# THE RELEVANCE OF SYSTEMATIC DATA ACQUISITION AND LARGE DATABASES IN THE IMPLEMENTATION OF LARGE PROJECTS

Carlos R. Valenzuela*, Alfonso Zuluaga**
*CLAS/ITC, Cochabamba, Bolivia
Director
CLAS@UMSS.EDU.BO
**Federación de Cafeteros de Colombia

**KEY WORDS**: Systematic data acquisition, Digital Elevation Models, GIS, Spatial databases, Cadastre

## ABSTRACT

Systematic data acquisition at a national level constitutes an important activity for planning development activities and implementing large projects. The National federation of Coffee Growers in Colombia aware of the need of present and future information needs about the coffee growing areas has developed a coffee land information system. The system includes a comprehensive data acquisition procedure at a local level; the data are captured digitally and transferred to a regional database. The regional database is used for regional planning activities, later after a generalization procedure, the data are sent to the national database in Santa Fé de Bogotá. More than 30,000 black and white aerial photographs (1:10,000 scales) and 1,300 topographic maps (1:25,000 scale) are used to delineate the spatial units. The current land use and the area of each parcel are captured in the field using field computers and stored in the regional database, which in turn is used in conjunction with a GIS. The GIS technology will provide, among other things, dynamic support in digital cartography and basic georeferenced information for crop forecasting, environmental protection, infrastructure projects, and more importantly, a permanently updated database, without the need for a coffee census.
Large national projects without adequate and systematic data acquisition procedures, generally, are more expensive and the less viability to be successfully implemented, such is the case of the land clearing and rural cadastre in Bolivia. Several consulting companies executed the land clearing and rural cadastre in various areas of Bolivia, however, because of a lack of a systematic procedure and adequate guidelines, the results of their work are not compatible. Compatibilization and national use of the data will require a new project with the logical cost in money and time. Lack of a systematic data acquisition procedures and an inadequate planning activity has generated social unrest and the termination of an international contract endangering a large water project of close to 200,000 US$ investment. It was supposed to provide drinking water to a population of close to 1 000, 000 inhabitants, irrigation water for a large prime agricultural area and electricity to complement the rural electrification plans of the region

## 1 INTRODUCTION

In most developing countries, the degradation of natural resources represents a critical connection between the increase in population density and rural poverty. As population density increases in rural areas, natural resource use intensifies in order to meet the rising food and fuel (mainly wood) requirements of the population. This immediate need for greater production promotes a change from traditional practices that maintained and conserved resources to more productive but resource depleting agricultural practices. In the absence of practices for the sustainable management of the natural resources, the intensification of their use leads to their degradation. In turn, the resulting loss of productive capacity of natural resources prevents the population from meeting its needs, thereby aggravating the initial conditions of rural poverty. Greater poverty again promotes increased pressure on the natural resource base, causing more degradation. Thus, a vicious cycle is established whereby poverty leads to land degradation, which accentuates poverty further.

Poverty is one of the major obstacles to the development of democracy because it promotes polarization of positions, corruption and intransigence. Rural poverty is leading to the emergence of populist-racist extremist groups, which hold the premise that interracial exploitation is the cause of the misery affecting much of the population.

There is an urgent need for new forms of integration. An alternative methodological basis should be developed for regional decision-making and management to create opportunities for harmonized ecological and economic development. To accomplish this, it is necessary to have a methodological framework, designed and developed to satisfy local conditions, it should include a realistic set of variables and parameters, their thresholds and weights of evidence together with the most appropriate operator and a good validation procedure. It should, of course, maintain the balance between easily data acquisition procedures, updateable and consistent database and a robust model that represents reality accurately.

Examples of large developing projects in Colombia and Bolivia illustrate the advantages of proper data acquisition procedures and systematic databases and the problems of not having them

## 2 THE COLOMBIAN COFFEE INFORMATION SYSTEM

The coffee zone is situated along the three mountain chains that form the Andes system in Colombia, between 10 to 100 N and 71$^{o}$ to 78$^{o}$ W. Coffee is cultivated at elevations ranging from 1000 to 2000 masl. The coffee belt covers around 7 million hectares, of which 900,000 hectares (2000) distributed in 600 municipalities are planted with coffee. Most coffee lots are smaller than three ha and the great majority of the coffee farms are smaller than 10 ha. The temperature ranges from 18$^{o}$C to 22.5$^{o}$C and the annual rainfall ranges from 1,000 to 3,000 mm. The main soil classes (USDA) are Fulvudands Melanudands, Hapludands, Dystropepts, and Troporthents.

