GEOGRAPHICAL INFORMATION SYSTEMS AS A TOOL TO EXPLORE LAND CHARACTERISTICS AND LAND USE

with reference to Costa Rica

Promotoren:

Dr. Ir. J. Bouma, hoogleraar in de bodeminventarisatie en landevaluatie, speciaal gericht op de (sub)tropen

Dr. Ir. L.O. Fresco, hoogleraar in de tropische plantenteelt, met bijzondere aandacht voor de plantaardige produktiesystemen

GEOGRAPHICAL INFORMATION SYSTEMS AS A TOOL TO EXPLORE LAND CHARACTERISTICS AND LAND USE with reference to Costa Rica

Jetse J. Stoorvogel

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
Dr. C.M. Karssen,
in het openbaar te verdedigen
op woensdag 18 oktober 1995
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

CIP-gegevens Koninklijke Bibliotheek, Den Haag

Stoorvogel, J.J., 1995.

Geographical information systems as a tool to explore land characteristics and land use with reference to Costa Rica / J.J. Stoorvogel. - [S.1. : s.n.].
Thesis Wageningen. - With ref. - With Summary in Dutch
ISBN 90-5485-449-9
Subject headings: GIS / land use inventory / soil survey / Costa Rica / sustainability

Printed in the Netherlands by Drukkerij Modern, Bennekom

Propositions

1. A soil survey database can only be structured and organized effectively after a thorough analysis of future applications.
   - this thesis

2. For the analysis of land use dynamics, Markov chains need probability modifiers to interpret the observed sequences.
   - this thesis

3. A modular approach to link GIS and external models based on data exchange is preferred above full integration.
   - this thesis

4. In contrast to the increasing attention for special-purpose soil surveys at detailed scale levels, it can be argued that general-purpose soil surveys are essential for screening potential study areas and planning sampling schemes.
   - this thesis
   - Burrough, P.A., 1991. Soil Information Systems. In: D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Eds). Geographical information systems. Longman Scientific & Technical, Harlow, United Kingdom, 153-169.

5. The statement of Yearsley *et al.* (1994) that dual GIS architectures - i.e. architectures with separate storage of geometric data and attribute data - are clearly undesirable originates from a purely technology driven point of view.
   - C. Yearsley, M.F. Worboys, P. Story, D.P.W. Jayawardena and P. Bofakos, 1994. Computational support for spatial information handling: models and algorithms. In: M.F. Worboys (Ed.). Innovations in GIS. Taylor & Francis Ltd., London, United Kingdom, pp: 77
   - This thesis

6. The development of practical tools to define and identify scales in multi-scale GIS studies requires a clear definition of scale within GIS applications.
   - H.M. Hassan and C. Hutchinson (Eds), 1994. Natural resource and environmental information for decisionmaking. The World Bank, Washington, D.C., USA.
   - W. Andriesse, L.O. Fresco, N. Van Duivenbooden, and P.N. Windmeijer, 1994. Multi-scale characterization of inland valley agro-ecosystems in West Africa. *Netherlands Journal of Agricultural Sciences* 42: 159-179.

7. Simple deterministic models to estimate sustainability indicators (e.g. NUTDEP, QUEFTS, USLE) are essential to support regional planning exercises.
   - J.J. Stoorvogel, E.M.A. Smaling, B.H. Janssen, 1993. Calculateing soil nutrient balances in Africa at different scales. I Supra-national scale. *Fertilizer Research* 35: 227-235.
   - B.H. Janssen, F.C.T. Guiking, D. Van Der Eijk, E.M.A. Smaling, J. Wolf, and H. Van Reuler, 1990. A system for quantitative evaluation of the fertility of tropical soils (QUEFTS). *Geoderma* 46, 299-318.
   - W.H. Wischmeijer, and D.D. Smith, 1978. Predicting rainfall erosion losses. A guide to conservation planning. Agric. Handbook No 537. USDA, Washington, D.C., USA

8. Including only sustainable land use systems in the analysis of alternative land use scenarios can not yield realistic scenario results.
   - F.R. Veeneklaas, H. Van Keulen, S. Cissé, P. Gosseye, and N. van Duivenbooden, 1994. Competing for limited resources: options for land use in the fifth region of Mali. In: L.O. Fresco, L. Stroosnijder, J. Bouma, and H. Van Keulen (Eds). The future of the land: mobilising and integrating knowledge for land use options. John Wiley & Sones Ltd, Chisester, United Kingdom, 227-247
   - Jansen, D.M., J.J. Stoorvogel, and R.A. Schipper. 1995. Using sustainability indicators in agricultural land use analysis: an example from Costa Rica. *Netherlands Journal of Agricultural Science* 43: 61-82.

9. The general opinion that pesticide use in the Atlantic Zone of Costa Rica has strong negative effects on the environment is not supported by any research data.

10. Short term research contracts will lead to more diverse researchers and more dynamic research departments. Additionally, it improves interaction between departments and universities due to exchange of researchers.

11. Do not refrain from modelling. It is just a formalisation of what everybody has been doing for centuries.

Propositions accompanying the Ph.D thesis 'Geographical information systems as a tool to explore land characteristics and land use with reference to Costa Rica'. Jetse J. Stoorvogel, Wageningen, October 18, 1995.

Preface

When Christopher Columbus reached Costa Rica in search for the wealth of the Indies, the need for geographical information systems (and especially global positioning systems) was already evident. Nevertheless, the knowledge based systems used by the conquistadors led to a rapid inventory of the earth surface and form the basis for present day geography. With the high pressure on land in many parts of the world, land use planning requires a formalisation of the knowledge based systems with clearly defined data and relations available for every user.

Almost 500 years after Columbus, I landed on the American continent in search for tropical soils and their relation with present day land use. Impressed by Columbus' mistakes and successes in combination with the rapid developments in information sciences, I started this thesis research. I would like to thank my promotors Johan Bouma and Louise O. Fresco for initiating this research and their continuing support. Their enthusiasm and encouragement were highly motivating. Despite the large distance, they were always willing to comment and discuss articles and parts of this thesis.

The USTED methodology was developed together with Don Jansen and Rob Schipper. They are greatfully acknowledged for the stimulating interdisciplinary research. The discussions with André Nieuwenhuyse on soils and land use in the Atlantic Zone were highly motivating. Many M.Sc. students came to Costa Rica for their practicals or thesis research. Jeroen van Alphen, Randy Benjamins (GIS database), Marleen Belder (land use inventory), and Jacomijn Pluimers (biocide modelling) are acknowledged for their contributions to specific parts of this thesis. Luis Guillermo Valverde and Luis Guillermo Quirós are greatfully acknowledged for data collection. Hans Jansen is acknowledged for both coordination at the Atlantic Zone Programme and comments on parts of this thesis. Peter Burrough is greatfully acknowledged for commenting on a previous version of this thesis. Rob Sevenhuysen, Olga Carvajal, Fernando Cambronero, Celia Alfaro, Miguel Astua, and Edgar Alfaro are acknowledged for the logistic support at the Atlantic Zone Programme. This research would not have been possible without the endless support of Marjon Oostrom who took care of the logistics in my personal life. Both *paranimfen* Eric Smaling and Wim Andriesse are acknowledged for their friendship and teaching me the basic concepts of science in the beginning of my career. I am very thankful to the co-authors of my papers for the fruitful discussions that lead to the specific papers. Finally, I thank my family for their moral support and for trying to understand my work.

Table of Contents

1. Introduction

1.1 Soil and land use problems as a source for improving geographical information systems

Although the world's total area under crop land and permanent pasture has increased only slightly by 2% per year during the last two decades, there has been a significant increase in agricultural production (FAO 1974, 1992). The resulting intensification of world's agriculture leads to a growing concern about the sustainability of agricultural production and its environmental effects. Agriculture is challenged to deal with an increasing demand for agricultural products in future and more stringent environmental constraints. The goals of farmers, who are the final decision makers in agriculture, do not necessarily correspond with the general objectives of policy makers and environmentalists. The decisions of these farmers, however, can be directed by agricultural policies, regulations and incentives to match them more adequately with regional and national objectives (e.g. Lutz and Daily 1991). Additionally, agricultural sciences can provide alternative technologies, which, if suitable for the farmer's situation, may be adopted.

In Costa Rica, increases in the agricultural production were, until the early eighties, mainly the result of an expansion of the production area rather than an increase of productivity (Hartshorn *et al.* 1982). Deforestation rates have decreased since then to less than 0.2% per year (Kaimowitz 1994) and actually most primary forests are found in the protected areas (national parks, forest reserves) which cover approximately 20% of the Costa Rican territory (Ramirez and Maldonado 1988, Alvarado *et al.* 1993, Fournier 1993, Sader and Joyce 1988). As a result only a 5% expansion of the agricultural area took place in the last decade (estimate based on Lizano 1993 and Fournier 1993). Nevertheless, annually a 4% increase in agricultural production was reached in the last decade mainly as the result of alternative varieties, higher inputs and changes in the cropping pattern (estimate based on data from Lizano 1993, Anonymous 1994).

Although the agricultural production is increasing, it is lagging behind population growth in many parts of the world. In this rather hazardous situation, there is no room for a process of trial and error to develop agriculture. Agricultural sciences, therefore, need to develop tools for an *ex ante* evaluation of policies and regulations. Also, alternative technologies should be thoroughly tested at both field and farm level before supplying them to large groups of farmers. This *ex ante* evaluation requires a basic understanding of natural resource processes and the driving forces behind land use changes.

Fortunately new tools like Geographical Information Systems (GIS[1]), simulation models and linear programming (LP) models are being developed and continuously improved. These tools allow land use planners to explore different land-use options. In addition, they support the evaluation of incentives and measures to direct land use changes according to various policies.

GIS are already routinely used to store, manage and analyze spatially related data (Hassan and Hutchinson 1992). Nevertheless, no standard procedures have been developed to include GIS in disciplinary methodologies dealing with spatial data. For instance, a standard procedure for soil surveying developed by Soil Survey Staff (1951) has been adapted and updated by different authors (Dent and Young 1981, Landon 1991). Studies have been carried out to include GIS technology in soil survey procedures and GIS based soil survey databases like SOTER (Van Engelen and Wen 1995) and STATSGO (Soil Survey Staff 1993) are being developed. However, no procedures have been adopted as a new standard by the different surveyors. Methodologies for land cover and land use inventories are currently emerging (Turner *et al.* 1994). Due to the regular use of satellite imagery, which comes in digital format, these inventories may employ GIS technology. However, similarly to soil surveys, no standardisation of methodologies takes place. For regional soil surveys and land use inventories, GIS technology is often only employed as a computer-aided mapping tool.

Policy makers are becoming increasingly aware of the environmental effects of agricultural production, and sustainability is increasingly becoming a policy objective (Farshad and Zinck 1993). The analysis of soil survey and land use inventories is, therefore, often focused on sustainability related topics. Many definitions for sustainable development and the sustainability of agricultural production exist in the literature (FAO 1993, Lélé 1991). In general, only few relevant and quantifiable indicators can be operationalized in agricultural land use analysis (Jansen *et al.* 1995). The inventories of both land and land use should enable a geo-referenced analysis of these indicators to allow for the incorporation of these parameters in land use planning. This may require a linkage between the models estimating these sustainability indicators and GIS.

"Information sciences" develop GIS technology almost independently from the applications. Commonly used GIS packages like PC Arc/Info[2] provide relatively few tools for spatial analysis (only 5% of all PC Arc/Info commands are related to spatial

[1] Following Bonham-Carter (1994), the acronym GIS is used for either a single geographical information system, or several systems, or to the field of geographical information systems as a whole.

[2] Arc/Info and PC Arc/Info are registered trade marks of Environmental Systems Research Institute, Inc., Redlands, CA., USA.

analysis). This probably originates from the wide range of applications that different disciplines might give a GIS, leaving companies like ESRI with an impossible task to include all required operations. The packages do, however, increasingly facilitate links with external models, which can be developed by different disciplines.

From both sides, GIS users and GIS developers, the developments can be characterized as "technology driven". It is necessary to take a more "application driven" approach where disciplines focus on the application and adoption of standard GIS packages, made available by the information scientists. Additionally, they should identify the requirements of GIS for their applications and feed them back to information scientists. This interaction is crucial to avoid sterile, purely "technology driven" approaches. Information scientists should certainly continue to do basic work but the efficiency of their work would increase when fed by problems of the real world.

In contrast to a general impression that GIS is high-tech and unsuitable for developing countries (Taylor 1991), in practice an increasing use of these systems can be observed in these countries. GIS technology is found in most Costa Rican organizations dealing with spatial data. Almost all organizations use commercial GIS packages like PC Arc/Info and occasionally IDRISI[3]. The organizations focus on GIS-supported applications and do not deal with the internal organization of the GIS. Also in the Costa Rican context a technology driven use of GIS can be observed where GIS is mainly used to make sophisticated maps. Less or no attention at all is paid to data quality and to the systematic analysis of spatial data.

1.2 This thesis

The use of GIS for the inventory and analysis of land characteristics and land use receives increasing attention in literature (e.g. Bonham-Carter 1994, Michener *et al.* 1994, Maguire *et al.* 1991). However, the proposed techniques are often not applicable using commercial GIS packages as they require adaptions to the internal organization of the GIS. This thesis deals with the use of commercial GIS software for the storage and analysis of land characteristics and land use. Specific research topics for the study include:

- the optimization of data storage on the basis of possible use,
- the quantification of temporal dynamics in land use, and

[3] IDRISI is a registered trademark of the IDRISI project/Clark University, Worcester, MA, USA.

- the integration of GIS and other tools and procedures for the analysis of land use scenarios.

Several aspects related to the use of GIS are outside the scope of this thesis, although they may have a significant importance:

- As indicated by, for example, Bregt (1992), GIS can be used to support the optimization of sampling schemes. Most studies make use of geostatistics and apply to detailed scale levels. At smaller scales, different mapping units are often delineated on the basis of aerial photographs. The present study deals mainly with existing databases at regional scales, where the inventories of land and land use are carried out using aerial photographs.
- The uncertainties and quality of data may influence significantly the results of any study. Few independent parameters have been developed to indicate the accuracy of spatial data and its effect on modelling exercises. Studies like for example Heuvelink (1993) may contribute significantly in the operationalisation of these parameters.

In this thesis, approaches to the use of GIS for land and land use related studies are developed. The approaches are explored and illustrated with examples from the Northern Atlantic Zone in Costa Rica (Figure 1.1). Special attention is paid to the sustainability related side of land and land use analysis. The research is part of an interdisciplinary research programme named the Atlantic Zone Programme (AZP). This programme is a cooperation of the Tropical Agricultural Research and Higher Education Centre (CATIE, Costa Rica), the Ministry of Agriculture and Livestock (MAG, Costa Rica), and the Wageningen Agricultural University (WAU, The Netherlands).

Commercial GIS software often has a dual architecture, with separate storage of spatial and attribute data. For users of the software, the structures for the spatial data are often fixed, and only the attribute data can be structured user specifically. Therefore, database structures in this thesis focus on structures for attribute data. Structures for soil survey data are often complex due to the occurrence of soil associations and soil complexes. In Chapter 2, alternative database structures for soil survey data are proposed and evaluated based on a general data model and different indicators for the efficiency of databases for queries. Applying the database for different modelling approaches to estimate biocide leaching, these structures are found to be rigid in terms of the level of detail they provide. Therefore, decision rules were developed enabling different applications of soil survey data at different levels of detail.

Land cover and land use databases have relatively simple legends. Consequently the thematic database structures can be relatively simple. Due to their great temporal

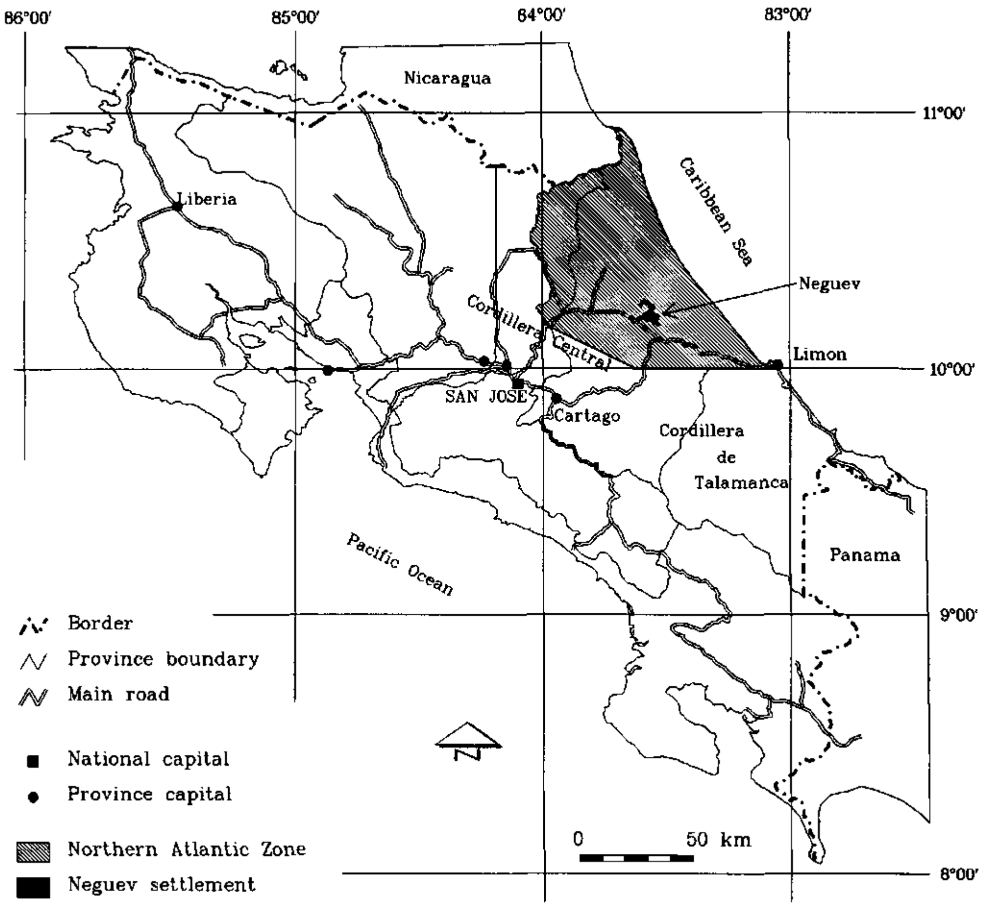


Figure 1.1 Location of the Northern Atlantic Zone and the Neguev settlement in Costa Rica

variation, however, land cover and land use data have an additional dimension. On a large scale this variation comprises cropping sequences, whereas on a small scale it includes broad land cover changes. The quantification of land use dynamics using both standard and new indicators is discussed in Chapter 3.

