raisteki

GEOGRAPHICAL INFORMATION SYSTEM (GIS) AND ITS STATISTICAL APPLICATIONS IN EGYPT

Mostafa Mohamed Ahmed[1]

This paper deals with the application of Geographical Information System (GIS) technology to statistical and censuses activities in Egypt. GIS is an essential information technology component for future management and decision making. It can underpin both the preplanning phase and the subsequent data compilation, analysis, dissemination and display. Also GIS can reduce the data blindness due to the multidimensional and multivariate complexity of statistical and census data. Two case studies from the ongoing GIS-applications in CAPMAS illustrate these points. The first presents the use of GIS in all stages of economic census in Egypt (selecting pretests area, field works, visualising and analysing results). The second describes the use of GIS in governorates statistical applications such as Population, Housing, Agriculture, Industry and Infrastructure.

1. INTRODUCTION

The Central Agency for Public Mobilisation and Statistics (CAPMAS) is considered the official source for providing all the country organisations, universities, research centres, individuals and international organisations with data, statistics and information that help in planning, development, evaluation, policy formulation and decision making purposes. Also CAPMAS carries out all different kinds of censuses, (population, housing, establishments and economic at the national and local levels).

Due to the rapid progress achieved in information technology and in order to develop CAPMAS' activities so as to enable it to cope with the rapid advances in the technology of storage, retrieval, analysis and dissemination of information, CAPMAS undertook in 1989 to establish the National Statistical Information Network (NSIN) to deal with Integrated Information Systems. For this purpose, we chose one of the advanced technologies in computer systems; it is the Geographical Information System (GIS) in addition to Communication Systems.

---

[1]General Manager of GIS Centre, Central Agency for Public Mobilisation and Statistics, (CAPMAS), Cairo, Arab Republic of Egypt.

GIS is a group of tools, means and automated procedures that can be efficiently used for data transfer, coding, storage and retrieval. GIS has also capabilities for analytical methods to explore relations among data files, and to display and disseminate information which bears a relation to geographical information. GIS uses geographical location for integrating and analysing information and issuing reports as maps to facilitate the deducting of information.

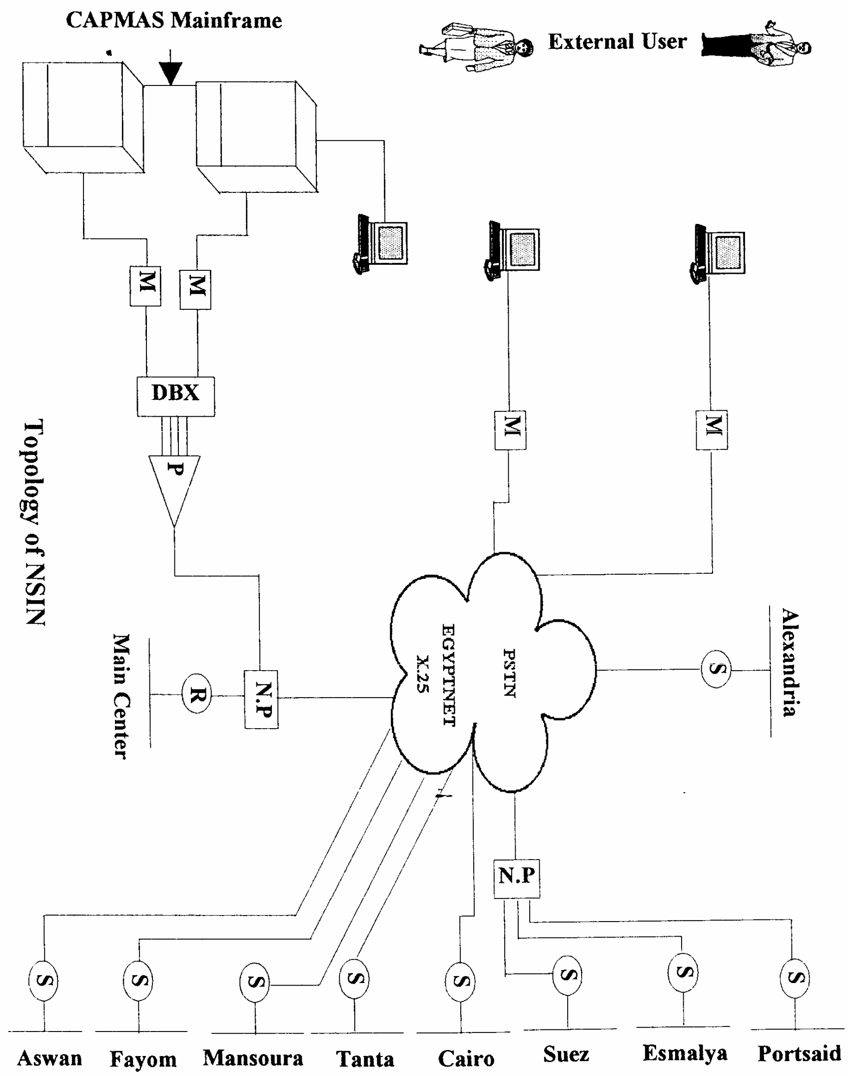
CAPMAS has many computers of different sizes, power and brands (IBM, ICL, HP). Each of these computers stores and processes different statistical data sets, related to a specific sector and/or activity using different Data Base Management Systems (DBMS). Also different applications and databases were developed such as Population and Housing, Foreign Trade, Marriage and Divorce, Economic Census, etc.