Colombia is the second largest producer of coffee in the world, as well as the principal grower and exporter of mild arabica, generally acknowledged the most aromatic and flavorful coffee variety to coffee consumers all over the world.

The National Federation of Coffee Growers of Colombia is a private, non-profit, non-political organization founded in 1927 to improve the cultivation of coffee and to help raise the standard of living of coffee farmers. The Federation s activities have helped shape the coffee industry into a stable element of the Colombian economy. More than 300,000 farmers are engaged in coffee cultivation. At the local operating level, the Federation works through its municipal committees grouped in departmental committees. These committees carry out the Federation s programs for technical assistance, education, health, welfare, and infrastructure. The governing board of the Federation is the National Coffee Congress comprised of delegates from each coffee department

The Federation and ITC have had a permanent agreement on technical cooperation since 1987 [Valenzuela, 1988]. As a result, in 1992 both parties subscribed to an accord in which ITC provides the Federation with technical and scientific support in the fields of remote sensing and GIS for the development of the coffee-growing region and the Federation provides ITC with its experience for improving these fields of expertise. Three Federation staff members have received training at ITC, and ITC consultants have visited the Federation every year.

In 1998, Colombia produced 13,637,000 bags of coffee (each weighing 60 kg) and exported 13,598,000, which is 14% of the world production. The coffee industry constitutes 8% of the working population in Colombia. The Federation has always been aware of the need for updated information of the structure of the coffee land and the social and economic conditions of the coffee farmers. Several coffee censuses have been carried out since 1927. The first, in 1932, found around 356,245 ha planted with coffee. This census was updated with a national statistical sample in 1955/56 [Cepal_FAO, 1958]. The next two censuses took place in 1970 and 1980. During their execution, current technology was used, particularly in the 1980s when topographic maps and aerial photographs (100,000 black-and-white photos, 1:10,000 scale) were intensively utilized, as well as computer support. For the 1990s, the Federation undertook a large project to build up a coffee land information system involving GIS technology [Fedecafe, 1993].

The National Coffee Congress, aware of the needs for a dynamic information source in the coming years, asked the Federation to develop a coffee land information system (CIS) as a basic tool for planning and research and to provide data on the social, economic and technical structure of the coffee zone. Based on the NCC, two components of the CIS are being implemented: the attribute and georeferenced databases. The latter includes digital elevation models (completed in 1993) and the coffee lots georeferenced database (Completed in 1996).

During the period 1993 1997, the Coffee Federation carried out the National Coffee Census (NCC). Under the supervision on the Federation, several Colombian specialized companies [Fedecafe, 1993] worked in 564 municipalities covering 15 departments in the country. The coffee area covers 3.6 million hectares and they used 30,000 aerial photographs at a scale 1:10,000 and 1,300 topographic maps at a scale 1:25,000.

The census data included georeferenced data of close to 500,000 Agricultural Production Units, 1 200,000 coffee lots and the social economic data of the coffee growers, their families and housing as well as the basic data of the inhabitants of the coffee growing area.

The NCC constitutes the basis of the Coffee Information System (CIS) that will provide the Federation basic and strategic information to design and implement policies and programs. These policies and programs should improve the competitivity and sustainability of the Colombian coffee industry and the welfare of the coffee growing families.

### 1.1 WHAT IS CIS

The CIS is a dynamic information system that includes the main variables of the coffee infrastructure (coffee lots, areas, number of plants, varieties, age, illumination, altitude above sea level), the social-economic conditions of the coffee growing families and the indicators of basic needs of the families living in the coffee growing area.

The CIS has two main components:

- The alphanumeric database
- The georeferenced database

**1.1.1 The alphanumeric database.** This database (current version of February 2000) contains the data collected during the CIS; it includes 64 relational tables and has management and continuous updating facilities. It is implemented in Visual Fox Pro in personal portable computers, it has user-friendly interfaces with access to Internet and allows data transfer, both ways, with the georeferenced database, in ILWIS, [Meijerink, et al, 1988], of coffee farms and lots.