When the tools for spatial analyzes are not provided by a commercial GIS package, the GIS can be linked to external models. This is generally the case when sustainability related topics are included in the analysis. Chapter 4 provides general structures for the GIS-model link and illustrates them with an example, where GIS is linked with a LP model. In addition, two examples are further elaborated to i) optimize the distribution of land use in a given area to reduce soil nutrient depletion, and ii) analyze alternative land use scenarios through systems integration. The latter forms the basis for an exploratory methodology with which the effects of policies and incentives can be estimated.

Chapter 5 evaluates the use of GIS databases and data needs for land use analysis. Four practical examples from the Atlantic Zone of Costa Rica are presented, covering respectively: i) identification of potential areas for maize cultivation, ii) problems with sustainability, iii) the risk of ground and surface water contamination with a commonly used nematicide, and iv) the analysis of alternative land use scenarios.

Chapter 6 lists future challenges and presents general conclusions.

1.3 The study area

The study area comprises the perhumid tropical lowlands in the northern part of the Atlantic Zone of Costa Rica measuring approximately 5,450 km^2. An extensive spatial database that comprises data on the natural resources, agricultural land use, and the human environment is available for the area (Stoorvogel and Eppink 1995). This makes the area extremely useful for this specific study. Land use data for the whole study area are only available for 1984. Studies to land use dynamics (Chapter 3) are, therefore, carried out for parts of the area where aerial photographs for different years were available. More detailed studies presented in Chapter 4 are carried out for the Neguev settlement (Figure 1.1). This settlement comprises 47 km^2 and is located on the footslopes of the Turrialba volcano.

<u>The Northern Atlantic Zone of Costa Rica</u>

The perhumid tropical lowlands in the northeast of Costa Rica form the continuation of the Nicaragua basin, a subsidence basin filled with alluvial and marine deposits. The basin is limited in the southwest by the Central and Talamanca Mountain Ranges (Figure 1.2). Active volcanism is found in the Central Mountain Range. In the north, a number of basaltic cones are found. Along the coast, marshy backswamps are located.

The climate is characterized by water excess all year around with less precipitation in February and March and a mean annual rainfall between 3300 and 7000 mm (Stoorvogel and Eppink 1995). Temperatures vary little throughout the year with an average annual temperature of approximately 24° C in the lowlands decreasing with 0.42 °C per 100 m rise in altitude (Herrera 1985).

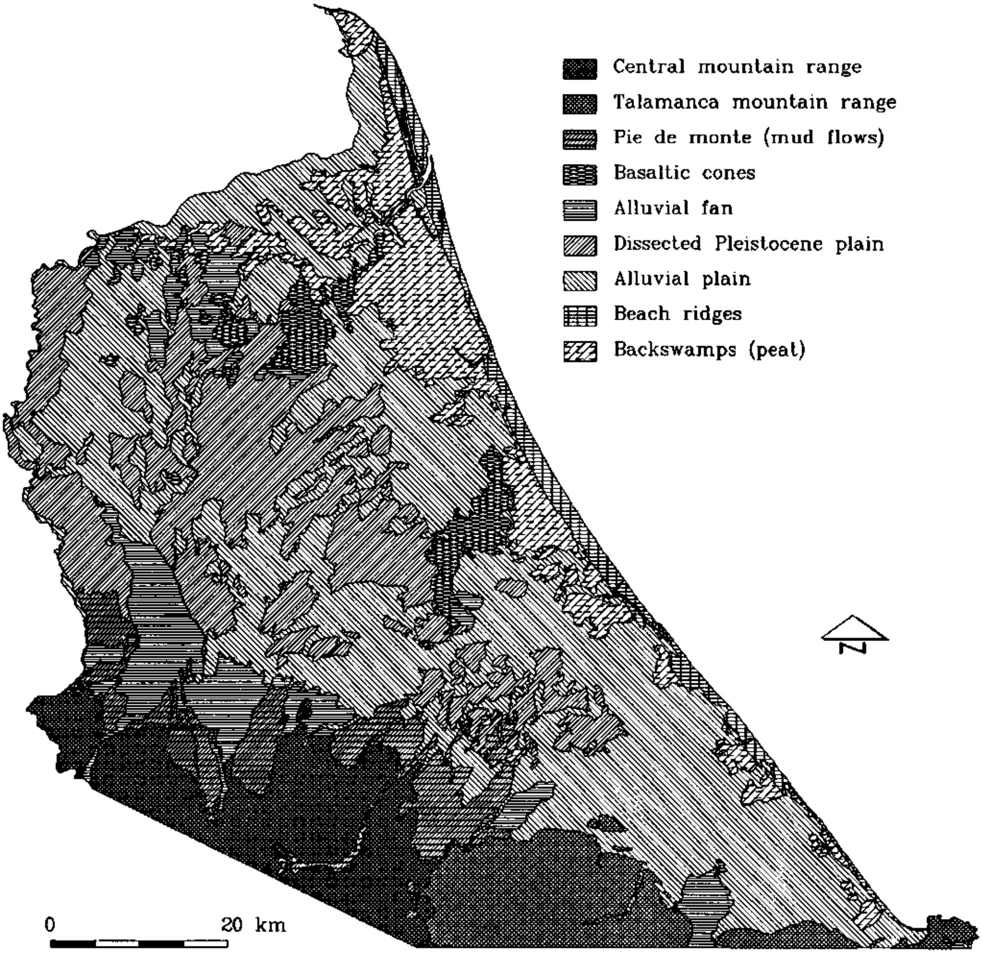


Figure 1.2 Main geomorphological units in the Atlantic Zone

The soils in the Atlantic Zone (Figure 1.3) can be described and classified (Soil Survey Staff 1994) as:

S1: old, strongly weathered, clayey and well drained soils (oxic Humitropept and Haploperox) on mudflows both on the footslopes and as remnants in the alluvial plains (23.4%),

S2: old, moderately weathered, sandy and moderately well drained soils (aeric Tropaquept and aquic Humitropept) developed in sedimentary rock in the Talamanca Mountain Range (3.5%),

Figure 1.3 General soil groups in the Northern Atlantic Zone (based on Wielemaker and Vogel 1993)

S3: young, well drained soils (andic Tropopsamment and andic Dystropept) developed in young alluvial deposited sediments of volcanic origin (25.5%),

S4: young, slightly weathered poorly drained soils (Tropaquept) developed in sandy, volcanic sediments from the Central Mountain Range (2.1%),

S5: young, slightly weathered poorly drained soils (Eutropept) developed in fine textured sediments (24.7%), S6: young, slightly weathered moderately well to well drained soils (Dystropept and Tropaquept) developed in sandy to loamy sediments from the Talamanca Mountain Range (4.9%),

S7: soils (Hydrudand) developed in volcanic ashes under extremely humid conditions (3.3%), and

S8: peat soils (Histosol) developed in the coastal backswamps (12.6%).

A number of soils occur in very small areas. A more generalized classification was therefore based on soil fertility and soil drainage. The latter identifies fertile, well drained soils, fertile poorly drained soils, and infertile, well drained soils. These three groups correspond roughly with group S3, S5 and S1, respectively.

In previous centuries a dispersed Indian population was found in the area. However, during Spanish colonization the area was found to be practically inhabitated. Major colonization in the area started with the construction of the railroad in 1865. The railroad was constructed for the transport of the coffee harvest from the higher areas and was located on the boundary between the footslopes and the alluvial fan. Besides being the most suitable area for the construction of the railroad, it crossed the area with fertile, well drained soils, which was suitable for banana production. Starting on the footslopes, colonization took place mostly in northern direction into the alluvial plains. At the moment most of the Atlantic Zone outside the protected areas is colonized.

Agricultural land use in the Atlantic Zone (Figure 1.4) ranges from extensive cattle raising and breeding to intensively managed plantations for banana and palm heart production. Farms vary between big plantations and small farms. The latter are often organized in settlements schemes of the Institute for Agricultural Development (IDA *Instituto de Desarrollo Agropecuaria*).

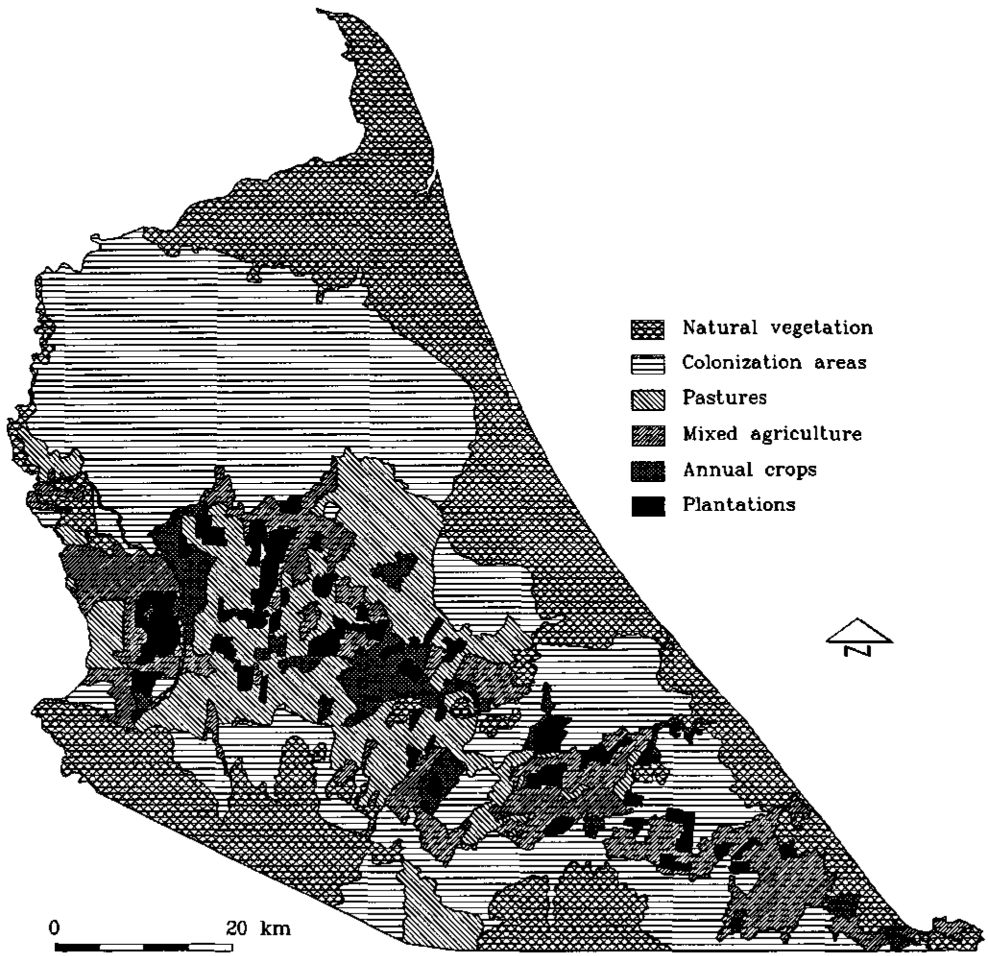


Figure 1.4   1984 land cover in the Northern Atlantic Zone

The Neguev settlement

The Neguev settlement is located on the footslopes of the Turrialba volcano, north of Guápiles-Limon highway (Figure 1.5). The settlement is managed by IDA, which is the main organization dealing with the reorganization and management of agricultural settlements. IDA settlements cover almost 20% of the northern part of the Atlantic Zone (Stoorvogel and Eppink 1995). A full description of the Neguev settlement is given by De Oñoro (1990). The spatial database for the Neguev settlement comprises

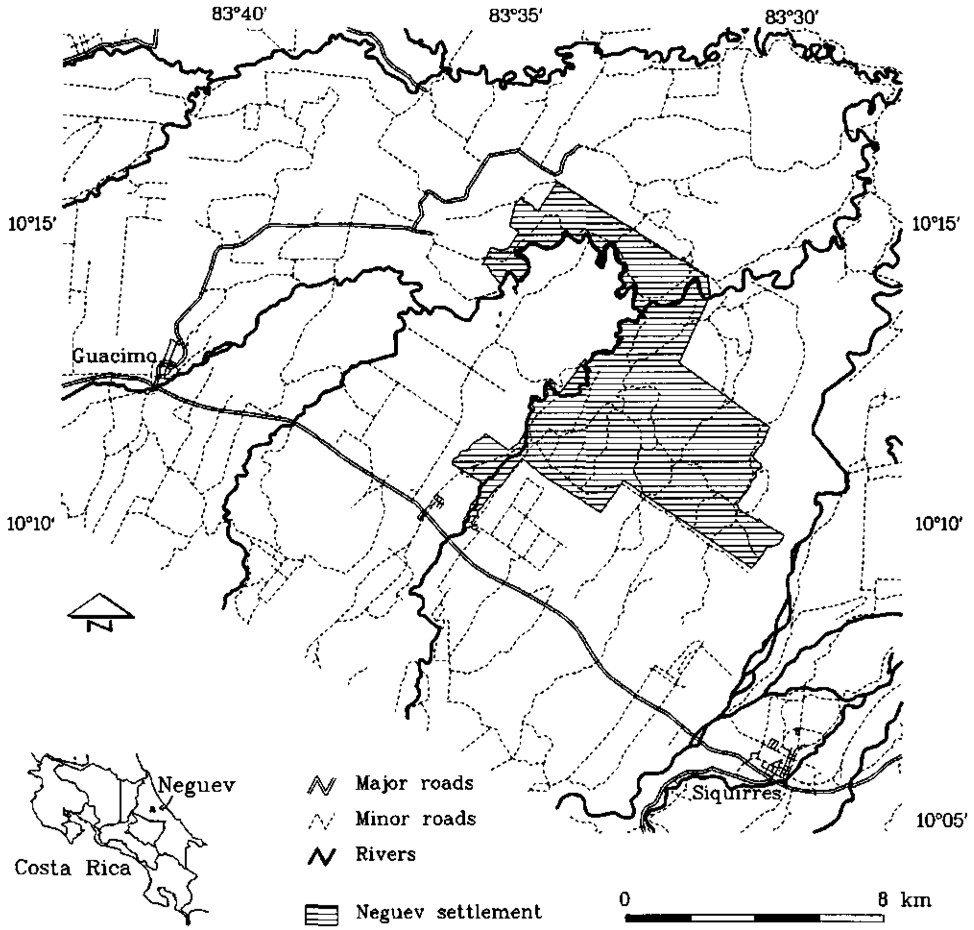


Figure 1.5 Location of the Neguev settlement

a 1:20,000 soil survey, a map with the location of the different farms and for the southern part of the settlement a land use map for 1986.

A semi-detailed soil map of the settlement (1:20,000) was made by De Bruin (1992). The soil map was generalized on the basis of soil fertility and drainage (Figure 1.6). The resulting four soil groups can be described and classified (according to the Soil Survey Staff 1994) as

SFW: young, well drained volcanic soils with a high soil fertility (andic Eutropept, typic Udivitrand), SFP: young, poorly drained volcanic soils with a high soil fertility (aquandic Tropaquept),

SIW: relatively old, well drained soils with a low soil fertility developed on mud flows and Pleistocene alluvial deposits (oxic Humitropept and Haploperox), and

P: swamps.

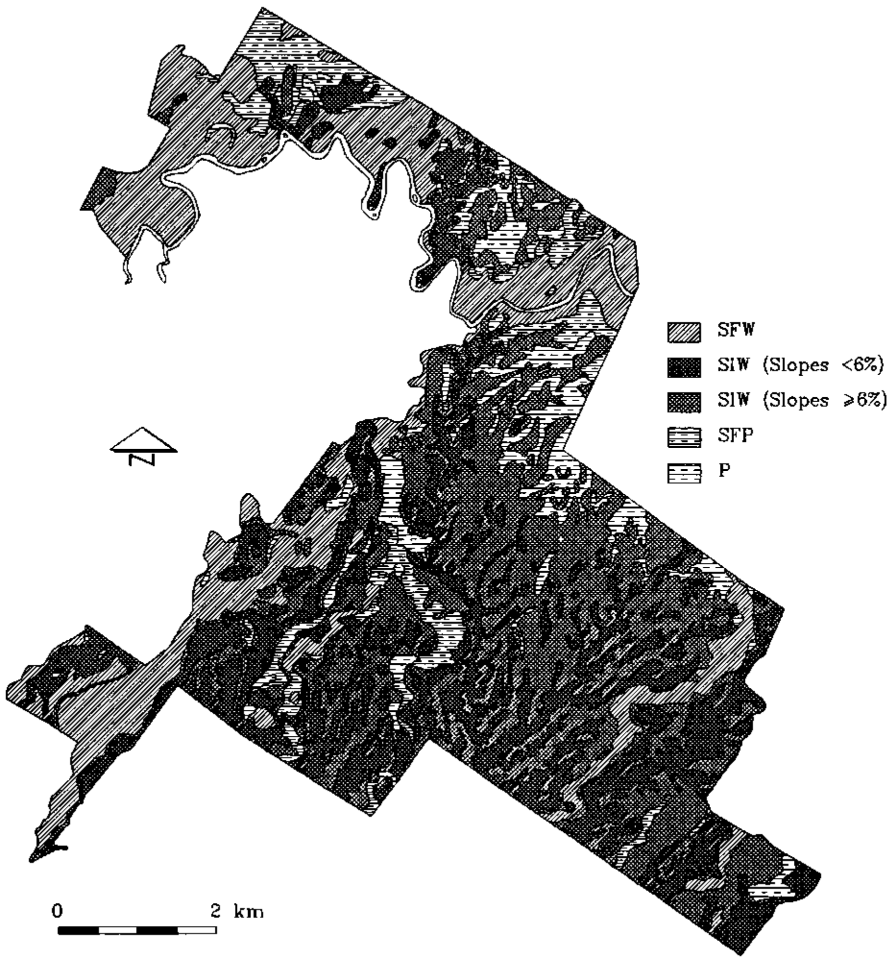


Figure 1.6 General soil groups in the Neguev settlement (after De Bruin 1992)

Originally the Neguev settlement was a large cattle ranch with some smaller parts under forest. In 1979 the settlement was occupied by settlers after which IDA intervened, bought the farm and took care of parcelling. Although the settlement scheme was established a decade ago, pasture used for extensive cattle breeding (with a cattle density of one head per ha) and to a smaller extent forest still dominate land use in the area (Figure 1.7). Yet, smaller areas are presently cultivated with maize, red pepper, tubers, coconut, cacao, plantain and fruit trees. The annual and perennial crops are scattered on small parcels throughout the area. Recently the cultivation of palm heart expanded rapidly.

