CHAPTER 16

Spatial integration of environmental and socio-economic data

Anna Lleopart and Jordi Marturià
Institut Cartogràfic de Catalunya, Barcelona, Spain

Introduction

It is not possible to develop studies to analyse and model the interactions between human activities and the environment without taking into account the geographic location and the spatial relations that are set up. For tackling the challenges to sustainable coastal development, it is necessary to possess the knowledge of a set of elements that have a geographic component, such as the location of the most vulnerable zones, pollution, land-use demands, deforestation, and population movements.

To obtain a comprehensive viewpoint of the entire system, it is necessary to implement a two-stage approach. The first stage is to gather the widest possible range of information within a common database. The second is to provide tools to model and analyse the system. The diversity and complexity of spatial information under consideration calls for its integration on a computer-based system, which guarantees efficient management. The GIS (geographical information system) provides the capability for storing, managing, and analysing such data. It has developed into an ideal tool for studies of territorial management and sustainability and covers both stages of the approach referred to earlier.

This chapter aims to go through the first stage, using the case study of Goa as a means, and describe the spatial integration in a common database of environmental and socio-economic data. The second stage, the decision support tools, will be covered in Chapter 20 of this book.

Bases for spatial integration

During the last few decades, powerful tools (including GIS) have been developed to manage geographic information by linking descriptive information (what things are) with geographic information (where things are). These can be defined as computer systems to store, manipulate, analyse, retrieve, and present information attached to a spatial location. In other words, the GIS is an information system specifically designed to manage data with geographical component, and adds new, specific functionalities (such as spatial analysis and modelling and geographic data visualization and presentation) to the common ones (such as database queries and statistical analysis).

The quality and value of information obtained as a result of a GIS analysis depends to a great extent on data availability and the accuracy with which the reality is modelled in the database. It is important to select the correct data to develop the study, design a proper data model of the database, and define the required criteria to guarantee spatial integration of the data involved. The bases for this integration are the definition of a common geographic reference framework for all datasets and the establishment of the suitable geographic component for each set.

For the Goa case study, this integration must include, in a coherent manner, all the compiled biophysical, social, and economical datasets, representing the actual environmental conditions and the socio-economic pressures. The main aspects of the spatial integration and the spatial database definition are the following.

- *The geographic reference framework* The selection of a geodetic reference system (ellipsoid and datum) and a cartographic representation system (cartographic projection) is the first basic element of spatial integration. Once the framework is defined, it is necessary to have the tools to convert the data that does not correspond to any of these criteria.
- *The accuracy of the database* This aspect is not a secondary one, as it will allow or avoid the integration of data from different sources.
- *The geographic referencing of datasets* The second basic element of spatial integration is the establishment of the most suitable geographic component for each dataset, which has to be done in agreement with its nature.

To make this spatial component determination possible, and to spatially integrate all the datasets homogeneously, some spatial reference data must be collected. This dataset is called the cartographic database and most of the thematic datasets are geographically referenced to it. The way this is done depends on their nature and compilation methodology. The general guidelines of this geo-referencing process are as follows.

Data with direct spatial reference (geographic coordinates)

1 *Data compiled by fieldwork with Global Positioning System* This data is directly integrated once the coordinates are converted to the defined geodetic reference system and cartographic projection.
2 *Data obtained from existing sources of information (maps)* This data has to be digitized using the cartographic database as the geographic reference dataset.
3 *Temporal series of satellite images* These have to be geo-codified using the cartographic database as the geographic reference dataset.
4 *Derived data* Data derived from spatially integrated data can be integrated directly into the database.

Data with indirect spatial reference (administrative division code, postal code, etc.)

