GIS INCORPORATING INDIGENOUS KNOWLEDGE:
THE CASE OF LOCAL SOIL CLASSIFICATION AND UTILIZATION IN COSTA RICA

Thesis submitted in partial fulfilment of the requirements for the award of
Master of Science degree
in
Geographic Information Systems for Rural Applications
July 1994

by

RHODORA M. GONZALEZ

Wageningen Agricultural University

International Institute for Aerospace Surveys and Earth Sciences

To my father,
who did not live to read this report
but whose memories inspired me to its conclusion.

ACKNOWLEDGMENT

I wish to express my gratitude to the people and institutions that made this study possible and my stay in The Netherlands a fruitful and pleasant one.

Mr. Horacio R. Morales, President of the Philippine Rural Reconstruction Movement, allowed me to embark and learn the GIS technology while The Netherlands Fellowships Programme extended financial support.

Our Directors of Studies: Drs. Johan C. de Meijere has been a source of encouragement, wisdom and cheer-- and a soulful cellist. Ir. Jan Hein Loedeman's enthusiasm with this thesis guided me to the right path. He also led the class through to Wageningen, even opened his doors for us-- at home for Sinterklas and in the office for the hexagon table. And Dr. Gerrit F. Epema kept the spirit alive.

My advisers: Drs. Sjoukje A. Osinga has been an amiable mentor and a superb mezzo soprano. I learned so much about the concepts of soil taxonomy from Dr. Willem G. Wielemaker and from Ir. Sytze de Bruin who knows the Neguev soils by heart. Dr. Cees Leeuwis helped me in the initial conceptualization of this thesis. Ir. Jetze J. Stoorvogel of the Atlantic Zone Programme provided the much needed maps. Ir. Theo Guiking imparted insights on farming systems and gave me the necessary Neguev reports which made me pursue the topic.

To the Landmeetkunde and its staff: from Chairman Prof. Martien M. Molenaar, to Ir. Rene van Schans, to Ir. Sylvia de Hoop who deepened my understanding of GIS. Ir. John Stuiver has been a very user-friendly interface to Arc/Info. Hans van Leeuwen, Nanna Surayana, John van Smalen and Wim Droesen have been very helpful and cheerful. Mr. Dick de Wit's technical assistance and happy tunes filled the corridors and livened our days. And of course, the LMK secretaries Anke Hoeneveld, Corrie Geurs and Minke Wierstra have been very friendly and helpful with the things we needed.

I will never forget my colleagues-- Jose Antonio M. Casasnovas, Abebe B. Assefa, Kibira M. Sula and Hambai F. Meetwa for the endless discussions and exchange of ideas and most of all for the warmth and camaraderie that sent us through the duration of the course. Even for so short a time Emilio Ascaso, Patricia Eindthoven, Ana Sanchez and Tomas Laguna shared us a lot of knowledge and fun.

Special thanks to my husband, Aldo and children Julian, Katrina and Domenico for the sacrifice and understanding while I'm away from home.

7 July 1994

Dear Dr. Wielemaker,

Thank you so much for the invaluable advice and inputs -- all the knowledge that you shared that made this work possible.

Sincerely yours,
Rhodora.

CONTENTS

Contents (cont'n)

1 INTRODUCTION

1.1 On The Value of Indigenous Knowledge

Transfer of technology and information not adapted to the unique socio-cultural and environmental circumstances of farmers has contributed to degradation of agricultural resources and often to a decline in the economic and social well-being of the intended beneficiaries. Development plans are drafted at too high a level of abstraction and generalization (Richards, 1986). Oftentimes natural resource surveys, which form the basis of management decisions, rely on foreign taxonomies which are not always useful for the end user farmers. The problem is rooted to the lack of coherence between the information a farmer needs and the information transferred by the taxonomic system (pers.con. Wielemaker, 1994). Furthermore, rural small-holder farming is a unique system and can be thought of as a big laboratory where the local people are learning by doing or "trial and error" (REPPIKA, 1990). From astute observations and experiences, empirical relationships are developed and therefore the local people themselves possess a wealth of knowledge (henceforth called indigenous knowledge) of time-tested resource management methods.

Studies indicate that a lack of understanding of the local population's socio-cultural and epistemological systems coupled with the lack of local participation hindered the effectiveness of the innovators and the programs which they had designed (Warren, 1975). This conclusion has prompted a re-examination and re-orientation of many research and extension programs so that recommendations are consistent with the circumstances of the farmers (Altieri, 1983). There is now much evidence and understanding that when resource-poor farmers do not adopt technology, it is usually not from ignorance but because the technology does not fit their physical, social and economic conditions (Chambers, 1985).

The Rio Declaration of the 1992 World Summit on the Environment formally endorsed the value of indigenous knowledge and recommends it as a basis of development efforts. And since many Western trained environmentalists and resource managers end up advising and sometimes even managing resources of other lands and cultures, much damage could be avoided if they understood the cultural, ecological foundations behind non-Western systems of resource management (Klee, 1980).

These trends in recognizing indigenous knowledge (IK) as valuable resource for sound and sustainable development is faced with the fact that IK is being lost due to modernization, urbanization, population growth, ecological change and mass media (REPPIKA, 1990). Few examples have been methodically recorded, and fewer still have been studied with the purpose of developing an integrated approach to solving agricultural and rural problems (Brokensha, et. al 1980).

It is unfortunate that there exists only a few examples of these types of studies since the farmers have to have much experience with their lands that can be used (for example by soil scientists) and the final product can be taken as plus value for the exchange of information (van Uffelen, 1990). Beginning agricultural development work with indigenous soil classification systems can make research and development efforts more effective and more economically efficient.. a useful methodology on both a practical and theoretical level for fostering sustainable agricultural development (Pawluk, et. al. 1992).

1.2 Defining Indigenous Knowledge

Indigenous knowledge (IK) is used synonymously with "traditional" and "local" knowledge to differentiate the knowledge developed by a given community from the international knowledge systems or "Western" knowledge systems generated through universities, government research centers and private industry (Tick, 1993). Information is collected from observing the local environment to solve problems of agricultural production and is passed on through many generations and thereby becomes refined into a system of understanding of natural resources and relevant ecological processes that may lie in the form of principles embodied in stories and religious teachings or in the form of taxonomies (Pawluk, et. al. 1992).

1.3 Indigenous Knowledge of Soils and Classification Systems

Classification systems are contrivances made by men to suit their purposes.. they are not themselves truths that can be discovered (USDA Soil Survey Staff, 1975). Classification systems everywhere reflect distinctions and priorities that are relevant to the creators of the system (Pawluk, et. al. 1992). Indigenous knowledge systems relate to the ways members of a given community define and classify phenomena in the physical/natural, social and ideational environment (Tick, 1993).

Soil classification systems, local or scientifically based ones, provide a common means for talking about soil. They simplify complexity and continuum of the real world into more understandable discreet classes. This simplification is based on criteria biased toward its intended use such as agriculture, engineering or soil evolution (Tabor, 1990).

The objective of soil taxonomy is to have hierarchies of classes that permit us to understand the relationship between soils and the factors responsible for their character (USDA Soil Survey Staff, 1975). One way by which soil scientists classify soils is according to the parent material ("soil genesis") which is based on clay mineralogy and in that context, the soil forming factors (e.g. climate, relief, flora and fauna, lithology, etc.) are important determinants. From these factors, the soil scientists distinguish the different types of soils and assign them their taxonomic class. The more specific the description of the soil forming factors are, the more detailed the soil characteristics can be described. At higher level of classification soils are only characterized in terms of the soil forming process they had undergone (i.e., illuviation, rubefaction, weathering, etc.). At lower level of classification texture class, depth, drainage class, reaction (pH) class can also be specified (pers.con. Wielemaker, 1994).

Land classification systems that are developed by farmers separate soils by characteristics important to a farmer like fertility, manageability, and flooding period if any. They are not based on laboratory analysis but are based on practical (day-to-day) work with their land. The soils that are identified by the farmers closely resemble those of the Soil Series (very detailed classification), in some cases making finer distinctions than would normally be made (Tabor, 1990). On the other hand, they may also make generalizations according to capability criteria or land suitability (pers.con. Wielemaker, 1994).

1.4 Indigenous Knowledge of Soils and GIS

Soil is one of the most important natural resources for the planet. It forms the foundation of agricultural systems and vegetation communities, which in turn affect human and wildlife population. Many developing countries do not have a comprehensive soil resource inventory. In some cases, complete surveys are available but ignored because of lack of communications (or lack of coherence of information), differences in technical language, or in disciplinary approaches (SRIG, 1981).

Incorporating IK of soils in such surveys adds more end-user information and therefore could provide planners a better description of the resource base. Local systems can help the soil scientists identify agricultural interventions that will most economically improve the soil's productivity--- by identifying soil characteristics that are most limiting to present and future management practices. This approach to soil surveys can provide better insights into the farming system, which in turn can better guide agricultural research and development planning (Tabor, 1990). By knowing local priorities, less effort need be spent providing information that is less relevant and more effort spent providing information and assistance that are useful (Pawluk, et. al 1992).

Through soil surveys, large quantities of data are gathered from different sources like image and photo-interpretation, field observations and determinations, and laboratory analyses. Terrain features from the real world are identified and described. These are called geographic data or spatial data. Efficiency in storing and processing these primary data and in displaying the derived information are considerably increased when using computer-assisted procedures operating on appropriately structured databases (Zinck and Valenzuela, 1992).

A Geographic Information System (GIS) provides a certain structure for the description of spatial data and tools for its storage, retrieval, analysis and display (Burrough, 1986). It is an organized collection of computer hardware, software, geographic data and personnel designed to efficiently capture, store, update, manipulate, analyze and display all forms of geographically referenced information (ESRI, 1992). GIS may also be defined as a combination of human and technical resources, together with a set of procedures, which produces information to support decision makers (Tulladhar, 1992).

This report is about the development of a GIS for a specific application-- that of incorporating indigenous knowledge of soils for improving soil suitability classification in soil resource surveys in order that agricultural development planners become more aware of the situation they hope to improve. GIS allows the storage and integration of data from different sources within a single system and therefore offers a consistent framework for analysis.

1.5 Organization of the Report

After the short introduction to indigenous knowledge, the context of the problem to be addressed and the logic of building a spatial information system, this report proceeds to illustrate the step-by-step development of a prototype system. The processes are shown in the following chapters:

Chapter 2 "What Is To Be Done?": The Research presents the objectives and methodology of the research as well as the general background of the case study area.

Chapter 3 "To Be Or Not To Be" : Information Analysis discusses the information requirements in the light of available information for the creation of the database.

Chapter 4 "I Think, Therefore I Am" : On Knowledge Bases explores the concepts behind knowledge-based systems and how it links with GIS.

Chapter 5 "Through The Looking Glass" : Database Modelling demonstrates the approach to object abstraction using object-oriented modelling and the development of the final Entity-Relationship diagram for the case study. It shows the transformation of the data model into formats dictated by requirements of the softwares ARC/INFO and ORACLE and the final database design.

Chapter 6 "As You Like It" : System Implementation presents the results of running the system prototype, the comparison between the scientific soilmap and the one generated by applying indigenous criteria, and the problems that were encountered.

Chapter 7 Evaluation and Conclusion is the analysis of the results, the insights and perspectives. Recommendations for a more thoroughgoing and a more efficient implementation of the system is expounded.

List of References The books, documents, reports used for the realization of the research are listed.

Appendix The Data Dictionary, AMLs and batch-files are presented for reference.

2 "What Is To Be Done?" : The Research

2.1 Research Objectives

The general objective of this research is to develop a spatial database that incorporates indigenous knowledge of soils by means of GIS techniques in order to assist in defining soil suitability classification in soil resource surveys. To realize this, the research set the following **specific objectives:**

To determine the farmers' criteria in classifying their soils and to systematize storage and retrieval of such knowledge of soils.

To compare the similarities and differences of the farmers' land suitability classification with that of the scientists.

To implement the system to a test case.

To generate a report.

2.2 Research Questions and Hypotheses

Fundamental questions and hypotheses that have to be answered and tested in the course of the research are the following:

Question 1: Can IK be incorporated in a GIS design? If yes, how? If not, why?

Question 2: Can incorporating IK in a GIS design help improve land suitability classification? If yes, how? If not, why?

Hypothesis 1: IK can be incorporated in knowledge based systems where the knowledge is represented in the form of empirical rules and relationships.

Hypothesis 2: Indigenous criteria for soil classification and aggregation form the bases for the definition of indigenous soil units.

2.3 The Study Area

2.3.1 Background

The Atlantic Zone of Costa Rica (see Fig.1) is an area with a large diversity of soils varying from fertile alluvial soils to very poor wastelands. The Neguev area is located in the districts of Germania and Cairo (of the canton of Siquirres) and in the districts of Pocora and Rio Jimenez (of the canton of Guacimo) in the province of Limon. The temperature doesn't vary much during the year and has an annual average of 30.5°C maximum, 20.9°C minimum, and 25°C mean. The rainy months are July and December and the average annual precipitation is 3666 mm (de Bruin, 1992).

The Neguev settlement area occupies 5,340 hectares and is the result of an occupation organized in the 1970s by a union of small farmers, the *Union de Pequeños Agricultores de la Region Atlantica* (UPAGRA), in the Northeast part of the province of Limon at the boundary of Siquirres. The property was owned by a livestock firm called

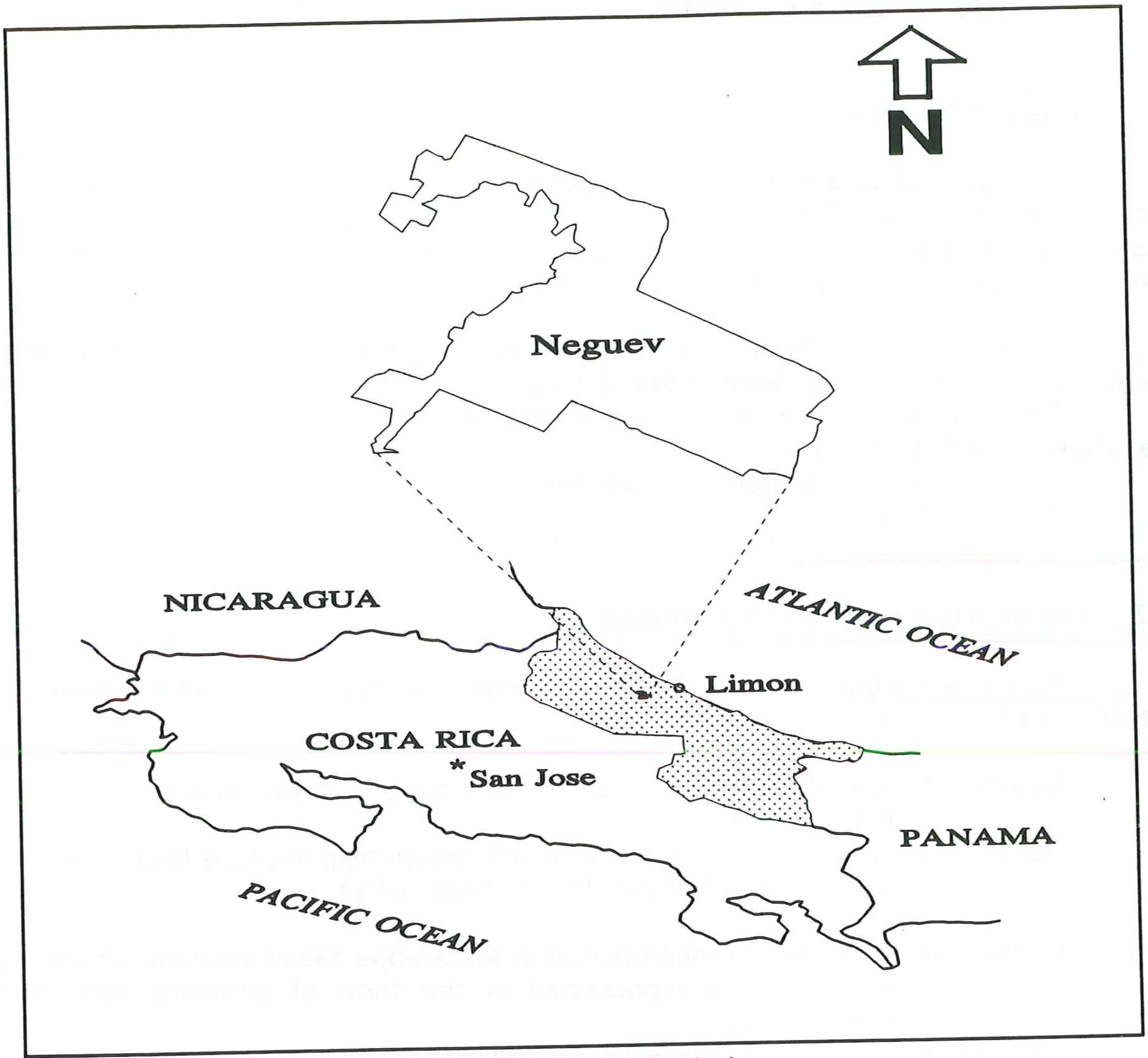


Fig.1 Location of the Study Area

Industrial Neguev, S.A. which is a subsidiary of *Inmobiliaria Agromercantil Caribe, S.A.* Occupations like these were tolerated by the Costa Rican government for as long as they remain within a tolerable scale. But when the situation became hostile in 1979, the *Instituto de Desarollo Agrario* (IDA)-- the government agency in charge of agricultural development intervened. IDA bought large tracts of land (one of them the Neguev area) in order to quell the invasion of forests and estates by the dislocated immigrants in search of jobs in nearby banana plantations that hire people only on a temporary basis (Huising, 1993). The Neguev land was subdivided into 313 parcels of 10 to 17 hectares each and were awarded to farmers by lottery. IDA is the official authority for the management of the Neguev settlement. This settlement project was included in the so called "0-34" program financed by a convention with USAID-- the U.S. Agency for International Development (van den Berg and Droog, 1992; van Uffelen, 1990).

Neguev is divided into five sectors: La Lucha, Milano, Bella Vista, El Silencio, and El Peje. Every sector has its own communal center with a school, sports green, some small shops and a small public center (van Uffelen, 1990).

2.3.2 Choice of Area

The choice of this area stems from the fact that the research question--originally related to another region in the world --was most conveniently answered by the use of available data from the Atlantic Zone. Previous studies reported on the existence of a local soil classification system in the Atlantic Zone of Costa Rica which had to be defined in terms of the survey team's scientific classification system (e.g., Mucher, 1991; Veltman, 1991) and another study on the criteria that the farmers use in classifying and determining the aptitude of their land (i.e., van Uffelen, 1990). These reports contain data which, although limited in scope, precisely fit the purpose of the research presented here. The availability of a soilmap and a parcelmap at the right scale was also a decisive factor.

2.3.3 Benefits From the Study

The output of this study is a methodology to formalize and incorporate non-quantitative local knowledge into a database that usually require quantitative input. Specifically, manipulation and reclassification of a database of soils using local criteria is put forward. In that way, a combination of theoretical formulations of scientists and practical experience of farmers is realized to come up with a soil suitability classification that can better describe the actual conditions of an area. This could benefit a development planning process that formulates projects and gives recommendations to farmers on how to effectively use their land.

2.4 Research Methodology

2.4.1 Literature Review and Interviews

Intensive research of books, reports, and various publications was done to determine the nature of the complex subjects indigenous knowledge, knowledge-based systems and soil classification systems. Personal interviews with the soil scientists and agronomists who worked in the Neguev area were also undertaken in order to understand the concepts behind the soil taxonomy and to gain more insight about the study area and its people. Translation of several documents and references from Spanish to English was also carried out.

2.4.2 Defining Soil Resource Survey

Soil resource survey, the field and subsequent laboratory study of soils, is based on grouping individual soils into units defined by their characteristics, properties, and evolution, elements which permit the expression of their specificity, the role they play in ecosystems, and their possibilities for utilization. Mapping is performed to show the spatial distribution of these defined units (SRIG, 1981).

2.4.3 Information Analysis

An assessment of the available data was done and methods to transform those that have to fit the requirements of the study were developed. Indigenous soil classification criteria were obtained from reports on its existence in the study area. Problems with the presently used land suitability classification were also gathered from several reports like de Bruin (1992); van Uffelen (1990); and Brink and Waaijenberg (1987).

2.4.4 Knowledge Representation

A methodology to represent knowledge and criteria for indigenous soil classification systems was developed in accordance with concepts of knowledge-based systems. This concerns knowledge acquisition from experts (in this case the farmers) and representation of such knowledge in the form of rules to be used in defining soil units.

2.4.5 Conceptual/Logical Database Modelling

Selected terrain features were considered for modelling the real world. Object abstraction using object-oriented modelling approach (Molenaar, 1993; Egenhofer and Frank, 1989) was used to define terrain features. An Entity-Relationship Diagram (Chen, 1976) was drawn and the corresponding skeleton tables were made, showing all keys and foreign keys.

2.4.6 Physical Database Modelling

The design of a relational database in ORACLE, which was the available database management software, was found to be adequate. The skeleton tables were used as guides in creating tables to be accommodated by the softwares used on a UNIX workstation platform:

ARC/INFO (v.6.1.1)
ORACLE (v.6.0.34.2.1)
SQL*PLUS (v.3.0.11.1.1)

2.4.7 System Test

Pre-processing and editing of maps were also done to conform with the software Arc/Info requirements. A case on soil suitability mapping for maize and banana was implemented to test the system. A simulated case on application for credit was also done. Statistics were calculated to help in analyzing the changes after the reclassification process. Maps were produced to display the results and help in analyzing the effectivity of the designed database in incorporating IK of soils.