It is clear that CAPMAS has a unique and heterogeneous environment. In the meantime, Its strategy is to open all databases for general use by government, public and private sectors. Meanwhile, advanced communication systems were adapted to ensure providing the user with required information in due time and place by using modern communication equipment through EGYPTNET (X.25) and the public switching telephone network (PSTN). So NSIN realises the on-line Communication of external users with databases that are built on CAPMAS different computer systems.

The NSIN project achieved the initiation of eight branches in Alexandria, Esmailya, Portsaid, Suez, Dakahlia, Gharpia, El-Fayom and Aswan. The initiation of NSIN branches in the remaining governorates will be carried out gradually.

The topology of the NSIN project is shown in following the diagram.

Topology of NSIN



1.1. The Objectives of the NSIN Project

1. Utilising GIS-technology for:
   - The provision of base maps that are needed in undertaking the national censuses.
   - Linking the statistical data to its geographic features for better presentation and for spatial and network analysis.
   - Building Digital Base Map Database (DBMD) and making it available to governmental and private sectors.
   - Building regional geographical statistical DB for each governorate to help in decision making.

2. Shifting from centralised data processing to decentralised data processing, which means distributing databases in the sites where information comes from, and making data retrieval easy for decision makers.

3. Shifting from the old method of handling information through paper to the electronic method (on-line data accessing).

4. Adopting modern communication systems to facilitate on-line service for the external user.

5. Training of new academic staff members in the latest developments occurring in the GIS field.

2. GIS AND OTHER INFORMATION SYSTEMS

Advances in computer technology have provided the basis for the development of the geographic information system (GIS). A GIS not only provides for the maintenance of spatially related data, but also provides associativity between the data elements. This allows comprehensive data analysis and the creation of reports on an ad hoc basis. The capability to establish and maintain associativity among features, and to create reports based on that associativity, sets GIS apart from other systems that only store and maintain geographic data. The flexible data analysis provided by a GIS permits the creation of reporting windows within which a variety of questions can be framed in response to the immediate needs of the user.

GIS can be described through its ability to carry out spatial operations and linking data sets using location as a common key. The characteristic feature of a GIS is that the data is spatially indexed, allowing the information from maps, graphs, and tabular data to be combined and displayed in a variety of formats and media.

The relationship between GIS and computer-aided design, computer cartography, database management and remote sensing information system is important to establish a definition of GIS. It is sometimes argued that GIS is a subset or a superset of these systems as shown in the following diagram.

Computer-aided design (CAD) systems were developed for designing and drafting new objects. They are graphic-based and use symbols as primitives to represent features in the interactive design process.

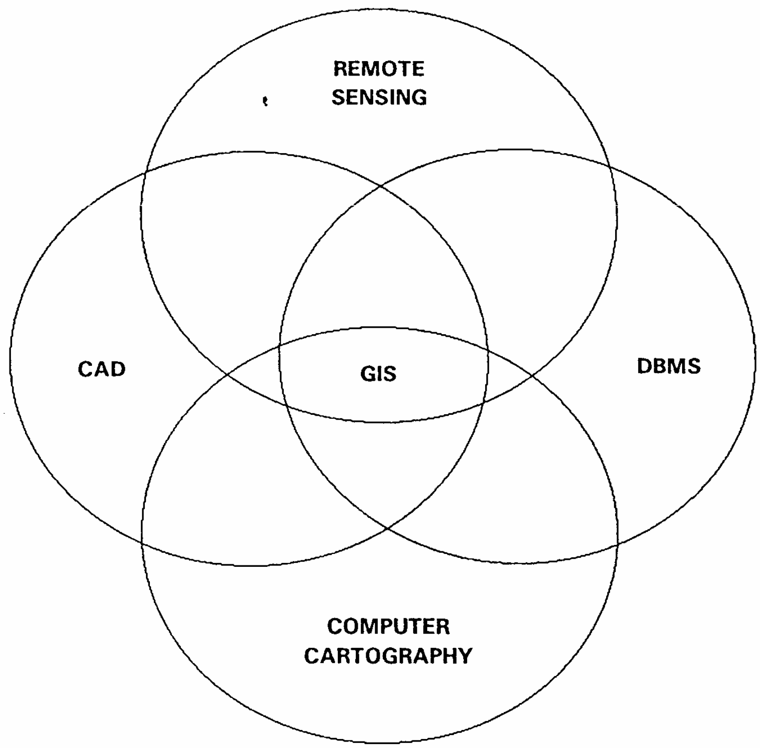
Computer cartography systems focus on data retrieval, classification and automatic symbolisation. They emphasise display rather than retrieval and analysis. Computer cartography systems utilise simple data structures which lack information on topology.