- This term refers to data without an explicit geographic component (like tables) but which uses as spatial reference another entity included at the cartographic database, which indeed has an explicit geographic component. This is the case with socio-economic information, referring mainly to village boundaries.
- *The data model* The detailed design of the data model must take into account the data exploitation needs, which determine the data model type and, accordingly, the data collection and its incorporation into the database. At present, most GIS software can handle two main data models.
  1 *Vector data model* Data is stored in points, lines, and polygons, which carry the geometry of the geographic features and have linked attributes.
  2 *Raster data model* Data is organized in a regular grid, with each point or cell of the grid having a location and a value.

Both models have advantages and disadvantages; their use is determined principally by the kind of information that is to be stored and the way it is to be used. Generally, it could be said that the vector model is the most suitable for describing discrete elements while the raster model is more efficient at describing those elements that have continuous variation in space.

Geographic reference framework and accuracy of the Goa Spatial Database

To establish a more suitable reference system and working scale for the case study, it is necessary to consider the geographic dimensions of the area and the level of data accuracy required. The study area covers about 600 hectares in Goa, going from 500 m of the water coastal strip, through the beaches, and to the associated drainage sub-basin. It includes, totally or partially, the following villages from Bardez *taluka*: Assagao, Anjuna, Arpora, Parra, Calangute, Nagoa, Marra, Candolim, Nerul, Reis-Magos, and Saligao. Considering the type of data to be collected, analysed, and viewed, as well as the cartography of reference, the more appropriate working scale would be around 1:50 000.

The Survey of India is the official organization responsible for the elaboration of Indian cartography and publishes the 1:50 000 topographical sheet series. For this reason, it has been decided to establish the same geographic reference framework for the GSD (Goa Spatial Database).

Cartographic database

As mentioned earlier, the cartographic database is used as the spatial data of reference that makes possible homogeneous spatial integration of all the required thematic data, coming from several sources and via several collection methods.

Sources of information

As coastal areas are populated and exhibit dynamism with continuous territorial change, no accurate topographic data was available. To obtain an actual cartographic database of the area – the GCD (Goa Cartographic Database) – the approach developed is

based on the use of a high-resolution, geo-coded satellite image as the main reference. Satellite image series are systematic, continuous, and frequently updated sources of spatial information. For the purpose of this project, the panchromatic images of IRS 1D (taken in 1999 and with 5.8-m resolution) were selected.

Data model

Once the primary source of information is decided, it is important to define properly the basic needs of cartography to avoid abortive efforts as well as inadequate information generation. It has been determined that the spatial data of reference needed for the case study is the population nucleus, the communication net, the beach areas, the hydrographical net, the contour lines, the village limits, and the geo-codified satellite images. This database combines vector and raster data. The resultant data model can be seen in Figure 1.

Figure 1
Data model of the Goa Cartographic Database

Collection method

The interpretation of the infrastructure and other identifiable elements on the images is a useful and quick way to collect an updated topographic database. Information on the population nucleus, the communication net, the beach areas, and the hydrographical net was compiled by this method. Only in the case of information related to administrative boundaries and contour lines, this method is not applicable; such data was digitized from existing documents. The procedures for data incorporation into the GCD are shown in Figure 2 and briefly described here.