2.5 Assumptions and Limitations

2.5.1 On the Scope of the Study

This undertaking is limited to the outcome of previous studies conducted in the Neguev settlement area and had to rely on the various reports from such studies. The intention is to develop a methodology to structure and formalize indigenous knowledge in order to be incorporated in a GIS design. Therefore, uncertainties and fuzziness of classifications were not dealt with.

2.5.2 On the Data at Hand

Personal interview with the farmers in order to elicit IK of soils was not possible because of limitations in time and resources. The IK identified in a previous study in the area (i.e., van Uffelen, 1990) was deemed enough for the purpose because the main aim is to show how IK could be formalized and structured in a GIS design. But owing to the differences in the objectives of the previous studies and those of the present, the data were limited to those that could be found fit for the purpose of showing the basic principles that underlie the proposed methodology.

The socio-economic data were lifted from the document *Base de Datos de Una Encuesta de Caracterizacion de Fincas Realizada en el Norte de la Zona Atlantica de Costa Rica, 1987*, or simply called the *Encuesta General*. The basic maps were supplied by the Atlantic Zone Programme and benefited from verifications with the original soilmap (scale 1:20,000) that was prepared by S.de Bruin and W.Wielemaker and from their extensive knowledge of the area.

2.5.3 On IDA

Knowledge of the information requirements of IDA was culled from general descriptions of its objectives and functions as mentioned in almost all of the reports about the area, for it being the chief administrator of the Neguev settlement. Feedbacks from the people of Neguev regarding the performance of the IDA (as documented in the *Encuesta General*) were also considered for the present study.

3 "To Be Or Not To Be" : Information Analysis

3.1 Introduction

Information analysis is the analysis of the information use and the determination of information requirements of an organization (McMenamin and Palmer, 1984). It is the first step in building an information system after defining the problem (and after deciding about the necessity of building one). On the basis of the problem defined, the necessary information that the database will have to contain are also defined. This is compared with an assessment of the availability of these data. The comparison is very important because as a rule, especially in third-world countries, availability of data is a severe limiting factor upon the effectiveness of GIS applications. The output of this stage is an initial list of systems requirements.

The system to be developed must contain only the essential requirements--those that are truly necessary to fulfil its purpose, *regardless* of how the system is implemented (McMenamin and Palmer, 1984). This means that information requirements are identified independent of the technology or softwares to be used. The processes and flow of information in performing organizational functions are analyzed. Drawing techniques like the Data Flow Diagrams (DFD) which show the processes and the exchange of data among them, and the Entity-Relationship diagram (Chen, 1976) are used to facilitate the analysis.

3.2 Neguev Situationer

For ease in administration, IDA divided Neguev's 313 parcels into five sectors, namely: La Lucha, Silencio, Milano, Bella Vista, and El Peje. A parcel may have different land uses, depending on what the farmer decides or on what the IDA promotes for production. As the chief administrator of the settlement area, IDA's field technicians monitor the production of every parcel and prescribe ways to improve production yields. A farmer may avail of credit facilities provided by the IDA or other creditors. Farmer organizations exist in the area and a person may become a member of any of these organizations. An information system that could assist in these functions is to be designed. Basic queries in locating soils for specific crops should be answered more in accordance with reality. This is to assist in activities similar to the following:

Some years back, the IDA organized a maize project. Maize is generally grown on black soils but the project also included farmers that had experience with maize on red soils. The selection criteria was based on good drainage and high pH. But analysis of the yield showed that the type of soil was a more decisive factor than simple parameters like soil pH and that the knowledge of the farmers on the type of soils make valid instruments for determining soil capacity (van Uffelen, 1990).

The IDA has been using the *Centro Cientifico Tropical* (CCT) classification system based on the *Manual Para La Determinacion De La Capacidad De Uso De Las Tierras De Costa Rica* (or simply the *CCT Capacidad de Uso*) for determining soil suitability classification. The *CCT Capacidad de Uso* divided the land into 10 classes:

Class 1 - For Annual Cultivars (Very High Yield)
Class 2 - For Annual Cultivars (High Yield)
Class 3 - For Annual Cultivars (Moderate Yield)
Class 4 - For Permanent or Semi-Permanent Cultivars
Class 5 - For Intensive Pasture
Class 6 - For Extensive Pasture
Class 7 - For Tree Crops
Class 8 - For Intensive Forest
Class 9 - For Extensive Forest
Class 10 - For Protection

These classes are indicated for each of the scientific soil units that the IDA makes use of in formulating its projects. The CCT definition for annual cultivars include the following crops: rice, beans, maize, tomatoes, chili, cabbage, onions, peanuts, yam, tobacco, etc. Semi-permanent and permanent cultivars include sugarcane, coffee, black pepper, banana, platano, papaya. Tree crops are mostly cocoa, coffee, citrus, coconut, and macadamia (CCT, 1985). In this classification, lower categories of soils cannot be used for purposes of the higher categories; but higher categories of soils can be used for purposes of the lower categories. The problem with this classification system is that it assigns a land class to a group of crops altogether. A soil suitability classification combined with general crop groups such as annuals, semi-permanent or permanent cultivars, is of little use when it is necessary to plan for a specific crop since each crop has its own particularities (de Bruin, 1992). As a result, problems in giving the proper advice to the farmers occur and are manifested in some of their unfavorable remarks obtained during the *Encuesta General* (Brink and Waaijenberg, 1987):

"A lot of IDA projects are a failure, and so we don't have faith in their technicians."

"The technicians are novices in practice, but they think as if they know everything."

"Perhaps a farmer has more experience than them."

"They make recommendations without taking to account the climate."

"It costs a lot, and IDA has inappropriate ideas-- very advanced and costly."

3.3 Farmer Perception and Scientific Formulation

Most soil classification systems give information on the genesis of the soils as reflected by objective and measurable criteria. At the highest category level, which is the Order, a number of principal genetic soil types are recognized and are subdivided to criteria which are order-specific. The classification goes on to Sub-order, Great Group, Sub-group, Family, and Series. This means that the criteria for classification do not convey consistent information other than genetic properties. In many soil surveys, even at semi-detailed scale (1:50,000) soils are just classified down to Sub-group level. Therefore, only general information like depth, drainage class, and texture (as in the case of vertisols) are transferred to the user of the soilmap. Because classification is based on soil genesis, it does not give information on potential land use. That is why a translation of data (like land evaluation which considers agroecological information) from soil taxonomy is a necessary step to estimate the most appropriate land use. And generally, the needed information is found at the level of detail equivalent to the Soil Series (pers.con. Wielemaker, 1994).

The farmers, on the other hand, do not perform objective measurements on their soils but assess its qualitative performance based on the crop yields and some constraints like flooding and soil manageability. From the findings of van Uffelen (1990), the farmers of Neguev settlement associate observable characteristics of their soils with the behavior of their crops. Color is an important criteria to distinguish soil types as indicated by their local names:

1. Tierra Negra (black soils)
2. Tierra Bermeja (brown soils)
3. Tierra Colorada (red soils)
4. Tierra Muy Roja (very red soils)
5. Suamposa (swampy or poorly drained soils)

The brown, red and very red soils are an elaboration of the indications for the red soils.

Other criteria according to farmers are: altitude or location (i.e., high or low places), humidity (i.e., wet or dry), texture (i.e., sandy or clayey), and structure (i.e., soft or hard). These criteria are related with the color criteria that in high places, there are generally red-colored soils which are dry and clayey and the *capa-dura* (hard and dry layer) phenomenon is common. And that black soils which are more fertile, soft and humid are generally found in lower areas and near the river banks. The Neguev farmers have also evaluated their lands according to the performance of their most important crops:

1. Cropgroup one (corn, beans, rice, banana, platano and pasture) performs better on black soils.
2. Cropgroup two (palmito and cocoa) can be grown on either black or red soils.
3. Cropgroup three (pineapple and chili) performs better on red soils.

3.4 The Proposed Method

If the qualitative and quantitative approach of the farmers and the scientists can be integrated, planners can take advantage of the combination of information. The knowledge of the farmers account for what is left out by the *CCT Capacidad de Uso* (van Uffelen, 1990). How to go about this? The proposed methodology follows (see Fig.2):

With this method, to consider the "world view" of the farmers is to define the landscape with their criteria (the IK rules). This will be facilitated by aerial photographs and a topographic map. To articulate the farmers' classification of soils with that of the scientists, it is necessary to define some data in common from the farmers and the soil scientists (Furbee, 1990). This will be done with the IK criteria in describing soils in the landscape and will be shown in Chapter 4. The resulting indigenous soilmap will be further evaluated using the farmers' criteria for evaluating the soils. The outcome will be an indigenous soil suitability map which can then be compared with the scientific soil suitability map that used the *CCT Capacidad de Uso* for its process (dotted ellipse in the figure). The output is a report that planners can consider in promoting agricultural production-- one wherein scientific formulation and farmers' perception are integrated. Future close examination of the two should give a contextualized view of similarities and differences in classes in the farmers' and scientists' systems.

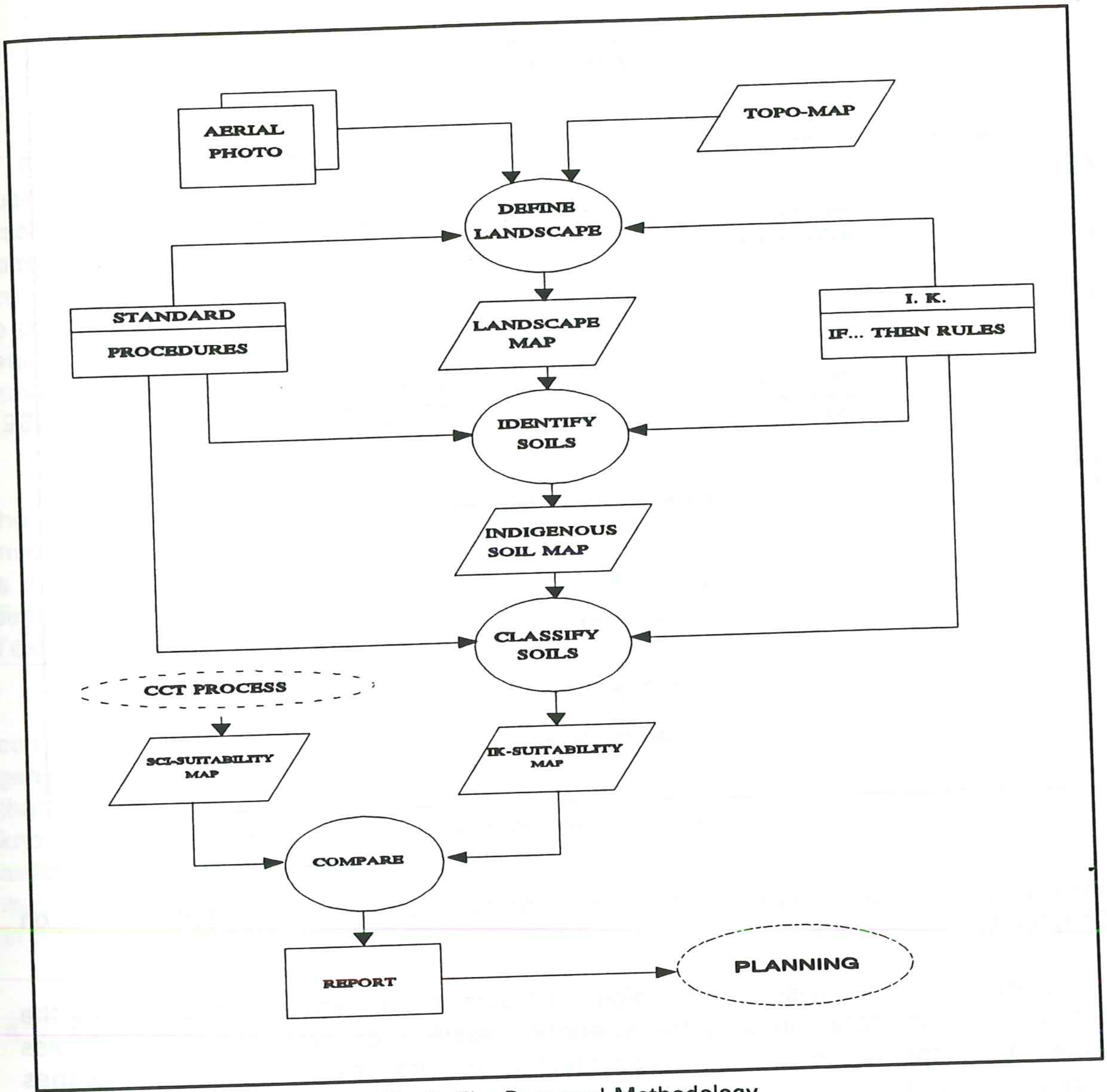


Fig.2 The Proposed Methodology

The problem with this schema is that aerial photographs of the right scale are not available. Scale of 1:10,000 to 1:20,000 are preferred in order to come closer to the approximation of the farmers' view and be able to delineate the landscape according to their criteria. These indigenous criteria for defining the landscape are very necessary but are not available from the reports at hand . It is expected that the farmers would identify objects in the landscape different from those that the soil scientists would (pers.con. de Bruin, 1994). A different approach had to be formulated and is shown in Figure 3.

The revised method is to reclassify the existing scientific soilmap using indigenous criteria for classifying soils. A *de facto* farmers' soilmap would result and later evaluated according to their criteria in assessing their crop yields. The created indigenous soil suitability map can then be compared with the scientific soil suitability map. GIS techniques will be very useful in selecting individual soil units and

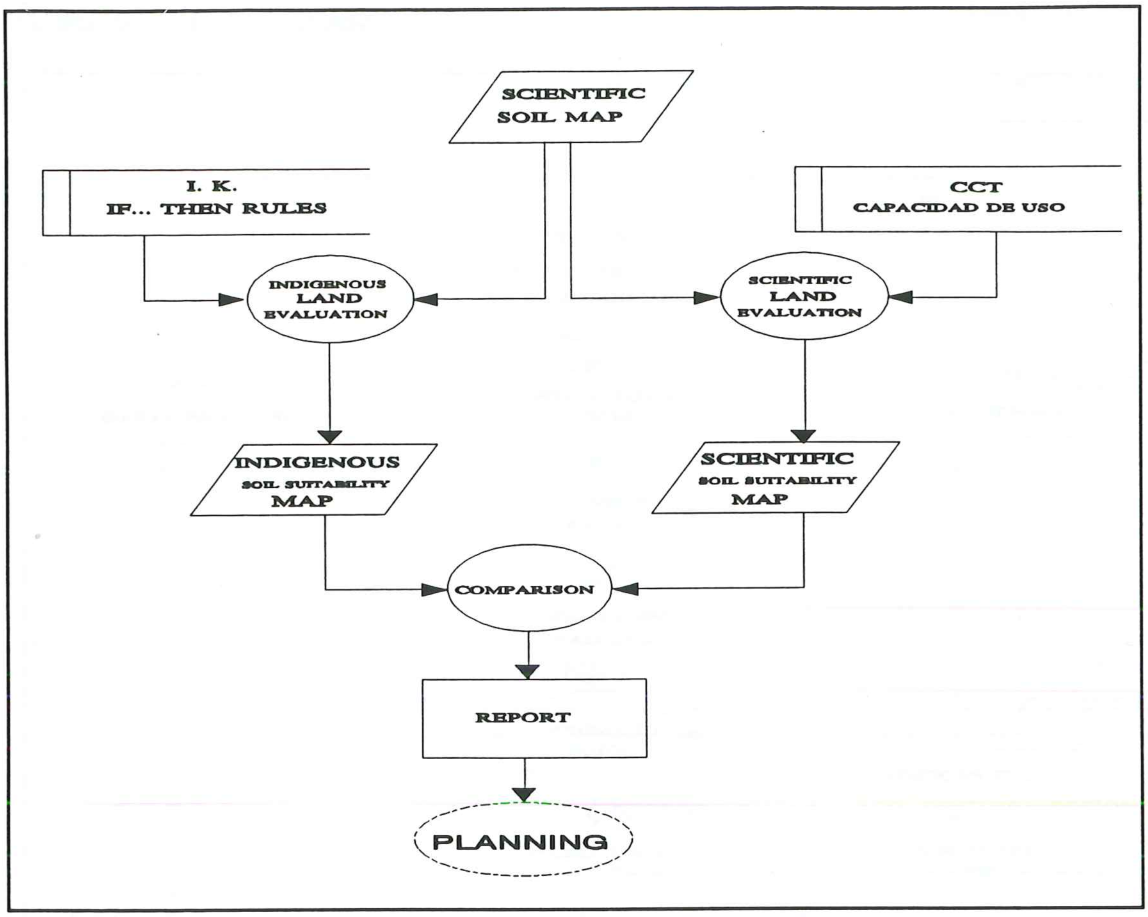


Fig.3 Revised Schema

reclassifying them according to the farmers' criteria. The process will be elaborated on in Chapter 6.

Considering the revised methodology in Figure 3, the basic data required are the scientific soil units (determined by the scientific classification system) and the IK rules collected from the farmers. As for administrative purposes, the parcel boundaries provided by the parcelmap and the list of farmer-owners are the most important information. Tables 3a and 3b on page 26 show the list of entities identified to satisfy the information requirements of the system to be built.

The data necessary to define and describe the soils of Neguev are available from the original soilmap made by S.de Bruin and W.Wielemaker at 1:20,000 scale-- just the right level of detail for the parcel-based study. Digitized maps, including the parcel map and the land use map were obtained from the Atlantic Zone Programme. The socio-economic data are culled from the *Encuesta General* prepared by Brink and Waaijenberg (1987) and from another report prepared by Veltman (1990). Delineation of the scientific soil units are straightforward digitizing (from the existing soilmap), while those of the indigenous soil units are defined by the farmers' criteria in classifying the soil-- the location on the landscape and the color (van Uffelen, 1990). These criteria are also used to create the knowledge base for IK rules. The process of creating a knowledge base is dealt with in the following chapter.

4 "I Think, Therefore I Am" : On Knowledge Bases

4.1 Introduction

Philosophers, writers, books and entire courses have attempted to answer the question "What is knowledge?" Recent advances in computer science have even made research and development in Artificial Intelligence (AI). As a result, computers (or computer programs) are now able to mimic specific areas of human reasoning. These are called expert systems and expert system shells-- computer programs that are able to utilize expert knowledge and perform inference procedures in solving problems that require expertise for their solution. These were made possible by capturing and integrating an expert's knowledge into such programs (McGraw and Harbison-Briggs, 1989).

Then, again what is this "knowledge" that has to be identified and captured? For the purposes of this study, knowledge is all that somebody knows about something-- a model of the world which can be created or modified by new information. Information is the transferrable knowledge-- a structured collection of data; and data is the result of direct observation of events-- the values of the attributes of objects (Molenaar, 1989; Teskey, 1989).

The correctness of data with respect to the real world can be objectively verified by comparison with repeated observations. But knowledge includes abstractions and generalizations which are typically less precise and cannot be easily objectively verified that's why we look for experts to provide them (Israel, 1986). This is because the knowledge of an expert includes heuristics (from the Greek word *heuriskein* or Eureka! meaning, "I have discovered"). Heuristics are rules, but mostly "rules-of-thumb" and may be difficult to make explicit. The expert simply knows it from his or her experience in his/her field of expertise (Furbee, 1990; Rich and Knight, 1991).

4.2 Knowledge-based Systems

Knowledge-based systems (KBS) utilize expert knowledge and are designed to reach the same conclusions that a human expert would be expected to reach if faced with a comparable problem. An expert is a person who possesses the expertise or the desired knowledge and the task of eliciting this knowledge is done by a "knowledge engineer" (McGraw and Harbison-Briggs, 1989).

Unlike ordinary computer programs, KBS attempts to solve problems in specific disciplines using reasoning based on combination of rules and definitions with facts to draw conclusions. The process relies heavily on theories of logical deduction developed by mathematicians and philosophers and adapted to particular applications by engineers, scientists, planners and managers across a wide range of disciplines (Kim, 1991).

A knowledge-based system has three basic components: the knowledge base, an inference engine and a user interface (McGraw and Harbison-Briggs, 1989; Robinson and Frank, 1987). The knowledge base contains facts and rules expressing an expert's heuristics for the topic or domain. Such domain specific facts and rules are separated from the procedural language (the inference engine) used for controlling program execution. This makes it much easier to encode and maintain facts and rules. The separation of knowledge and procedures of applying the knowledge is one of the main characteristics of expert systems.

The inference engine is made up of rules that are used to control the use of the rules in the knowledge base as to when and how specific problem-solving knowledge is used. PROLOG, LISP and SMALLTALK, are symbolic programming languages which are preferred by professional programmers as they offer great ease and flexibility. Any other programming language that is designed to apply and process the knowledge to solve a particular problem may be used. FORTRAN, C, PASCAL, etc. are some of the examples that are used to write original programs. Another option is to use commercially available expert system shells in order to free the programmer and/or knowledge engineer of the tediousness of writing original programming codes. The user interface like menus, displays, and prompts allows communication or interaction between the expert system and the end user.