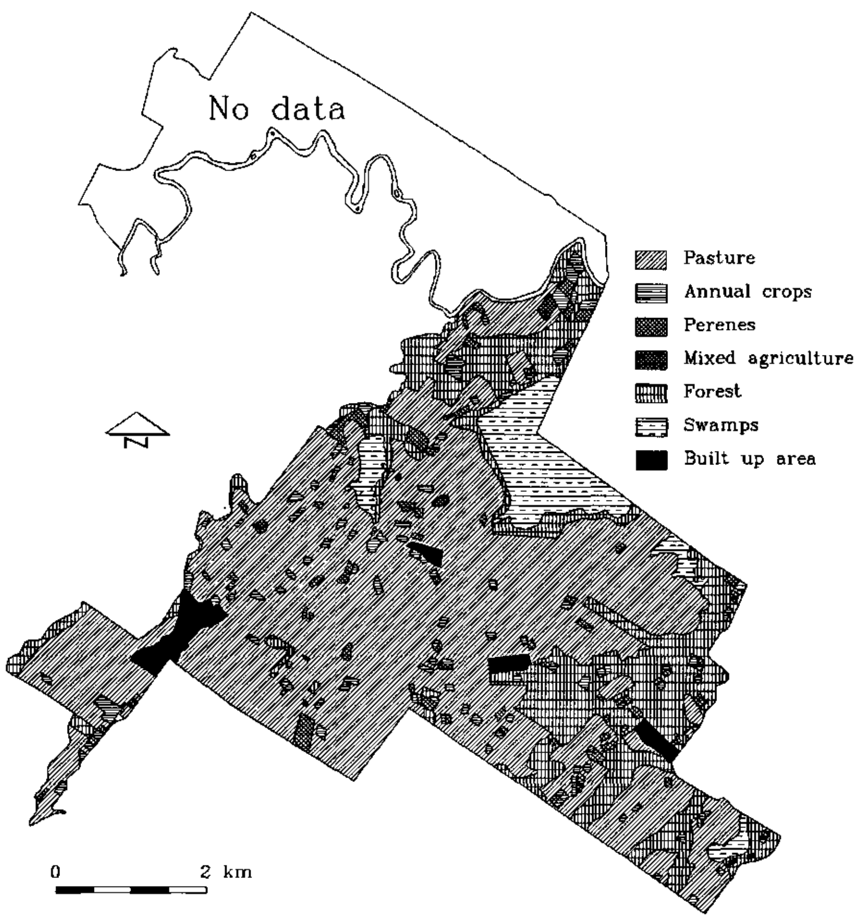


Figure 1.7 1986 land use for the southern part of the Neguev settlement (Overtoom *et al.* 1987)

2. The storage of soil survey data

The Sections of this chapter are based on the following publications:

Section 2.2: Stoorvogel, J.J., and M. Molenaar, 1995. Soil associations and complexes in GIS based soil and terrain databases: a Costa Rican case study. Submitted to *International Journal on Geographical Information Systems.*

Section 2.3: Stoorvogel, J.J., and J. Bouma, 1995. A multi-level soil information system to estimate biocide leaching in Costa Rica. Submitted to *Soil Science Society of America Journal.*

parameters are required that indicate the efficiency and flexibility of the database structure for certain queries.

In this section, a data model for the attribute data of soil and terrain databases is presented. Alternative database structures and quantitative parameters to determine the efficiency of queries are used to put the data model into practice. The setup of a soil and terrain database is illustrated with a five step approach for soil survey data of the Northern Atlantic Zone. The 1:150,000 soil map with its database (based on Wielemaker and Vogel 1993) is stored in PC Arc/Info and serves for the different disciplines of the AZP besides other research projects and organizations.

2.1.2 The data model

GIS database structures are partially determined by the GIS architecture. Two major GIS architectures, raster and vector, can be distinguished. Raster based systems are established by a link between position and thematic data. For vector based systems the object is described by geometric and thematic data (Molenaar 1991). The geometric and thematic object descriptions can be handled independently and are linked by a feature identifier. For soil survey data, the objects are the delineated areas of the soil map. The shape and position of the delineated areas are described by the geometric data, whereas the soil distribution and additional properties (like geology and geomorphology) are described by thematic attributes. Compared to the traditional soil map, the feature identifier is the identification soil legend (Soil Survey Staff 1951) that links delineated areas to the thematic object description within the database of non-spatial attributes (Burrough 1986). This section is focused on the vector approach and deals with the organization of the thematic object description. In most cases, the structure for geometric data is fixed in commercial GIS packages. Similar attribute structures can, therefore, be used for grid based systems.

In the data model proposed for soil and terrain databases (Figure 2.1) the delineated areas (DAs) are classified as mapping units (MUs), i.e. DAs with similar attributes (e.g. soil distribution). If the MUs comprise soil associations, terrain units (TUs) are described. These TUs are individual components of the association. TUs are described by one or several sub-units like for instance a geological unit (GeU), a geomorphological unit (GmU) and a soil unit (SU). In soil and terrain databases, emphasis is placed on the latter with more general descriptions for GeUs and GmUs. The properties of SUs will be described in detail with reference to representative profiles for the different soil series and descriptions of the soil horizons. GeUs are, for

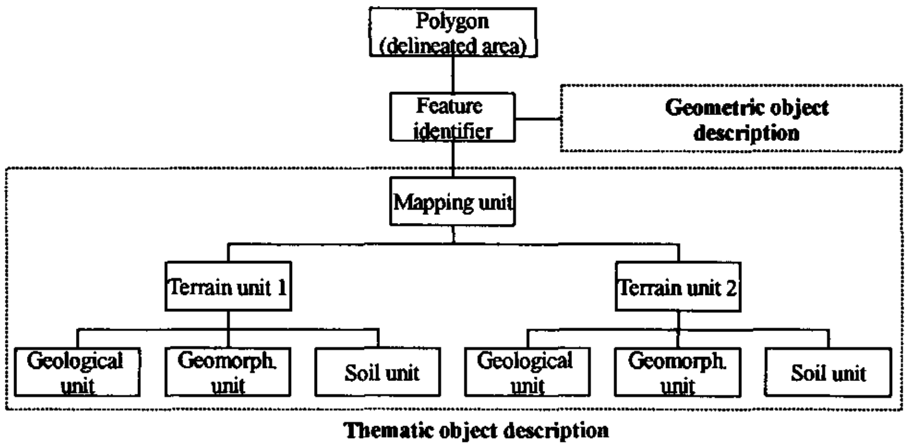


Figure 2.1 Data model for a soil and terrain database

instance, described in terms of the age and type of deposit, whereas GmUs are described in terms of slope, drainage and stoniness of the terrain. The way sub-units are described largely depends on the available data, the scale and the purpose of the soil map.

2.1.3 Alternative database structures

Several database structures for soil survey data have been developed. Examples are SOTER (Van Engelen and Wen 1995) and STATSGO (Soil Survey Staff 1993). The setup for the relational database containing the thematic information of soil and terrain data depends mainly on the dataset and on the objectives of the database. On the basis of the data model (Figure 2.1), an inventory of alternative database structures is carried out. Three alternative database structures will be described according to the entity-relationship model (Chen 1976).

<u>Structure A</u>

Figure 2.2 shows a database structure proposed by Baumgardner and Van de Weg (1989) and Wielemaker and Vogel (1993). Different entity sets can be distinguished for delineated areas (E_{DA}), mapping units (E_{MU}), terrain units (E_{TU}), geological units (E_{GeU}),

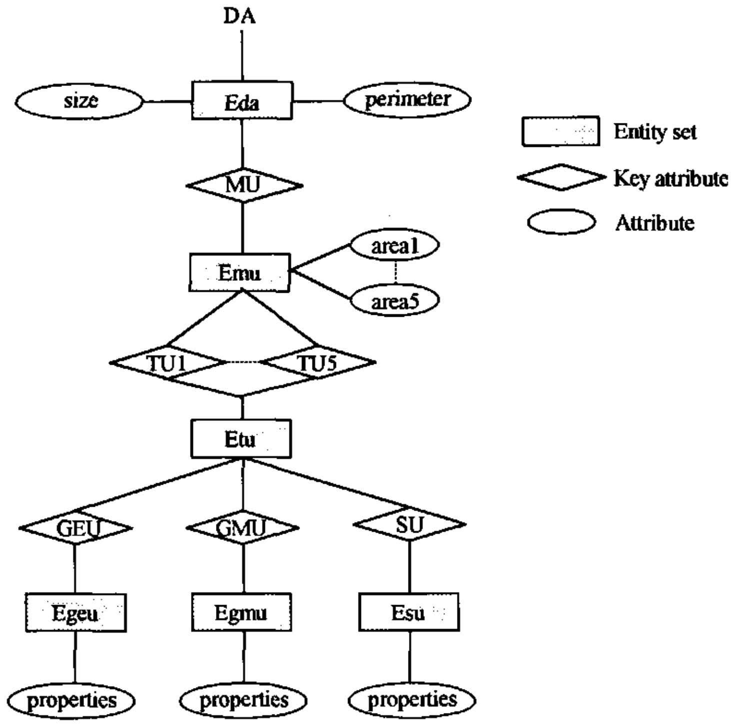


Figure 2.2 Entity relationship model for database structure A

geomorphological units (E_{GmU}) and soil units (E_{SU}). All the entity sets are stored in separate two dimensional tables in the database.

E_{DA}, which in Arc/Info coverages is named the polygon attribute table, describes the classification of the delineated areas into mapping units. Two key attributes are present: the feature identifier for the link with the geometric data, and the MU identifier as a link to E_{MU}. Additionally two topological attributes for the area and perimeter of DAs are added by Arc/Info.

E_{MU} has, besides the MU identifier, five key attributes relating the MU to five different terrain units. Five additional attributes indicate the percentage of the mapping unit covered with the corresponding terrain units. The five key attributes for the different terrain units are all linked to one key attribute in E_{TU}.

Terrain units are described by different sub-units in E_{TU}. Besides the TU identifier, E_{TU} contains key attributes to the entity sets of these sub-units: E_{GeU}, E_{GmU}, and E_{SU}. The different entity sets for the sub-units describe the specific properties included in the survey.

The database structures presents a complex relation between E_{MU} and E_{TU}. If, for instance, one is interested in a specific surface stoniness, a quintuple relation (in case of five terrain units) between E_{MU} and E_{TU} is necessary to check whether one of the five terrain units presents surface stoniness. Databases organized according to Structure A will contain many empty fields for mapping units where less than five terrain units are described. The databases will therefore occupy relatively much disk space.

Structure B

Database structure B (Figure 2.3; e.g. Van Engelen and Wen 1995) differs from Structure A by the definition of E_{MU}. E_{MU} in Structure B has two key attributes for the MU and the TU. An additional attribute indicates the percentage that the TU occupies in the MU. A mapping unit with, for instance, four terrain units will now occupy four records in E_{MU}. With a varying number of terrain units, the number of records describing the MU will change correspondingly. In this way, no empty fields are found in the database. At the same time, the complex relation between E_{TU} and E_{MU} is avoided. Mapping units with surface stoniness can now be selected by a single query.

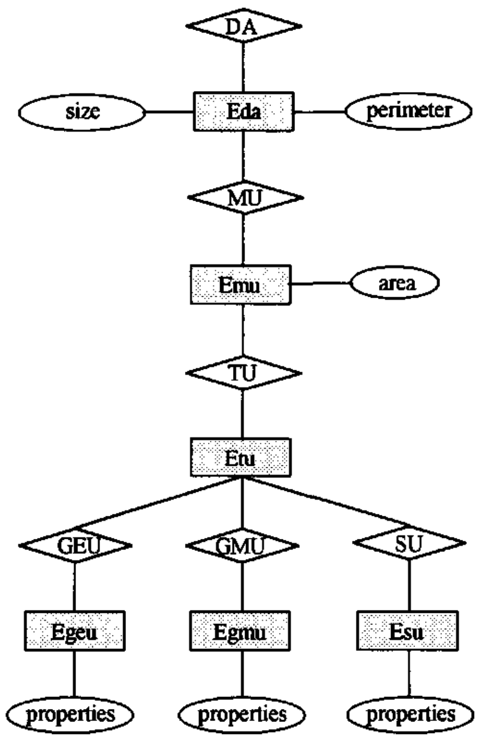


Figure 2.3 Entity relationship model for database structure B

Structure C

Structure A and B can indicate a variability which not necessarily is relevant for users. A mapping unit can be homogeneous for geomorphology but at the same time may be composed out of several soil types. Users interested in geomorphology still have to deal with this variability. In structure C, however, GeU, GmU and SU are treated as spatially independent and separate terrain units for the different sub-units are defined (Figure 2.4). To avoid that the spatial dependency is lost completely, a separate index file is included indicating the spatial dependency for the different sub-units. The definition of this index file may be simple derivations of E_{MU} and E_{TU} as defined in Structure A or B.

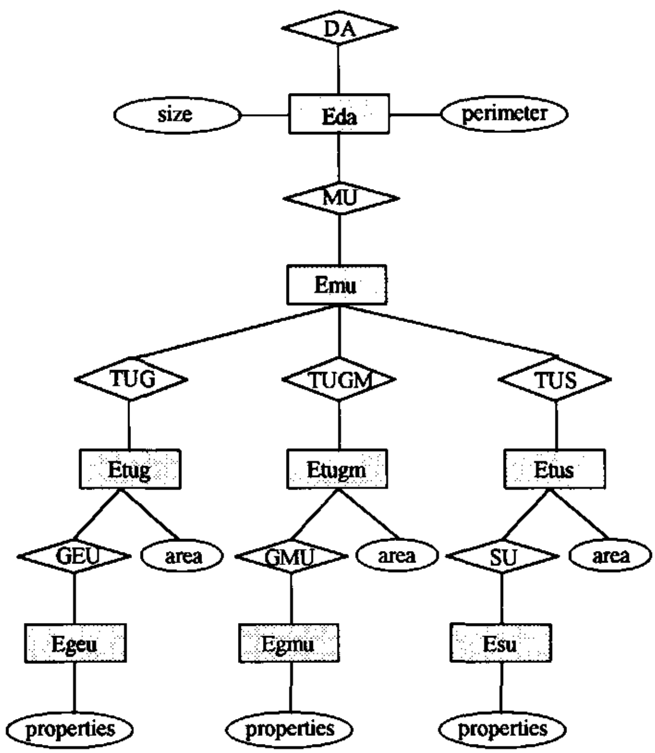


Figure 2.4 Entity relationship model for database structure C

2.1.4 Queries

Queries may be feature oriented or thematic oriented. Feature oriented queries comprise queries that look for the properties of a certain terrain feature. A feature oriented query checks for instance whether fertile well drained soils occur in a delineated area. Thematic queries are based on the values of one or more attributes. The terrain features that fulfill these criteria are selected. A thematic oriented query may be the selection of all delineated areas with fertile soils. The way the two queries can be put into practice is completely different. In most databases thematic oriented queries have to pass several "one (or few) to many" relations as feature oriented queries have to pass several "many to one (or few)" relations.

The efficiency of a database structure merely depends on the requested information. However, a certain number of boundary conditions like normalisation of the database should always be fulfilled. Depending on the database and its use the most suitable normal form has to be found. These normal forms guarantee for relational databases that problems of redundancy and anomalies do not occur (Ullman 1982). The optimisation of the database structure is based on the efficiency and the flexibility. The efficiency stands for a minimum of algorithmic steps during the query. On the other hand the flexibility indicates whether a wide range of queries, thematic oriented or feature oriented, is possible. To evaluate possible database structures, it is necessary to quantify these criteria. One of the possible indicators is the average number of fields that has to be read in the various tables for thematic and feature oriented queries. The number of columns and records for the different tables that are involved in the query are determined after which Equation 2.1 estimates the total number of fields to be read.

$$\bar{X} = \sum_{i=1}^{F} (R_i * C_i) \qquad \text{Equation 2.1}$$

In which:

	$\bar{X}$	=	average number of fields
	F	=	total number of tables
	R_i	=	number of records in table i
	C_i	=	number of relevant columns in table i

Of course commercial database management systems have advanced querying possibilities which will increase the efficiency of querying. Nevertheless, equation 2.1 will give a good indication of the efficiency of the database to a certain query.

Besides the number of fields, the number of files and the amount of disk space required determine the efficiency of the database. The number of files, generally, corresponds with the number of entity sets.

2.1.5 The Costa Rican case

For the creation of the database structure five different steps can be followed:
1) Identification of the dataset
2) Setup of a data model
3) Creation of alternative database structures
4) Inventarisation of common requested queries
5) Evaluation of the database structures

These five steps will be described and evaluated for the soil and terrain database of the Northern Atlantic Zone.

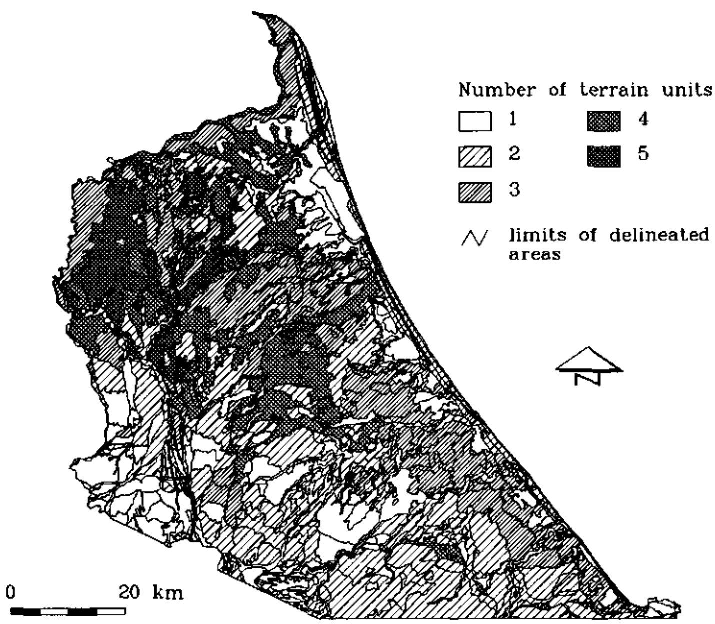


Figure 2.5 Complexity of the delineated areas in the Atlantic Zone soil survey

Step 1 includes the inventarisation of the database, which is the result of the 1:150,000 soil survey of the northern part of the Atlantic Zone. It comprises an area of 5,450 km^2 described by 753 delineated areas. For the 1:150,000 soil map, 154 different mapping units have been described. The database is a compilation of more detailed information at scale levels varying between 1:20,000 and 1:1:100,000. As a result, 67% of all polygons and 82% of the area are described by associations (Figure 2.5).