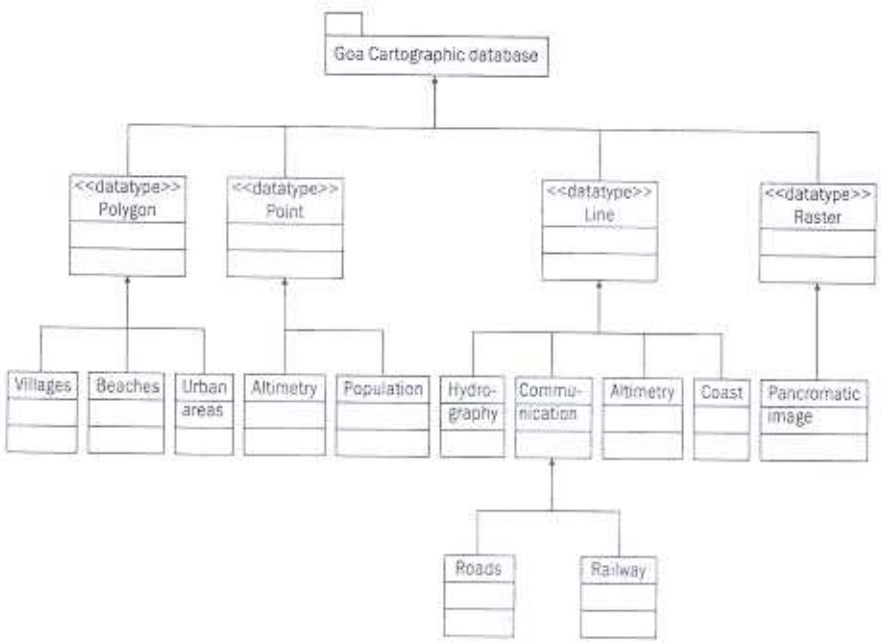

Figure 2
Data compilation procedures

Geo-coded satellite image

The satellite image is geo-coded at the geodetic reference system and the cartographic projection defined. This procedure gives geographic coordinates to the digital image, and is based on the identification of corresponding points on the image and on the territory, coordinates

for which are obtained through the GPS (Global Positioning System). The data obtained is the geo-coded panchromatic satellite image.

Interpreted data

Data compilation is done by means of screen digitization of the different elements identified by interpreting the satellite image and with the help of tourist maps and other available documents. The data compiled by this method includes the population nucleus, the communications net, the beach areas, and the hydrographical net.

Digitized data

In the case of information not directly interpretable from the satellite image, it is necessary to rely on cartographic maps and schemes. First, the source maps are treated by techniques similar to those used for the image geo-codification, so a scanned image of paper documents must be generated and geo-coded. Finally, all relevant

Figure 3
Goa
Cartographic
Database

data can be digitized on screen. The data compiled through this method includes the altimetry and the village boundaries.

Now the GCD, composed by satellite imagery, the topographical data, and the administrative boundaries (Figure 3) is completed and available. This accomplishes the necessary first step for spatial integration of thematic datasets.

Thematic data integration: Goa Spatial Database

To achieve the final objective of creating a spatial database with the necessary information for the case study (the GSD), once the common geographic reference framework is established and the cartographic database is available, the integration of thematic data is essential; this comprises all the data on socio-economic dynamics, land cover, groundwater quality and availability, superficial water quality, and geological features.

This thematic data is varied and obtained through different methods: analysis of existing cartographic data and production of new cartography; generation of data by interviewing people from the area; compilation of statistical data from existing census; generation of data on a GIS environment through the analysis of available information, and so on. To make the integration task possible, this data should be compiled using the given guidelines on geographic referencing of datasets.

The thematic datasets included in the GSD have been compiled following the specified guidelines and spatial integration has been achieved. The integration process has used the following procedures.

Data compiled by fieldwork with GPS

The result is a table containing the attribute data compiled for settlement locations and the coordinates of the position obtained with GPS, properly converted to the established geographic reference system. Data integrated in this way includes field surveys related to vegetation and water status.

Data obtained from existing sources

Such data is collected using the digitized data compilation process described earlier, based on the use of existing cartography. In this

case, the compiled data is mainly environmental information (for example, components of the DRASTIC map).

Temporal series of satellite images

Such data is processed using a method similar to geo-coded satellite imagery compilation process described earlier. The geo-coding process uses, in this case, the panchromatic satellite image included in the GCD to determine the coordinates of the corresponding points. A set of two temporal series of multi-channel satellite images has been included in the GSD.

Derived data

Such data is generated from existing data using GIS analysis tools. Three kind of derived data have been included in GSD.

1 *Classified data*: land cover images derived from geo-coded satellite images applying classification algorithms
2 *Overlaid data*: DRASTIC maps obtained as a result of the application of arithmetic operators to a set of components
3 *Interpolated data*: GALDIT map, generated by applying interpolation algorithms to point data.