Database management systems (DBMS) are well-developed systems optimised for storing and retrieving non-graphic attribute data. They have limited graphical retrieval and display capabilities.

Remote sensing systems are designed to collect, store, manipulate and display raster data typically derived from scanners mounted on aircraft or satellite platforms.

Thus it appears that GIS is best viewed as a subset and not a superset of the other types of information systems. The emphasis on spatial analysis in GIS raises the question of the difference between GIS and statistical analysis systems. Goodchild (1988) makes a useful contribution defining spatial analysis as that set of analytical methods which require access to both the attributes of the objects under study and to their locational information. Conventional statistical analysis systems such as ASA and Minitab are oriented toward the analysis of a spatial data and lack appropriate capabilities for spatial analysis and modelling (Anselin, 1989). These statements about GIS as spatial analysis systems must be viewed with caution, for as Rhind (1988) indicates, virtually all GIS developments thus far have resulted in

The relation between GIS and other Information Systems



“data retrieval and sifting” engines; modelling work has not yet been brought together with this technically accomplished sub-structure.

The various ideas about GIS can be synthesised and presented in the form of three distinct but overlapping views (Maguire, 1991). These can be termed the map, database and spatial analysis views.

GIS can be applied to many types of problems. Rhind (1990) sets out a general classification of the types of generic questions which GIS is frequently used to investigate. These classifications are:

Condition	What is it ....?
Location	Where is it ....?
Trend	What has changed ....?
Routing	Which is the best way .....?
Pattern	What is the pattern ....?
Modelling	What if ....?

3. GIS AND CENSUS

In all the previous censuses most computing was centred around mainframe computers in centralised locations accessed mainly through data entry stations. Also there was a complete absence of mapping software and diffusion of statistical packages was very limited. These shortcomings did not prevent analyses of census data; in its crudest form, an analysis of the patterns within the data would take the form of a printing of all the tables for each enumeration district (ED) and filing them for later manual analysis.

Afterwards, several statistical analysis packages, such as SPSS and SAS, added graphics and mapping functions for general usage. Although there were mapping programs, these did not have a complete topological data structure to represent the real world; they were only able to plot maps on a printer, colour plotter, and display vector features on computer screen. Nevertheless, these early often crude developments laid down the foundation of computer-aided design (CAD) and GIS. The GIS revolution changes the entire context in which users see and use census data. It is noted that all good GIS packages now contain the databases and mapping functions that 10 years ago would have required the linking of other mapping packages.

GIS also trivialises many previous ‘hard’ census analyses and display tasks. Their present limitations are their focus on the visual rather than the statistical and the difficulty (really inconvenience) of linking the two aspects of analysis. However, in many ways the greatest strength of a GIS in a census analysis context is not the mapping or database functionality, or its ability to provide another way of accessing and mapping census data, but the tools it contains for manipulating census data together with other geographical information, various spatial queries such as what the population of 0-10 year olds living within 1 km of a particular road is. This is useful. Much more useful is its ability to re-engineer census geographies in non-straightforward ways, e.g., to create zoning systems that are composed of aerial entities that are similar in size or shape or social heterogeneity. This is a basic prerequisite for sensible geographical analysis and is one of the greatest gifts that GIS can give the census user: freedom from the tyranny of fixed arbitrary census geography.

3.1. GIS Analysis of Census Data

Analysis of census data in a GIS can take on various forms and various levels of complexity, from simple data retrieval to complicated models. Data retrieval may be spatial or aspatial. Aspatial retrieval refers to the selection of features or data. This usually takes the form of simple rule-based data boundaries as:

- Find all the EDS with a population greater than 100 people.
- Find all EDS with a mean household size of five people.
- Find all EDS with a population greater than 100 and over 50% council property.