The database allows the following operations:

- Updating data on farmers, farms, lots and plants
- Reports on rural education and extension workers to coffee growers and their farms
- Assign areas to extensions workers, according to the spatial division of the Federation
- Perform queries on the status of the coffee plantations in any territorial or administrative zone from tables and graphs
- Perform data transfer electronically from local to national level and vice versa

Currently, the database is installed 500 portable computers operated by extension workers in the field.

**1.1.2 The georeferenced database.** Cartographic component and aerial photographs has thirteen hundred (1,300) topographic maps scale 1:25,000, elaborated by the Geographic Institute Agustín Codazzi, constitute the basic cartography of the CIS. Topographic maps scale 1:100,000 and 1:500,000 are also available. Twenty five thousand (25,000) white and black aerial photographs, scale 1:10,000, taken in several years since 1980 constitute the effective data source for the CIS. Additionally, there are close to 50,000 aerial photographs to support stereoscopic and multitemporal work from previous studies. The topographic maps and the aerial photographs are located in each of the regional of departmental offices. These documents are used in the field in the continuous updating process of the area cultivated with coffee.

The coffee lots are delineated on the 1:10,000 aerial photographs. The minimum area of a coffee lot is 1.5 ha (400 coffee plants). A dot on the photograph indicates all coffee lots smaller than 1.5 ha. In the coffee municipalities of major importance (50), all coffee lots are identified and delineated.

**1.1.3 Georeferenced database.** The cultural, environmental, agronomic, and economic knowledge of the Colombian coffee culture is associated with a spatial location. The spatial data allows digital data handling using communication modern techniques and geographic information systems. Data with these characteristics have an added value, not only because of their quality, but also for their digital analysis capability.

It is important to note the cooperation with the Geographic Institute Agustín Codazzi and the International Institute for Aerospace Survey and Earth Sciences (ITC) in the introduction and conceptualization of the information system technology. It should also be noted the continuous collaboration with the developing group of ILWIS.

**1.1.4. Organization of the digital georeferenced data.** The CIS spatial data has the following components:
*The digital elevation model*: The DEM is necessary to accurately delineate the coffee lots using the monoplot (automatic rectification) facilities of ILWIS. It was generated by interpolation of the 50-meter interval contour lines; the area covered is between the 700 and 2,100 meters above sea level zone. The contour lines were digitized [14] from the 1:50,000 topographic maps and a rigorous dense network of control points ensure the quality of the DEM [Bargagli, 1991, Fedecafe, 1982, Kiernan, 1991, Walsh et al, 1987].

*Georeferenced coffee lots database:* has the purpose to maintain constantly the accurate location, spatial distribution, and area of the coffee lots. Currently, it has 1 151, 000 georeferenced lots in 564 municipalities.

*The attribute database:* It includes all the alphanumeric data and the respective link with the georeferenced database. It supports the spatial analysis in the administrative units of the Federation (Veredas, Districts, Municipalities, Sectionals, and Departments), and facilitates the elaboration of thematic maps.

## 1.2 HOW DOES THE CIS WORK?

Once the Coffee census was completed in 1997, the spatial (georeferenced and attribute) database was delivered to the Coffee regional in the Departments and Municipalities to initiate the coffee farms, coffee lots, and coffee growers updating process

The updating process includes verification in the farm of changes, adjustments, or additions of the coffee farms, lots, and farmers attributes. This is done in forms especially designed and on the aerial photographs.

Field updating is a daily process in the field, digitizing of the aerial photographs is done twice a year. Coffee lots area estimation is performed in the field using conventional procedures.

Data transfer is performed monthly (first ten days of the month). The extension worker informs of all activities to the local office; this informs the departmental office and from there, data are sent to main office in Bogotá de Santa Fé. The data are processed and consolidated in each instance and sent back to the extension worker.

The updating process is one of the most important activities of any information system, it requires to design, and implement a series of procedures and norms for the capture, processing, quality control, and maintenance of the data. This ensures that the data are perceived under the same optic and with precise functioning schemes.

In the CIS, professionals from the Coffee Growers federation extension service perform this process (approximately 1,200). They worked in almost all the coffee municipalities of the country; they perform educational and technical supervision to coffee growers. Each professional is trained and once training is completed, he/she gets the database of his/her area and the necessary software with the portable computer.