Table 1 Goa Spatial Database contents and type of dataset geographic referencing

Dataset	*Geographic referencing**
Socio-economic data	
Percentage share of income (by industry of origin)	5
Percentage share of income (by sector)	5
Income (by industry of origin)	5
Percentage of literates to total population	5
Percentage of male/female literates to total male/female population	5
Statistics on workers distribution (by sector)	5
Statistics on sex ratio	5
Statistics on workers and non-workers	5
Statistics on population	5
Statistics on population density	5
Statistics on migration rates	5
Statistics on average annual growth rate in population	5

Continued...

Table 1 Goa Spatial Database *(contd...)*

Dataset	Geographic referencing*
Estimated pressures	
Fertilizer/pesticide use (kg)	5
Dispersal of potash fertilizers (kg)	5
Land use for industries and tourism (m^2)	5
Dispersal of nitrogenous fertilizers (kg)	5
Dispersal of pesticides (kg)	5
Dispersal of phosphatic fertilizers (kg)	5
Sewage waste generated by domestic population and tourism	5
Solid waste by domestic population	5
Water and groundwater consumption (by use in domestic, livestock, agriculture, industries, and tourism) (m^3)	5
Socio-economic scenario	
Environmental loads by domestic and livestock	
Water demand	5
Sewage generation	5
Solid waste generation	5
Environmental loads due to the tourism sector	
Water consumption (per room per day) by hotel type	5
Sewage generation	5
Land requirement	5
Land space per tourist (m^2)	5
Environmental load due to the industry sector	
Water demand	5
Sewage generation	5
Land requirement	5
Environmental loads due to agriculture	
Percentage irrigated area and water requirement per crop	5
Land under agriculture	5
Estimated fertilizer consumption	5
Dispersal of agricultural chemicals	5
Household survey	
Percentage of households with economic activities	5
Quantity of firewood and twigs/dung consumed in the households	5
First source of income of the household	5
Area of residential land, paddy fields, coconuts, and aquaculture (m^2)	5
Residential land, paddy fields, coconut plantations, and barren land (by type of property)	5
Previous land use by present land use (number and percentage of households)	5
Land not cultivated because of salt-water intrusion	5
Household involvement in tourism (by type of activity)	5
Type of waste disposal in households	5

Continued...

Table 1 Goa Spatial Database *(contd...)*

Dataset	Geographic referencing*
Tourism survey	
Built-up area of the infrastructure and total area occupied by the infrastructure	5
Number of units that introduce changes in physical structure (% within category)	5
Use of diesel energy (% within category)	5
Use of gas by category by village (% within category)	5
Use of kerosene (% within category)	5
Use of solar energy and biogas (% within category)	5
Use of wood (% within category)	5
Use of electricity (% within category)	5
Hotels with garden (% within category)	5
Land use prior to construction (% within category)	5
Sewage treatment facilities offered by the hotels (% within category)	5
Waste generation of the tourism units in season (kg)	5
Waste water disposed (% within hotel category)	5
Number of employees on season, trained workers, and percentage of family workers in unit	5
Land cover	
Geo-coded satellite images (1989, 1990, 1999, 2000)	3
Land cover (1990)	4
Land cover (2000)	4
Changes in land cover	4
Station location	1
Data on mangroves	1
Mangrove vegetation status	1
Structure and composition of mangroves	5
Coastal stresses and drivers	5
Surface water	
Water quality, with respect to metal pollution	1
Water quality, with respect to nutrient pollution	1
Measurement location	1
Ground water	
Depth to water table	2
Net recharge	2
Aquifer material	2
Soil type	2
Topography (slope map)	4
Impact of the unsaturated zone	2
Hydraulic conductivity	2
Drastic maps	4
GALDIT map	4

Continued...