Spatial data retrieval may be the interaction between the user and the computer screen using the mouse or other pointing device to point at an object or feature and ask "what is this?".

Spatial queries may be based on the topological or geometrical relationships between objects or features.

- Ask GIS to retrieve all EDS adjacent to or within another type of spatial unit such as a national park.
- Ask GIS to retrieve EDS within 1 km of a motorway.

In terms of the analysis of census-based information, six primary functions are defined (Openshaw, 1995).

3.1.1. Geo-referencing

A GIS allows the computation of relationships between spatial entities, e.g., assigning the census data to the relevant spatial units or geographic features on a digital map. Aggregation and cross-area referencing of census data become straightforward in a GIS. For some spatial analysis such as auto-correlation and zoning, the contiguity of spatial entities is needed; only a GIS can present the geographic features in a realistic way.

3.1.2. Spatial measurement

A GIS provides various geometric functions which allow the measurement of distance, area and density, and also calculate the shortest distance between a point and a line. These functions can be used to obtain the basic measurements, such as the size of built-up areas and travel-to-work distances, which may be required by further spatial and statistical analysis.

3.1.3. Spatial representation and visualisation

In GIS, the distribution of population can be represented by symbol points (of different sizes) or by density shades (of different colours), so that the relationships between spatial objects can be visualised on a graphic terminal or a colour plotter. Since the objects and their attributes are interrelated, spatial queries can also be made by selecting the objects on the screen or by entering the attribute value.

3.1.4. Spatial aggregation

Because spatial objects and their relationships are identifiable or separable in a GIS, aggregation or disaggregation can be easily achieved. In most GIS packages, the commands can group any spatial objects which are adjacent and have one or more attribute value in common. In recent years, many statistical models have been incorporated into a GIS.

3.1.5. Spatial overlay

GIS can create a relationship between the spatial features held in two separate map layers. The overlay functions allow a GIS to store the multiple datasets at different spatial levels or for different spatial units.

3.1.6. Cell-based modelling and surface modelling

The cell-based models can accurately portray continuous surfaces. They can process any discrete data and transfer them to other forms of spatial features, e.g., from raster to vector, and from grid cells to lattices or a triangular irregular network (TIN). For census analysis, a GIS can generate a three-dimensional surface to visualise the results and this can be a useful way of complementing more traditional map displays.

3.2. Visualising Census Data

It is one of the latest rediscoveries in science. It may be that census maps can become alive by merging computer visualisation with virtual reality concepts. As humans we are blinded by the multidimensional and multivariate complexity of the census databases we can now create as we often cannot see the patterns for the detail (Openshaw, 1995). Visualisation is a step towards reducing this data blindness. Simple graphs and charts, maps and diagrams are included in this category as they can contain only a few variables, even if they are of most elaborate design. This is not visualisation, it is a process of making visible what was hidden in huge tables of figures, often too large to be printed on paper. It involves the creation of new solutions for a dilemma of our times: having too much information to comprehend when using conventional analytical techniques.

4. GIS AND STATISTICAL ACTIVITIES

A GIS depends on connecting statistical information with its geographical locations specially when a huge amount of information from different resources is available. The statistical information is presented on 1:5000 digital maps. It is connected with its geographical locations in the form of different layers, each being designated for a certain type of information.

Statistical analysis can be undertaken by processing and merging many layers to get more specific information and help in accurate decision making.

GIS is composed of three basic axes:

1) Perfect Geodetic Co-ordinate Network

This aims to make available a consolidated system of co-ordinates to link information with the real geographic location. We utilise the Geodetic Egyptian Network which depends on the UTM system.

2) Geographic Digital Database for the Base Map

We chose our base map on a scale of 1:5000 because we found that this scale was the most suitable one for statistical and census information. We found that these maps were produced in 1977 and available to only five governorates (Cairo, Suez, Esmailya, Portsaid, Alexandria). A geographic digital database (DDB) was built for each of these governorates using these maps. At the same time CAPMAS began a huge project for initiating new maps for these governorates by aerial photography to update the geographic digital database and production has already begun for the first phase of this program which includes Cairo, Alexandria, Suez, Portsaid, Esmailya; the second phase will deal with the maps of Dakahlia and Gharpia governorates. Maps for the remaining governorates will be produced afterwards in succession.

3) The layers of required information were built and linked with their geographical locations in the base map.

The main roles of a GIS in statistical activities operations are:

- Selecting pre-tests areas

Using geographical maps in selecting the pre-tests areas and samples for different surveys, a GIS has an effective role in selecting the sample representing the universe to give a more comprehensive and expressive image of the reality of the universe and its characteristics. Also a GIS can add spatial characteristics not available before in the statistical frame such as the distance of the sample units from rivers, seas, mountains, services, etc. These characteristics have a great effect on measuring the required statistical indicators.