4.3 Knowledge Acquisition

Knowledge acquisition refers to the transfer and transformation of problem-solving expertise from a knowledge source (whether a human expert or a document) to a program. This describes the dual process of extracting and translating expert knowledge or heuristics into rules (McGraw and Harbison-Briggs, 1989). Knowledge acquisition is the first step in building a knowledge-based system after defining what the system has to do. This may be repeated several times as the expert is consulted for verification. This is the phase that distinguishes knowledge-based systems from most conventional information systems. This is also the phase that can make or break the whole system because the resulting knowledge base depends on the success of acquiring the knowledge from the knowledge source. For these reasons, knowledge acquisition is the biggest bottleneck in the development of knowledge-based systems (Olson and Rueter, 1987).

Studies in expert systems have developed a number of ways to approach knowledge acquisition. There are direct methods like interviews, questionnaires, observations, and think alouds in which the expert explains what knowledge he/she uses (assuming that the expert can explain and express himself/herself) while the knowledge engineer simply takes note. Indirect methods like multi-dimensional scaling, repertory grid analysis or triad test allow the expert to perform intermediary sample tasks in order for the knowledge engineer to recognize hierarchies and structures of knowledge and to articulate what is believed to be occurring in the session. Sometimes, the knowledge engineer becomes an expert him/herself, relying on introspection to articulate the requisite knowledge. A combination of different methods or approaches to knowledge acquisition is usually done to effectively capture and represent the knowledge of an expert (McGraw, 1989; Olson and Rueter, 1987).

4.4 Knowledge Base and Knowledge Representation

The result of an exhaustive knowledge acquisition is a knowledge base. It differs from a database in that a database is a collection of data representing facts while a knowledge base contains information at a higher level of abstraction. It differs from conventional computer programs and database management systems because of the treatment of facts and rules as "data" in the knowledge base. In conventional computer programs rules are imbedded in the program itself. Hence, it is difficult to separate the rules from the procedural or control mechanism of program execution (Israel, 1986).

All the factual and empirical knowledge of an expert (or experts) which are important for problem-solving in a specific area of application are (in principle) contained in a knowledge base. It is a representation of the application domain knowledge or a "slice of reality". It is made up of facts, rules and procedures relevant to an application domain. There are three main methods of representing formalized knowledge: semantic networks, which is a network of relation links and object nodes; frames, which is a modular organization of knowledge per topic; and production rules, which are simple condition-conclusion statements (Robinson and Frank, 1987).

This study will make use of production rules to represent expert knowledge. These rules are in the form of IF <premise> THEN <conclusion and/or action>. In the premise part, questions are asked about the logical links between the characteristics of the objects. In the conclusion part, new facts and characteristics are added to the knowledge base and/or actions are executed (McGraw and Harbison-Briggs, 1989; Robinson and Frank, 1987; Olson and Rueter).

4.5 Indigenous Knowledge as a Knowledge Base

Farmers don't usually know any of governing equations in the field as they don't need exact numerical measures. They need only qualitative descriptions. Yet they are perfectly capable of establishing empirical relationships of objects and phenomena around them. They develop rules-of-thumb as a result of their day-to-day experience with their land. These rules are passed on for generations and are refined into a system of understanding of the world around them-- a "world view" (Pawluk, 1992). They can also be represented in the form of IF <premise> THEN <conclusion and/or action>. For example:

IF <the soil's color is very red> THEN <it is not fertile>.

or sometimes, there are more conditions for the premise:

IF <the soil's color is very red>
AND <is located on high ground>
THEN <it is good for pineapple>.

The challenge in building an indigenous knowledge base is in understanding and reasoning with abstract, qualitative observations. The common-sense knowledge and heuristics applied by farmers is part of what must be modelled. Why bother? Because a formal model of the local farmer's "world view" would result. Then a comparison of the local model with the Western models would be possible. Foreign advisors, as well as local officials could then better understand the folk-constructed world of reality within which decisions are made and behavioral patterns produced (Warren, 1975).

4.6 Caveats

Experts-- folk and otherwise --seldom realize how much they know, much less how they structure that knowledge cognitively. They frequently believe that they make their judgments on the basis of relatively small number of rules, when in fact an expert system is likely to require hundreds of rules to model a rather minor domain of behavior (Furbee, 1989). They are also not usually good at explaining what they know. They simply know how to do or decide under different circumstances, but they may not be able to articulate about it. And more often than not, experts are poor with percentages and factors. They are influenced by the more recent developments or experiences in their field of expertise and therefore their judgment on statistics can be unreliable. Besides, experts can make mistakes, too. For these reasons, the knowledge engineer has to take care and be creative in extracting knowledge from experts. And therefore, a KBS cannot replace an expert because it cannot contain as much knowledge. A KBS is made just to assist in solving a problem that requires some expertise and make the knowledge easily available to other people (pers.con. Osinga, 1994).

4.7 KBS and GIS

There are a number of areas where GIS are expected to benefit from the application of KBS technology-- geographic data input, geographic database management, cartographic designs and geographic decision support systems. Many systems appear to have relied on the cartographic and geographic knowledge resident in journals and textbooks, thus avoiding the time, effort and expense of extracting knowledge from human cartographers, surveyors, geographers and regional scientists (Robinson and Frank, 1987), much less from local resource managers like farmers and tribal elders.

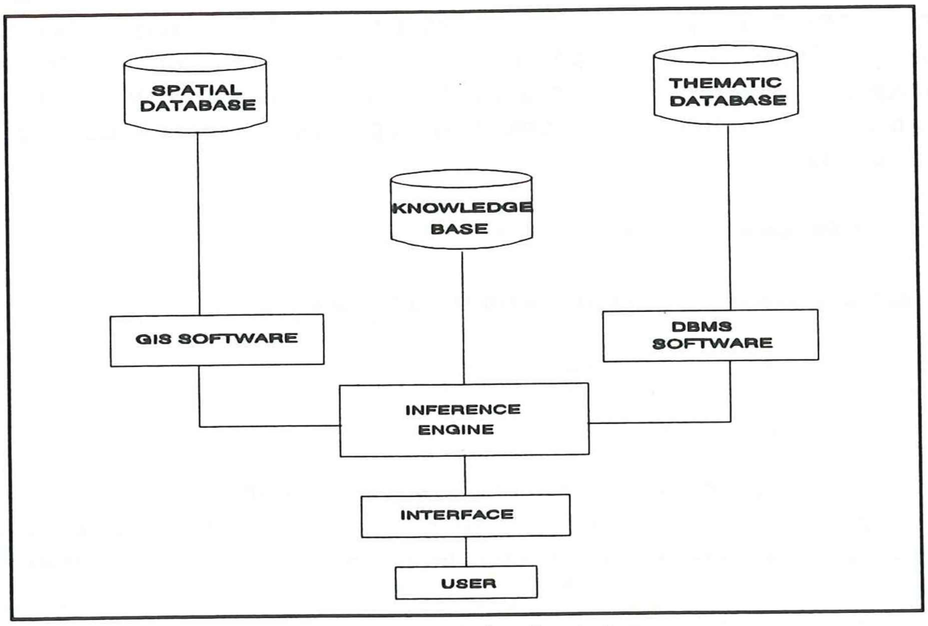


Fig.4 KBS-GIS Configuration

This study will use KBS in geographic data input-- that of incorporating IK of soils in a GIS design and the configuration would be as in Figure 4. The "inference engine" would be a combination of query language statements and macro language programs. This will be facilitated by a GIS and a database management software link-up to manipulate the databases. The interface are batch files that activate a pull-down menu system.

GIS has four major components: data acquisition component, database management component, data manipulation and analysis component, and data display component. In each of these, KBS can play a role for creating an intelligent GIS because of the knowledge involved (Burrough, 1992). Incorporating IK of soils into a spatial database concerns the process of data acquisition. This is the surveying stage in which terrain features are identified, geometric descriptions like shape, size and position are made, and thematic attributes such as land-use, soil type and elevation are evaluated. These data are often unusable in their original form and have to undergo pre-processing so that relevant information are extracted for further processing. The final product can be presented in graphic format (maps), alpha-numeric (reports) or a database (disks). Figure 5 shows the outline of information flow in GIS and the influencing factors such as the available technology and the methods used, not to mention the "world view" of the surveyor.

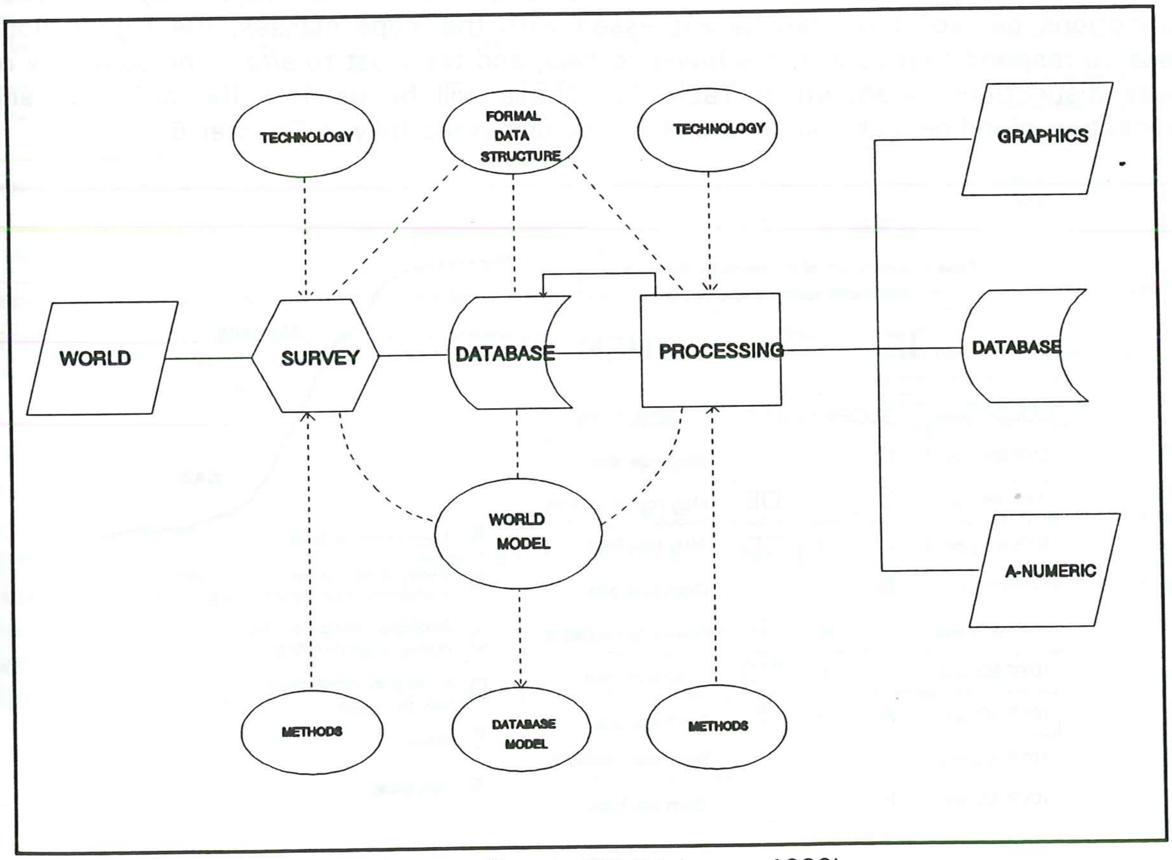


Fig.5 Data Flow in GIS (Molenaar, 1989)

4.8 Farmers' Knowledge from Neguev

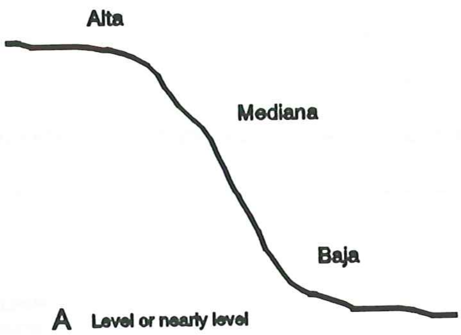
Traditional data sources for GIS are existing maps, field surveys, aerial photographs, satellite images, census data and other socio-economic data. The methodology presented in Chapter 3 would make possible the integration of traditional data sources and local knowledge. And since the local source are local farmers who possess "world views" that are different from Western views, the first task is to learn that one is talking about the issue in terms congruent with those of the persons under study. This concerns asking the right questions from the point of view of the farmers and defining objects according to their criteria (Furbee, 1989). The summary of IK rules derived from van Uffelen's interview with the farmers of Neguev is shown in Table 2 and forms part of the indigenous knowledge base. The types of red soils identified by the farmers were reported to correspond with the scientific soil classification as follows (van Uffelen, 1990):

Farmers' red soils		Scientists' soil units	Color Code
Tierra bermeja (brown soil)	=	Milano soils (Mi)	10YR 4/3
Tierra colorada (red soil)	=	Neguev soils (Ne)	10YR 3/3
Tierra muy roja (very red soil)	=	Silencio soils (Si)	5YR 4/4

The color code is derived by soil scientists from the Munsell Soil Color chart (see Munsell, 1969). While red soils are generally found in high places, each type is distinguished further by their relative location in the terrain. Hence, the local names such as *tierra bermeja alta*, *tierra bermeja mediana*, and *tierra bermeja baja*, etc. Such distinctions per soil type can be expressed with the slope classes: the higher slope class correspond to *mediana*; the lower to *baja*; and the least to *alta*. The summary of these distinctions is shown in Table 1. These will be used in the definition and delineation of indigenous soil units and will be discussed fully in Chapter 6.

Table 1. Summmary of IK Criteria for Red Soils Articulated with Scientific Soil Units

IF					THEN
COLOR	and/or	SLOPE CLASS			SOILTYPE
5YR 4/4	and	C			Muy roja alta
5YR 4/4	and	E	or	DE	Muy roja mediana
5YR 4/4	and	D	or	CD	Muy roja baja
10YR 3/3	and	B			Colorada alta
10YR 3/3	and	D	or	E	Colorada mediana
10YR 3/3	and	C	or	CD	Colorada baja
10YR 4/3	and	A	or	B	Bermeja alta
10YR 4/3	and	F			Bermeja mediana
10YR 4/3	and	E			Bermeja baja



A Level or nearly level

B Gently sloping or very gently sloping; Undulating or gently undulating

C Sloping or strongly sloping; Rolling or gently rolling

D Moderately steep (single); Hilly (complex)

E Steep

F Very steep

Table 2. Summary of IK Rules (*from van Uffelen, 1990*)

IF			THEN					
COLOR	**and/or**	**LOCATION**	**CROP**	**FERTILITY**	**HUMIDITY**	**CAPADURA**	**FLOOD**	**NAME**
Black	and	high	1	high	high	No	1	Tierra negra alta
Black	and	middle	1	high	high	No	2	Tierra negra mediana
Black	and	low	1	high	high	No	3	Tierra negra baja
Black	and	swampland	***	***	***	***	4	Suamposa
Brown	and	high	2	moderate	low	Yes	0	Tierra bermeja alta
Brown	and	middle	2	moderate	moderate	Yes	0	Tierra bermeja mediana
Brown	and	low	2	moderate	moderate	No	0	Tierra bermeja baja
Red	and	high	2	low	low	Yes	0	Tierra colorada alta
Red	and	middle	2	low	low	No	0	Tierra colorada mediana
Red	and	low	2	low	moderate	No	0	Tierra colorada baja
Very Red	and	high	3	low	low	Yes	0	Tierra muy roja alta
Very Red	and	middle	3	low	low	No	0	Tierra muy roja mediana
Very Red	and	low	3	low	moderate	No	0	Tierra muy roja baja

<u>Cropgroup</u>
1: maize, beans, rice, banana, platano pasture
2: palmito, cocoa
3: pineapple, chili

<u>Flooding</u>
0: No problem
1: Some days after some years
2: Some days every year
3: Some weeks every year; may occur early in May
4: Always in water

The knowledge acquisition process undertaken may not have been as exhaustive as it should be (because of differences in the objectives of the studies), but additional information were obtained from the soil scientists (W.Wielemaker, S.de Bruin) that worked in the area. This was done especially in the case of articulating the black soils identified by the farmers with those of the scientific soilmap units (see Appendix C). These translations will be needed later in building the spatial database of soils.

5 "Through The Looking Glass": Database Modelling

5.1 Introduction

The term **model** is used to mean a set of relationships or information about the real world (Aronoff, 1989). **Information** is a structured collection of data. And a **data model** provides a formal means for representing information and a formal means of manipulating such a representation (Peuquet, 1984). **Data modelling** is the process of abstraction and documentation using a data model. It is used to describe the structure of a database-- the data types, relationships and constraints on the data (Elmasri and Navathe, 1989).

Conceptual modelling is the first step in converting features in the terrain into a database. It involves the recognition, abstraction and definition of spatial features of the real world. The complex world is simplified as relevant information on selected features are recorded. This selection process creates a conceptual model of the real world-- an understanding of what it is and how it behaves (Aronoff, 1989). The output of this stage is the "universe of discourse" or the area of interest or concern (Avison and Wood-Harper, 1990).

Logical modelling, on the other hand, constitutes the core of database design. It includes definite decisions about the aggregation level of analysis and the descriptors of the identified objects for a well-defined "universe of discourse". This process is constrained by the chosen database management system to be used-- sometimes called the data structure (Elmasri and Navathe, 1989). The final Entity-Relationship (ER) Diagram (Chen, 1976) is drawn. Skeleton tables are also created to facilitate the creation of the final tables for implementation.

The last step is physical modelling. It describes how the database is organized for physical storage and accessed in such a way that the logical requirements can be met. At this point, the conceptual/logical data model is translated into formats to be accommodated by the particular software to be used (Elmasri and Navathe, 1989). This chapter discusses the development of the database design for the case study.

5.2 Conceptual Modelling

There are two ways to approach conceptual modelling (see Fig.6): the first is the field approach wherein a terrain feature is represented by a position (e.g. *x,y*) and is described by thematic attributes like height, soiltype, etc. The second, and the one that will be used in this study, is the object approach-- an object is identified and then thematic data as well as geometric data like shape, size, and position are assigned to it as descriptors. In this way, an entity of whatever complexity can be represented by exactly one object with thematic and geometric attributes (Molenaar, 1990; Egenhofer and Frank, 1989).

5.2.1 Object-oriented Modelling

There are several types of relationships the terrain objects can be identified with. And these are the four basic concepts of object abstraction that object-oriented

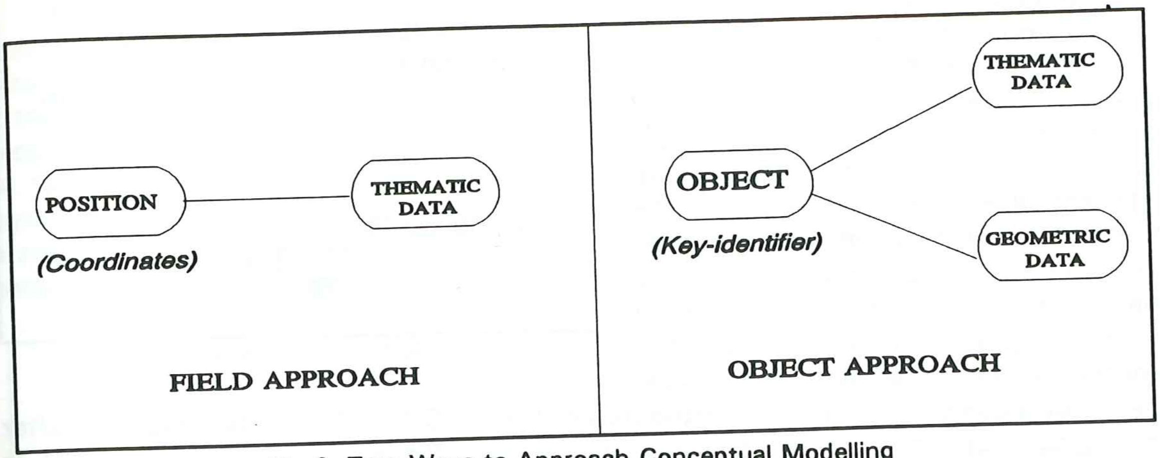


Fig.6 Two Ways to Approach Conceptual Modelling

approach in modelling is built upon: classification, generalization, association and aggregation (Molenaar, 1993; Egenhofer and Frank, 1989). These concepts were used in defining the relevant objects for this study as shown in the following examples:

Object classification (Fig.7) is the mapping of several objects to a common class. Each class has a class label and a class attribute list. Every object is an "instance of" a class and has its own individuality reflected by its attribute values. For example, in this study, the class SCISOIL UNIT has the attribute list [SciSoil-ID, Slope class, Color]. A single instance such as SciSoil 10 with a Slope of C and color 10YR 4/4/ is an "instance of" the class SCISOIL UNIT.

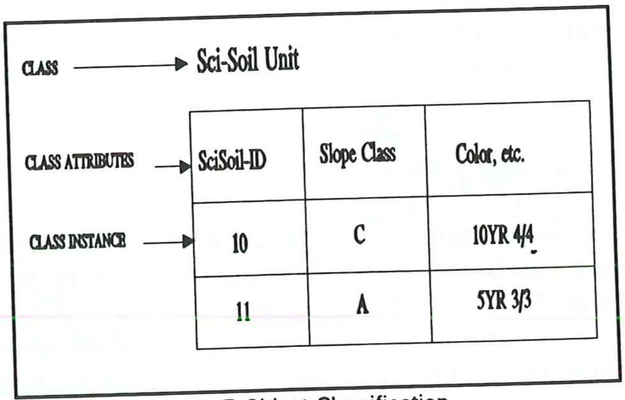


Fig.7 Object Classification

Object generalization (Fig.8) groups several classes of objects, which have some properties in common, to a more general superclass. Generalization may have an arbitrary number of levels (classification hierarchies). These hierarchies define vertical relationships between objects and object classes. In upward direction these are called "is a" relationships. In downward direction they represent specialization steps until the lowest level of objects or terrain features themselves. For example, the class RED SOIL and the class BLACK SOIL can both belong to a superclass SOIL. It is important to note that superclass or subclass are abstractions for the same object and do not describe two different objects. The Very Red Soil, for example, is at the same time an instance of the class RED SOIL and the superclass SOIL.