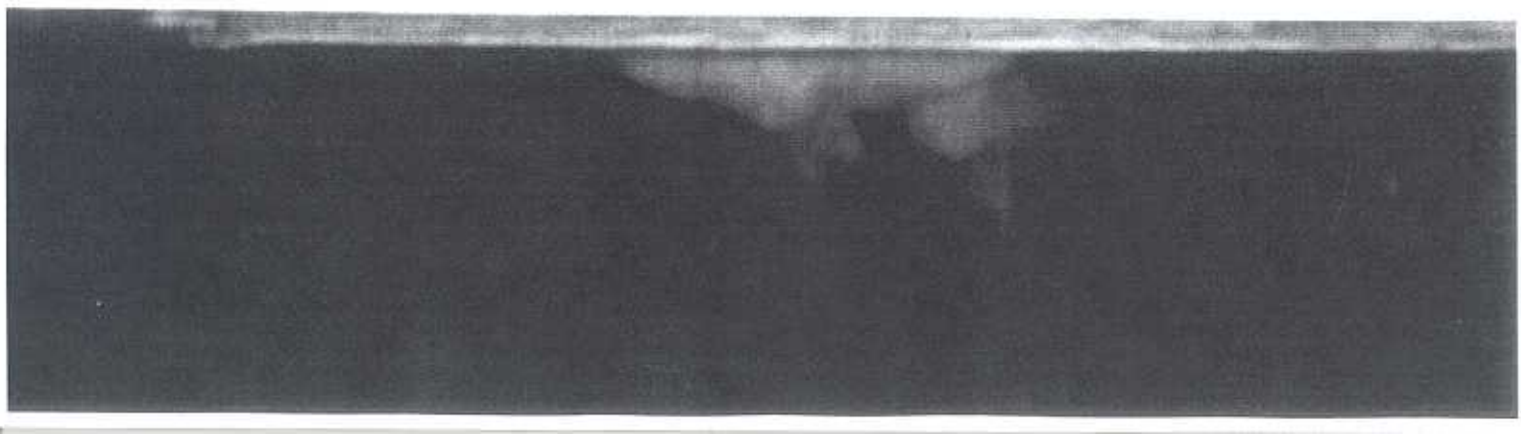

Table 1 Goa Spatial Database *(contd...)*

Dataset	*Geographic referencing**
Required well-head protection zones	4
Well location	1
Water quality	1
Cartographic Database	
Geo-coded panchromatic satellite image	a
Population nucleus	b
Communications net	b
Hydrographical net	b
Beach areas	b
Contour lines	c
Village boundaries	c

* The codes correspond to the ones established earlier under *Thematic data integration: Goa Spatial Database* and *Cartographic database: collection method*)

Tabular data

Such data does not have an explicit geographic component; spatial integration is done by means of its linkage to a geographic entity of the database. So, a field is added to the table by which a link can be established with a geographic entity of the database. This is the case with socio-economic data, where the added field contains the village code.

The content of the Goa Spatial Database and the geographic referencing used for each dataset are summarized in Table 1 and illustrated in Figures 4–7.

Conclusion

The GIS is the most comprehensive system for developing studies to analyse and model the interactions between human activities and the environment. Available data, to a large extent, influences the quality of the information obtained as a result of a GIS analysis. It is vital to integrate all the data into a unique spatial database. The fundamental aspects to bear in mind when establishing the database are the right selection of necessary data, the design of a proper database data model, and the definition of the required criteria to guarantee the spatial integration of the data involved. The bases for the spatial integration are the definition of a common geographic reference framework for all datasets and the establishment of the suitable geographic component for each set.

