Slope inclination (degree)	0-5	0.000	0.181
	5-10	0.109	
	10-15	0.081	
	15-20	0.145	
	20-25	0.100	
	25-30	0.159	
	30-35	-0.022	
	>35	0.002	

CONCLUSIONS

The database concept with graphical and attributive interpolation is a very powerful technique. We have developed a new information system named GEIS as a powerful assistant for researchers in the Saga University. The concept and character of the original data are aimed at treating all of the data in the same rank on the multidimensional space (x, y, z, and t) will help to generalise the analysis level. The program makes use of all kinds of data, such as position data in space, geological qualitative and quantitative data, and time series data that can be transformed easily in a position on the multidimensional space. Besides multivariate analysis of the physical attributes, any combination of data that are chosen by user's intention can be displayed as a mutual relation figure or time series chart. Secondary data are induced from the original data as a representative of area or time. Some kinds of data combination, say typical geological database layer, were profitable for analysing slope stability, groundwater, sub-surface flow, land subsidence, and even the earthquake ground motion. The GEIS database system is written in simple Visual BASIC and runs easily on the Windows system in a PC with 100 or higher MHz speed.

Table 2: Parameters for analysis

Item
Drainage density
Mean elevation (metre)
Landuse
Surface geology
Slope inclination (degree)

REFERENCES

Iwao, Y., Horita, A., and Sumita, M., 1986, The ground data processing system and its practical use by personal computer- A case in Saga plains. Soils and Foundation. Jour. of Japan Soc. of Geotech Engg, No.1595, pp. 29-33.

Iwao, Y., 1986, Effectiveness of using the ground data processing system for the alluvial plains. 5th Int'l IAEG Congress/Buenos Aires, 5.2.5, pp. 1553-57.

Izumiyama, R., 1998, *Study of Multidimensional and optimization of ground databsase*. M. Eng Thesis, Civil Engg., Saga University.