The CIS does not only involve the Coffee Growers federation, but also the farmers, who are the main actors. They generate and provide data to input into the system and are the direct beneficiaries of the CIS with updated and precise information for better and timely decision making.

### 1.2.1 Status and applications of the CIS

The CIS will be working during the second semester of 2000, 100 local offices of the 15 Departments with coffee growing areas. The process of updating, data transfer and consolidation will be operational. Currently, the process is operating in 15 pilot areas with satisfactory results. Training of 1,200 extension workers has been completed.

The federation has detail and precise information of the Colombian coffee structure as well as the financial situation of coffee producers and their families. This information is being used to improve strategic policies of technological, social, economic, and environmental improvement of the coffee area region. For instance:

- Measurement and georeferencing coffee lots in the country
- Local and national statistics of producers, farms, lots and plants
- Elaboration of detailed maps with agricultural, social, and economic variables
- Monitoring of data to forecast the coffee production in Colombia
- Support to other projects in the zone, for instance, reconstruction of the coffee belt zone after the 1998 earthquake
- Support to the Territorial Ordering of the coffee growing zone

## 3 BOLIVIA URBAN AND RURAL CADASTRE AND THE MISICUNI PROJECT

### 3.1 Urban and rural cadastre

The Bolivia National Institute of Cadastre has the mandate to regulate, coordinate, supervise, update, and maintain the urban and rural cadastre in the country, implementation, and execution of individual cadastre projects is done at a municipal level. Land clearing activities and the emission of land titles is a prerogative of the National Institute of Agrarian Reform (INRA). Because of the nature of the land clearing process, rural cadastre constitutes the initial activity of INRA, to determine the exact location and area of the parcel in question.

Because of a lack of a methodological framework, with a well defined data model and standard set of procedures, every company, both national and international, executing rural cadastre in various parts of the country, implemented their own methodologies or improvised one. In most cases, international donations or loans finance land-clearing activities; INRA coordinates and supervises all these activities. These, of course, are difficult tasks; quality control and efforts to make all systems compatible constitute an expensive project by itself.

The negative results, very few land titles issued by INRA, in more than four years of work and hundreds of millions of dollars expended are a consequence of a lack of appropriate data acquisition procedures and an adequate design and implementation of a systematic large database, together with a lack of human resources.

Similar situation happens with urban cadastre, where municipalities are considering implementing expensive cadastres with the main objective to increase their income by taxation. However, there is no cadastre tradition in Bolivia. This lack of experience is translated in expensive and unnecessary long projects, usually, with great technical deficiencies.

Coordinated efforts of the International Institute for Aerospace Surveys and Earth Sciences (ITC) of Enschede, The Netherlands, and the Center for Aerospace Surveys and GIS Applications for the

Sustainable Development of Natural Resources (CLAS), of Cochabamba, Bolivia, and various international donors are geared to assist in developing a cadastre urban and rural cadastre that can be used as a model to implement in all municipalities. This concept includes a data model, database structure and data quality control procedures. The national institute of cadastre and INRA will coordinate and supervise the implementation of the rural and urban cadastre respectively.

### 3.2 Misicuni project

The Misicuni project is a 40-year-old aspiration of the people from Cochabamba, a city of close to 1 000,000 inhabitants. It is located in the middle of Bolivia on an old tectonic depression surrounded by mountains and a semiarid climate with water deficit almost every month of the year. The Misicuni project is an ambitious plan to bring water from a large watershed almost 1500 meter higher in the mountains. The multipurpose project has three components: drinking water, irrigation water, and a hydroelectric plant to generate energy. The total cost of the project is estimated at 300 000,000 US$, international loans, own money, and an increase in the fees of the current drinking water supply would finance the project.

The Project has three phases; the first includes the construction of a tunnel of approximately 20 Km across the mountains, construction of a 100 m high dam and the corresponding reservoir. The second and third phase consist in the construction of two additional tunnels to capture water from two adjacent rivers to increase the capacity of the reservoir. The three phases project is planned to be completed in 2003 and only the construction of the first tunnel is completely financed. To insure financing of the construction of the dam and reservoir, distribution of drinking water, and two other tunnels, the project was internationally tendered and awarded to a multinational company.