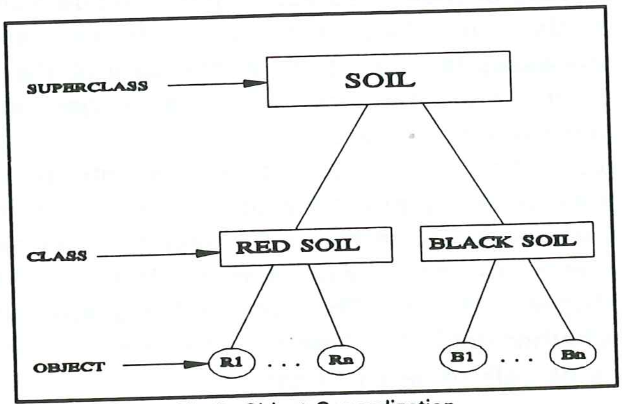


Fig.8 Object Generalization

Object aggregation (Fig.9) defines composite objects which are built from elementary objects. In upward direction this is called "part of" relationship and in downward direction this is called "consist of" relationship. When considering the aggregate, details of the constituent objects are suppressed. The fact that elementary objects can be aggregated into composite objects implies that their attribute values may also be aggregated and that the geometry is also changed (e.g., size, shape). For example, if one of the attributes of the SCISOIL UNIT is its area, then after reclassifying according to IK-criteria and aggregating, the area of the IKSOIL UNIT is easily calculated. However, when disaggregation is done, it will be difficult to come back to the attribute values of the constituent object. Inheritance of attribute values is only in the upward direction.

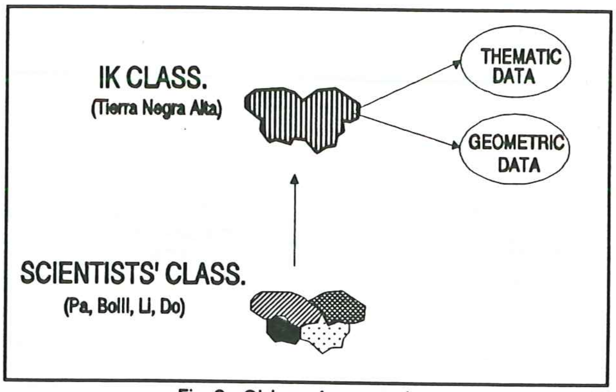


Fig.9 Object Aggregation

Object association are just sets of objects which do have some characteristics in common or a functional or administrative relationship. This is loosely defined and found by means of search operations based on attribute values and/or geometric relationships. For example, set of parcels with area < 3 hectares; or set of parcels along the river.

5.2.2 Semantic Aspects

At this point, the case at hand seems to look better than it is. But problems exist with the definitions of the objects to be modelled owing to differences in the "world views" of the farmers and the soil scientists. Both may be describing the same object-- the soils of Neguev, but their purposes differ hence the difference in the aspects to be described. Soil scientists identify landscape facets. A facet is a homogeneous zone where soil forming processes are considered the same (using profiles of their observation pits). Aside from the history (genesis) of the soil, its future is also described in terms of fertility and sensitivity or suceptibility to occuring processes (e.g., erosion) at the time of the survey (pers.con. de Bruin, 1994). On the other hand, farmers have a longer observation period (maybe passed on for generations) and are constantly monitoring fertility (as their prime concern) in terms of production yields, but not necessarily aware of ongoing soil processes. In Neguev, they generally associate observable characteristics like color and landscape formations with the soil's productivity (based on the performance of the crops)-- that soils in high places are generally red and less fertile than soils in the lower areas (van Uffelen, 1990; Mucher, 1991). Their object for abstraction is implied as the landscape formation-- whether upland or lowland or midland. The IDA personnel may be looking at the area as parcels to be administered.

This explains why in a database, object descriptions have their own particular context (Molenaar, 1993; Teskey, 1989) which depends on the users' purpose for such a database. And if the aim is to provide relevant information to another user (who is not the creator of the database) then there has to be a translation of definitions or meanings for an effective conveyance. In Chapter 4, IK criteria for soil classification were articulated with those of the scientists in order to come up with an approximation of the farmers' "world view". IK's landscape formation was translated into slope classes of the scientists. The result is a translation of the object definition into that of the farmers-- a context transformation (see Fig.10). And within the new context, aggregation is carried out. After aggregation, the IK-soil units inherit the attributes of its aggregate Sci-soil units. It can then be evaluated using IK-rules to assess any other characteristics according to farmers' criteria.

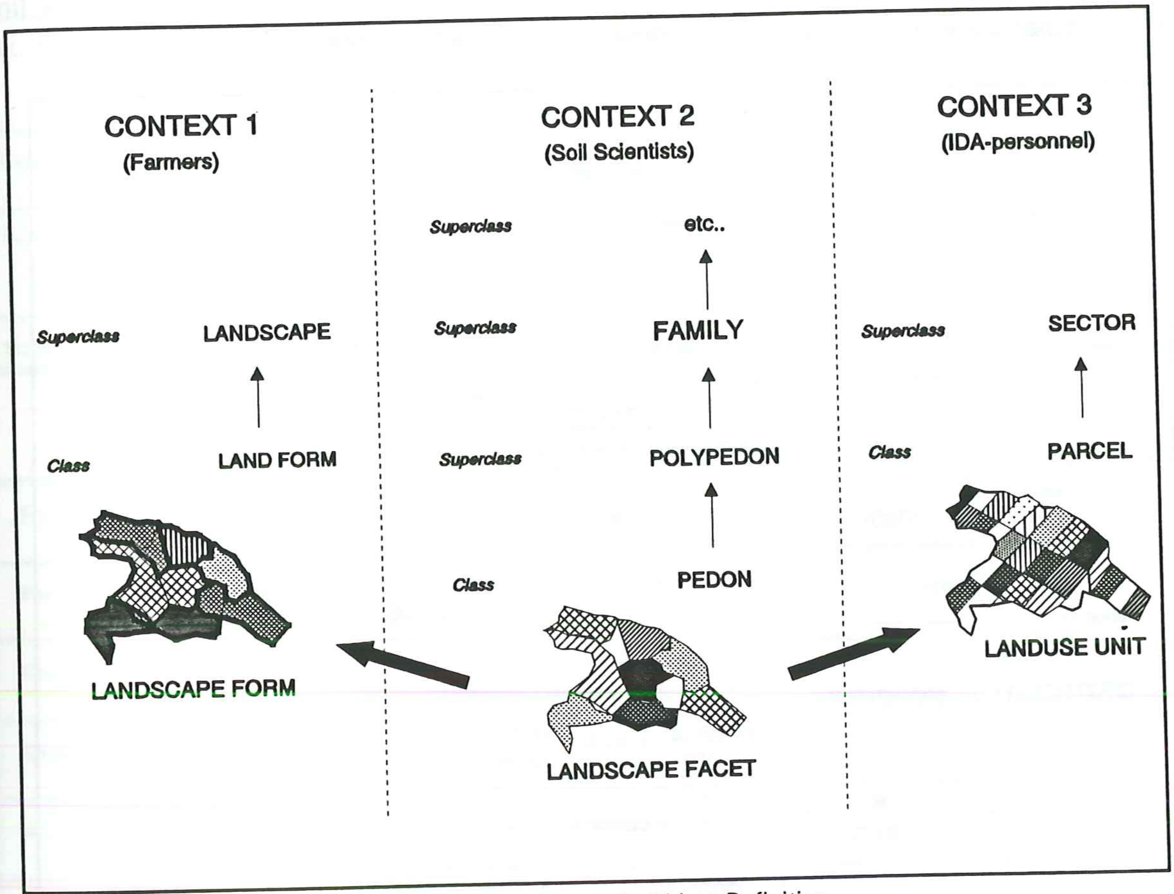


Fig.10 Users' Context in Object Definition

5.2.3 Modelling Spatial Data

Once the object of abstration is defined, there are two options in modelling: either to represent its geometric component as a raster model or a vector model (Aronoff, 1989; Peuquet, 1984; Burrough, 1986). In the raster model, the space is regularly subdivided into cells (usually square in shape) and the location of geographic objects or condition is specified by the row and column position of the cells they occupy. In the vector model, which is the one used for this study, objects or conditions in the real world are represented by point, line and polygon features. Spatial information is represented using homogeneous units of these elementary features. Vector data models can either be a spaghetti data model-- wherein features are represented by a file of geographic coordinate strings (x,y) with no inherent structure; or a topological data model-- where the basic logical entity is the arc. An arc is a series of points that start and end at a node. A node is an intersection point of two or more arcs. It could also be the end of a dangling arc (as a dead end of a road). A polygon is an area enclosed by arcs. Point objects are represented by their point coordinates. Topology among these features are defined and stored in the model. Topology is the mathematical method used to define spatial relationships (e.g., contiguity, connectivity) among these geometric primitives (Aronoff, 1989). This study will use the topological vector model which allows spatial analysis using topological data and representation of objects by its geometric primitive (point, line, polygon).

Figure 11 shows the conceptual model showing the relevant spatial features and how they will be represented in the database in terms of geometric primitives (in this particular case, polygons). The supporting explanations are found in Table 3a.

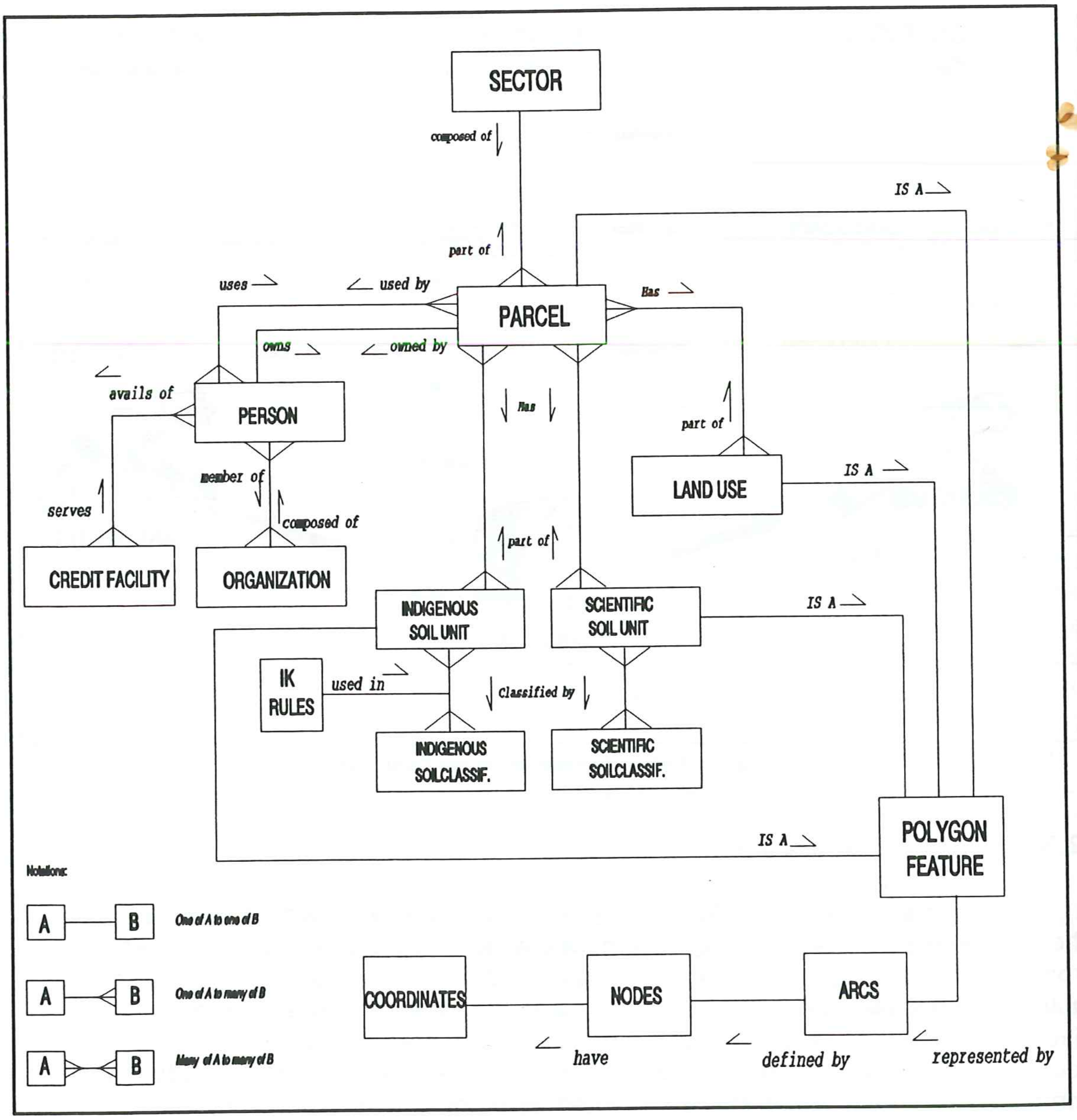


Fig.11 The Conceptual Model

Table 3a. Identified Spatial Features

Feature	Definition	Geometric Feature	Thematic Attributes
Sci. Soil unit	Smallest unit of soil area delineated by scientists to be of homogeneous properties	Polygon	Code, Name, pH, texture, slope class, etc.
IK. Soil Unit	Smallest unit of soil area delineated by Neguev farmers to be of homogeneous properties	Polygon	Code, Name, color,fertility, texture, etc.
Land use	Purpose for which a piece of land is used; also the crop planted on it	Polygon	Code, area
Parcel	Subdivided land for individual ownership	Polygon	Parcel No., area, perimeter
Sector	Local administrative boundary	Polygon	Name

Table 3b. Additional (non-spatial) Entities for the Model

Entity	Definition	Thematic Attributes
Person	Person living within the Neguev settlement	Name, sex, origin, etc.
Credit Facility	Banks and other financial institutions lending money to farmers	Name, type
Organization	Grouping of persons formed for a purpose	Name, type
IK. Soil Classification	Neguev farmers' system of differentiating soils	Name
Sci. Soil Classification	Scientists' system of differentiating soils	Name

5.3 Logical Modelling

The conceptual design formulated in the preceding section is translated into a specific database structure that can be accommodated by a specific software to be used. This also defines how the data will be stored. There are three database structures to choose from: hierarchical, network or relational. A fourth type, the object-oriented structure, is at its development stage at the moment. This study would utilize the relational database structure in building the database because of the availability of a corresponding GIS software (Arc/Info) which is capable of supporting a relational and topological model.

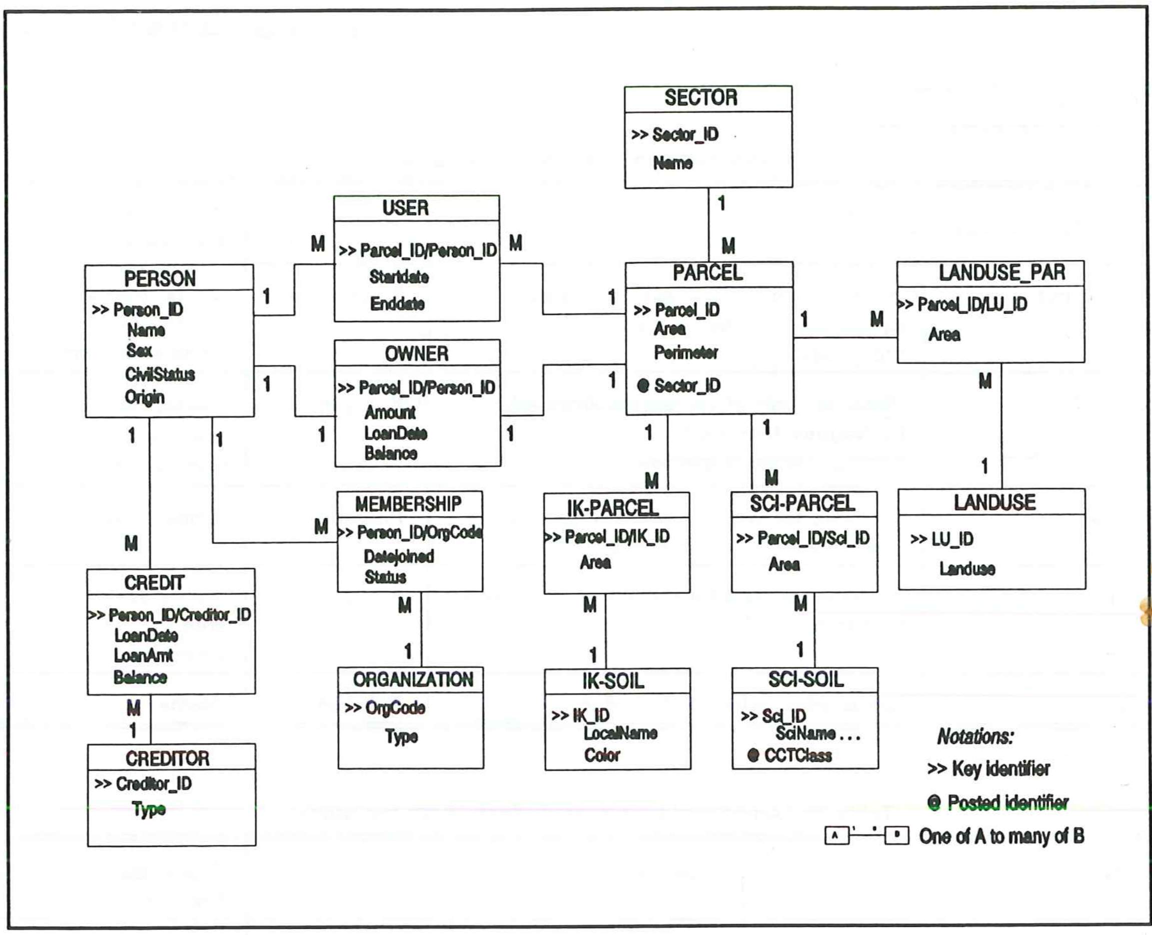


Fig.12 The Logical Data Model

5.3.1 The Relational Structure

In a relational database structure, the data are stored as a collection of values in the form of records, called tuples. Each tuple represents a fact (i.e., a set of permanently related values). These tuples are grouped together in two-dimensional tables. The table represents the relationships among all the attributes it contains, and so it is often termed a relation (Aronoff, 1989). Each identified entity is represented as a table with all its attributes appearing as columns and each row representing a single record.

5.3.2 Identifying Entities and Relationships

An entity is anything you want to keep records about (e.g., parcel, person, land use) which are identified during the conceptual modelling stage. A relationship is an association among entities such as "persons own parcels". It is impossible (and, perhaps unnecessary) to record every potentially available entities and relationships (and the information concerning them) which are to enter into the design of a database (Chen, 1976). This process of forming entities and relationships can be very subjective, that's why a good understanding of the problem situation is required in order to decide which entities and relationships are important (Avison and Wood-Harper, 1990). An exact description is also not necessary.. it is enough that we construct models of the environment which satisfy the information requirements (Bregt, 1993).

The identified entities are deemed enough for the requirements of the system. There were no available data on households, and therefore this entity is not considered. The conceptual model is now transformed into a logical model (see Fig.12) that which take to account the data structure to be used. All necessary data items are grouped together in a table. Normalization (see Date, 1990) is used to ensure a nonredundant model of data. Normalization is useful in analyzing data, removing unwanted redundancies and forming logical records prior to physical database design. The process of normalization and its drawbacks (for one, it could lead to the creation of too many tables that reduces computing performance) are described in detail in Date (1990). In this case, additional tables (of the relationships) are constructed to normalize the tables and avoid redundancies. Also, since only one scientific classification system (USDA) and one IK classification system are being used (Neguev farmers'), no table was created for SCI-SOILCLASSIF and IK-SOILCLASSIF . Instead, the CCT-class was posted in the Sci-SOIL unit table.

5.3.3 Some Modelling Guides

The types of relationships (whether one-to-one, one-to-many, many-to-many) are determined by enterprise rules (see Howe, 1989). These are rules that govern the area of concern. In this case, the enterprise rules are as follows:

1. A SECTOR is composed of many PARCELs.
2. A PARCEL must belong to a SECTOR.
3. A PARCEL is owned by one PERSON.
4. A PERSON may own one PARCEL.
5. A PARCEL is used by at least one PERSON.
6. A PERSON may use many PARCELS.
7. A PARCEL is used for at least one LANDUSE.
8. A LANDUSE may be found in a PARCEL.
9. A PARCEL has at least one IK-SOIL UNIT.
10. An IK-SOIL UNIT may be found in a PARCEL.
11. A PARCEL has at least one SCI-SOIL UNIT.
12. A SCI-SOIL UNIT may be found in a PARCEL.
13. A PERSON may be a member of an ORGANIZATION.
14. An ORGANIZATION is composed of many PERSONs.
15. A PERSON may avail of a CREDIT FACILITY.
16. A CREDIT FACILITY may serve many PERSONs.