The data model (**Step 2**) follows the setup of Figure 2.1. The mapping units are described by one to five TUs. A MU, for instance, comprises three terrain units for small remnants of old eroded mud-flows and young alluvial deposits which are poorly drained along streams. Although the survey was focused on soils, descriptions for geology and geomorphology are also included.

A large number of database structures can be created (**Step 3**). Figures 2.6, 2.7, and 2.8 show fragments of the soil and terrain database for the Atlantic Zone structures according to Structure A, B, and C respectively. Up to 5 different terrain units are identified. In Structure A, 54% of the fields in E_{MU} are empty, because only in few cases five terrain units are actually identified. No empty fields occur in Structure B resulting in less disk space required (E_{MU} occupies in Structure A 7.7 kB compared to 5.1 kB in Structure B). However, the number of records in E_{MU} increases significantly from 154 (the number of MUs) in Structure A to 332 in Structure B, whereas the number of columns decreases. In Structure C, the spatial dependency between the different sub-units is excluded, leading to a significant decrease in the number of associations defined for the geological and geomorphological unit (52 and 98 respectively). The soil unit forms the origin of most associations and yields 127 different associations.

Although projects are often unaware of queries that will be requested most, a general inventory of thematic and/or feature oriented queries can always be carried out. The inventory of queries (**Step 4**) can, however, be repeated and the database structure may be adapted. Four main groups of queries, which are all frequently requested at the Atlantic Zone Programme, can be identified:

1) thematic single-property queries (e.g. locate mapping units with possible occurrence of specific clay minerals),
2) thematic multi-property queries (e.g. locate flat or almost flat areas with infertile, well drained soils),
3) feature oriented single-property queries (e.g. give a general impression of geological features in the Atlantic Zone), and

4) feature oriented multi-property queries (e.g. indicate stoniness, soil depth, and soil pH for a banana plantation).

Eda

DA	MU	size
4	30	6.7
5	11	1.8
6	27	12.3
7	44	137.8
8	47	5.0
9	31	2.5
10	40	4.6

Emu

MU	TU1	A1	TU2	A2	TU3	A3	TU4	A4	TU5	A5
11	4	100	-	-	-	-	-	-	-	-
27	73	50	6	30	8	20	-	-	-	-
30	70	100	-	-	-	-	-	-	-	-
31	72	50	12	50	-	-	-	-	-	-
40	12	70	71	30	-	-	-	-	-	-
44	7	50	1	20	47	20	13	10	-	-
47	3	60	7	40	-	-	-	-	-	-

Etu

TU	GEU	GMU	SU
4	4	102	0
6	4	99	64
8	4	93	66
12	2	105	70
70	2	104	68
71	2	104	47
72	2	104	33
73	1	106	3

Egu **Egmu** **Esu**

Figure 2.6 Fraction of the soil and terrain database for the Northern Atlantic Zone, structured according to database structure A

Eda

DA	MU	size
4	30	6.7
5	11	1.8
6	27	12.3
7	44	137.8
8	47	5.0
9	31	2.5
10	40	4.6

Emu

MU	TU	A
11	4	100
27	73	50
27	6	30
27	8	20
30	70	100
31	72	50
31	12	50

Etu

TU	GEU	GMU	SU
4	4	102	0
6	4	99	64
8	4	93	66
12	2	105	70
70	2	104	68
71	2	104	47
72	2	104	33
73	1	106	3

Egu **Egmu** **Esu**

Figure 2.7 Fraction of the soil and terrain database for the Northern Atlantic Zone, structured according to database structure B

Eda

DA	MU	size
4	30	6.7
5	11	1.8
6	27	12.3
7	44	137.8
8	47	5.0
9	31	2.5
10	40	4.6

Emu

MU	TUGE	TUGM	TUS
11	4	31	1
27	6	62	84
30	2	33	51
31	2	41	62
40	2	42	81
44	13	71	93
47	7	51	59

Etug

TUGE	GU	A
2	2	100
4	4	100
6	1	50
6	4	50
7	1	60
7	4	40
13	4	70
13	1	20
13	2	10

Esu

Etugm

TUGM	GMU	A
31	102	100
33	104	100
41	104	50
41	105	50
42	104	70
42	105	30
51	109	60
51	99	40

Egmu

Etus

TUS	SU	A
1	0	100
51	68	100
59	3	60
59	65	40
62	33	50
62	70	50
81	70	70
81	74	30

Esu

Figure 2.8 Fraction of the soil and terrain database for the Northern Atlantic Zone, structured according to database structure C

The last step (**Step 5**) evaluates the different database structures using the number of files, required disk space, and the number of fields to be read for the different queries calculated with Equation 2.1 (Table 2.1). The number of files and required disk space for the different structures are rather similar for the different database structures and are not likely to present any problem with the common hardware configurations. However, large differences in the number of fields occur for the different queries and database structures. Structure A is especially efficient when only the dominant terrain unit is requested because E_{MU} has relatively few records. When all terrain units need to be checked the structure is less efficient due to the large number of empty fields. Structure B gives best results for the queries with more than one property requested. Structure C is efficient when the query only checks one sub-unit. If more sub-units are required the index file is necessary and the structure becomes less efficient.

It can be questioned whether one should have a fixed database structure. Relational database management systems provide possibilities to change the structure easily when the requested information or the user group changes. When the data are not updated regularly, different database structures can be generated for different queries. In cases of databases that are regularly updated, this may not be advisable as it may generate inconsistencies in the database. For the Atlantic Zone Programme, Structure B has been selected as the most appropriate database structure, because it allows for a wide range of different queries.

Table 2.1 Analysis of database structures and queries (Queries refer to the examples in the text under Step 2)

	Database structure		
	A	B	C
Number of files	6	6	5[1]
Memory required for the storage of the files (in kB for storage as DBase IV files)	121.3	101.7	140.9
total number of fields read per query			
1. Thematic (clay minerals)	2922	2516	2238
2. Thematic (flat infertile soils)	3290	2984	3466
3. Feature (geology)	2158	2800	2022
4. Feature (banana plantation)	4053	3059	3541

[1] Excluding index files.

2.2 An application oriented soil information system

2.2.1 Different stakeholders dealing with biocide leaching

With an increasing number of environmental problems related to biocide use in high input agriculture, leaching of biocides from the soil and the resulting contamination of ground water receives increasing attention. The cost of sampling programs often hampers studies on the effect of these biocides on the environment. Soil survey data provide, especially for studies at the regional level, one of the few readily available data sets that can be used to efficiently plan additional measurements or to run simple models. Although the introduction of GIS improved the availability of soil survey data, the structure and semantics used for data storage still determine to a high extent the usefulness of the data. Depending on objectives and resources, projects follow different approaches and make use of different models, varying from simple expert systems to complex deterministic models (Addiscot and Wagenet 1985). The type of model determines the data from the soil survey that can be used. Ideally, the database provides soil survey data at different levels of detail, so that different procedures can be linked to the database.

In the perhumid tropical lowlands in the Atlantic Zone of Costa Rica, the area under high-input agriculture expanded rapidly from 223 km^2 in 1984 to 466 km^2 in 1992 as estimated by aerial photograph interpretation. At the same time, ground water is used for drinking water and high-input agriculture occurs close to nature reserves. Biocide leaching is therefore receiving increased attention from different stakeholders. They try to draw the attention of policy makers to the effects of biocide use on the environment and human health. Depending on the objectives and resources of the organizations, they follow different approaches:

- The National University (UNA) started their research with hardly any data available (e.g. Hilje *et al.* 1987, Castillo *et al.* 1993). They focused their research on monitoring biocide concentrations of surface waters and sediments. Sampling, however, took place more or less randomly, due to the lack of data to structure the sampling schemes. Showing figures on the occurrence of different biocides in the environment, they drew attention to the problem. Nevertheless, the effect of the measured biocide concentrations on e.g. biodiversity is still unknown and data obtained are therefore hard to interpret.
- The national corporation of banana producers (Corbana) carried out leaching experiments to determine the amount of biocides leached from banana plantations. Banana production is seen as one of the main environmental dangers within the

Costa Rican lowlands. Corbana, therefore, tries to develop management practices which lead to less biocide leaching.

- A project of the International Union for the Conservation of Nature (IUCN) developed strategies for a sustainable development of the Tortuguero plains in the Northern Atlantic Zone. They estimated the area of banana plantations for each watershed and related it to the discharge of rivers draining these catchments (IUCN, non published data). The study indicated that potentially contaminated rivers were flowing into the nature reserves. Although the study was carried out with minimum data, it drew much attention and formed the basis for the official recognition that the area around the park should be included in the Tortuguero conservation area, implying restricted use of biocides.

UNA, Corbana, and IUCN all make use of the 1:150.000 soil survey for the Northern Atlantic Zone. Standard database structures for soil survey data (e.g. Baumgardner and van de Weg 1989) limited the use of these data due to the rigidity in the level of detail with which the data are stored. Consequently, there is a need to develop more flexible structures that allow for an efficient retrieval of data at different levels of detail. This section proposes a multi-level database structure for soil survey data. The database structure is the result of an analysis of questions being asked by different users which is illustrated by different approaches to deal with contamination of ground water with a commonly used nematicide. Soil survey data are structured to fulfill data requirements for the different approaches.

2.2.2 Using soil survey data to deal with biocide leaching

In the Northern Atlantic Zone, the high average annual rainfall in combination with high input agriculture results in a high risk for biocide leaching. Different organizations are dealing with the environmental effects of biocides. Each of these organizations would like to quantify the environmental effects of biocides by measuring biocide behaviour in all the soils series. However, they have limited resources and can only make general estimates. The risk for biocide leaching can be studied by different procedures, each with its own data requirements and at a specific level of detail. On the basis of objectives and available funds of users, the model that is most appropriate can be selected (Bouma *et al.* 1993). The soil survey for the Northern Atlantic Zone is useful as a basic data set for the different procedures. Soil survey data are included in a regional GIS which also includes data on land use and climate. The available soil

Presenting the results with the GIS for the zone, only the dominant soil series in each mapping unit are used to indicate potential hazard areas (Figure 2.10[a]). The decision tree classified 24% of the area as potentially hazardous. UNA can now focus on these areas, or even further reduce the area on the basis of land use data. For the selected areas more intensive sampling schemes can be employed thus using available funds more efficiently.

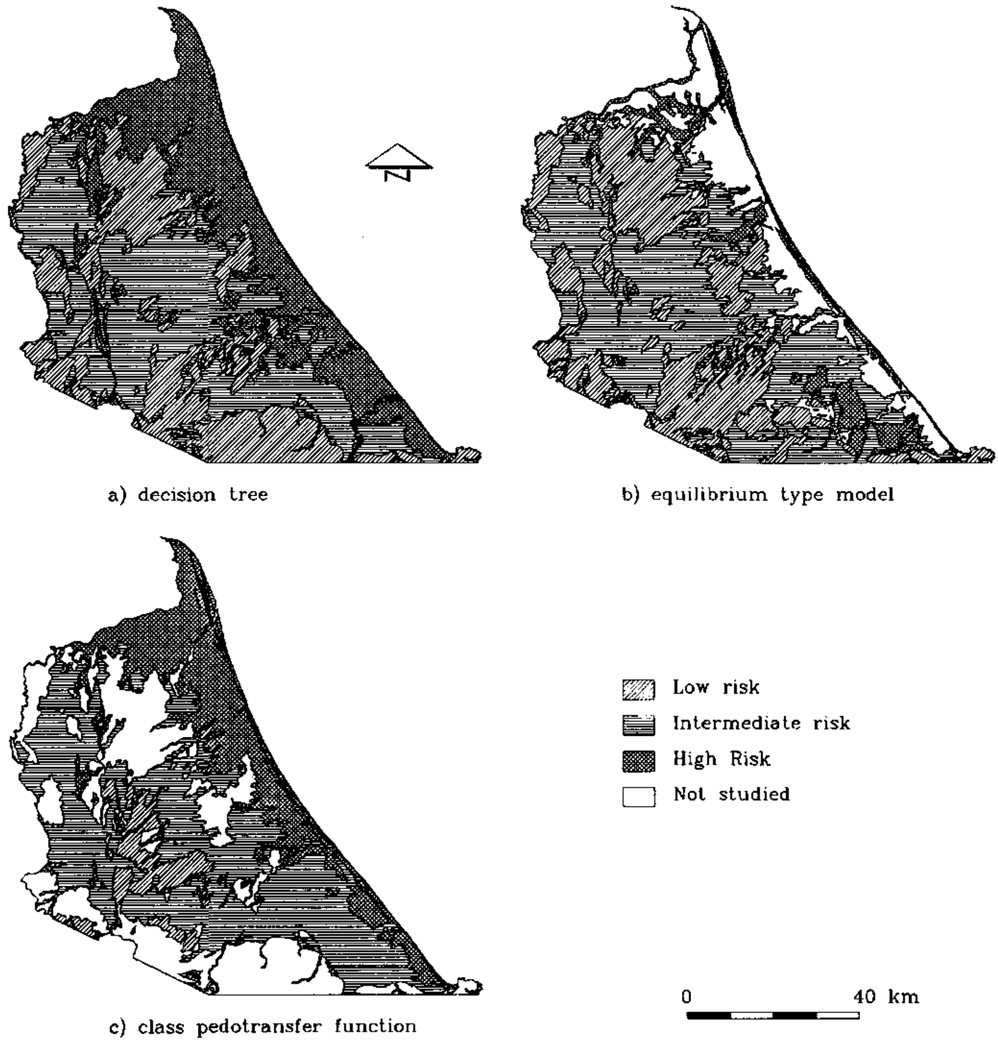


Figure 2.10 Risk for Ethoprop leaching in the Atlantic Zone of Costa Rica determined by different procedures

Pedons

Corbana carries out leaching experiments for the development of alternative management systems. Soil survey data can be used to find appropriate locations for measurements and to extrapolate the data measured for some pedons to others for which no measurements could be carried out. The need for a more quantitative approach to the prediction of environmental impacts may originate in economic implications (Bouma *et al.* 1993). A simple linear equilibrium type model is used to estimate biocide leaching from the different soil series in the Atlantic Zone. Generally, 10 kg/ha,yr Ethoprop is applied in banana plantations. On the basis of the Pesticide Environmental Fate One Liner Database (U.S. Environmental Protection Agency, Environmental Fate & Effects Division), the following relation for Ethoprop was established (r^2 = 0.76, n=5):

$$Kd = Koc * OM + 0.21 * Clay \qquad \text{Equation 2.2}$$

In which:	Kd	=	ratio between Ethoprop concentration at the soil surface and in soil solution
	Koc	=	equilibrium parameter for Ethoprop fixation on soil organic matter
	OM	=	soil organic matter content (%)
	Clay	=	clay content (%)

The soil is divided into successive 10 cm compartments for which the Kd values are calculated on the basis of the above pedotransfer function. The model assumes steady-state one-dimensional water flow through the soil (the average daily precipitation minus average daily evapotranspiration). Additionally, it is assumed that the equilibrium between sorbed and dissolved forms of the biocide is reached instantaneously and that Ethoprop has a half life time of 40 days (Pesticide Environmental Fate One Liner Database). Results are presented in Figure 2.10[b]. Although results are mostly based on literature from outside Costa Rica (especially for biocide behaviour), they are quantitative and can present additional insight in differences among management practices in different regions. When additional leaching experiments are carried out by Corbana and linked to the procedure, the procedure can be calibrated and actual biocide leaching can be estimated more accurately.

Horizons

In contrast to UNA and Corbana, whose final results are based on additional quantitative measurements, the IUCN did not carry out any additional measurements. The qualitative character of the decision tree (Figure 2.9) will, therefore, not yield sufficient detail. For more detailed results insight in the adsorption of Ethoprop will be necessary, and especially variation between the different soil series. General-purpose databases on biocide characteristics (e.g. Pesticide Environmental Fate One Liner Database) list data for some soils on biocide behaviour, but mostly for temperate climates and for non-volcanic soils. Adsorption of Ethoprop may, however, differ significantly due to the presence of allophane. A procedure for a relatively easy and fast appraisal of potential ground water contamination with biocides can, however, be based on soil survey data and on one additional soil property: biocide fixation in the soil (Figure 2.11). The procedure is based on the comparison of potential Ethoprop fixation to the soil matrix above the ground water and the Ethoprop application rate. On the basis of two threshold values for the fixation/application ratio, soils are classified in terms of their potential hazard to contaminate ground water. To extrapolate a limited number of Ethoprop fixation measurements, two approaches can be followed. One groups the horizons of the different soil series into a limited number of functional horizons and assumes that each sample is representative for a functional horizon. The alternative is to extrapolate the results of the batch experiment to the other soil horizons on the basis of a continuous pedotransfer function which in this case is a relation between Ethoprop fixation and known soil characteristics (r^2 = 0.86, n=15):

$$E_{fix} = 1.2 * OM + 0.21 * Clay \qquad \text{Equation 2.3}$$

In which: E_{fix} = potential Ethoprop fixation (mg/kg)
OM = organic matter content (%)
Clay = clay content (%)

Both procedures yield Ethoprop fixation for each soil horizon. Potential Ethoprop fixation in the soil matrix above the ground water table can now be calculated on the basis of the Ethoprop fixation, bulk density and stoniness. When Ethoprop application exceeds the potential fixation, Ethoprop leaching is certain to occur and the soil is classified as "high risk". An admittedly arbitrary boundary is set for low risk soils, when potential Ethoprop fixation exceeds 20 times the application. Other boundaries may be selected. The grouping of soil horizons into functional horizons and a pedo-

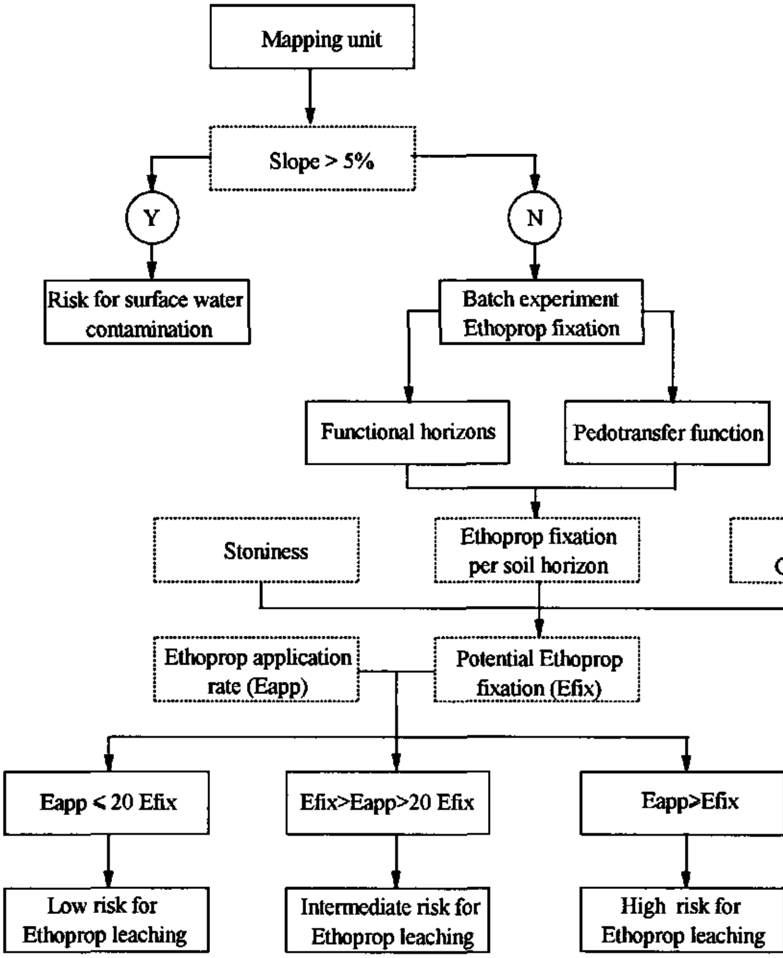


Figure 2.11 A framework for a fast appraisal of the risk for pesticide leaching

transfer function to estimate Ethoprop fixation resulted in similar results presented in Figure 2.10[c]. Organizations like the IUCN can use these data to differentiate banana plantations, and areas within plantations, on the basis of biocide leaching.