Figures 4–7 Contents of the Goa Spatial Database

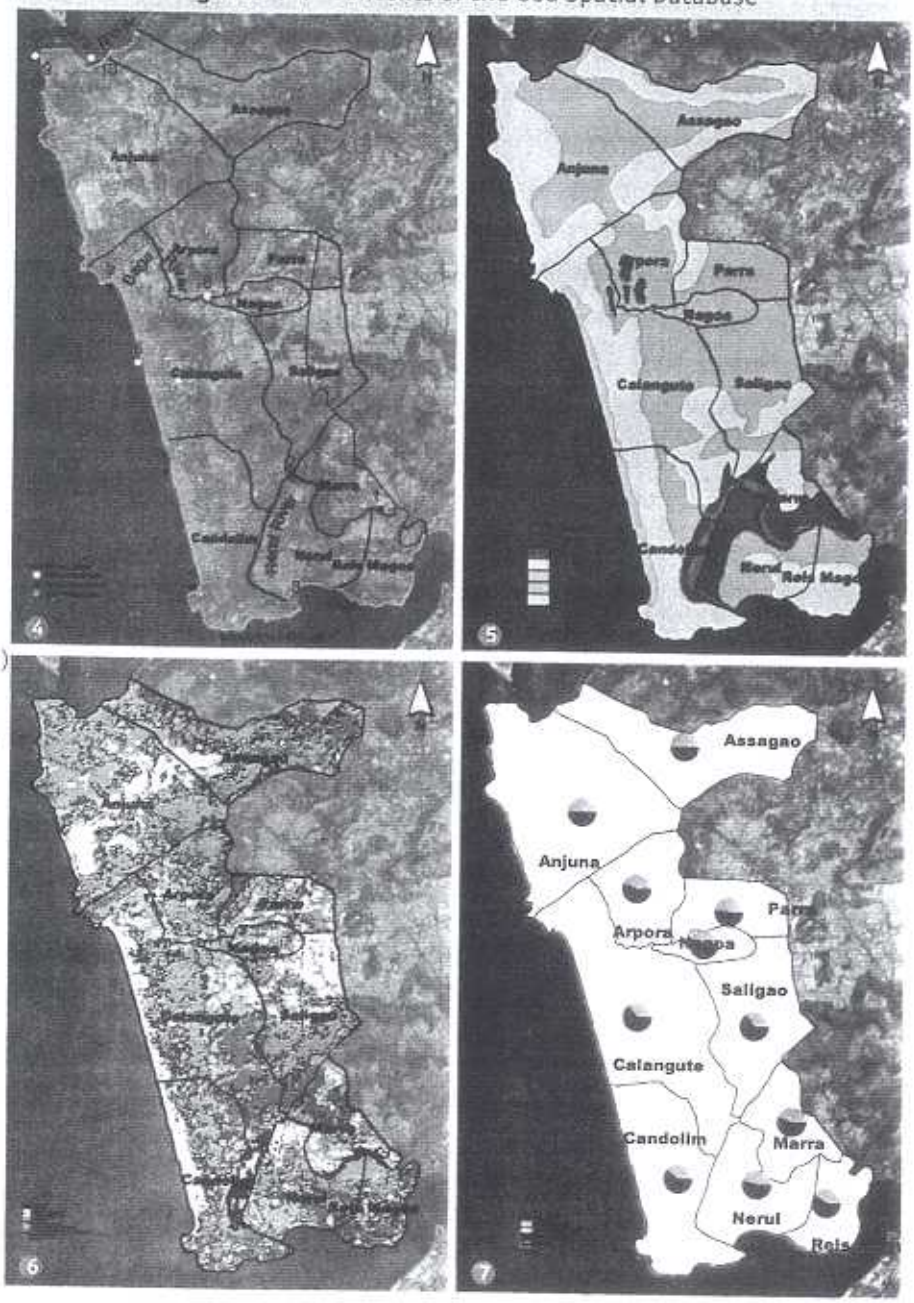

Figure 4
Surface water (data referenced using Global Positioning System)

Figure 5
Soil type (data digitized from existing documents)

Figure 6
Land cover (data derived by classification of multi-temporal satellite images)

Figure 7
Economic data (tabular data linked to villages)