Skeleton tables are used to represent the structure of a real table. They indicate the elements to describe each entity as well as the key identifier (underlined) and posted identifier (broken line) if there's any. The following are the skeleton tables used to guide the creation of the physical model:

SECTOR	(Sector_ID, Name)
PARCEL	(Parcel_ID, Area, Perimeter, Sector_ID)
LANDUSE	(LU_ID, Landuse)
LANDUSE_PAR	(Parcel_ID, LU_ID, Area)
IK-SOIL	(IK_ID, LocalName, Color, Fertility,...)

SCI-SOIL	(Sci-ID, SciName, Color, ..., CCTClass)
IK-PARCEL	(Parcel_ID, IK_ID, Area)
SCI-PARCEL	(Parcel_ID, Sci_ID, Area)
PERSON	(Person_ID, Sex, Origin, ...)
USERS	(Parcel_ID, Person_ID, StartDate, EndDate)
OWNER	(Parcel_ID, Person_ID, DateAcquired, AcqMode..)
ORGANIZATION	(OrgCode, Type,..)
MEMBERSHIP	(Person_ID, OrgCode, DateJoined, Status..)
CREDITOR	(Creditor_ID, Name, Type..)
CREDIT	(Person_ID, Creditor_ID, LoanDate, LoanAmt..)

5.4 Physical Modelling

A knowledge of how the data items are organized and stored in physical storage is not required for users. However, it is often helpful for users to have some knowledge of physical data storage in order to respond to constraints in the logical representation that may be imposed by physical storage. For example, denormalization of tables are sometimes necessary so as not to sacrifice the efficiency of the software when tables get very large (Elmasri and Navathe, 1989). And more importantly, in handling spatial databases care should be taken so as not to disturb the stored topological data or this would ruin the whole system.

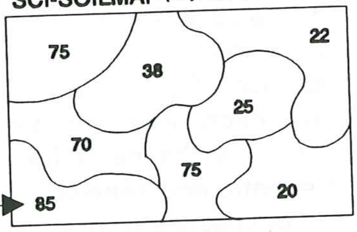
5.4.1 Data Structuring with Arc/Info

Arc/Info (v.6.1.1) is a GIS software that is capable of accommodating both vector and raster data models using a relational database structure. It is composed of two main systems-- the ARC part, which maintains the geometric data; and the INFO part, which stores the thematic data in tabular format which the user can maintain.

In the ARC, generating data is done mainly by digitizing maps. Each created map layer is kept in a storage called coverage. Digitized objects are translated into Arc/Info's measuring primitives: arcs, nodes, polygons, and label points. Arc/Info automatically creates the INFO data file at the background of these graphics when topology is established for a coverage. These are the polygon attribute table (CoverageName.PAT), arc attribute table (CoverageName.AAT), point attribute table (CoverageName.PAT) and node attribute table (CoverageName.NAT) as illustrated in Figure 13. They are a special kind of INFO tables that contain certain attribute information about coverage features which are automatically created in a specific order. Relations among these tables are ensured by two key-identifiers: one assigned by the software as internal pointer (#), and another assigned by the user (user-ID). Once the topology is corrected (by the command BUILD or CLEAN) and the user-IDs are assigned and verified, additional thematic data can be added (using the TABLES module) and related to these feature attribute tables (using the RELATE command). This allows the linking of attribute data with the geometry of the objects. The use of these commands will be dealt with in the next chapter.

INFO is a DBMS program in Arc/Info that controls the thematic and topologic files in tables. These tables are legible and can be edited (e.g., adding another column;

SCI-SOILMAP *(Object/Object Identifier)*



SCISOIL *(Thematic Data)*

SCISOIL-ID	SlopeClass	pH-Class	Drainage-Class	CCT-Class
70	2	1	3	3
75	2	3	3	3
80	2	2	3	3
85	2	6	2	2

(Geometric Data)

SCISOIL.PAT

RECORD	AREA	PERIMETER	SCISOIL#	SCISOIL-ID
1	-56715328.17651	60837.65806	1	0
2	7108.57129	382.51609	2	42
3	832801.19141	9052.92949	3	85
4	40963.18848	942.49708	4	75
5	33404.10596	896.65435	5	42

SCISOIL.AAT

RECORD	FNode#	TNode#	LPoly#	RPoly#	Length	SCISOIL#	SCISOIL-ID
1	1	2	2	3	167.56947	1	39
2	3	4	4	3	642.78455	2	38
3	7	6	3	5	726.39330	3	37
4	14	10	3	13	173.87948	4	43
5	15	11	9	6	62.01356	5	21

SCISOIL.NAT

RECORD	Arc#	SCISOIL#	SCISOIL-ID
1	819	1	1000
2	819	2	1005
3	820	3	1010
4	821	4	1019
5	822	5	1001

Fig.13 Physical Database Model for Spatial Entities in Arc/Info

assigning the user-ID) and queried by the user. But here is where care must be taken so as not to disturb the standard columns created by Arc/Info (the Area, Perimeter, internal-ID# for the .PAT; FNode#, TNode#, LPoly#, RPoly#, Length, internal-ID# for the .AAT; and the Arc#, internal-ID# for the .NAT). It is also wise to add another column which is identical to the user-ID column (an alias-column) to be used especially after performing several overlay operations. This is because of some imperfections in the software (pers.con. J.Stuiver, 1994). For polygons, each record in the .PAT actually represents the label point of the polygon. This explains the need for only one label point per polygon. By means of the internal-ID#, thematic and topologic data are linked to the geometric data in ARC because the tables are accessible from other modules.

5.4.2 Table Links for Spatial Objects

The spatial objects or entities identified in the preceding chapter are represented in the Arc/Info environment as exemplified in Figure 13 for the case of the scientific soil unit. In this case, a soil unit is linked to its thematic data using its key-identifier 85. The same key-identifier is used to link the soil unit to its geometric data (at least directly to the polygon attribute table or PAT). Topological data can be traced using the hash-attribute (#) of each feature attribute table. This is the internal identifier assigned by the software to keep tract of the relationships among the features. In this example, the Scisoil.PAT gives the Scisoil# (3) of the polygon defined by Scisoil-ID 85.

The arcs that define polygon 3 can then be traced using the topologic information from Scisoil.AAT. These are the arcs where polygon 3 appears in the LPoly# column (left polygon number) or RPoly# column (right polygon number). Furthermore, the nodes that define these arcs are given in the columns FNode# (from node number) and TNode# (to node number). More information about the individual nodes can be found in the node attribute table (Scisoil.NAT).

5.4.3 Table Links for Non-spatial Objects

The non-spatial objects or entities in the database are also interlinked using the key-identifiers and posted-identifiers earlier defined in the skeleton tables. INFO stores these tables and keeps tract of the relationship as shown in Figure 14. With Arc/Info's TABLES module, posted-identifiers can be added as shown in the table PARCEL where another column for Sector_Name is added. In the figure, tables with thick borders have corresponding map layers and are created in the ARC module (by digitizing or overlaying procedures), while the rest may be done using the TABLES module or other database management systems (DBMS) softwares.

Connecting with other DBMS softwares like ORACLE, INFORMIX, RDBMS and INGRESS is possible in Arc/Info. That means a database may be created in any of these softwares and later imported into Arc/Info environment (or the other way around--- the tables created in Arc/Info may be exported to other DBMS-softwares). This link-up makes the system more powerful in terms of data manipulation because of enhanced capabilities in querying. For this case, ORACLE is the available software. It is a relational DBMS that uses Standard Query Language (SQL) to retrieve/manipulate data from created databases. More discussions on how these are done will follow in the next chapter on implementation of the prototype system.

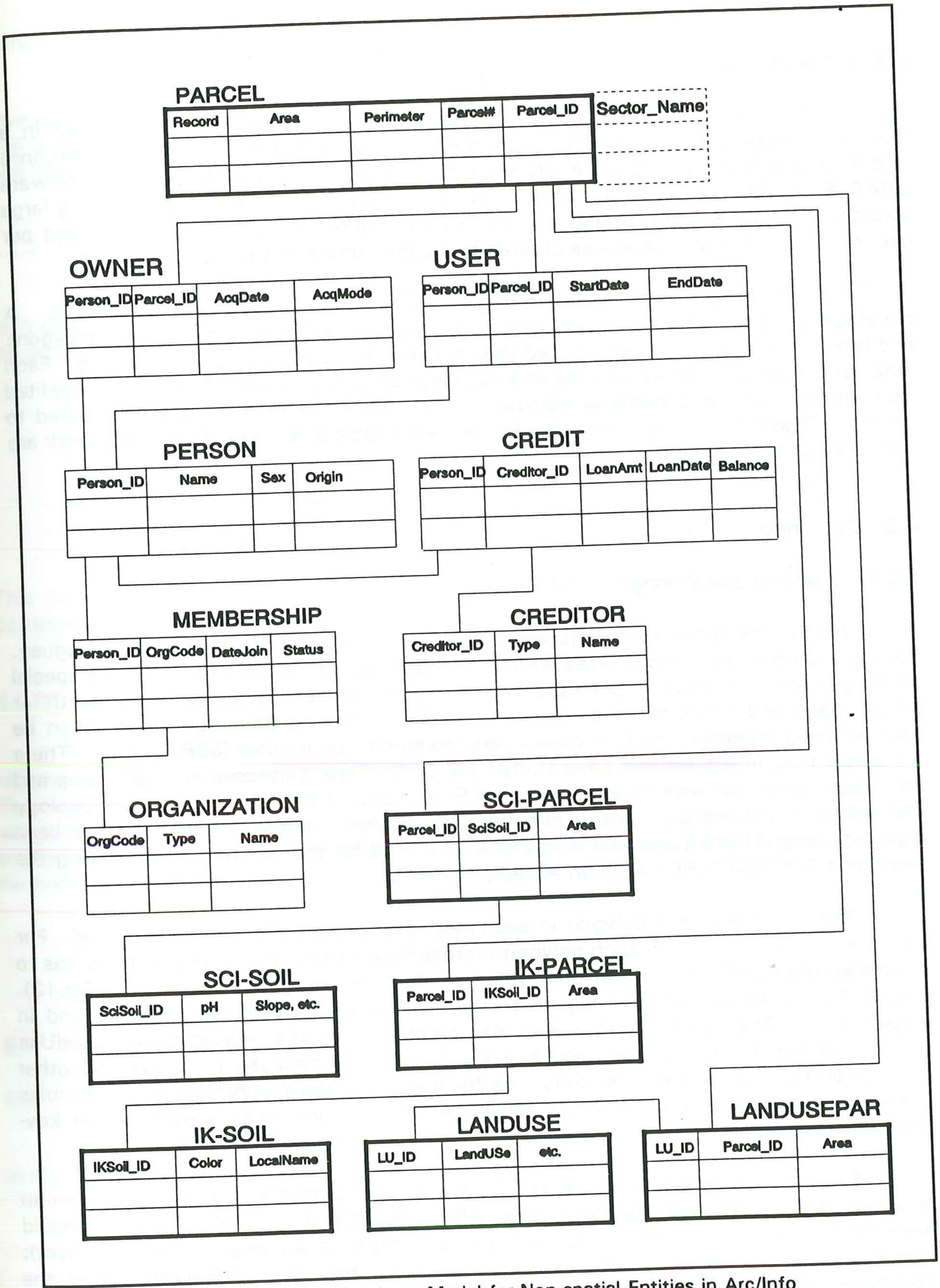


Fig.14 Physical Database Model for Non-spatial Entities in Arc/Info

6 "As You Like It" : System Implementation

6.1 Introduction

The physical database model of the previous chapter was implemented in a prototype system using the data from the study area and a GIS software Arc/Info (v.6.1.1) which is running on a UNIX platform and connects with a DBMS software ORACLE. UNIX is a computer workstation environment for handling and sharing large volume of data. Storage spaces are allocated per department (or workgroup) and per user/terminal. A workspace was created to accommodate this prototype system.

In Arc/Info, spatial data are organized in map layers called coverages. A coverage consists of topologically linked geographic features (arcs, nodes, polygons and label points) and their associated descriptive data stored as automated map. Each map used in this study was stored in a coverage as it is digitized. This stage benefited from the availability of maps in digitized format, but some data pre-processing had to be done. The procedures taken to make the data usable and make the system work are described in this chapter.

6.2 Data Input

6.2.1 Checking and Editing

Firstly, the reference points were identified from a topographic map of Neguev. Corresponding ground coordinates were registered into a tic-coverage. This is a special coverage created to contain points (called tics) that represent locations on the earth for which real-world coordinates are known. Each succeeding coverage should then be created using these tics so that map layers register to each other (ESRI, 1992). There were four map layers for this case study-- the SciSoilmap, Landusemap, Parcelmap and the IKsoilmap which was created using the SciSoilmap as in the proposed methodology and will be explained later in this chapter. The landuse map was found to have been digitized using different tics and was therefore edited for the tics to coincide (using the command EDITFEATURE LABEL in Arcedit module).

The labels for each polygon in each map layer were correspondingly coded. For example, in the SciSoilmap each polygon represents a particular soil unit and this has to be represented with its soil unit code (the key-identifier as explained earlier in Fig.13). The SciSoil table, actually the legend to the soilmap of the Neguev area, is found in Appendix C. The LandUse table was also created using the legend to the LandUse map. The Parcelmap did not need any additional table because there was no other description needed for it in this study. As for the IKsoil, the descriptions in the IK rules (Table 1) were used in its thematic table. Another column was added for the key-identifier.

Each map coverage was checked for the following: that all the polygons contain only one label point (with the command LABELERRORS); that all features that should have been digitized were really digitized (by verifying with the original soilmap sheet); that features that should connect actually do (with the option ERRORS ON at the command DRAWENVIRONMENT); that all features that are there should be there (no extra data).

The Arc/Info commands BUILD and CLEAN were used to construct topology after each editing session. Editing and reconstructing topology was done over and over until errors are spotted and editing is done. When all the errors were eliminated, the map layers were ready for performing geographic analysis on the model.

6.2.2 Establishing Relations

The link between the thematic tables and the polygon attribute tables (PAT) of each map was made possible using the command RELATE. For example, in the case of the scisoilmap.PAT and the Scisoil thematic table:

arc:> relate add		*(instruction to make a relation)*
Relation Name:	soilrel	*(the name for the relation table)*
Table identifier:	Scisoil	*(the table to be related to)*
Database Name:	INFO	*(database in which the table is found)*
INFO item:	Scisoil-id	*(the attribute name in the Scisoil table)*
Relate column:	Scisoil-id	*(the column name in the .PAT table)*
Relate type:	linear	*(instruction to relate records sequentially)*
Relate access:	auto	*(instruction for automatic relation)*
arc:> relate save basgis.rel		*(saving the relation in a relation file)*

This sequence of commands was done for each .PAT table in order to link with its corresponding thematic table (as in Fig.14).

6.2.3 Creating Tabular Data

The database was completed with the creation of the tabular data for the remaining entities Person, User, Owner, Creditor, Credit, Organization, Membership. This was done with the Tables module of INFO (see ESRI, 1992) and after deciding whether to store the data as numeric or character type. There was no real data available for credit information but simulated data was used to run the system.

For example,

```
Enter command:> define PERSON
Item Name:    Person-ID
Item Width:   4
Output Width: 5
Item Type:    B
```

6.3 Data Manipulation

6.3.1 Creating the IKsoil Map

Using the methodology presented in Chap.3 (Fig.3), the IKsoil map was created in the following manner:

1. SciSoil coverage was copied into a new coverage called IKsci

```
arc:> copy scisoil iksci
```

The created IKsci coverage has exactly the same contents as Scisoil but the column names were automatically changed by Arc/Info to IKsci (i.e., IKsci-ID, IKsci#).

2. A column called Scisoil-ID (the link key) was added to the IKsci.PAT in order for it to be related to the Scisoil thematic table.

```
arc:> tables
tables:> additem IKsci.PAT  Scisoil-ID 4 5 b
tables:> calculate Scisoil-ID = IKsci-ID
```

With the last command CALCULATE, the contents of the column IKsci-ID was copied into the column Scisoil-ID, and therefore the link between the IKscimap and the Scisoil thematic table was established.

3. Reclassification of the Scisoilmap was then performed (using the IK criteria in Table 2) on the IKsoil coverage (with the result being the *de facto* IKsoilmap).

```
arc:> arcedit
arcedit:> mapextent iksci
arcedit:> editcoverage iksci
arcedit:> editfeature label
arcedit:> select for soilrel//color = '5YR 4/4' and soilrel//slope# = 6
arcedit:> calculate iksci-id = 2006
```

Because of the established relation between the .PAT table and the thematic table, the SELECT command was able to perform the selection procedure on the thematic table (using soilrel//color). The last two commands were repeated for each IK-rule, which was automated in an Arc Macro Language (AML) file (see Appendix B)-- a user-created program in Arc/Info which was the virtual inference engine for the IK knowledge base. It makes the selection based on the IK rules earlier presented in Chapter 4. This stage was the realization of transformation of contexts.

4. The next step was to merge together adjacent areas with the same key identifier (user-ID). Arc/Info's command DISSOLVE was used. It dissolved the arc that separates two polygons having the same value for a specified item, in this case the user-ID. Then, it created an output coverage (IKsoil) where merged polygons were assigned only one label point (the user-ID).

```
arc:> dissolve iksci iksoil iksci-id poly
```

It should be noted at this point that IKsci remained to be a coverage itself representing the integration of IK and scientific soil classification which can be used later in querying both thematic tables.

5. Topology was constructed on IKsoil with the command BUILD. IKsoilmap was then ready for use.

6.3.2 Overlaying Operations

The map coverages used in this study were all polygon coverages (Parcelmap, SciSoilmap, IKsoilmap, Landusemap). The physical model in Fig.14 shows the relationships among these coverages. These relationships were realized using overlay operations. Polygon overlay is a spatial operation that overlays one polygon coverage on another to create a new polygon coverage. The spatial locations of each set of polygons, and their polygon attributes, are joined in the output coverage. This is done with Arc/Info's command IDENTITY. It keeps the input coverage and adds information from the identity coverage to create an output coverage (see ESRI, 1992). For example, input coverage Parcelmap was overlayed with identity coverage Landusemap to create the output coverage LandusePar. The effect was that the original coverage (e.g., the extent) of Parcelmap was maintained and additional information about the Landuse coverage was added to create LandusePar. Similar overlay operations were done with Parcelmap and Scisoil to create Sci-Parcel; and Parcelmap with IKsoil to create IKParcel. At this point, the logical relationships among the spatial data were realized.

6.4 Data Display and Analysis

6.4.1 Test Case Maize

A test case on soil suitability mapping for maize was implemented. An automated and interactive AML-program was created for the purpose of displaying the results on screen (see Appendix B.2). The scientific suitability map based on the CCT Capacidad de Uso was first displayed. This was done by the following selection query (see Maps 1A and 1B for clarity):

```
arcplot:> reselect scisoil poly soilrel//cap# < 4
          polygonshades scisoil 3
```

The farmers' soil suitability mapping was also displayed with another color.

```
arcplot:> resel iksoil poly ikrel//cropgroup = 1
          polygonshades iksoil 7
```

The overlapping areas corresponding to where both the farmers and the scientists agree was displayed with another color.

```
arcplot:> resel iksci poly ikrel//cropgroup = 1 and scisoil//cap# < 4
      polygonshades iksci 50
```

The areas where the farmers are actually planting maize (from the Landuse map of 1990) was also displayed with a different color.

```
arcplot:> resel landuse poly landuse-id = 1
          polygonshades landuse 4
```

Indeed it showed that the farmers were planting maize areas which they had classified to be good for maize. It was noted however, that they were only planting in small areas compared with the vastness of the area delineated as good for maize. This was explained by reports on the problem of plant diseases in the area (from Brink and Waaijenberg, 1987). Other reasons like market problems were beyond the scope of the study. The landuse map obtained did not include the northernmost area of Neguev.

The next thing to do was to analyze the limiting factors for each system. Flooding was the limiting factor for the CCT classification. This was applied on the IKSoil map and the result showed that the areas where scientific classification did not agree with that of the farmers were those areas with flooding problems:

```
arcplot:> reselect iksci poly ikrel//cropgroup = 1 and soilrel//flood# >1
          polygonshades iksci 5
```

But this was explained by the information that farmers know about when to expect the floods. They also know the areas where flooding risk is high-- where it could occur early in May (see Map 1B):

```
arcplot:> resel iksci poly ikrel//cropgroup = 1 and ikrel//flood = 3
          polygonshades 25
```

Statistics on the selected soil units was computed (with the command STATISTICS), and we see the reclassification of Scisoil units 61, 62, 81, 85, 90, 91 which were classified for pasture but was considered in the IK-classification for maize. The details follow:

MAISTATISTICS

SOIL#	FREQ.	SUM-AREA	CCT-#
50	8	2,894,870.064453	2 (for annual cultivars; high yield)
42	4	284,072.525879	3 (for annual cultivars; moderate yield)
75	2	233,994.664551	3 "
80	5	611,573.659668	3 "
64	2	96,446.203613	3 "
65	1	33,130.144531	3 "
43	1	47,433.497559	3 "
70	1	1,221,640.524414	3 "
85	2	966,214.172852	6 (for extensive pasture)
90	6	968,821.188477	6 "
81	13	792,482.022461	6 "
61	6	2,687,483.490723	6 "
62	7	1,748,020.795898	6 "
91	2	199,819.562500	6 "

6.4.2 Test Case Banana

Another case on suitability mapping for banana was carried out. A corresponding AML-program was created for the purpose (see Appendix B.3). The results were displayed on the screen as the following commands were carried out (see Maps 2A and 2B for clarity):

```
arcplot:> reselect scisoil poly soilrel//cap# < 5
          polygonshades scisoil 3
```

Then, again the farmers' soil suitability mapping was also displayed with another color.

```
arcplot:> resel iksoil poly ikrel//cropgroup = 1
     polygonshades iksoil 7
```

The overlapping areas corresponding to where both the farmers and the scientists agree were displayed with another color.

```
arcplot:> resel iksci poly ikrel//cropgroup = 1 and scisoil//cap# < 5
     polygonshades iksci 50
```

The map showing the limiting factors (Map 2B) explained the difference in the selected areas. While the CCT classification discriminates against flooding, the farmers on the other hand, seriously consider the presence of *capa-dura* or hard layer of soil as it pose problems in soil manageability and yield (van Uffelen, 1990). This limiting factor was applied on the scientific suitability mapping, and it showed that the areas where the farmers' classification did not agree with that of the scientists were those areas with *capa-dura* problems:

```
arcplot:> resel iksci poly soilrel//cap# < 4 and ikrel//capadura = 'y'
          polygonshades scisoil 2
```

The soil scientists had considered this phenomenon in their mapping, but the weakness of the CCT classification system had prevailed (as discussed in Chapter 3.2).