Although the results of the different procedures vary as shown in Figure 2.10, similar areas with high risk are identified. The equilibrium type model excluded the swamp areas, which is identified as high risk with the other procedures. The class pedotransfer function excluded the sloping areas, which are indicated to be low risk with the other procedures. All three procedures indicate the alluvial plains, which is the major agricultural area, to have an intermediate risk for Ethoprop leaching.

Additional measurements

Typically, only a limited amount of measurements can be carried out. These measurements can be for different units: mapping units, pedons, or soil horizons.

Mapping units

At the mapping unit level biocide concentrations in drainage water from plantations or watersheds (e.g. UNA, Rosales *et al.* 1992) can be measured. These data can be used to calibrate models and expert systems with which biocide leaching under alternative forms of management can be estimated. Additional measurements can be planned using soil survey data. Which procedures are being followed strongly depends on users. Because both biocide concentrations and discharges are often characterized by a peaky variation, many samples may be necessary for one mapping unit to get insight in the biocide leaching process. With limited resources, it may be necessary to select a restricted number of mapping units. For organizations like Corbana only mapping units where banana plantations occur are important. For others like the Costa Rican Ministry of Planning (Mideplan) all mapping units which are potentially suitable for banana cultivation are important to decide on future expansion of banana plantations. In the case of the Northern Atlantic Zone a selection of mapping units in which banana plantations occur reduced the number of mapping units to be considered from 154 to 58, resulting in an approximate 60% reduction of sampling costs. If sampling of 58 mapping units still exceed the available funds, the mapping units have to be generalized by e.g. listing only the dominant soil type or by listing only soil types covering more than a certain acreage. The number of mapping units suitable for banana production is estimated at 71, although this depends on the criteria used.

Pedons

At the pedon level measurements are mainly based on leaching experiments with lysimeters, which are carried out by e.g. Corbana. Due to the costs of the lysimeter experiments only few measurements can be carried out. Again only the soil series where bananas are actually found can be included in the analysis resulting in a reduction from 75 to 37 soil series. Even lysimeter experiments for 37 soil series may exceed the resources of most projects and, therefore, a selection or grouping of soil series will be necessary. Using generalized functional soil types based on texture, fertility and andic properties (Section 1.3), leaching experiments can be carried out for

8 instead of 75 soil types. Projects may include replicates to identify the variation in each of the soil groups allowing a check on the quality of the results.

Horizons

At the horizon level, measurements will be focused on biocide behaviour in the soil matrix e.g. adsorption and half life time. Soil adsorption can be measured for the soil horizons of the relevant soil series, or for a more limited number of functional horizons, which are determined on the basis of texture and soil organic matter content and in some cases on mineralogy (to separate horizons with high contents of allophanes, which are likely to have significantly different adsorption characteristics). For the batch experiment, the 291 soil horizons from 75 soil series in the Northern Atlantic Zone were grouped into 16 different functional horizons on the basis of texture, organic matter content and andic properties. For each of the functional horizons the Ethoprop fixation was measured. Soil horizons that occur below the ground water table and C-horizons that do not allow water transport are excluded from the measurements leaving 12 soil horizons to be sampled. 10 ml of a 226 ppm Ethoprop solution was added to 30 g of a 1:2 soil-water suspension. The soil samples were not dried between sampling and analysis to avoid irreversible drying effects. The soil moisture content was determined on separate samples. The Ethoprop concentration in the solution was measured after 30 minutes shaking. The Ethoprop fixation varied between 2% for a sand C-horizon and 48% for a well developed andic epipedon.

<u>Integration of different types of analysis</u>

The alternative types of analysis and sampling schemes do not have to be separated. Van Lanen *et al.* (1992) present a study where qualitative expert knowledge is used to screen potential problem areas. For areas where the qualitative analysis does not yield an answer, quantitative analysis may follow. This procedure may involve substantial savings. As can be seen in Figure 2.10, most procedures predict high risks for about 5-35% of the study area. More complex modelling approaches (e.g. quantitative mechanistic models) with the necessary sampling schemes, can be focused on these areas or, depending on user objectives, on both high and intermediate risk areas.

2.2.3 The soil survey database

The 1:150,000 soil survey for the Northern Atlantic Zone of Costa Rica (Wielemaker and Vogel 1993) has been carried out following the directives of Soil

Survey Staff (1951). Based on an aerial photograph interpretation, 753 delineated areas were identified in the 5,450 km^2 area, which were subsequently classified in terms of 154 mapping units. Each mapping unit is described by one to five terrain units that are unique combinations of soil series and phases (i.e. properties which are not included in the description of the soil series like, for example, surface stoniness and slope). In the study area, 75 different soil series were distinguished and described by 123 representative profiles.

Although the soil survey has been carried out according to standard concepts for surveying and data were stored in a more or less standard database format for soil survey data (see Section 2.1), users of the database were not satisfied with the structure and semantics of data storage. Even though queries enabled users to select specific data and to structure these data according to their needs, the database only provided data at a single level of detail. The different procedures to estimate biocide leaching require, however, data at different levels of detail for each of the hierarchial levels. On the basis of experiences in environmental modelling as illustrated in the previous section, the database structure was adapted to make it more flexible and functional for a larger group of users.

The new database structure (Figure 2.12) provides data at different hierarchical levels: the mapping unit, the pedon, and the horizon level. Users at the farm level (e.g. a plantation owner) or community level generally need information at the level provided by the 1:150,000 soil survey. However, users working at the regional level (e.g. planners from government agencies), request generalized data. Most users are not familiar with the study area nor with soil science and, as a consequence, they are not able to carry out these generalizations. Therefore, to increase the flexibility of the database, decision rules were developed to allow generalization at each of the hierarchical levels. No decision rule base was developed to generalize data belonging to the terrain units due to lack of user interest.

Mapping unit

Mapping unit descriptions consist of soil associations and complexes as a consequence of field variability and the 1:150,000 scale of the soil map. Three decision rules for the generalization of the mapping units were included: MU1) selection of the dominant soil series in the mapping unit, MU2) grouping mapping units with a similar soil distribution, and MU3) generalizing the soil series. The latter is particularly useful when soil series that characteristically occur together in an association or complex are combined.

For the Northern Atlantic Zone, MU1 resulted in a reduction of 154 to 64 mapping units. A grouping of mapping units with similar soil distributions (MU2) resulted in 87 different mapping units. Generalizations of the soil series (see Section) will also result in generalizations of the mapping units (MU3). The identification of 8 soil groups (decision rule P2) resulted in 59 different mapping units, whereas a further generalization in 4 soil groups yielded 46 mapping units. The objectives and resources of users determine which decision rule is the most appropriate for their particular application.

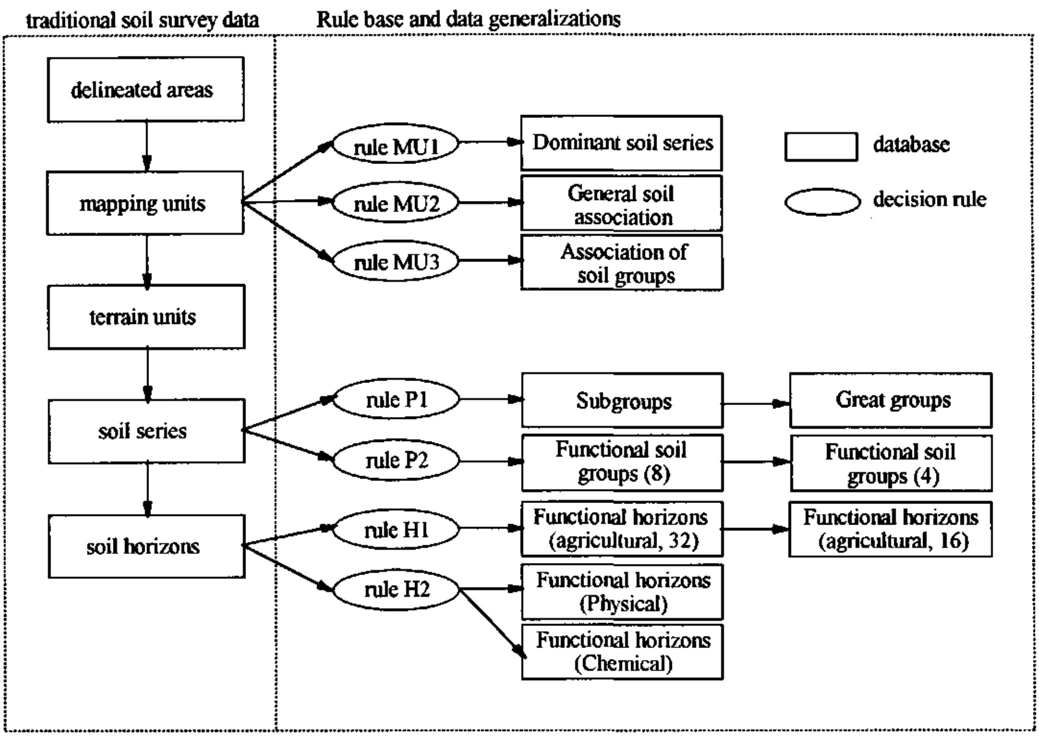


Figure 2.12 Database structure for the soil and terrain database of the Northern Atlantic Zone extended with a rule base for generalizations

Pedon

At the pedon level two generalizations were included: P1) according to the hierarchy given by Soil Taxonomy (Soil Survey Staff 1994), and P2) on the basis of the agricultural potential of the soil series. Using Soil Taxonomy, the 75 soil series were classified in terms of 40 subgroups, 18 great groups, 10 suborders, and 5 orders. Although the hierarchy given by Soil Taxonomy may be useful for specific pedological studies, users from other disciplines than soil science prefer a generalization on more functional criteria. For applications within the AZP, the soil series were grouped in 8 functional soil groups with a similar agricultural potential (See Section 1.3). A further generalization on the basis of soil fertility and drainage took place resulting in 4 soil groups. The latter is often used in combination with the Costa Rican system for land evaluation (SEPSA 1992), which is obligatory for legal transactions involving e.g. credit and subsidies. For Corbana the number of soil series coincides with the number of measurements. Reductions in the number of soil series will substantially reduce costs.

Horizons

Soil horizons are normally described as specific soil layers, occurring in a soil series. The 75 different soil series identified in the Northern Atlantic Zone were described with, on average, three different horizons resulting in a total of 291 different soil horizons. Many of these soil horizons had similar soil physical, chemical and mineralogical properties. A grouping of the soil horizons based on their texture, fertility and mineralogy (H1) resulted in 32 functional horizons. Specifically for biocide leaching, these functional horizons were generalized on the basis of texture, organic matter and andic properties in 16 functional horizons. To provide functional horizons for other studies, functional horizons were also defined on the basis of soil physical properties and soil chemical properties (H2), resulting in 11 and 13 functional horizons respectively.

Generalizations at low hierarchical levels, such as the soil horizon, are likely to have implications at higher hierarchical levels as was illustrated for the effect of defining soil groups on the number of mapping units. Similarly, a generalization of soil horizons in soil physical functional horizons will result in different soil series with identical functional horizons. Including the full description of 32 functional horizons and their respective depths will reduce the number of "functional soil series" from 75 to 65. Only considering the sequence of functional horizons without their respective depths will even further reduce the number of "functional soil series" to 48.

In many cases data requirements can be fulfilled by querying the database, when the setup of the database structure follows some general rules like normalization. Expert knowledge, however, is required to determine the decision rules for the generalizations. Many users of the soil survey database, are not soil scientist and are not able to carry out such generalizations, which can be done best by the original surveyor. The final product of the soil survey should, therefore, include a rule base for generalizations at each hierarchical level as discussed in this chapter. The rule base is a formalisation of the surveyors knowledge. Due to the multifaceted nature of mapping units, pedons, and horizons the rule base for generalizations are often not in a strict hierarchy with PARTOF relations as e.g. proposed by Molenaar (1993) who ranks spatial objects in terms of classes and superclasses. Although the hierarchy may be useful for one specific decision rule, each decision rule will have its specific hierarchy resulting in, for example, soil physical functional horizons for water balances and soil chemical functional horizons for biocide adsorption.

2.3 Considerations for the setup and use of a soil information system

Typically, small scale soil surveys yield soil associations. Although database structures become complex due to associations and complexes, a wide range of queries is possible. Special attention is needed for the development of database structures and queries. Although still problems exist with the presentation of the complex objects, they are valuable supplements to the database and good indicators for the spatial variability.

For the creation of efficient databases a five step approach is proposed. The steps include the analysis of the data and the development of a data model, an inventory of alternative database structures, an inventory of possible queries, and the analysis and selection of a database structure. This framework was useful for the development of the database structure at the Atlantic Zone Programme.

Database structures can be developed through a sequence of problem analysis, definition of data requirements, and database structuring. Although general information-theory provides general rules for the setup of database structures, the use of the database will determine the efficiency of these structures. Structuring the database requires a thorough analysis of user requirements to develop the appropriate structure and the associated decision rules.

The presented multi-level soil information system is based on a general database structure for attribute data and a rule-base containing the decision rules for the generalizations at each of the hierarchical levels. Although general theories can be used for the structuring of the database, including normalisation and hierarchical classification structures, users require the rule base to allow operational generalizations. Most users are not able to generalize data on the basis of their limited regional or disciplinary knowledge.

Available soil data enable a screening for potential hazard areas, a good planning of sampling schemes and simple modelling. This may significantly reduce costs by avoiding expensive measurements or model calculations for areas that can already be classified as being highly susceptible using simple procedures and available data.

3. Temporal analysis of information

This Chapter is based on the following publication:

- Stoorvogel, J.J., and L.O. Fresco, 1995. Quantification of land use dynamics, an illustration from Costa Rica. Submitted to *Land Degradation and Rehabilitation*.

3.1 Land use dynamics

Land degradation is often initiated by changes in land cover resulting from human land use rather than natural change. Changes in land use are, therefore, increasingly seen as an important issue at global as well as regional scales (Turner *et al.* 1994). Studies of land cover conversion - i.e. the change from one cover class to another - as well as land cover modification - i.e. changes within a given land cover class - are now carried out in many areas (e.g. Brouwer and Chadwick 1991, Houghton *et al.* 1991, Reiners *et al.* 1994, Garrity and Agustin 1995). Nevertheless, many aspects of land use changes remain poorly understood, although knowledge of the bio-physical drivers of land use, based on well-known methods from land evaluation (Van Diepen *et al.* 1991) is much more developed than that of socio-economic drivers (Bilsborrow and Okoth-Ogendo 1992, Veldkamp and Fresco 1995). In particular, adequate indicators to quantify land use dynamics - i.e. land cover conversion and modification over time - are hardly available. As a result it becomes very difficult to compare rates of change between different areas and periods.

Fortunately, aerial photographs, satellite imagary and GIS now enable more frequent inventories and bring the analysis of land use within reach of many researchers. In view of the surging interest in spatial and temporal analysis in GIS (Fortheringham and Rogerson 1994), it is all the more surprising that little or no work has been done thusfar on land use dynamics indicators.

3.2 Indicators for land use dynamics

Based on a case study in Costa Rica, this chapter reviews an existing method of dealing with temporal changes, elaborates two alternative methods and compares the outputs. The three methods are as follows:

1. A single-time analysis of spatial patterns based on qualitative knowledge of temporal patterns of land use evolution. This method is new, and has been included to demonstrate the potential of a dynamic analysis based on a single data set.
2. A more or less standard approach using Markov chains, to describe the probabilities of successive land use systems. This method is somewhat refined by including a stratification with soil type as a probability modifier.
3. A new adaptation of the Markov chain method by including a geographical analysis to relate changes in land use to probabilities based on the shape and size of polygons and land use in neighbouring polygons. This method reflects the observations that (a)

land cover conversion of an area often starts from adjacent areas, and (b) there is a relationship between polygon shape and land (e.g. riverline forests) and between shape and speed of conversion ("narrow strips" are "broken through" more frequently).