- In field work;
- By producing general and detailed maps with different scales suitable for statistical work and censuses. Using a GIS, areas and distances can be calculated to help in implementing statistical work;
- Visualising and analysing the data.

Visualising is the study of information through graphical means freed from manual restriction by the advances in scientific computing. Analysis of statistical and census data can take various forms and levels of complexity as discussed above.

5. CASE STUDIES

5.1. GIS and National Censuses

In Egypt, a GIS was used in the economic census of 91/92 and in the pretest of the 1996 population, housing and establishment censuses.

5.1.1. Economic Census Application: Establishment of Health Services and Medical Aid

We chose one of the activities of the economic census of 91/92 which is health services and medical aid. A GIS was applied in all the census stages:

- Selecting the economic survey samples as required.
- Producing geographical maps with different scales for field work.
- Visualising and analysing census data.

All census data were linked to their geographical locations through the map of the country and detailed maps for all governorates. Census data include:

a) Health Services

- Number of medical establishments according to kind of sector.
- Number of beds in the establishments.
- Number of workers in the establishments (doctors, nurses, aid staff, others).

b) Public Medical Aid

- Number of centres.
- Number of workers (doctors, aid personnel, others).
- Medical equipment and instruments.

- Number of operations according to kind of accident (traffic accidents, burns, poisoning cases, etc.).

All the previous data are presented for each governorate according to its administrative areas to the level of the districts and sheiakha.

Comparisons among different governorates with respect to health services can be performed by displaying the country map and selecting the specific governorates for comparison according to any item.

A GIS allows inquiries on:

- Location of the establishments and their serviced areas.
- The nearest medical or aid centres to a certain point and the shortest way to reach it.
- The establishment and aid centre with some specific characteristics and locations.
- Selecting alternative places for establishing a new medical or aid centre according to certain rules or criteria.

5.2. Governorates Statistical Geographical Applications

A GIS was applied to the governorates of the first phase (Alexandria, Cairo, Portsaid, Esmailya, Suez) and a number of applications (layers) were proposed.

In implementing these applications we use information engineering (IE) methodology in the analysis stage, which involves two steps:

- Information strategy planning (ISP);
- Business area and analysis (BAA).

Then we developed the digital database following these steps:

- Building the database by using the digital base map database (DBMD) which we built in CAPMAS and then getting the attribute data into the digital database.

- Performing geographical analyses which include data manipulation, display and analysis.
- Presenting the result and report generation.

As a result of this implementation a number of layers were proposed:

- Population
- Housing
- Agriculture
- Industry
- Food security
- Infrastructure (telephone, drainage, water, roads, electricity).

This is beside building some other special layers for each governorate, such as administrative boundaries, important landmarks, tourism, health and education.

6. CONCLUSION

It is evident that statistics and national censuses data are considered as the basic components of planning, development, policy formulation and decision making. A GIS has the capability to deal with the collection, management and analysis of large volumes of spatially referenced data. So using a GIS with statistical and census databases reduces the data blindness due to the multidimensional and multivariate complexity of those data.

CAPMAS has built a digital base map database (DBMD) for 5 governorates on a scale of 1:5000 and is planning to complete all the other governorates consecutively. The digital base map database is available to all governmental and private organisations. CAPMAS depends on standard specifications for effective sharing data sets, interchange formats, database designs and applications, to reduce cost, risks and implementation times.

Co-operation and co-ordination must be considered among sectors to ensure integrating information from different departments for planners and decision makers. Management co-operation is required to ensure that common GIS standards and procedures are developed and followed. Co-operation among the technical staff is required to share technical knowledge because GIS technology is currently minimal.

REFERENCES

Anselin, L. 1989. *What is Special about Spatial Data? NCGIA*. Technical Paper 89.

GIS World. 1994. GIS Technology 90: Results of the 1990 GIS World geographic information systems surveys. *GISWorld*. Fort Collins.

Goodchild, M. F. 1988. T*owards an Enumeration and Classification of GIS function.*

Maguire, D. J., and Dangermond, J. 1991. The functionality of GIS. In Maguire D. J., Goodchild, M. F., and Rhind, D. W. (eds.) *Geographical Information System: Principles and Applications.* London: Longman.

Mounsey, H. 1988. *Building Databases for Global Science.* London: Taylor & Francis.

Openshaw, Stan. 1995. *Census Users' Hand Book.*

Rhind, D. W. 1990. *Global Databases and GIS.* London: Taylor & Francis.