STATISTICS was also computed for the selected soil units:

BANANSTATISTICS

SOIL#	FREQ.	SUM-AREA	CCT-#
50	8	2,894,870.064453	2 (for annual cultivars; high yield)
42	4	284,072.525879	3 (for annual cultivars; moderate yield)
75	2	233,994.664551	3 "
80	5	611,573.659668	3 "
64	2	96,446.203613	3 "
65	1	33,130.144531	3 "
43	1	47,433.497559	3 "
70	1	1,221,640.524414	3 "
32	13	1,246,007.002441	4 (for semi-permanent or permanent cultivars)
31	6	1,001,967.291504	4 "
33	1	15,574.069824	4 "
85	2	966,214.172852	6 (for extensive pasture)
90	6	968,821.188477	6 "
81	13	792,482.022461	6 "
61	6	2,687,483.490723	6 "
62	7	1,748,020.795898	6 "
91	2	199,819.562500	6 "

Results show Scisoil units 31, 32, 33 being classified by CCT for semi-permanent and permanent cultivars (which includes banana), but the farmers regard this soil units as problematic because of the *capa-dura* phenomenon.

6.4.3 Simulated Case Application for Credit

Since data on individual credit was not available, a hypothetical case was carried out to simulate a credit application procedure. IDA gives credit incentives to farmers and these are influenced by the type of crops that IDA is promoting for a particular period. A farmer comes in to apply for credit and the parcel number is keyed into the system. The location of the parcel on the map as well as information about its soil (whether according to scientific classification or according to IK classification) are displayed. In addition to these, information on the credit history of the parcel-owner and parcel-user are also displayed on screen. Based on these information (and the criteria set by the organization), the person in charge of the facility can then decide if the applicant is qualified for credit or not. This procedure was facilitated by a program created using ORACLE/SQL which was linked with Arc/Info (see Appendix B.4):

arcplot:> &run credit
Type Parcel Number: <205>

==

INFORMATION ON PARCEL No: 205
Actual Landuse:

Landuse	Area (sq.m.)
Pasture	108,608.07839
Fruit trees	7,838.62628
Banana-coconut intercrop	6,161.53589

Potential Use (IK classification):

SoilType	Area (sq.m.)	Cropgroup	Capa-dura
Tierra bermeja alta	4,914.67754	2	Y
Tierra bermeja baja	11,225.87501	2	N
Tierra Negra mediana	107,084.11760	1	N

Cropgroup			
1:	maize, rice, beans, banana, platano, pasture		
2:	palmito, cocoa		
3:	pineapple, chili		

Potential Use (CCT classification):

Soil Type	Taxonomic Class	Area (sq.m)	CCT-class
Destierro var.	Andic Eutropepts, loamy	107,084.11760	Extensive pasture
Milano plains	Andic Humitropepts, fine	4,870.36569	Permanent cultivars
Milano escarped	Andic Humitropepts, fine	11,123.15577	For protection

PERSONAL INFORMATION:

Owner	Acquisition Mode	Origin	Organization
Gonzalez	purchase	Guapiles	S.P.A
Owner	**Creditor**	**Credit Date**	**Credit Balance**
Gonzalez	IDA	1988	0
User(s)			
Andres	B.C.R.	1990	100,000.00

= =

6.4.4 The IK-Suitability Map

The over-all IK-suitability map was displayed and compared with the original CCT-suitability map to analyze the similarities and differences (see Maps 3A and 3B). Soil units (of the black soils) which were originally classified for forest and protection (CCT-# 8,9,10) were retained as such since there were no IK-rules at hand regarding the matter (and to respect forest conservation measures). Statistics were computed using Arc/Info command STATISTICS and the results are as follows:

SOILSTATISTICS

SOIL#	FREQ.	AREA	CCT#	IK-CROP#	Assessment
42	4	284,072.525879	3	1	retained
85	2	966,214.172852	6	1	upgraded
75	2	233,994.664551	3	1	retained
96	54	3,963,648.442383	9	n.a.	
22	39	1,560,027.355469	6	2	upgraded
90	6	968,821.188477	6	1	"
21	91	9,616,610.054199	6	2	"
11	2	148,920.026855	9	3	upgraded
24	36	1,441,042.601563	7	2	retained
80	5	611,573.659668	3	1	"
23	3	659,535.926270	7	2	"
81	13	792,482.022461	6	1	upgraded
100	10	135,161.692383	10	n.a.	
25	42	5,563,835.747559	8	n.a.	
12	15	1,029,197.386719	9	3	upgraded
61	6	2,687,483.490723	6	1	"
15	1	5,285,787.446289	9	3	"
13	2	65,217.676758	9	3	"
62	7	1,748,020.795898	6	1	"
63	4	146,109.428711	10	n.a.	
64	2	96,446.203613	3	1	retained
65	1	33,130.144531	3	1	"
32	13	1,246,007.002441	4	2	downgrade
35	3	221,828.537109	10	n.a.	
95	1	44,717.753906	9	n.a.	
91	2	199,819.562500	6	1	upgrade
34	6	396,400.680664	8	n.a.	
31	6	1,001,967.291504	4	2	downgrade
41	1	139,527.020508	10	n.a.	
43	1	47,433.497559	3	1	retained
14	1	485,300.795898	9	3	upgrade
33	1	15,574.069824	4	2	downgrade
70	1	1,221,640.524414	3	1	retained
50	8	2,894,870.064453	2	1	"

Scisoil units with CCT-classes for maize (2 and 3) maintained its class for maize in the IK-classification. Scisoil units with CCT-class 6 (for pasture) were classified into IK-cropgroup 1 which is also for pasture but also includes maize, rice, beans and banana. The Scisoil legend shows that these are predominantly the areas with moderate flooding (several days in majority of the years), but the farmers must have known the period of the year when the flooding is expected and have adjusted their cropping pattern, hence the consideration of these areas which is considerably vast (1,851 hectares in total).

Scisoil units with CCT-class 7 (for tree crops) agree with the IK-classification for cropgroup 2 (palmito and cocoa). On the other hand, Scisoil units with CCT-class #4 (for semi-permanent and permanent cultivars) has to be reconsidered for certain crops which are sensitive to the *capa-dura* phenomenon prevailing especially in soils with high clay content. *Capa-dura* is the compaction of the superficial layer of soils brought about by the trampling of the field or work with heavy machines before the establishment of the settlement (van Uffelen, 1990). According to the same report, the farmers consider this as a very important factor. Some cultivars like corn, banana and beans produce nothing on such soils. As for IK-cropgroup 3 (pineapple and chili), Scisoil units that fall into this category are classified by CCT as class #9 (for forest).

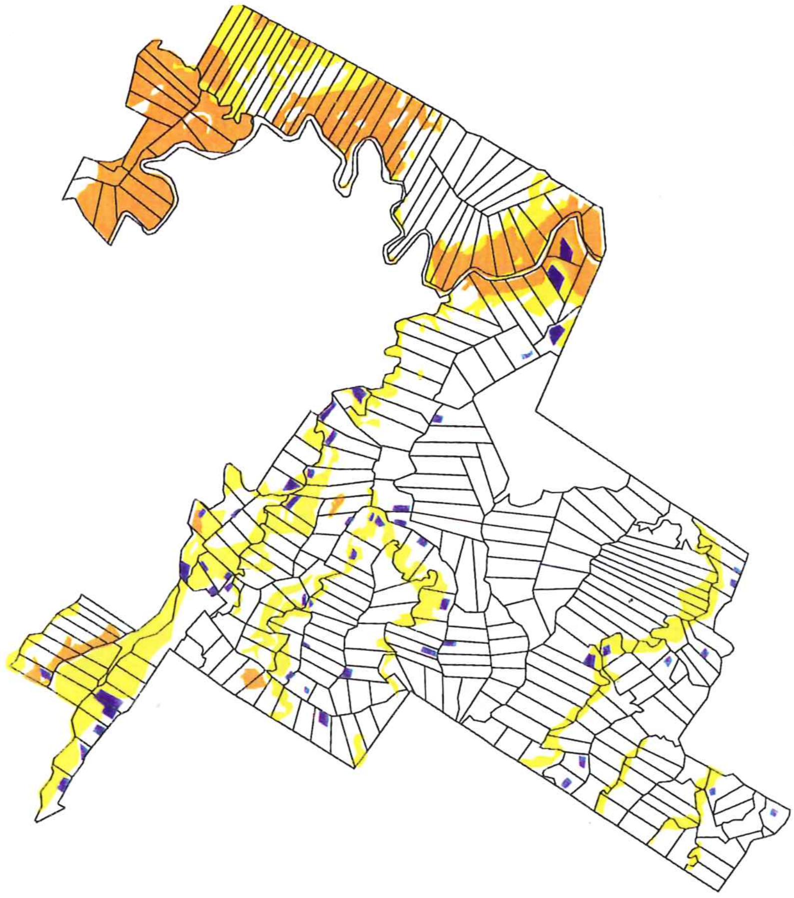
Looking back at the maize and banana test cases, we see a new soil suitability mapping (Maps 1B and 2B). Areas where farmers and scientists agree could be rated highly for the crop since what is expected corresponds to what is tried and tested. While those affected by limiting factors (in this case flooding and *capa-dura*) might need some intervention to be considered for the purpose and could be given a lower rating. For the actual landuse statistics in 1990 (see below): 40% of the total area of Neguev is being used for pasture. Further look into the landuse per soil unit (Appendix D) shows that even soils classified for annual cultivars are under pastures. This reveals some insights in the prevailing farming systems-- that farmers are not just considering physical soil properties in deciding what to make use of their land. Factors like market or agro-ecological ones are also considered, and they require further research.

LANDUSE#	FREQ.	AREA	DESCRIPTION
11	516	21,428,125.447948	pasture
12	351	9,241,172.301888	forest
13	61	2,505,300.002994	swamp
14	25	800,820.094436	village
1	120	662,856.989896	maize
9	79	353,699.097147	tubers
8	59	296,434.278605	banana
7	54	276,786.524257	cocoa
5	68	234,534.355796	chili
78	21	139,559.743771	banana-cocoa
10	15	116,370.965056	fruit trees
89	16	95,639.524437	banana-tubers
6	21	93,851.881977	pejibaya
4	20	88,534.457987	coconut
49	4	62,016.519763	coco-tubers
81	6	43,553.221398	maiz-banana
2	11	37,354.777635	beans
3	6	15,537.667727	pineapple
84	3	9,419.476154	banana-coconut

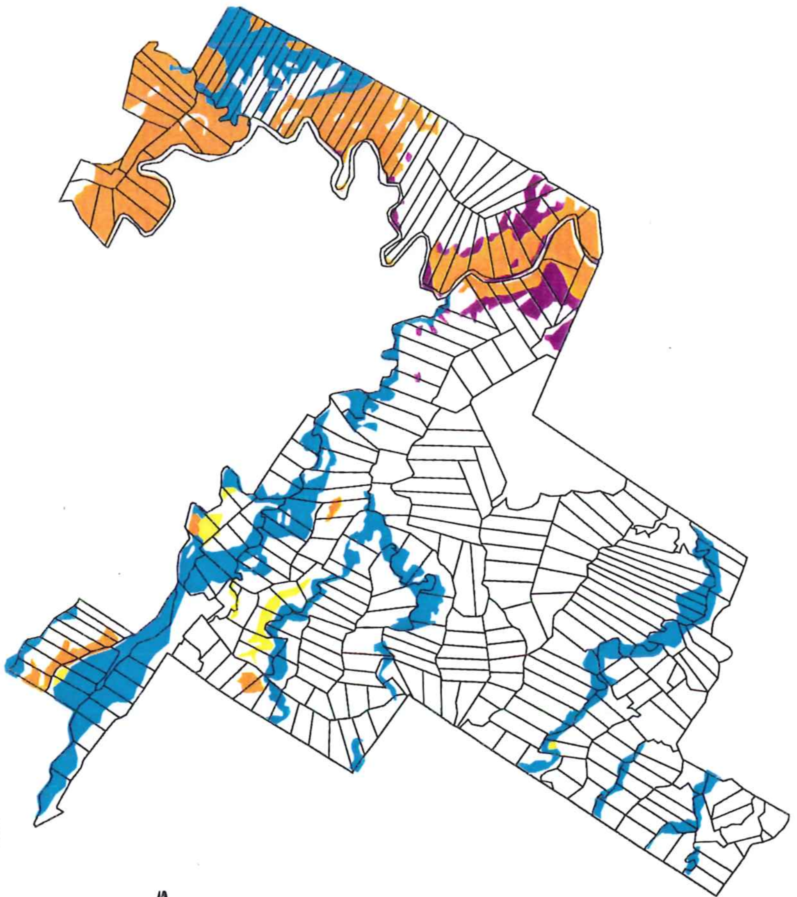
AREAS CONSIDERED FOR MAIZE PRODUCTION
NEGUEV, COSTA RICA

Scale = 1:100,000

- Not suitable using both criteria
- Actually used
- Flooding problems
- Farmers' criteria
- Flooding May
- Farmers' & scientists' criteria



Map 1A: All Possible Areas

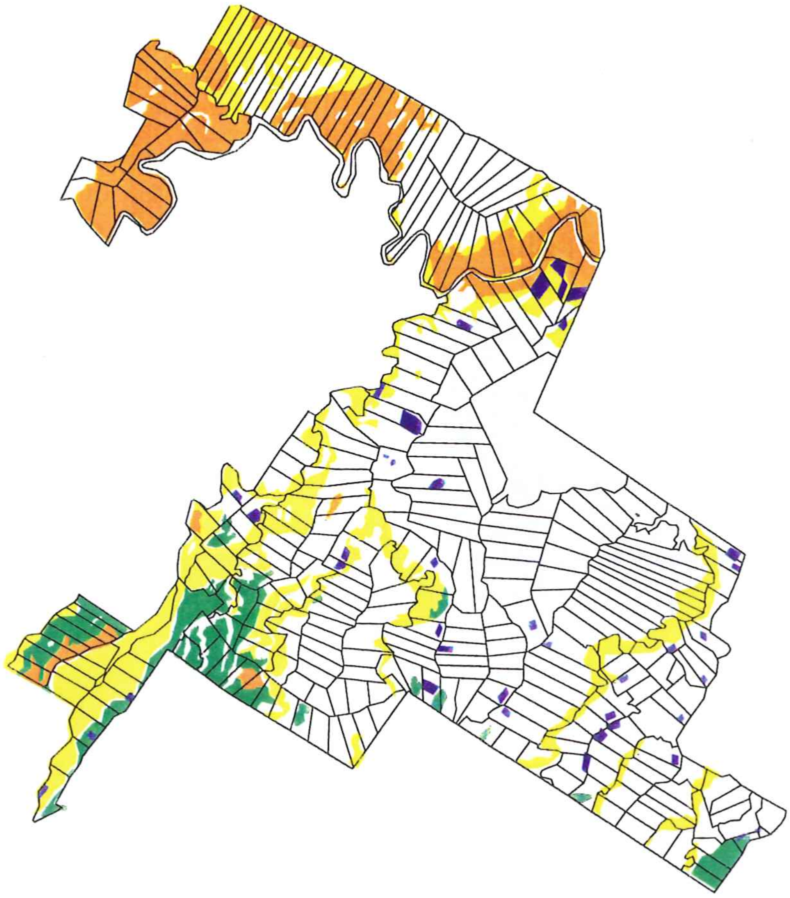


Map 1B: Considering Limiting Factors

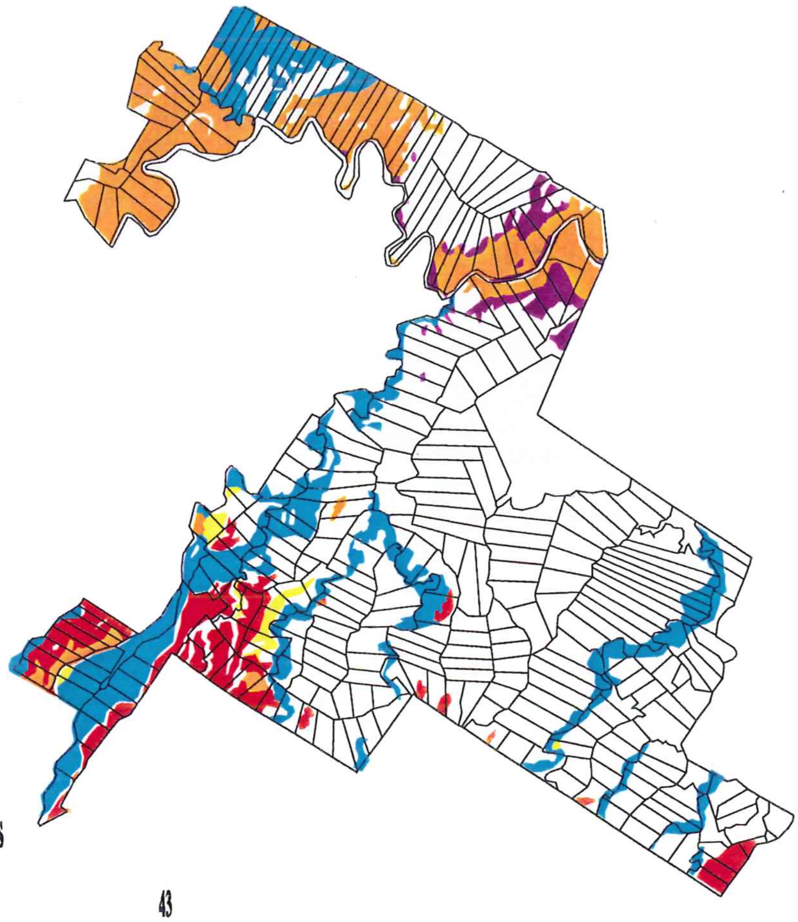
AREAS CONSIDERED FOR BANANA PRODUCTION
NEGUEV, COSTA RICA

Scale = 1:100,000

- Not suitable using both criteria
- Capa-dura problems
- Scientists' criteria
- Actually used
- Flooding problems
- Farmers' criteria
- Flooding as early as May
- Farmers' & scientists' criteria