The three methods are elaborated for a pilot area in the perhumid tropical lowlands in the northeast of Costa Rica. Because of rapid rates of demographic and land cover change in this region, the case study area provides an appropriate testing ground for land use dynamics. Colonization in the area started approximately 100 years ago and continues until present. As a result most of the primary forest has now disappeared. Both land cover conversion and land cover modification through human land use continue to take place. Although several studies to land use dynamics have been carried out for the area, they were mainly focused on deforestation trends (Sader and Joyce 1988, Veldkamp *et al.* 1992). The successive colonization of the region induced by the construction of the railroad can be characterized by a number of typical land use modifications:

- Primary forest is converted into secondary forest and extensive agriculture. In the present context extensive agriculture comprises patches of primary and secondary forest in combination with extensively managed pastures for cattle breeding and arable farming for home consumption (Sader and Joyce 1988).
- After deforestation, infrastructure is improved and an intensification (in terms of inputs per unit area and time) of agriculture takes place, coinciding with the cultivation of annuals and perennials as cash crops.
- On fertile, well drained soils, banana plantations are developed, partially directly after deforestation, but mostly as conversions after pastures. More recently other plantation crops like ornamental plants and palm heart (*Bactris Gasipaes*) have been introduced.

Although these land use modifications in general terms are described in literature, the process and the dynamics have never been quantified.

The study area comprises 2950 km^2 for which aerial photographs at a scale of 1:60,000 are available for 1992. For a smaller area of 150 km^2, aerial photographs at scales of 1:60,000 and more detailed, with approximately 10 year intervals have been used (1948/1952, 1960, 1973, 1984, 1992) to draw five sets of land use maps. Neither for 1948 nor for 1952 a full coverage of the study area was available. Therefore, a combined interpretation was made, hereafter referred to as 1950. Ground truthing has been carried out in 1992, to verify the 1992 photo interpretation. The classification procedure developed for 1992 has been used to interpret the photographs of the previous years. A 1992 Landsat TM image was available for the Atlantic Zone to determine the colonization frontier in the areas outside the study area. The 1:150,000 soil information system as described in Chapter 2 was used to derive soil data. All data were processed using PC Arc/Info version 3.4.2.

3.2.1 A single-time analysis of spatial patterns

If only a single set of land use data is available, which is often the case in developing countries, it is normally impossible to draw any conclusions on temporal dynamics. However, if qualitative knowledge on land use evolution is available it may be combined with the land use data set, in order to derive insight in the evolution of spatial patterns. In the study area, the geographical position between the old railroad and the actual deforestation frontier is related to the earliness of colonization, and in particular land cover conversion from forest cover. Thus, all geographic positions in the area can be related to a certain stage in the land use sequence starting from the earliest, i.e. closest to the railroad. By interpreting the spatial pattern as a chronological sequence, land use patterns in time may be deducted. However, this pattern may be modified by soil type and other physical parameters (e.g. altitude and slope). These modifications can be studied by stratifying the area on the basis of these parameters.

<u>Calculation procedure</u>

The study area comprises two main agro-ecological zones, a mountainous area in the south and large alluvial plains in the north (Figure 1.2). The mountainous area includes the footslopes of the Cordillera de Talamanca and the Cordillera Central and the higher part of the alluvial fans. The plains include the lower part of the alluvial fans and the large alluvial plain with some remnants of older mudflows and some Pleistocenic volcanic cones. The railroad, which was the starting point of the colonization of the Atlantic Zone, has been constructed on the limit of the two areas. The actual deforestation frontier in the north was identified on the 1992 aerial photographs. The northern part of the Atlantic Zone is not included in the study area due to the lack of aerial photographs. Nevertheless, to determine the location of the colonization zones in the study area, it is necessary to know the location of the deforestation frontier. The frontier was identified on the basis of satellite imagery (1991 Landsat TM).

Assuming linearity in the colonization pattern between the railroad and the deforestation frontier, 8 zones (N1 to N8) were identified between the railroad and the actual colonization frontier in the alluvial plains. South of the railroad, in the mountainous area, the colonization frontier is closer to the railroad and 4 zones (S1 to

S4) were identified. Because the colonization zones do not represent equidistant zones around the railroad, the area between the railroad and colonization frontier was subdivided in zones that were characterized by a specific ratio between the distance to the railroad and the distance to the colonization frontier. The colonization zones vary in width between 1.5 km in the east to 5 km in the west (where colonization took place more rapidly).

The soil map has been generalized into four main soil groups: (i) well drained soils with a relatively high soil fertility, (ii) poorly drained soils with a relatively high soil fertility, (iii) the well drained soils with a relatively low soil fertility, and (iv) the peat soils in the swamps. These soil groups can be classified according to Soil Taxonomy (Soil Survey Staff 1994) as (i) Andic Eutropept and Typic Udivitrand, (ii) Aquandic Tropaquept, (iii) Oxic and Andic Humitropept, and (iv) Histosols. The peat soils only occupy a very small area in the colonized area and were excluded from the analysis. In part, the poorly drained soils are artificially drained and as consequent similar to the fertile, well drained soils.

The 1992 land use map was generalized into five broad land cover types: forest (F), extensive agriculture (Ex), pastures (Pa), mixed cropping (Mi) and plantations (Pl). Both the soil map and land use map are characterized by many complex mapping units representing combinations of different soils and land use. Due to the scale of the original data, spatial desaggregation of soil or land use is impossible. For soils, the dominant soil type in each of the mapping units is considered as being the driving force behind land use changes. For land use this resulted in general classes of land use associations.

The colonization zones are projected on an overlay of the soil map and the land use map, and consequently land use in each of the zones was determined.

<u>Results</u>

Figure 3.1 presents the 1992 land cover map created for the study area. The subdivision in colonization areas is presented in Figure 3.2. The overlay of the land cover map, the soil map (Figure 1.3) and the colonization areas resulted in the percentage area coverage of the five broad land cover types for three major soil groups as a function of the geographic position (colonization zones) (Figure 3.3).

It can be observed that the fertile well drained soils in the plains are rapidly deforested after colonization and replaced by extensive agriculture. Extensive agriculture is slowly replaced by plantations and pastures. In the mountainous areas deforestation seems to proceed at slower rates, probably due to the sloping landforms

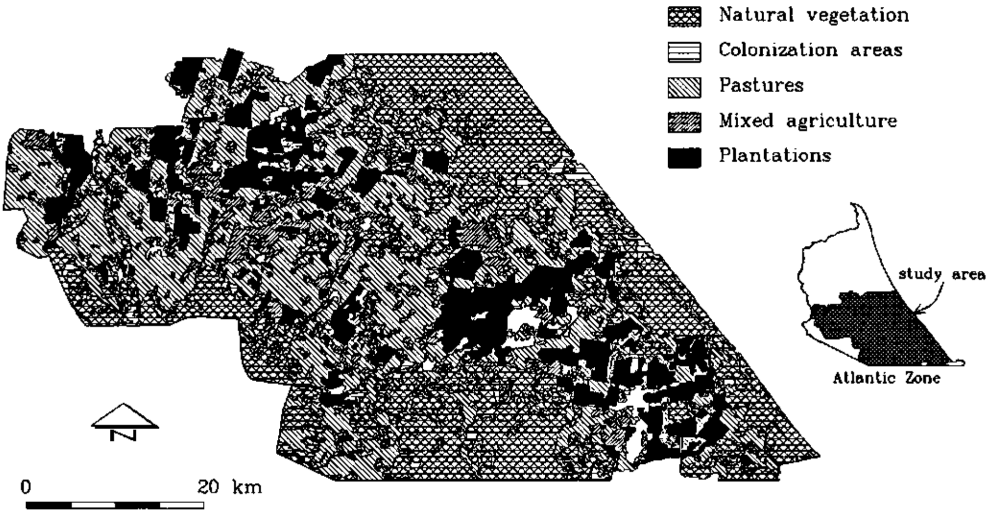


Figure 3.1   1992 land cover in the study area

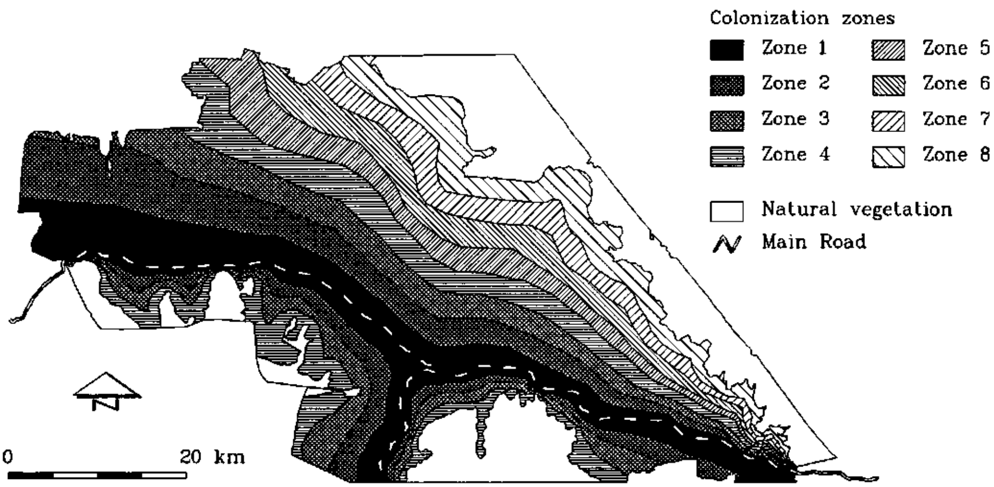


Figure 3.2   Colonization zones in the study area

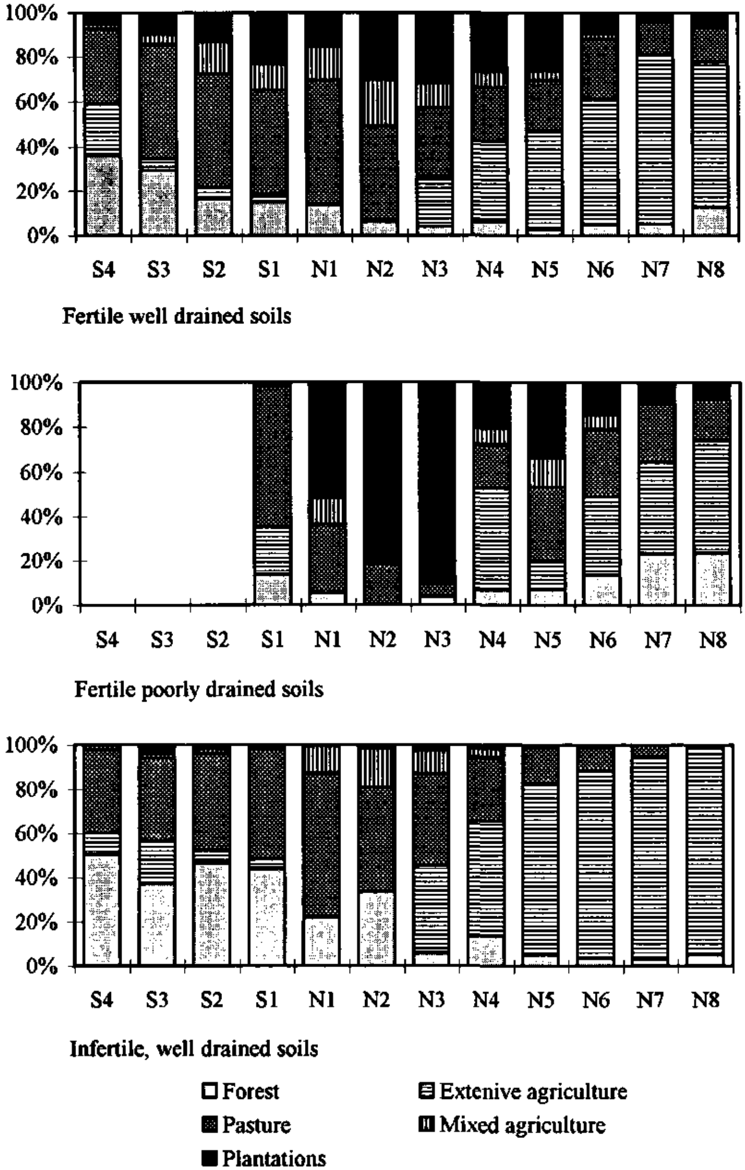


Figure 3.3 Land cover per soil type and colonization zone

and possibly accessibility. Coinciding with the deforestation, the area under plantations, mixed cropping and pastures slowly expands.

On the fertile, poorly drained soils, land cover modifications are less pronounced. The importance of plantations (up to 90% in N3) on these soils is striking. Because deep artificial drainage is standard practice in the plantations, water excess is not a severe limitation. In addition, plantations require large areas (at least 100 ha) that are united. The fertile well drained soils are all developed and hard to obtain, whereas the poorly drained soils are mostly under pasture and extensively used. In the southern mountainous areas, fertile, poorly drained soils are almost absent.

Infertile, well drained soils are dominated by extensive agriculture in recently colonized areas in the north (N5-N8). Extensive agriculture is mostly replaced by pastures in the older colonized areas (N1-N2) around the railroad. South of the railroad, in the mountainous areas, considerable forest cover (almost 50% of the area) is still found.

The results show that the temporal sequence of land cover transformation and modification through human land use can be deducted from the actual spatial distribution in relation to the railroad and the deforestation frontier. Large differences between the land use sequences on the different soils are seen, indicating the importance of such a stratification parameter.

3.2.2 Standard Markov chains with a soil type modifier

If the distribution of land use is known for more than one year, a transition matrix may be established. This transition matrix is used for the calculation of probabilities of land use sequences or Markov chains (Cox and Miller 1965). At a detailed temporal scale the sequences correspond to land cover modification or crop rotations (Jansen 1994) At a coarser temporal scale, the sequences correspond to land cover conversions, such as the conversion of forest to pasture or arable lands. and to larger time lags.

With relatively long time lags, several land use conversions can take place within one time lag. Two sequences that at first sight look different may well be the same when several modifications take place within one time lag. For example, both sequences Forest-Pasture-Mixed and Forest-Extensive-Mixed can be found when the actual sequence is Forest-Extensive-Pasture-Mixed, and land use modifications take place fast compared to the temporal resolution.

The analysis should be time-independent, i.e. sequences are the same starting at t_1 or t_2. This means that the sequence F-F-Ex-Pa equals F-Ex-Pa-Mi. For a single step probability or with small time lags, this problem will not occur.

Markov chains are based on the calculation of the probability that one land cover changes into another. In the modified Markov approach, soil types are introduced to reflect the fact, that for biophysical reasons, land covers are closely linked to soil types (e.g. banana plantations occur only on fertile soils).

Calculation procedure

Firstly, five sets of aerial photographs (1950-1992) were interpreted using the classification key developed for 1992. The interpretations was corrected geometrically to a single geographic basis. An overlay of the five land use maps was produced resulting in land use sequences for each of the newly created polygons. This resulted in a four-dimensional transition matrix where each cell corresponds to the total acreage of a certain land use sequence. For the study area, the time lags were relatively large (approximately 10 year steps). Due to the relatively small area, the sample of polygons is relatively small. In combination with the large number of possible combinations, almost each polygon yields an unique land use sequence. Therefore, the four samples, representing the land use modifications in each of the time steps have been combined yielding a single step Markov chains. As a result the analysis is time-independent: differences between 1950 and 1960 are treated similar to differences between 1960-1973. A soil type modifier is included by projecting the overlay on the soil map.

Results

Figure 3.4 presents the 5 land cover maps for the study area. The single step Markov chains (Table 3.1) indicate the probabilities for land cover conversions on each of the soil types. The diagonal represents the probability that no land cover modification takes place in a ten-year interval. The generally low probability, around 0.5, indicates that the area is highly dynamic. For example, it can be deducted that the area under forest may reduce every decade with 54% on the fertile well drained soils, rates that are also found by Sader and Joyce (1988) and Veldkamp *et al.* (1992). In 1950 almost 50% of the study area was under forest compared to 10% in 1992. The fertile poorly drained soils show the lowest deforestation rate, probably due to their low agricultural potential. As defined here the forest cover includes both primary (logged) forest and secondary regrowth. This explains why conversions from other land covers into a forest cover occur: these imply conversions into secondary forest types.

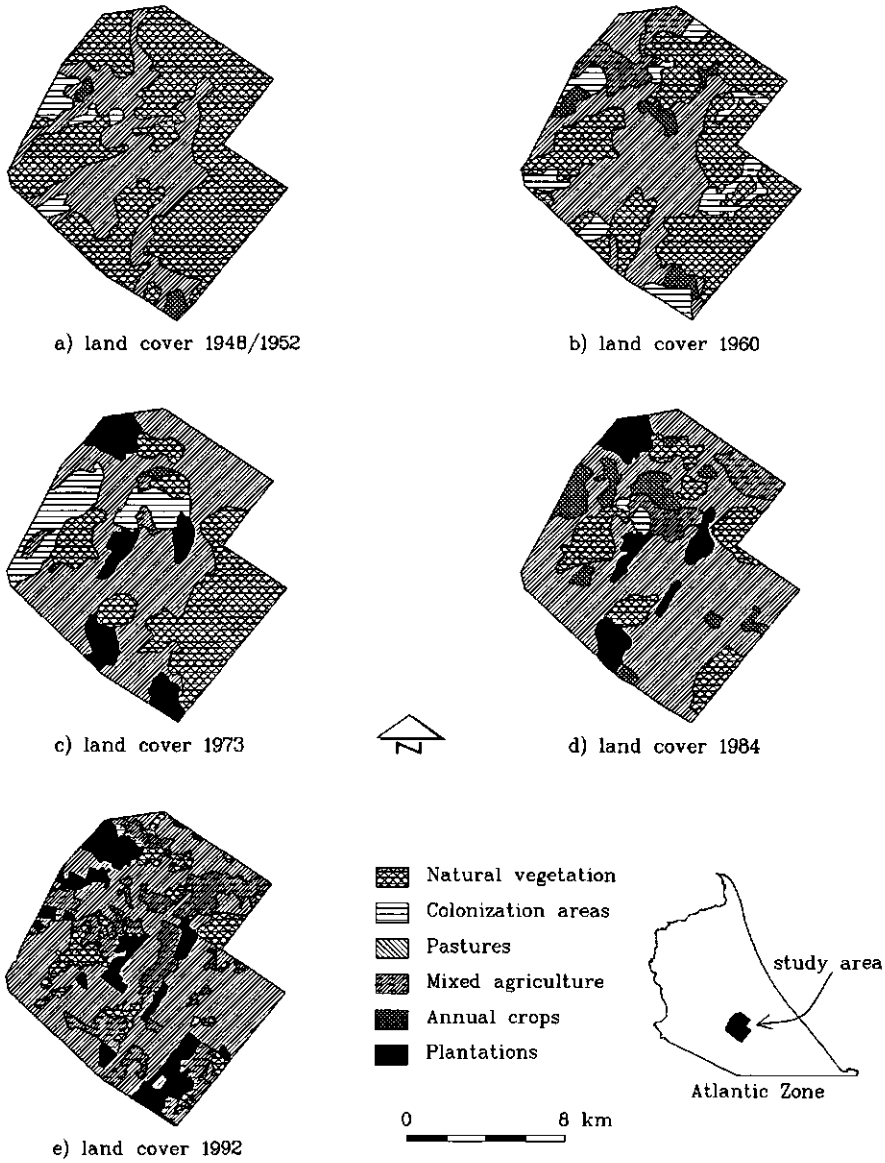


Figure 3.4 Land use in the study area between 1950 and 1992

Extensive agriculture is mostly converted into pastures and on the fertile soils into plantations. The area under a certain land cover does not influence the probabilities for land cover change. For instance, extensive agriculture is almost absent on the fertile poorly drained soils, reducing significantly the value and accuracy of the probability estimates for these soils. Only small differences occur between the soil types. Conversion of any of the land covers into plantations on the infertile, well drained soils is unlikely. Generally, on the fertile, well drained soil the probabilities to convert any of the land covers into mixed cropping or plantations is high, corresponding with the observed trends in the study area.