The construction of the tunnel began in 1999; it was planned to be completed in 2001 together with the dam and reservoir. The company constructing the tunnel based its budget and construction time on studies made over a period of 20 years, under different projects and at a cost of approximately 12 000,000 US$. Immediately construction began, the problems with faulty and/or missing geological data were apparent and several corrective measures had to be taken increasing the costs to a point that the constructing company decided to break the contract and go to an international arbitration procedure. This constituted the point of the iceberg, the water company increase the cost of drinking water up to 200 % in certain cases originating a week of a social unrest not seen since the time of the dictatorship more than 30 years ago. The social unrest caused two dead and tens of injured people and put in danger the fragile democracy of the country.

If the project is not completed or it losses its multiple purpose would be because of the lack of an adequate data acquisition procedure and a systematic database to store data and use them in various related applications. Duplication of efforts with the corresponding expenses is the result of missing or lost reports, which in all cases were printed documents. This is aggravated by the situation of being one of the poorest countries in South America.

ITC and CLAS (CLAS, 2000) are busy developing a geographic information system to help plan the irrigation component of the project. An inventory and diagnostic of the potential irrigation area is under way together with the conceptualization of the system.

## 4 CONCLUSIONS

The Colombian example illustrates a case where the implementation of an information system is and use of modern technology even at field level is possible. The CIS will imply that Coffee census is an activity of the past and it is replaced by a system that guarantees updated and high quality biophysical and social-economic data of farms, parcels, plants and most importantly the farmers, their families and inhabitants of the coffee belt. The application of the information system ensures better planning activities, alternative land use changes and in general improvement of the living quality of the inhabitants of the area.

In contrast, the Bolivian examples show the problems arising because of improvisation and lack of a coherent planning process resulting in projects that are longer in execution and more expensive. The outcomes of these projects are generally below expectations and may lead to social unrest endangering democracy.

Educating a critical mass in principles of regional planning and information systems is imperative to avoid problems like the ones of the Bolivian examples. Regional planning and land use planning are required in most developing countries; however, few universities in developing countries include these subjects in their curricula. These activities are left most of the time to economists or architects with none or very basic knowledge of the territorial planning and spatial information systems, let alone, systematic spatial data acquisition and large databases.

Planning agencies and international donors should consider implementing formal training components in spatial databases and information systems in all of their projects until a critical mass is created. This cadre of professionals would ensure the success of large projects by the use of appropriated methodological frameworks.

## 5 ACKNOWLEDGEMENTS

This article was written jointly by Alfonso R Zuluaga, responsible for the design and implementation of the georeferenced database of the National Coffee Land Survey, and Carlos R Valenzuela of ITC, with the participation of members of the Coffee Studies and Projects Department: Hernando Gonzalez M (head), Manuel Guerrero V (photo interpretation head) and staff members Jaime Castañeda, Luis A Vacca and Henry Avila. We also thank the experts working on this project: Iván Villegas, Diana P Ramírez, Rafael Mora, Mariana Bustos, and Nestor González.

## REFERENCES

Bargagli, A. 1991. Screen digitizing, geometric corrections and monoplotting for Spot, Landsat and normal serial photographs in the ILWIS environment. ITC, Enschede.

CEPAL-FAO. 1958. El café en América Latina. Mexico. 56 pp.

CLAS, 2000. Inventory and analysis of the potential for irrigation in the Valle Centra de Cochabamba. Study made for the Misicuni project. Misicuni internal document. 150 p. Cochabamba, Bolivia.

Federacafe. 1982. Censo Cafetero 1980. Bogotá, Colombia.

Federacafe. 1993. Sistema de Información Cafetera, Encuesta Nacional Cafetera. Manual de Procedimientos. Santafé de Bogotá, Colombia.

Federacafe. 1993. Base Altimétrica Digital de la Zona Cafetera. Dept Estudios y Proyectos Básicos, Santafé de Bogotá, Colombia.

Kiernan. B D. 1991. Quality control dictates conversion success. GIS World, Nov 1991, pp 90-93.

Meijerink, A, C R Valenzuela sud A Stewart (eds). 1988. The Integrated Land and Watershed Management Information System. ITC publ 7, Enschede.

Valenzuela, C R. 1988. Informe sobre la asesoría del ITC a la Federación Nacional de Cafeteros de Colombia sobre el uso de la tecnología SIG. Bogotá, Colombia.

Walsh, 5. D Lightfoot and D Buttler. 1987. Recognition and assessment of error in geographic information systems. PE&RS 53, pp 1423-1430.