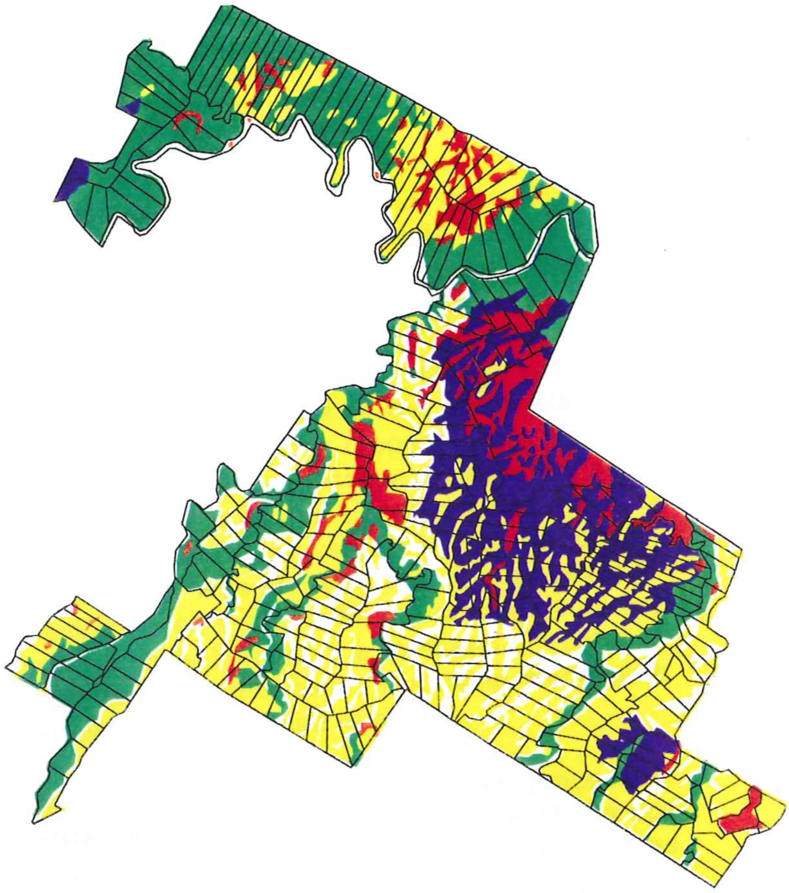
Map 2A: All Possible Areas



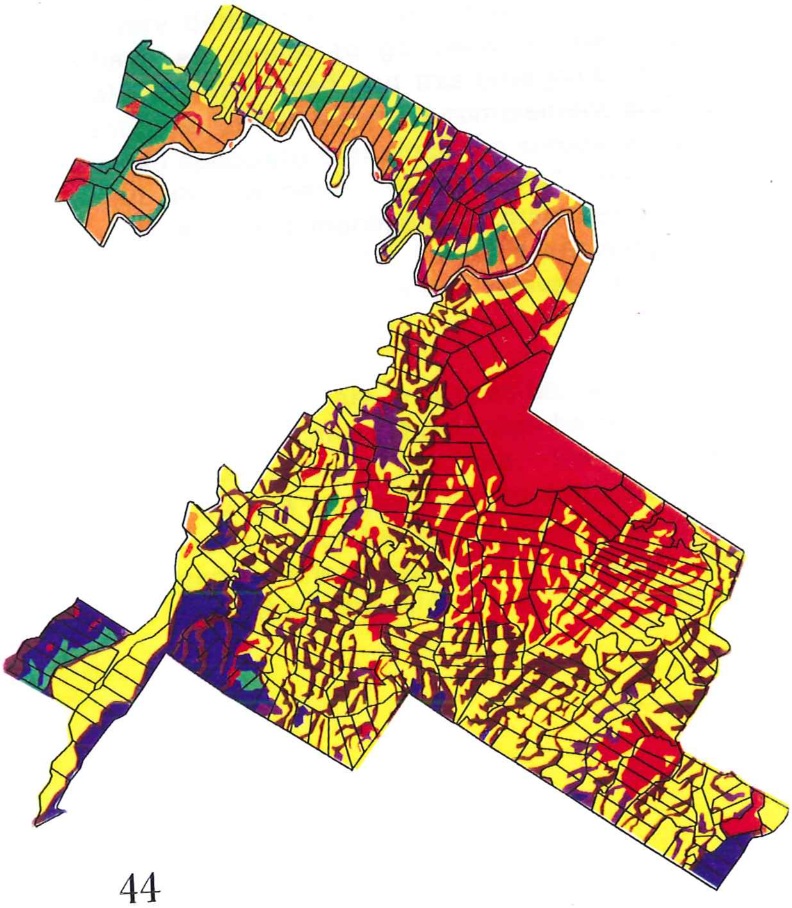
Map 2B: Considering Limiting Factors

SOIL SUITABILITY COMPARISO
NEGUEV, COSTA RICA

Scale = 1:100,000



Map 3A: IK Classification



Map 3B: CCT Capacidad de Uso

7 EVALUATION AND CONCLUSION

The case presented highlights several difficulties and problems which indicate areas for further study at the same time that they provide qualifications to the analysis earlier made. Things have to be considered. Firstly, it should be noted that the object of abstraction of farmers for the soil was not explicitly defined. This is attributed to the general unavailability of data-- the maps and photographs of the right scale and the inadequacies in the knowledge acquisition phase. In this case, their object of abstraction took the form of the reclassified and aggregated scientists' soil units, which does not exactly reflect the farmers' "world view". As in hypothesis 2, IK-rules for soil classification and aggregation were somehow used to reclassify scientific soil units. This was realized with the articulation of IK-criteria using those of the scientists (e.g., color and slope). Problems in semantics were addressed in this respect-- that we have to accept the different object definitions and descriptions because of differences in contexts and later find out how an object in one context can be translated into a similar object in another context. The IK-rules at hand may be very few and simplistic. Nevertheless, the proof of principle for hypothesis 1 was shown-- that it is possible to incorporate indigenous knowledge in knowledge based systems be it just an exploratory attempt. It still needs further studies in order to fully harness the foreseen advantages of KBS and GIS link-up in general, and in particular as this case, in moving along both classification systems, making use of the knowledge contained in each.

Secondly, the results presented have not undergone field verification processes. It is indeed fortunate that this study benefited a lot from informed opinions of the soil scientists and agronomists who had worked in the area. Agreements were reached as to the acceptability of the resulting maps.

Thirdly, we look at the applicability of the methodology in majority of cases where there are no available soilmaps of very detailed scale as the one used in this study. In such cases, soil surveyors have an option to go back to the original methodology presented in Figure 2-- to make the map itself, but this time jointly by the scientists and the local farmers. The intention is to find ways to complement and not to replace conventional soil surveys. We could speculate that since the survey will also revolve around indigenous knowledge and the farmers' "world view", the most important factors determining local resource value and management practices will be captured during the survey. As pointed out earlier in Chapter 1.4, "by knowing local priorities, less effort need be spent providing information that is less relevant and more effort spent providing information and assistance that are useful."

Fourthly, local soil names and local soil knowledge are area specific. Application on a regional scale is not feasible. Fuzziness in classification should also be recognized especially because the farmers of Neguev are not native to the area. Some settlers even come from Panama and Nicaragua. Neguev is a relatively young settlement (10 years old at the time of the study), the farmers may not have accumulated as much knowledge about the soils compared with indigenous knowledge present in much older villages. Inspite of this, they know much more than the IDA technicians who are supposedly from the area itself.

Lastly, this study is limited to the physical properties of the soil in determining its potential use as in soil resource surveys. But farmers are conceivably concerned with the economic and risk factors associated with land use decisions. The actual landuse shown in the maps and the corresponding statistics give some indications. Farmers are actually planting crops on areas classified for protection or forest by the CCT. But in other cases, farmers whose lands are classified good for annual cultivars are simply having pastures. There are far too many factors to consider in deciding for land use. The presented prototype system is but one input to a much more complex land use planning process. The simple model used in the present study can be improved so as to capture, for example, profit maximization or risk minimization behavior among farm producers.

At this point, we recall the objectives of the research and find that these have been achieved. A spatial database has been created and tested-- one that incorporates indigenous knowledge of soils and gives a description of soil resources that better reflects actual conditions. In that respect, the data acquisition component of GIS was being addressed. GIS techniques were used to handle and manipulate both descriptive and locational data about soils. More importantly, its spatial analysis capabilities like retrieval, classification and overlaying were utilized in articulating local knowledge of soils with that of the scientists-- a modest attempt to integrate theory and practice.

-oOo-

LIST OF REFERENCES

Aronoff,S. Geographic Information Systems: A Management Perspective. Otawa: WDL Publications, 1989.

Altieri,M. The Question of Small Farm Management: Who Teaches Whom? In: Agriculture, Ecosystems and Environment, No.9. Amsterdam: Elsevier, 1983.

Avison,D.E. and Wood-Harper,A.T. Multiview: An Exploration in Information Systems Development. England: Alden Press, 1990.

Bregt,A. Integrating GIS and Process Models for Global Environmental Assessment. Wageningen: DLO Staring Center, 1993.

Brink,M. and Waaijenberg,H. *Base De Datos De Una Caracterizacion De Fincas Realizada En El Norte De La Zona Atlantica De Costa Rica*. Costa Rica: AZP, 1987.

Brokensha, et.al. Indigenous Knowledge Systems and Development. Washington: University Press of America, 1980.

Burrough, P. Principles of Geographical Information Systems for Land Resources Assessment. London: Oxford University Press, 1986.

Burrough,P. Development of Intelligent Geographic Information Systems. In: International Journal of Geographic Information Systems, vol.6(1), 1992.

CCT. *Manual Para La Determinacion De La Capacidad De Uso De Las Tierras De Costa Rica*. Costa Rica: CCT, 1985.

Chambers,R. et.al. Farmer First: Farmer Innovation and Agricultural Research. London: ITP, 1989.

Chen,P. The Entity-Relationship Model-- Toward a Unified View of Data. In: ACM Transactions on Database Systems, vol.1(1), Massachusetts: ACM, 1976.

Date,C. An Introduction to Database Systems. vol.1. USA: Addison & Wesley, 1990.

de Bruin,S. *Estudio Detallado de los Suelos del Asentamiento Neguev.* Costa Rica: Atlantic Zone Programme, 1992.

Egenhofer,M. and Frank,A. Object-Oriented Modelling in GIS: Inheritance and Propagation. AUTO-CARTO Proceedings: Report 102, 1989.

Elmasri,R. and Navathe,S.B. Fundamentals of Database Systems. USA: Benjamin & Cummings Publishing, 1989.

ESRI. Understanding GIS: The ArcInfo Method. California: ESRI, 1992.

Furbee,L. A Folk Expert System: Soils Classification In The Colca Valley, Peru. In: Anthropological Quarterly, vol.62(2), 1989.

Hopgood,A. Knowledge-based Systems For Engineers And Scientists. Boca-Raton: CRC, 1993.

Howe,D. Data Analysis for Database Design. UK: Arrowsmith Ltd., 1989.

Huising, J. Land Use Zones and Land Use Patterns in the Atlantic Zone of Costa Rica. Wageningen: Ponsen & Loijen, 1993.

Israel,D. AI Knowledge Bases And Databases. In: On Knowledge Base Management Systems: Integrating Artificial Intelligence and Database Technologies. (ed. Brodie,M.) New York: Springer-Verlags, 1986.

Kim,J. et.al. Expert System Applications to Urban Planning. USA: Springer Verlag, 1991.

Klee,G. World Systems of Traditional Resource Management. London: Winston & Sons, 1980.

McGraw,K. and Harbison-Briggs,K. Knowledge Acquisition: Principles And Guidelines. New Jersey: Prentice-Hall, 1989.

McMenamin,S. and Palmer,J. Essential Systems Analysis. New Jersey: YourdonPress, 1984

Molenaar,M. Towards A Geographic Information Theory. Enschede: ITC Journal, 1989.
Molenaar,M. Status And Problems Of Geographical Information Systems: The Necessity Of Geoinformation Theory. In: ISPRS Journal Of Photogrammetry And Remote Sensing. Amsterdam: Elsevier Science Publishers, 1991.
Molenaar,M. and Janssen,L. Terrain Objects And Their Monitoring By The Integration of GIS And Remote Sensing. (Handouts for GIS Theory).
Molenaar,M. Object Hierarchies and Uncertainties in GIS or Why is Standardization So Difficult? In: GIS 4/1993.
Mucher,C. A Study on the Spatial Distribution of Land Use in the Settlement of Neguev. Costa Rica: AZP, 1991.
Munsell,A. and Birren,F. A Grammar of Color: A Basic Treatise on the Color System of Albert H. Munsell. New York: McGraw-Hill, 1969.
Olson,J. and Rueter,H. Extracting Expertise From Experts. In: Expert Systems, vol.4(3), 1987.
Pawluk,J. et.al. The Role of Indigenous Soil Knowledge in Agricultural Development. In: Journal of Soil and Water Conservation, vol.47, 1992.
Peuquet,D. A Conceptual Framework and Comparison of Spatial Data Models. In: Cartographica, vol.21(4), 1984.
REPPIKA. Regional Program for the Promotion of Indigenous Knowledge in Asia. Philippines: REPPIKA, 1990.
Rich,E. and Knight,K. Artificial Intelligence. Singapore: McGraw-Hill, 1991.
Richards,P. Indigenous Agricultural Revolution: Ecology and Food Production in West Africa. Colorado: Westview Press, 1985.
Robinson,V. and Frank,A. Expert Systems For Geographic Information Systems. In: Photogrammetric Engineering And Remote Sensing, vol.53 No.10. Virginia: ASPRS, 1987.
SRIG. Soil Resource Inventory Group: Proceedings of Workshops at Cornell University 1977-78. New York: Cornell University, 1979.
Smith,G. Rationale for Concepts in Soil Taxonomy. Cornell: SMSS, 1986.
Tabor, J. Ethnopedology: Using Indigenous Knowledge to Classify Soils. In: Arid Lands Newsletter, vol.30. Arizona: OALS, 1990.
Teskey,F. User Models And World Models For Data, Information And Knowledge. In: Information Processing And Management, vol.25 No.1. UK: Pergammon Press, 1989.
Tick,A.(ed.) Indigenous Knowledge and Development Monitor, vol.1(2), The Hague: CIKARD, 1993.
Tulladhar,A. Introduction to GIS. ITC-course handout, 1992.
USDA Soil Survey Staff. Soil Taxonomy. Virginia: USDA, 1975.
USDA Soil Survey Staff. Keys To Soil Taxonomy. Virginia: USDA, 1992.
van den Berg,P. and Droog,R. Quantification Of Farming Systems In The Neguev Settlement. Costa Rica: AZP, 1992.
van Uffelen, *Conocemientos Endogenos y Cientificos Para La Determinacion del Capacidad de Uso del Suelos de Neguev*. Costa Rica:AZP, 1990.
Veltman, B. Land Use In Rio Jimenez Linked To Soil Types With The Encuesta General And Aerial Photographs. Costa Rica: AZP, 1991.
Warren,D.M. Indigenous knowledge Systems For Activating Local Decision-Making Groups In Rural Development. Iowa: ISU, 1975.
Warren,D.M. Linking Scientific And Indigenous Agricultural Systems. In: The Transformation Of International Agricultural Research And Development. (ed. Lin Compton, J.) Colorado: Lynne Rienner, 1989.
Zinck,J.A. and Valenzuela,C. Soil Geographic Database: Structures and Application Examples. In: Introduction to Geographic Information Systems, ITC-1992.

APPENDIX A

DATA DICTIONARY

SECTOR	- local administrative boundary of Neguev
Sector_ID	- the key-identifier of the sector
Sector_name	- the name of the sector
PARCEL	- subdivided land for individual ownership
Parcel_no.	- the key-identifier of the parcel
LANDUSE	- the purpose for which a piece of land is used; also refers to the crop planted on the parcel
LU-ID	- the key-identifier of a particular landuse
LANDUSE-PAR	- the relationship table between the parcel and landuse showing which parcelsare being used for a particular purpose
IKSOIL	- indigenous soil unit; smallest soil unit delineated by farmers to be of homogeneous properties
IK_ID	- the key-dentifier of a particular indigenous soil unit
LocalName	- the term used by the farmers to refer to a particular soil unit
SCISOIL	- scientific soil unit; smallest soil unit delineated by soil scientist to be of homogenous properties
Sci_ID	- the key-identifier of a particular scientific soil unit
SciName	- the term used by the scientists to refer to a particular soil unit
CCTclass	- the classification rating of a scientific soil unit according to CCT
IK-PARCEL	- the relationship table between the parcel and the indigenous soil units showing which parcels contain a specific soil unit
SCI-PARCEL	- the relationship table between the parcel and the scientific soil units showing which parcels contain a specific soil unit
PERSON	- an individual living within the Neguev settlement
Person_ID	- the key-identifier of a particular person
Sex	- the sex of the person
Origin	- the place where the person came from before settling in Neguev
Birth_yr	- the year when the person was born
Civil_stat	- the civil status of the person
USERS	- the relationship table between the parcel and the person showing which parcels are being used by a particular person
StartDate	- the date when the person started using the parcel
EndDate	- the date when the person stopped using the parcel

Appendix A (cont'n)

OWNER	- the relationship table between the parcel and the person showing which parcel is owned by a particular person
DateAcq	- the date when the person acquired the parcel
AcqMode	- the manner by which a person acquired a parcel
ORGANIZATION	- grouping of persons formed for a purpose
OrgCode	- the key-identifier of an organization
Type	- the type of organization; also the purpose for which it was formed
MEMBERSHIP	- the relationship table between the person and the organization showing which organizations a person is a member of
DateJoined	- the date when a person joined a particular organization
Status	- the status of membership of a person
CREDITOR	- the credit facility from where a person applies for credit
Creditor_ID	- the key-identifier of a creditor
CreditorName	- the name of the credit facility
CREDIT	- the relationship table between the person and the creditor showing which creditor loaned money to which person
LoanDate	- the date when the loan was made
LoanAmt	- the amount of the loan
Balance	- the amount not yet paid from the loan

APPENDIX B

AML/SQL-Programs

B.1 AML-Program for Reclassifying Soils

```
/*Silencio
select for soilrel//color = '5YR 4/4' and
soilrel//slope# = 'C'
calculate test-id = 2006
nselect
select for soilrel//color = '5YR 4/4' and
soilrel//slope# = 10 or ~
 soilrel//slope# = 'DE'
calculate test-id = 2007
nselect
select for soilrel//color = '5YR 4/4' and
soilrel//slope# = 'D' or ~
 soilrel//slope# = 'CD'
calculate test-id = 2008
nselect
/*Neguev
select for soilrel//color = '10YR 3/3' and
soilrel//slope# = 'B'
calculate test-id = 2003
nselect
select for soilrel//color = '10YR 3/3' and
soilrel//slope# = 'D' or ~
 soilrel//slope# = 'E'
calculate test-id = 2004
nselect
select for soilrel//color = '10YR 3/3' and
soilrel//slope# = 'C' or ~
 soilrel//slope# = 'CD'
calculate test-id = 2005
nselect
/*Milano
select for soilrel//color = '10YR 4/3' and
soilrel//slope# = 'B' or ~
 soilrel//slope# = 'A'
calculate test-id = 2000
nselect
select for soilrel//color = '10YR 4/3' and
soilrel//slope# = 'E'
calculate test-id = 2001
nselect
select for soilrel//color = '10YR 4/3' and
soilrel//slope# = 'F'
calculate test-id = 2002
nselect
/*Suamposa
select for soilrel//soil-code = 'Tu' or
soilrel//soil-code = 'u'
calculate test-id = 2012
nselect
/*Negra
select for soilrel//flood# = 0 and soilrel//color
<> '10YR 4/3' and ~
soilrel//color <> '10YR 3/3' and soilrel//color
<> '5YR 4/4'
calculate test-id = 2009
nselect
select for soilrel//flood# = 1
calculate test-id = 2010
nselect
select for soilrel//flood# = 2
calculate test-id = 2011
/*create topology
save
q
build iksoil poly
```

B.2 AML-Program for Maize Suitability

```
clear
mapex soilmap
relate restore basgis.rel
killmap maiz
map maiz
mbegin
mapscale 65000
pageunits cm
pagesize 21 29.7
box 0 0 21 29.7
box .5 27.5 14.5 29.5
maplimits 1 1 21 25
textfont 93713
textquality kern
polys parcel
keyposition 1 27
keyshade maiz.key
textfont 93715
textsize 1 .75
move 1.5 28
text 'AREAS FOR MAIZE PRODUCTION'
move 16.5 25.5
textsize .75 .5
text 'Scale = 1:65,000'
move 15.5 24.5
text 'NEGUEV, COSTA RICA'
markerset municipal.mrk
markersymbol 128
markersize 3
marker 18 28
reselect soilmap poly soilrel//cap# < 4
polygonshades soilmap 3
&pause
reselect iksoil poly ikrel//cropgroup = 1
polygonshades iksoil 7
```

```
&type
&type 'Selecting overlaps (scientific and IK)...'
&pause
readselect overlaps
polygonshades soilmap 50
&type
&type 'Determining actual Land Use...'
&pause
reselect parceluse poly landuse2-id = 1
polygonshades parceluse 4
polys parcel
&message &off
resel soilmap poly soilrel//flood# > 1
&type
&pause
&type 'Selecting flooding problems...'
resel iksci poly ikrel//cropgroup = 1 or ~
  soilrel//cap# < 4
resel iksci poly soilrel//flood# > 1
polygonshades soilmap 5
polygonshades iksci 5
&type
asel iksci poly
resel iksci poly ikrel//cropgroup = 1 or ~
  soilrel//cap# < 4
resel iksci poly ikrel//flooding = 3
&type
&type 'Flooding could occur early in May...'
polygonshades iksci 25
&pause
&type
&type 'Selecting capa-dura problems...'
asel iksci poly
resel iksci poly soilrel//cap# < 4
resel iksci poly ikrel//capadura = 'y'
polygonshades iksci 2
polygonshades parceluse 4
mend
&message &on
&type
&pause
&type ====================
dbmsexecute oracle select * from table20,
table21, table22
&type
&type
&type 'MAIZE Production Information for 1990:'
&type
dbmsexecute oracle select * from table10
&type
&type 'Parcels that planted maize:'
&pause
dbmsexecute oracle select * from table11
asel iksoil poly
asel soilmap poly
asel parceluse poly
asel iksci poly
&return
```