Table 3.1 Probabilities for land cover changes in approximate 10 year time steps per soil type

Soil	Fertile, well drained					Fertile, poorly drained					Infertile, well drained				
	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl
Forest	0.46	0.14	0.15	0.12	0.13	0.56	0.09	0.16	0.08	0.11	0.40	0.18	0.19	0.20	0.03
Ext. Agr.	0.11	0.47	0.08	0.15	0.19	0.12	0.43	0.09	0.18	0.18	0.10	0.34	0.38	0.16	0.02
Pasture	0.10	0.12	0.48	0.14	0.26	0.13	0.08	0.41	0.25	0.13	0.16	0.09	0.37	0.30	0.08
Mixed	0.04	0.03	0.20	0.50	0.24	0.06	0.04	0.19	0.38	0.33	0.07	0.04	0.20	0.67	0.02
Plantation	0.05	0.10	0.27	0.12	0.36	0.13	0.07	0.26	0.19	0.36	-	-	-	-	-

3.2.3 Geo-referenced Markov chains

The sequences observed on the basis of the standard Markov chain analysis are refined by introducing probabilities based on geometric and geographic position. The geometric position reflects the size and shape of the polygon - as expressed by surface area and an index value for the shape of the polygon. Generally, land cover conversion starts at the polygon margins (Skole and Tucker 1993).

Calculation procedure

The calculation is based on four consecutive steps:

1) Similar to the standard Markov chains, the different aerial photographs are interpreted and geometrically corrected.

2) For each of the polygons the surface area and circonference are provided by the GIS package. Two additional parameters for *shape* and *neighbouring land covers* are calculated.

- An index value representing the shape of the polygon is calculated:

$$S = 4\pi \frac{A}{C^2} \qquad \text{Equation 3.1}$$

In which:	S	=	index for shape
	A	=	area of the polygon
	C	=	circonference of the polygon

  The index value S ranges from 1 for circular polygons to approaching 0 for extremely irregular shaped polygons and is independent of size. Polygons along the border of the map are excluded from the analysis due to their unknown shape.

- The length of the boundaries between the *nth* polygon and each of its neighbours is available in the line attribute table in PC Arc/Info. Land covers are ranked according to the management intensity and the observed conversion sequence: 1) Primary forest, 2) Secondary forest, 3) Extensive agriculture, 4) Pasture, 5) Mixed farming, and 6) Plantation. A boundary index N for the type of land use in neighbouring polygons is calculated by multiplying the fraction of the boundary with the corresponding rank number.

$$N = \sum_{Lu=1}^{6} f * Lu \qquad \text{Equation 3.2}$$

In which:	N	=	boundary index
	f	=	fraction of border neighbouring with Lu
	Lu	=	land cover rank

  The index value N will range between 1 when the polygon is completely surrounded by forest to 6 when the polygon is surrounded with plantation. In other words, the boundary index indicates the weighed average ranked land cover of neighbouring polygons.

3) An overlay is created of the 5 different land use maps.

4) The land use modification between t_i and t_{i+1} is calculated and stratified for size, shape and neighbour index. Depending on the size of the study area, the number of

observation dates and the number of polygons, the size, shape index and neighbour index can be stratified in a number of classes. Although the number of observations in the Costa Rican case study is relatively high, the area for which the aerial photographs were available is small, especially when the outside polygons have to be excluded. For illustrative purposes the effect of the three different parameters is presented separately.

Table 3.2 Probabilities for land cover changes stratified for polygons size, shape, and neighbours

Size	< 2 km^2					2-5 km^2					>5 km^2				
	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl
Forest	0.43	0.12	0.21	0.13	0.11	0.56	0.11	0.13	0.13	0.07	0.84	0.11	0.04	0.01	0.00
Ext. Agr.	0.12	0.37	0.11	0.20	0.20	0.12	0.43	0.09	0.18	0.18	0.15	0.79	0.03	0.02	0.01
Pasture	0.12	0.08	0.43	0.26	0.11	0.13	0.08	0.41	0.25	0.13	0.02	0.13	0.79	0.03	0.03
Mixed	0.07	0.03	0.10	0.53	0.27	0.06	0.04	0.19	0.38	0.33	0.06	0.08	0.09	0.57	0.20
Plantation	0.17	0.00	0.36	0.10	0.37	0.13	0.07	0.26	0.19	0.36	0.00	0.00	0.02	0.02	0.96
Shape	< 0.33					0.33-0.66					> 0.66				
	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl
Forest	0.46	0.05	0.27	0.19	0.03	0.53	0.11	0.11	0.15	0.11	0.56	0.10	0.23	0.11	0.00
Ext. Agr.	0.20	0.44	0.12	0.09	0.15	0.12	0.43	0.09	0.18	0.18	0.12	0.47	0.08	0.15	0.18
Pasture	0.16	0.03	0.47	0.25	0.09	0.13	0.08	0.41	0.25	0.13	0.08	0.14	0.41	0.34	0.03
Mixed	0.05	0.16	0.09	0.31	0.39	0.06	0.04	0.19	0.38	0.33	0.03	0.02	0.13	0.46	0.36
Plantation	0.08	0.02	0.33	0.27	0.30	0.13	0.07	0.26	0.19	0.36	0.04	0.07	0.22	0.15	0.52
Neighbours	< 2.33					2.33-3.66					> 3.66				
	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl	F	Ex	Pa	Mi	Pl
Forest	0.40	0.10	0.15	0.22	0.13	0.34	0.12	0.24	0.14	0.16	0.37	0.12	0.24	0.14	0.13
Ext. Agr.	0.16	0.33	0.16	0.19	0.17	0.15	0.38	0.09	0.18	0.19	0.14	0.35	0.12	0.18	0.21
Pasture	0.14	0.17	0.34	0.21	0.14	0.21	0.15	0.28	0.25	0.10	0.11	0.15	0.41	0.23	0.10
Mixed	0.11	0.13	0.16	0.27	0.34	0.13	0.03	0.17	0.32	0.35	0.16	0.04	0.16	0.34	0.30
Plantation	0.20	0.13	0.17	0.16	0.34	0.20	0.08	0.21	0.23	0.28	0.14	0.06	0.26	0.18	0.36

Results

The results with separate stratifications for size, shape and neighbours of the polygons are presented in Table 3.2.

The size of the polygons is classified in three groups, each with an approximately equal number of observations. The size of the polygons significantly influences the land cover modifications. This is clearly shown by the probability that the conversion of forest will take place, which ranges from 0.57 for small polygons to 0.16 for the large polygons. Similar results can be observed for extensive agriculture, pasture and plantations.

Although the shape of a polygon theoretically influences the probabilities, the results do not clearly confirm this effect. Nevertheless a slight positive relation can be observed. Disturbances by other parameters like soil type and polygon size, for which a clear relation has been found, may confound the influence of shape. A combined analysis may be necessary, especially because there may be a relationship between shape, land cover and soil type. Long narrow irregular shapes may, for instance, correspond with riverline or swamp forest on poorly drained soils. The shape, cover and land use will be related to drainage patterns and relief.

Similar to the size and shape index, the boundary index has been classified into three groups with equal numbers of polygons. Similar to the shape index no clear relations can be found.

3.3 Discussion and applications of the different land use dynamics indicators

Changes in land use and cover over time have been studied in three different manners. The merits and drawbacks of the three methods may be formulated as follows:

- Single time analysis:

The advantage of this method is that only one land use data set is necessary to obtain results. A clear disadvantage is that it can only be used in areas for which qualitative data on the colonization history are available. It is probably most suitable to deal with land cover conversion rather than with land cover modification, such as intensification - e.g. the change to high yielding varieties and use of higher inputs, which are too detailed to be captured. In the single time analysis, even when good insight in the colonization history does exist, supposed changes may actually involve different developments (or land use sequences) at different locations in the same zone.

Alternative approaches may be developed for other areas where clear spatial-temporal patterns of land use exist in relation to e.g. urbanization, mining or tourist industry developments.

- Classical Markov chains:

When more than one land use inventory is or can be carried out, Markov chains are an appropriate tool. The advantage above the singe time analysis is that for each polygon land cover changes are quantified in time. It is demonstrated that the inclusion of a soil modifier is an important adaption in the standard Markov chain approach.

- Adapted Markov chains:

Indicators to be combined with Markov chains are created by carrying out a spatial pattern analysis of the polygons themselves and the neighbouring polygons (border analysis). Combining the spatial pattern indicators with a stratification for soil type is possible, although it requires large datasets to have sufficient data for each class. In that case, the shape parameters should be determined before the overlay is made with the soil survey.

There is no conclusive evidence that shape and neighbouring relations are related with land cover change probabilities, although this may be likely on practical grounds (e.g. Skole and Tucker 1993). This merits further study. In addition, there is an increasing need for statistical analysis to support for example observed differences in Markov chains or between maps. Recent reviews on this topic do not yield any statistical test for this kind of analysis (see for example in Bailey 1994 and Bonham-Carter 1994).

The quantification of land use dynamics is essential in studies of land use and cover change to assess the impact on land degradation or the feedback to climate change models. This would improve the current approaches in land use inventory studies that focus on spatial analysis but remain qualitative and do not allow a comparison with other regions.

The proposed indicators for the description of land use dynamics can be determined with standard GIS packages, although for the adapted Markov chains a vector based system is preferable.

4. Linking GIS and models

The sections of this chapter are based on the following publications:

Section 4.1 Stoorvogel, J.J., 1995. Linking GIS and models: structures and operationalisation for a Costa Rican case study. *Netherlands Journal of Agricultural Science* 43: 19-29.

Section 4.2 Stoorvogel, J.J., 1993. Optimizing land use distribution to minimize nutrient depletion: a case study for the Atlantic Zone of Costa Rica. *Geoderma* 60: 277-292.

Section 4.3 Stoorvogel, J.J., 1994. Integration of computer-based models and tools to evaluate alternative land use scenarios, as part of an agricultural systems analysis. *Agricultural Systems*: in press. (Invited paper at the ICASA symposium 'Role of Agronomic models in Interdisciplinary Research', Annual ASA meetings, 13-18 November 1994, Seattle (WA), USA.)

4.1 Structures and operationalisation

4.1.1 The need to link GIS and models

Although GIS are powerful tools for the analysis of geo-information, commercial GIS packages do not always meet the specific needs of users (O'Kelly 1994). User requirements often comprise very specific disciplinary operations and user oriented shells. To a limited extent, GIS software enables the development of applications that may include (simple) models. Implementation of additional procedures into GIS packages or modifications of GIS packages usually coincides with high costs due to the complexity of GIS software and additional modelling systems (Abel *et al.* 1994). Therefore, GIS can be considered to be a closed system, i.e. no changes in the internal schemes of the software can be made. Specific disciplinary analysis like crop growth simulation need, therefore, external models, which work independently from the GIS and perform the analysis which the GIS package is unable to handle. For operationalisation, the GIS needs to be linked to these external models. Although the necessity to link GIS with models is generally recognized, many practical problems are known to occur (Burrough 1989, Abel *et al.* 1994). Part of the problems originate in the incompatibility of data formats, data organization or semantics which respectively requires reformatting, restructuring and data analysis before the GIS database can be used in combination with external models. No GIS architecture has yet been developed that conceptualizes the link between GIS and external models. At present, therefore, the link between models and GIS is often established in an ad-hoc manner (e.g. Meijerink 1989, Steyaert and Goodchild 1994). Specially designed structures may facilitate this link and can be included in the GIS for operationalisation.

In this section, the structure and operationalisation of the link between a LP model and GIS as it is used in the USTED methodology is presented. The USTED methodology (*Uso Sostenible de Tierras en El Desarrollo*; Sustainable Land Use in Development) aims at the analysis of alternative land use scenarios. Using these land use scenarios, the effect of agricultural policies can be analyzed. The methodology uses a LP model for the basic calculations, in combination with a GIS and crop growth simulation models.

Although the structure is developed for a specific case, it can form the basis for many different applications in which a GIS and external models are linked as illustrated in Sections 4.2 and 4.3.

4.1.2 A general structure for GIS-model linkage

Concepts

Large spatial databases are often organized in a layered structure, where each layer stores data on geographical features related to a specific theme and within a specific geographic area (Frank and Mark 1991). In the present context, each layer is referred to as a map. A GIS database from a land use planning project may thus contain layers with data on soils, climate and land use. The geographical features in each of these layers are polygons, which represent areas with a specific soil type, climate and land use respectively. Combinations of layers can be analyzed through map overlays. Van Oosterom (1990) indicates a hierarchical structure for the description of geographical features (Figure 4.1). Each feature is characterized by spatial data and thematic attributes, which are linked by an unique identifier. Spatial data are subdivided in geometric data and topological data. The geometric data comprise information on the position and shape of the features. On the basis of the geometric data of all the features in one map, the topological data for the individual features can be determined which indicate the spatial relationships (e.g. connectivity and adjacency) between the features. The thematic attributes include the characteristics of the geographical features (e.g. soil type and soil properties for a mapping unit in a soil map).

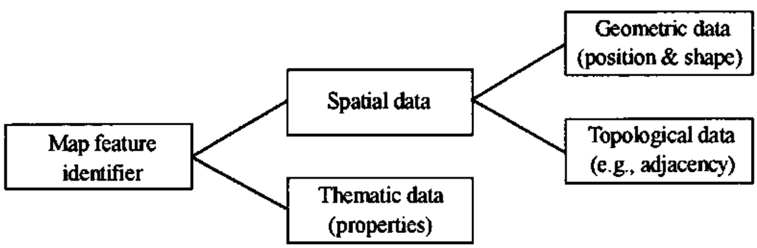


Figure 4.1 A hierarchical structure for data storage in a GIS (after Van Oosterom 1990)

A model is a formal relation between exogenous input parameters and derived endogenous output parameters. In the case of e.g. a crop growth simulation model, the exogenous input parameters comprise data on soil, crop and climate. On the basis of these input data, the model calculates e.g. the expected crop yield, which is the endogenous output parameter. Models, which are linked to a GIS, can use geometric, topological and thematic attributes as exogenous parameters and calculate new

geometric data and/or thematic attributes as endogenous parameters. It is not necessary to import topological data in the GIS, as they are a function of the geometric data and, therefore, can be determined within the GIS. Variations in the type of model input and output can vary considerably as shown in the following three examples:
- Crop growth simulation models can simulate the crop production for polygons representing objects at different aggregation levels (e.g. field, farm, agro-ecological zone) on the basis of climatic and pedological characteristics (e.g. Van Keulen and Wolf 1986). For the simulation, thematic attributes for the different geographical features (soil and climatic properties) are used. The simulated crop productions can be linked as new thematic attributes to the spatial objects represented in the map.
- A model for the infestation of a crop with pests requires data for the specific field which is being modelled (e.g. soil type, micro-climate). In addition, the occurrence of the host plants in surrounding fields may influence the risk of infestation. Hence, the input exists of topological data and thematic attributes of both the modelled and neighbouring fields. Although the model input exists of both topological and thematic attributes, the output only comprises thematic attributes, e.g. the risk for the infestation of a specific field.
- Models which simulate three-dimensional processes like ground water-flows, need spatial data (including geometric and topological data) as well as thematic attributes. During three-dimensional modelling a new geometry will be created by changing the geometry of existing objects and possibly the creation of new objects. When data are stored in raster format, spatial objects are represented by a set of fields. The spatial definition of the fields remains the same and, therefore, no changes in the geometry take place. In the case of vector-based maps, the changes in the geometry will lead in many cases to a new geometry. If the geometry changes, both geometric data and thematic attributes will be imported and a new map will be created in the GIS. However, if the geometry of the features is not altered, the model results can be imported and added as thematic attributes to the original map from which the data where exported.

Although models differ in their data requirements, they do not necessarily differ in the type of data which need to be imported in the GIS. All three models illustrated in the examples may yield only thematic attributes, which generally can be linked to an existing map in the GIS.