B.3 AML-Program for Banana Suitability

```
clear
mapex soilmap
relate restore basgis.rel
killmap maiz
map maiz
mbegin
mapscale 65000
pageunits cm
pagesize 21 29.7
box 0 0 21 29.7
box .5 27.5 14.5 29.5
maplimits 1 1 21 25
textfont 93713
textquality kern
polys parcel
keyposition 1 27
keyshade maiz.key
textfont 93715
textsize 1 .75
move 1.5 28
text 'AREAS FOR BANANA PRODUCTION'
move 16.5 25.5
textsize .75 .5
text 'Scale = 1:65,000'
move 15.5 24.5
text 'NEGUEV, COSTA RICA'
markerset municipal.mrk
markersymbol 128
markersize 3
marker 18 28
/* selection according to scientists
reselect soilmap poly soilrel//cap# < 5
polygonshades soilmap 3
&pause
/* selection according to farmers
reselect iksoil poly ikrel//cropgroup = 1
polygonshades iksoil 7
&type
/* displayign overlaps
&type 'Selecting overlaps Scientific and IK...'
&pause
reselect iksci poly soilrel//cap# < 5 and
ikrel//cropgroup = 1
polygonshades iksci 50
&type
/* displaying areas being used for the specific
crop (in 1990)
&type 'Determining actual Land Use...'
&pause
reselect landuse2 poly landuse2-id = 8 or
landuse2-id = 78 or ~
 landuse2-id = 89 or landuse2-id = 81 or
landuse2-id = 84
polygonshades landuse2 4
polys parcel
```

```
&message &off
&type
&pause
&type 'Selecting flooding problems...'
asel iksci poly
resel iksci poly soilrel//cap# < 5 or
ikrel//cropgroup = 1
resel iksci poly soilrel//flood# > 1
polygonshades iksci 5
&type
asel iksci poly
resel iksci poly ikrel//cropgroup = 1 or
soilrel//cap# < 5
resel iksci poly ikrel//flooding = 3
&type
&type 'Flooding could occur early in May...'
polygonshades iksci 25
&pause
&type
&type 'Selecting capa-dura problems...'
asel iksci poly
resel iksci poly soilrel//cap# < 5
resel iksci poly ikrel//capadura = 'y'
polygonshades iksci 2
polygonshades landuse2 4
mend
&message &on
&type
&pause
&type ===================
dbmsexecute oracle select * from table20,
table21, table22
&type
&type
&type 'MAIZE Production Information for 1990:'
&type
dbmsexecute oracle select * from table10
&type
&type 'Parcels that planted maize:'
&pause
dbmsexecute oracle select * from table11
&message &off
asel iksoil poly
asel soilmap poly
asel parceluse poly
asel iksci poly
&message &on
&return
```

B.4 AML-Program for Credit Application

```
mapex soilmap
pageunits cm
pagesize 21 29.7
relate restore basgis.rel
asel parcel poly
&label enter
disconnect oracle
connect oracle gonzales/rhodora@T:ds10:A
clear
map parcels
map end
&setvar x [response 'Type Parcel Number~
  (or hit ENTER to quit)' 0]
&if %x% > 312 &then &goto enter
&if %x% eq 0 &then &goto exit
reselect parcel poly parcel-id = %x%
polygonshades parcel 7
asel parceluse poly
resel parceluse poly parcel-id = %x%
&sys sqlplus -s gonzales/rhodora@T:ds10:A
@test.sql %x%
&message &on
&type
&type ===================
&type 'INFORMATION ON PARCEL NUMBER'
  %x%
&type
list parcel poly area perimeter
&type
&type 'Actual Land Use:'
&type
list parceluse poly userel//desc area
&type
&type
&type 'Potential use according to IK Evaluation:'
&type
dbmsexecute oracle select * from table3.
&type ===================
&type CROP GROUP 1 = Maize, Rice, Beans,
Banana, Pasture
&type          2 = Palmito, Cocoa
&type          3 = Pineapple, Chili
&pause
&type
&type 'Potential use (Scientific Evaluation:)'
dbmsexecute oracle select * from table5
&type ===================
dbmsexecute oracle select * from table2
&type
&type 'CREDIT History:'
dbmsexecute oracle select * from table6
&type ------------------------------------------
dbmsexecute oracle select * from table7
polygonshades parceluse userel//code landuse.lut
&pause
asel parcel poly
&goto enter
&label exit
&type 're-starting program...'
&return
```

B.5 SQL Link

```
drop table mais;
create table mais (soilmap_id, cap_desc, area) as
select soilmap_id, cap_desc, sum(area)
from ikscipat, sci_soil, ik_soil
where soilmap_id = soil_no_ and ikno = ik_id
and cropgroup = 1
group by soilmap_id, cap_desc;
commit;

drop table palmito;
create table palmito (soilmap_id, cap_desc, area)
as
select soilmap_id, cap_desc, sum(area)
from ikscipat, sci_soil, ik_soil
where soilmap_id = soil_no_ and ikno = ik_id
and cropgroup = 2
group by soilmap_id, cap_desc;
commit;

drop table chili;
create table chili (soilmap_id, cap_desc, area) as
select soilmap_id, cap_desc, sum(area)
from ikscipat, sci_soil, ik_soil
where soilmap_id = soil_no_ and ikno = ik_id
and cropgroup = 3
group by soilmap_id, cap_desc;
commit;

set verify off
drop table table1;
create table table1 (ikname, total_area) as
select distinct ikname, sum(area) from
ikparcelpat t1,
ik_soil t2 where t1.ik_id = t2.ikno and parcel_id
= &1 group by ikname;
commit;

drop table table3;
create table table3 as
select distinct t2.ikname, total_area, cropgroup,
capadura from ikparcelpat t1,
table1 t2, ik_soil t3 where t1.ik_id = t3.ikno
and t3.ikname = t2.ikname
and parcel_id = &1;
commit;

drop table table4;
create table table4 (soil_code, total_area) as
select distinct soil_code, sum(area) from
sciparcelpat t1,
sci_soil t2 where t1.soilmap_id = t2.soil_no_
and parcel_id = &1
group by soil_code;
commit;

drop table table5;
create table table5 as
select distinct t3.soil_code, description,
taxon_class, total_area, cap_desc
from sciparcelpat t1, sci_soil t2, table4 t3
where t1.soilmap_id = t2.soil_no_ and
t2.soil_code = t3.soil_code and parcel_id = &1;
commit;

drop table table6;
create table table6 (owner, creditor, credit_date,
balance) as
select l_name, name, c_date, balance from
person t1, owner t3, creditor t4, credit t5
where t3.parcel_id = &1 and t3.person_id =
t5.person_id
and t5.person_id = t1.person_id and
t5.creditor_id = t4.creditor_id;
commit;

drop table table7;
create table table7 (users, creditor, credit_date,
balance) as
select l_name, name, c_date, balance from
person t1, users t3, creditor t4,
credit t5 where t3.parcel_id = &1 and
t3.person_id = t5.person_id
and t5.person_id = t1.person_id and
t5.creditor_id = t4.creditor_id;
commit;

drop table table2;
create table table2 (owner, acquisition_mode,
origin, organization) as
select l_name, acq_mode, origin, org from owner
t1, person t2,
membership t3, organization t4
where t1.parcel_id = &1 and t1.person_id =
t2.person_id
and t2.person_id = t3.person_id and t3.org_id =
t4.org_id;
commit;
exit
```

B.6 Pull-Down Menu for the system

```
1 pulldown menu
RECLASSIFY
 Soilmap &r reclass.aml
COMPARE
 Suitability &r compare.aml
MAIZE
 IK-suitability &r maize.aml
BANANA
 IK-suitability &r banana.aml
CREDIT
 Application &r test1.aml
clear
quit
```

Appendix C. Soilmap Table *(from de Bruin, 1990)*

SciSoil_ID	Unit	Description	Color	Slope	Slope#	Text	Depth	LpH	UpH	Drain#	Stones	Flood#	Cap#
11	Si C	Silencio; highly ondulating	5YR 4/4	5-8 10-16	C	Ac+	>300	4	5	4	<0.01	0	9
12	Si D	Silencio; rough hills moderately escarped	5YR 4/4	10-16 20-30	D	Ac+	>300	4	5	4	<0.01	0	9
13	Si E	Silencio; escarped	5YR 4/4	20-30 45-65	E	Ac+	>300	4	5	4	<0.01	0	9
14	Si+u C-D	Silencio; complex highly ondul. to rough with undiff.groups	5YR 4/4	5-8 20-30 0	CD	Ac+	>300	4	5	4	<0.01	0	9
15	Si+u D-E	Silencio; complex rough to escarped with undiff.groups	5YR 4/4	10-16 45-65 0	DE	Ac+	>300	4	5	4	<0.01	0	9
21	Ne B	Neguev; ondulating	10YR 3/3	1-3 5-8	B	Ac+	>150	4.2	5.4	4	<0.01	0	6
22	Ne C	Neguev; inclined or highly ondulating	10YR 3/3	5-8 10-16	C	Ac+	>150	4.2	5.4	4	<0.01	0	6
23	Ne C-D	Neguev; highly ondul. to rough	10YR 3/3	5-8 20-30	CD	Ac+	>150	4.2	5.4	4	<0.01	0	7
24	Ne D	Neguev; rough or moderately escarped	10YR 3/3	10-16 20-30	D	Ac+	>150	4.2	5.4	4	<0.01	0	7
25	Ne E	Neguev; escarped	10YR 3/3	20-30 45-65	E	Ac+	>150	4.2	5.4	4	<0.01	0	8
31	Mi A	Milano; plains	10YR 4/3	0 1-3	A	FAc-Ac	80-120	4.6	5.8	4	<0.01	0	4
32	Mi B	Milano; ondulating	10YR 4/3	1-3 5-8	B	FAc-Ac	80-120	4.6	5.8	4	<0.01	0	4
33	MiK2 B	Milano; ondulating stony surface	10YR 4/3	1-3 5-8	B	FAc-Ac	80-120	4.6	5.8	4	0.1-3	0	4

Soilmap Table *(cont'n)*

SciSoil_ID	Unit	Description	Color	Slope	Slop#	Text	Depth	LpH	UpH	Drain#	Stones	Flood#	Cap#
34	Mi E	Milano; escarped	10YR 4/3	20-30 45-65	E	FAc-Ac	80-120	4.6	5.8	4	no.data	0	8
35	Mi F	Milano; highly escarped	10YR 4/3	45-65	F	FAc-Ac	50-80	4.6	5.8	4	no.data	0	10
41	Do2	Dos Novillos; superficial	10YR 3/1	0 1-3	A	Fa	25-50	5.5	6	3	<0.01	2	10
42	Do3	Dos Novillos; moderate depth	10YR 3/1	0 1-3	A	Fa	50-80	5.5	6	4	<0.01	1	3
43	Do3K2	Dos Novillos; stony surf, mod.depth	10YR 3/1	0 1-3	A	Fa	50-80	5.5	6	4	0.1-3	0	3
50	Pa	Consociacion Rio Parismina	10YR 3/2	0 1-3	A	Fa-Fl	60-120	6	6.5	4	<0.01	1	2
61	De	Destierro and variants	10YR 3/1	0 1-3	A	F-Fl	50-120	6	6.5	3.5	<0.01	2	6
62	De	Destierro; clayey variant	10YR 3/1	0 1-3	A	FAc	80-120	6	6.5	3.5	<0.01	2	6
63	De2 = =	Destierro;var.clayey stony, superficial	10YR 3/1	0 1-3	A	FAc	25-50	6	6.5	3.5	<0.01	1	10
64	De3 = =	Destierro;var.clayey stony, mod.depth	10YR 3/1	0 1-3	A	FAc	50-80	6	6.5	3.5	<0.01	1	3
65	De4 = =	Destierro;var.clayey stony, deep	10YR 4/1	0 1-3	A	FAc	80-120	6	6.5	3.5	<0.01	1	3
70	Li	Consociacion Ligia	10YR 3/3	0 1-3	A	F-FAc	>90	6	6.5	3	<0.01	0	3
75	LuIII	Consociacion LaLucha mod.good drainage stony surface	7.5YR 2/2	0 1-3	A	Fa-F	50-80	6	6.5	3	<0.01	1	3

Soilmap Table *(cont'n)*

SciSoil_ID	Unit	Description	Color	Slope	Slop#	Text	Depth	LpH	UpH	Drain#	Stones	Flo-od#	Cap #
80	BoIII	Consociacion Bosque; mod.good drainage	10YR 4/3	0 1-3	A	F-Fl	50-80	6	6.5	3	<0.01	1	3
81	Bo	Consociacion Bosque	10YR 4/3	0 1-3	A	FAc-Ac	50-80	6	6.5	2	<0.01	2	6
85	Lu	Consociacion La Lucha	7.5YR 2/2	0 1-3	A	Fa-F	50-80	6	6.5	2	<0.01	2	6
90	Wi	Consociacion Williamsburg	10YR 4/3	0 1-3	A	FAc-Ac	50-80	5.5	6	2	<0.01	2	6
91	Wi2	Williamsburg; stony, superficial	10YR 4/3	0 1-3	A	Fl-FAc	25-50	5.5	6	2	<0.01	1	6
95	Tu	Fluventic Troposaprists	*****	0 1-3	A	no.data	no.data	99	99	1	<0.01	2	9
96	u	Undifferenciated group	*****	0 1-3	A	no.data	no.data	99	99	0	<0.01	3.5	9
100	Fl	Consociacion Flores	10YR 3/2	0 1-3	A	a-aF	5-25	4	4.5	4.5	<0.01	2	10

<u>Drainage classes</u>

0: very poor
1: poor
2: imperfect
3: moderately good
4: good
5: excessive

<u>Flooding classes</u>

0: none
1: occasional (few days in some years)
2: moderate (several days majority of the years)
3: severe (some weeks every year)
4: under water

<u>Black Soils Corresponding with IK Soil Units</u>

Tierra negra alta - Pa, BoIII, Li, Do
Tierra negra mediana - De, Wi, Lu, Do2
Tierra negra baja - Bo, Fl

APPENDIX D

LANDUSE ON SOIL-UNITS

SOIL#	CCT#-class	IK#-class	LANDUSE#		AREA
11	9 Forest	3 Pineapple,chili	0	no data	148,920.02686
12			1	maize	13,021.56106
12			5	chili	4,475.09917
12			8	banana	3,862.62061
12			9	tubers	7,259.86472
12			11	pasture	248,564.54528
12			12	forest	383,935.27386
12			13	swamp	328,384.06355
12			49	coco-tuber	39,415.59694
12			89	banana-tuber	186.01082
13			5	chili	409.56571
13			11	pasture	21,398.25738
13			12	forest	22,006.42997
13			13	swamp	21,371.49940
14			0	no data	450.73355
14			4	coconut	10,410.18910
14			5	chili	4,227.65700
14			8	banana	6,237.47046
14			11	pasture	290,454.11147
14			12	forest	173,523.71004
15			1	maize	277.86567
15			4	coconut	22,208.79613
15			6	pejibaya	18,777.91982
15			7	cocoa	17,086.12736
15			8	banana	5,712.68105
15			9	tubers	14,008.76223
15			10	fruit trees	6,389.62139
15			11	pasture	3,993,450.93035
15			12	forest	534,354.85897
15			13	swamp	651,615.99574
15			14	village	4,314.71338
15			81	maiz-banana	17,722.07084
21	6 Ext.Pasture	2 Palmito,cocoa	0	no data	909,405.06212
21			1	maiz	133,764.13286
21			2	beans	5,863.20648
21			4	coconut	40,544.84871
21			5	chili	75,703.32305
21			6	pejibaya	62,146.93391
21			7	cocoa	130,271.73010
21			8	banana	65,604.51070
21			9	tubers	77,272.60922
21			10	fruit trees	50,255.68794
21			11	pasture	5,342,164.12229
21			12	forest	2,350,741.45447
21			13	swamp	53,921.05895
21			14	village	219,798.66169
21			49	coco-tubers	171.66306
21			78	banana-cocoa	78,837.61734
21			81	maiz-banana	680.12850
21			89	banana-tubers	17,905.71125
22			0	no data	286,022.28548
22			1	maiz	14,464.18367
22			2	beans	158.73277
22			4	coconut	5,905.73399

22				5	chili	10,549.67468
22				7	cocoa	8,926.53516
22				8	banana	18,918.07974
22				9	tubers	7,457.30830
22				10	fruit trees	3,882.80854
22				11	pasture	858,838.76761
22				12	forest	302,221.11664
22				13	swamp	40,281.89175
22				14	village	2,458.35279
23	7	Tree crops	2	0	no data	659,535.92627
24				0	no data	336,227.30273
24				1	maiz	7,612.75294
24				4	coconut	2,982.74788
24				5	chili	24,131.69883
24				6	pejibaya	104.29486
24				7	cocoa	12,914.34397
24				8	banana	8,954.37950
24				9	tubers	6,898.22381
24				11	pasture	695,297.64165
24				12	forest	259,280.41300
24				13	swamp	69,942.90639
24				78	banana-cocoa	13,075.88418
24				89	banana-tubers	3,663.38978
25	8	Int.forest	2	0	no data	4,562.99645
25				1	maiz	44,983.38216
25				2	beans	14,172.68253
25				4	coconut	285.95949
25				5	chili	49,807.26935
25				6	pejibaya	12,551.04685
25				7	cocoa	35,451.62733
25				8	banana	49,435.16343
25				9	tubers	90,978.56846
25				10	fruit trees	16,624.73141
25				11	pasture	3,109,026.16859
25				12	forest	1,955,228.47198
25				13	swamp	10,449.53660
25				14	village	86,239.72659
25				49	coco-tubers	11,426.94193
25				78	banana-cocoa	47,597.52844
25				81	maiz-banana	7,063.13376
25				89	banana-tubers	15,853.35496
31	4	Perm/semi-perm	2	0	no data	17,512.47391
31				1	maiz	63,684.90413
31				3	pineapple	11,661.52872
31				5	chili	15,623.98748
31				8	banana	10,330.80449
31				9	tubers	2,043.20533
31				10	fruit trees	625.90672
31				11	pasture	554,528.56277
31				12	forest	67,039.59232
31				14	village	257,584.48689
31				84	banana-coconut	718.09140
32				0	no data	3,905.44453
32				1	maiz	12,534.44093
32				2	beans	2,743.04754
32				4	coconut	6,196.18269
32				5	chili	28,755.41708
32				9	tubers	18,866.28234
32				10	fruit trees	28,299.72596

32			11	pasture	1,109,425.57066
32			12	forest	30,017.55912
32			14	village	2,877.83325
32			78	banana-cocoa	48.71382
32			89	banana-tubers	2,336.24003
33			11	pasture	4,617.41973
33			12	forest	10,956.65010
34	8 Int.Forest	2 Palmito,cocoa	0	no data	28.62034
34			1	maize	10,376.28648
34			5	chili	1,567.45112
34			11	village	328,064.20874
34			12	forest	20,797.59154
34			14	village	32,609.86737
34			89	banana-tubers	3,116.81010
35	10 For protection	2	0	no data	1,628.98439
35			1	maize	18,529.91917
35			3	pineapple	3,117.53315
35			10	fruit trees	3,997.18595
35			11	pasture	143,065.61463
35			12	forest	5,787.99485
35			14	village	40,214.50476
35			84	banana-coconut	5,273.96320
41	10 For protection	1 maiz,rice,beans	0	no data	2,963.82748
41			11	pasture	6,732.44759
41			12	forest	129,897.91000
42	3 Annuals	1	0	no data	90,521.01905
42			1	maize	5,947.13456
42			11	pasture	187,603.88965
43			1	maize	1,273.16821
43			11	pasture	46,158.55170
50	2 Annuals	1	0	no data	2,389,623.98743
50			1	maize	34,025.31507
50			7	cocoa	3,805.13767
50			8	banana	33,465.88651
50			9	tubers	70,391.77404
50			11	pasture	167,617.64180
50			12	forest	182,747.19858
50			89	banana-tubers	13,226.45367
61	6 Ext.Pasture	1	0	no data	5,708.94829
61			1	maize	121,291.15273
61			3	pineapple	758.60587
61			5	chili	2,079.38499
61			7	cocoa	22,005.31319
61			8	banana	20,703.69350
61			9	tubers	19.47348
61			10	fruit trees	3,296.38229
61			11	pasture	2,123,715.61061
61			12	forest	207,546.38319
61			13	swamp	10,568.09738
61			14	village	154,721.94771
61			81	maiz-banana	11,002.20131
61			84	banana-coconut	3,427.42155
62			0	no data	3,078.04077
62			1	maiz	59,199.26271
62			2	beans	14,417.10831
62			7	cocoa	23,267.12298
62			8	banana	8,748.70458
62			9	tubers	12,380.48443

			11	pasture	484,578.90640
62			12	forest	1,125,322.01521
62			13	swamp	16,927.78738
62			0	no data	230.64523
63	10 For protection	1	9	tubers	850.59071
63			11	pasture	128,358.84832
63			12	forest	16,819.36282
63			5	chili	4,737.33917
64	3 Annuals		11	pasture	91,708.58509
64			1	maize	3,696.88055
65			11	pasture	27,013.87419
65			12	forest	2,616.36117
65			0	no data	1,221,640.52441
70			0	no data	233,994.66455
75			0	no data	611,573.65967
80			0	no data	326,799.52148
81	6 Ext.Pasture	1	1	maize	67,740.43511
81			8	banana	54,774.71414
81			9	tubers	10,000.62190
81			11	pasture	194,221.15122
81			12	forest	104,152.00592
81			89	banana-tubers	34,813.78803
85			0	no data	966,214.17286
90			0	no data	196,533.50242
90			1	maize	14,037.89294
90			9	tubers	20,746.55094
90			11	pasture	232,388.28963
90			12	forest	497,381.95925
90			13	swamp	7,804.88810
91			1	maize	1,067.84703
91			5	chili	2,797.63460
91			11	pasture	193,480.36435
91			12	forest	2,465.74209
95	9 Ext.Forest	suamposa	11	pasture	157.23643
95			12	forest	44,536.41032
96			0	no data	903,982.56634
96			1	maize	35,328.47189
96			5	chili	9,668.85358
96			6	pejibaya	271.68653
96			7	cocoa	23,058.58652
96			8	banana	9,685.56987
96			9	tubers	13,696.19402
96			10	fruit trees	2,998.91485
96			11	pasture	845,420.46473
96			12	forest	802,844.05089
96			13	swamp	1,294,032.27773
96			49	coco-tubers	11,002.31783
96			81	maiz-banana	7,085.68700
96			89	banana-tubers	4,537.76581
100	10 For protection	1	0	no data	128,627.99131
100			12	forest	6,534.34105