Structures to link a GIS with external models

When a GIS is linked to an external model, the GIS provides the input data and the model subsequently determines the derived parameters for the geographical features or for the map, in cases where the model changes the geometry of the objects. Afterwards the GIS can be used to visualize and analyze the model results. The link between GIS and external models is two-way and is primarily based on data interchange. Figure 4.2 gives a general framework for the GIS-model interface, based on six main steps. In most cases, available data in the GIS have to be translated to the specific input parameters of the model (Step 1-2). The data are exported to the model (step 3), and

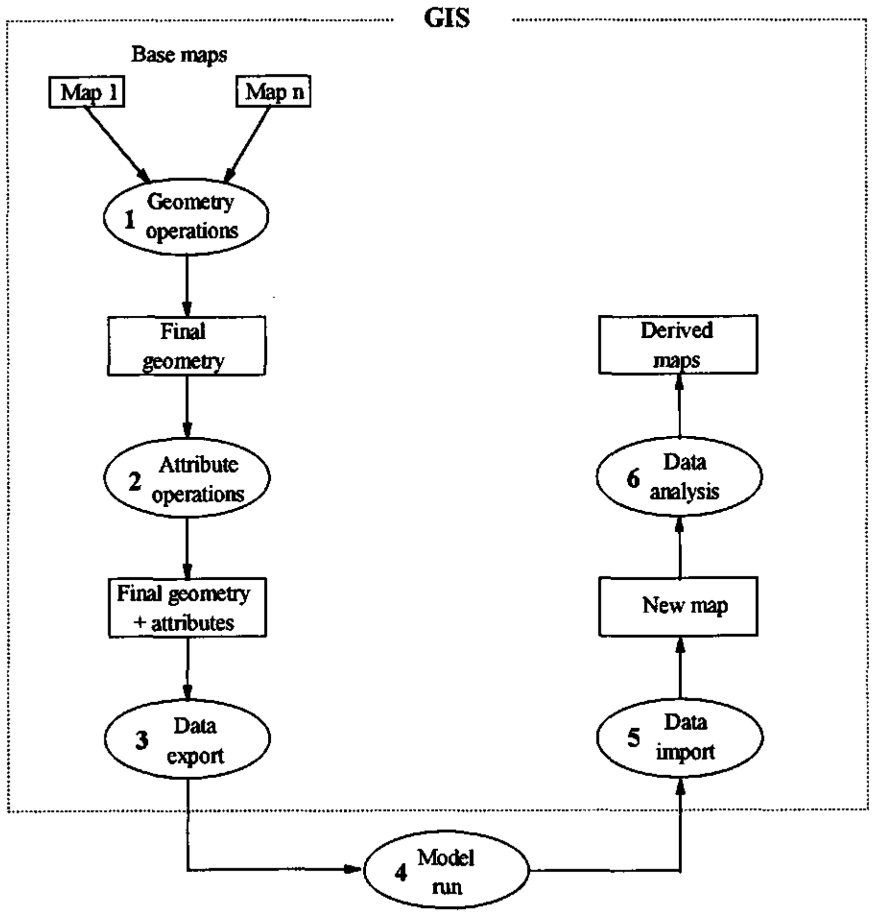


Figure 4.2 A six step approach (indicated by the ovals) for the GIS-model interface

after the model run (step 4), the model outputs are imported in the GIS database (step 5). The GIS can subsequently visualize and analyze the model results (step 6). The steps will be described and illustrated with an example for the link of a crop growth simulation model with a GIS to calculate the regional distribution of potential production.

Step 1 deals with the formulation of the geometric data. When the geographical features are not yet defined, they have to be created on basis of one or several base maps. Step 1 results in a map with the proper geometry, i.e. the features are the basic elements for the model. The map, however, may still lack the proper thematic attributes necessary for the model. The geometry operations used to define the geographical features include overlay operations for the combinations of maps but also buffer operations to generate zones around geographical features within certain spatial proximities, e.g. the area within a certain distance from a road, and the calculation of slope and aspect on the basis of a digital elevation model. For crop growth simulation models the map features should be characterized by a combination of soil type and climate. An overlay of the soil map with the climatic map yields polygons with a specific climate and soil.

In Step 2 the thematic attributes for the map are determined. A number of attribute operations, which may comprise mathematical and statistical calculations, have to be performed to acquire the correct variables. Queries to the database are necessary for the appropriate data structure. The result is a map with the appropriate map features and thematic attributes. In the case of the crop growth simulation model, queries can select the required climatic and soil properties for the different polygons.

Step 3 deals with the export of geometric, topological and/or thematic attributes from the GIS to the external model. In general, standard formats for data interchange (e.g. "comma separated value" files for thematic attributes and "digital line graph" files for geometry data) can be used. A well structured data exchange will enable automatization and thus operationalisation of the GIS-model interface. In the case of a crop growth simulation model, the identifier of the geographical feature and different climatic and soil properties are organized in a table and exported to the simulation model.

Step 4 comprises the actual model run, where on basis of the exogenous parameters, the endogenous parameters are calculated. Depending on the type of model, runs are carried out for the whole map or for the individual features. In the case of a crop growth simulation model the simulation will be carried out for each individual feature, or, if different features with the same thematic attributes occur, groups of features. During the model run the identifiers to the original geographical features should be preserved to link the model results to the original map.

The results of the model are imported into the GIS in **step 5**. This is the reverse procedure of step 3 and in most cases similar formats for data interchange can be used. If the geometry of the geographical features is not changed, the endogenous model parameters become new thematic attributes for the features of the base map. In the case of models which change the geometry, a new map is created on the basis of the model outcome.

Finally, in **step 6**, the model results are further analyzed. Subsequently, the model outcome may be analyzed in the GIS, through e.g. aggregation of geographical features with similar model results, or overlays with other maps. In the case of alternative production systems, the analysis may e.g. include a comparison with actual land use and production levels.

The structure as presented in Figure 4.2 mainly comprises a series of consecutive operations in the GIS environment, which is specific for each GIS-model link. Most GIS packages enable the development of simple applications. Although in many cases the models can not be defined in these applications, they can automate or support the GIS-model interface, resulting in the operationalisation of a series of standard GIS commands through a macro language.

4.1.3 The application of a GIS-model link

Land use scenarios can be analyzed by LP models (e.g. Schipper *et al.* 1995, Veeneklaas 1990). Typically, GIS does not provide tools for LP and no commercial LP software includes GIS facilities. Although LP models are not spatial (Chuvieco 1993), they can be linked to a GIS to relate the analysis to certain geographical features (e.g. farms, fields). The spatial presentation of the model results enables a quick interpretation of the LP results and a spatial analysis.

The USTED methodology has been developed for the analysis of land use scenarios. The LP model maximizes total net farm income for a sub-region (e.g. settlement or municipality), through a simultaneous selection of alternative land use systems for different farm types in the region. The farms are grouped into farm types according to size and soil type distribution. Alternative land use scenarios represent changes in the socio-economic or bio-physical environment of the region affecting the goal function, the constraints, and/or the alternative activities (land use systems and technologies, denominated LUSTs) of the LP model. The constraints indicate the availability of resources like land and labor, but may also comprise restrictions on other parameters for e.g. sustainability. The USTED methodology is operationalized for the Neguev settlement.

The link between LP model and GIS is established by the six step approach presented in figure 4.2. The results are given in Figure 4.3[a] (Step 1- 3) and 4.3[b] (Step 4-6).

Step 1: Two base maps, a 1:20,000 map of the farms (Anonymous 1981) and a soil map at the same scale (De Bruin 1992), were combined by an overlay procedure to yield a map with the farms and the soil limits. The thematic data of the map yield the soil types and the size of each of the 307 farms.

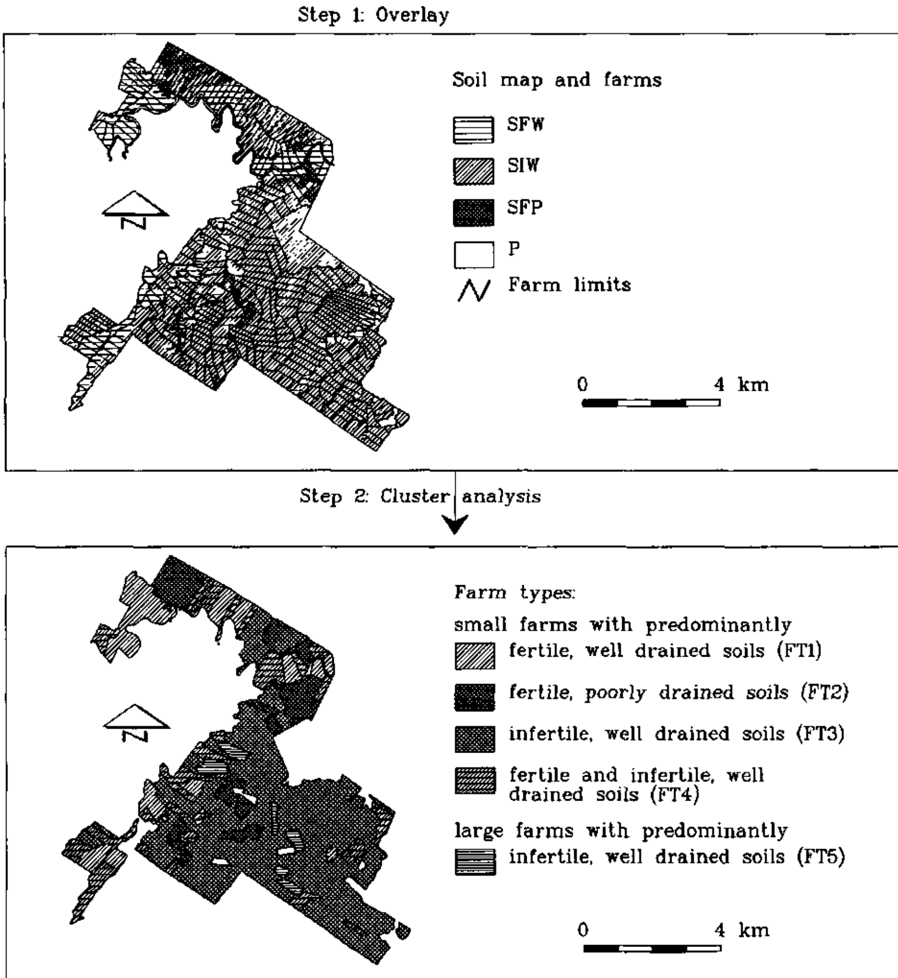
Step 2: The thematic attributes of the combined map comprise the farm identifier and a reference to a specific soil type. The LP model, however, does not deal with individual farms but considers farm types. Consequently, a farm classification was carried out by means of a cluster analysis (Schipper *et al.* 1995), resulting in five different farm types, each with a specific size and soil type distribution.

Step 3: The input parameters for the LP model, comprising the number of farms in each farm type and its average size and soil types, are exported to a file which can be read by the model.

Step 4: During the optimization with the LP model, LUSTs are selected for the soil types on each of the farm types. The selection of the LUSTs is based on the maximisation of the total net farm income, given the constraints which are defined for the model. The results of the LP model indicate the selected LUSTs for each soil type on the different farm types.

Step 5: The output of the LP model presents the selected land use systems for the different farm types. This data is linked to the map with the farm types, which was the result of Step 2. For each of the polygons (defined on the basis of farm type and soil type) the LP model can select several LUST. If more than one land use system is selected, LUSTs with high labor requirements are considered to be cultivated closer to the roads than crops with a low labor consumption. Using several buffer operations, the polygons of the map are subdivided in different zones, each between two distances to the road. The land use systems can now be distributed over the polygons. On the basis of the map with the optimal LUST distribution according to the scenario definition, a quick interpretation can take place.

Step 6: The analysis of the results is user dependent. Additional LUST characteristics can be linked to the map (e.g. biocide use in Figure 4.3[b]). This enables a spatial analysis of e.g. the sustainability of the scenario results. It may also yield data on



Farm type	Number	Area/farm (ha)	Soil type (in %) SFW	SIW	SFP
1	35	14.1	91	6	3
2	33	15.7	12	28	60
3	189	13.1	6	88	6
4	46	13.5	52	41	7
5	4	32.1	10	78	12

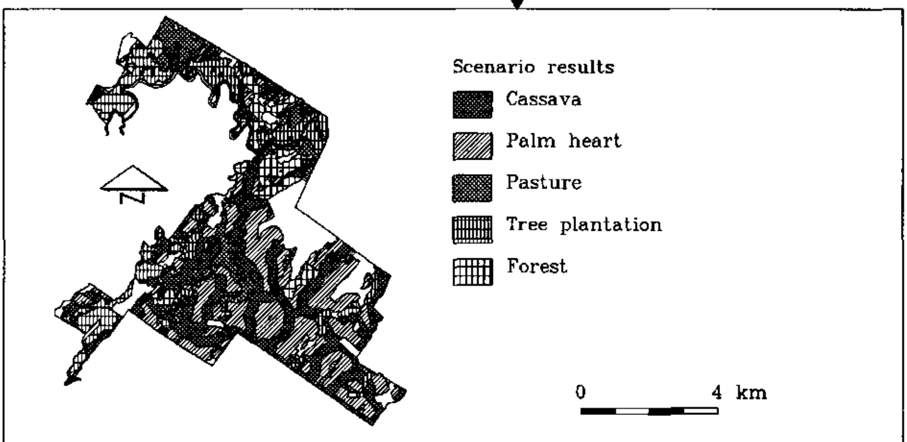
→ Step 4: LP model

Figure 4.3[a] The link between the GIS and the LP model for the Neguev case study, step 1-4

LP model →

Farm type	1	2	3	4	5
Forest	3.9	3.7	3.5	1.3	2.7
Tree plantation	1.3	3.0	4.6	11.0	0.3
Pasture	6.8	3.6	0.9	0.5	0.6
Cassava	0.0	0.5	1.1	1.3	1.0
Palm heart	2.7	21.5	3.8	0.6	8.5

Step 5: Data import



Step 6: Data analysis

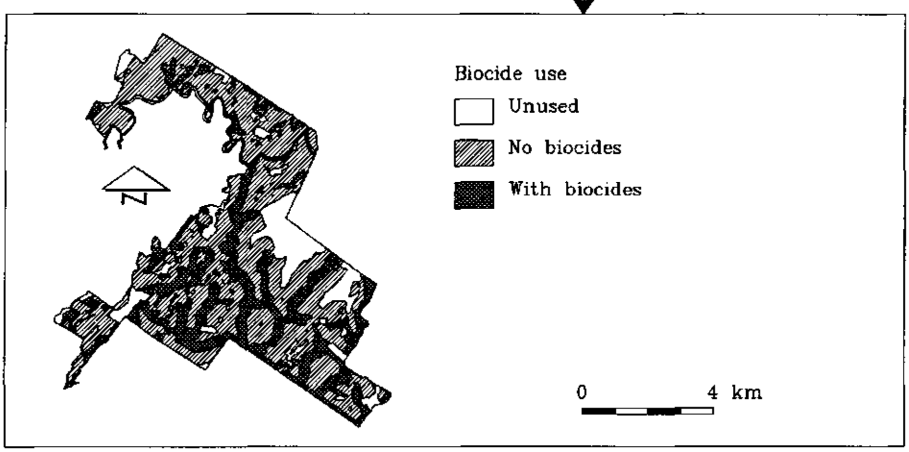


Figure 4.3[b] The link between the GIS and the LP model for the Neguev case study, step 4-6

spatial concentrations of specific productions, which may be important for the planning of specific services.

One of the advantages of scenario based studies as in the USTED methodology is the interactive way users can analyze the effect of changes in the socio-economic and the bio-physical environment on agricultural land use. Interactivity often results in a large number of alternative model runs and, therefore, requires a rapid interpretation of the results through the visualization of the scenario results and spatial analysis. The link between the GIS and the LP model determines the degree of interactivity. In the case of the Neguev settlement, the GIS allows for the development of applications through a macro language. The different steps described above are included as an application in the GIS. Therefore, although the LP model is not included within the GIS software, a highly interactive procedure is developed.

4.1.4 The operationalisation of the GIS-model link

GIS can play an important role in land use planning (Sharifi 1992, Despotakis 1991) and assessment of environmental projects (Campbell *et al.* 1989). Many applications need external modelling in combination with GIS. Nijkamp and Scholten (1993) emphasize a consistent use of GIS and models. Although the use of GIS and models for many applications is clear, linking is often a problem. Abel *et al.* (1994) provide a general structure for systems integration. They stress the importance of user interfaces for the GIS. However, with the use of commercial software packages, the user is restricted to the tools that come along with the package and, therefore, in the possibilities to link external models.

To avoid operational difficulties to implement specific disciplinary models in commercial software packages, the models can be linked to the GIS. Most GIS systems allow for the development of applications through e.g. macro-languages and can thus be used for simple models. The present Section shows the importance to focus the applications on the link between models and GIS. This can easily be operationalized for models which use topological and/or thematic attributes as input. In cases, where the model uses and changes geometric data, the model internally will make use of a GIS and integration may be necessary. The link between GIS and models of any kind may be useful to extend the standard possibilities of a GIS with additional operations. The structure for GIS-model linkage organizes the GIS-model interface in a number of relatively simple operations.

The six-step structure for the link between GIS and external models is generally applicable when users deal with one model and a GIS. In cases where additional database systems are involved the structure becomes more complex and more specific. Due to the clear definitions of the different steps, the structure can function as a good basis for the development of applications. Specific requirements for the operationalisation of the GIS-model link are:
- the availability of a GIS which allows for the development of applications,
- the availability of formats for data interchange which can be used by both GIS and model, and,
- users, who are aware of the limitations and the assumptions of the different procedures.

When these requirements are fulfilled, the operationalisation can be realized and in most cases automated. The efficiency of the operationalisation depends strongly on the number of user decisions which are necessary during the model run. Considering a relatively simple model, which determines the endogenous parameters on the basis of a set of exogenous parameters without any user decisions, the application can function as a new command within the GIS environment. The user does not have to be aware that an external model is being used. The risk of automatization of analysis and applications is that users unaware of the procedures may use the application as a black box and for datasets outside the range of validation. When users are relatively inexperienced with the models the link with GIS may be supported by a well designed interface to clearly show the interaction instead of full automatization.

4.2 Integration of GIS and linear programming models to minimize nutrient depletion

4.2.1 Land use planning on a sustainable basis

Land use planning is still too often seen as solely an optimization of agricultural production (e.g. Cooke 1982, Faber 1986). However, in the last decades awareness of sustainable agriculture is growing, demanding the development of techniques for land use planning which take the sustainability of the production into account (Fournier 1989, Lélé 1991, Sharifi 1992, Fresco *et al.* 1990). Therefore, the matching process between land units and land utilization types (FAO 1976) should not only take place on the basis of potential production but also include the sustainable basis of that production. One of many factors contributing to sustainable production is the soil nutrient balance, which, in a negative case, may lead to mining of the soil nutrient stock. Although the separate techniques to assess the nutrient balance and to perform land use planning are available, they are rarely combined into an operational method for land use planning. This section shows the importance of integrating the nutrient depletion in the land use planning procedure through a case study of the southern part of the Neguev settlement in the Atlantic Zone of Costa Rica. Current land utilization types have been redistributed over the area using a LP model, minimizing for the nutrient depletion without leading to a reduction in yield.

The study area of 36 km^2 is located in the southern part of the Neguev settlement scheme in the Atlantic Zone of Costa Rica (see Section 1.2) Originally the Neguev settlement was a large cattle ranch with some smaller parts under forest. Although the settlement scheme was established a decade ago, pasture and to a smaller extent forest still dominate land use in the area. Yet, smaller areas are presently cultivated with maize, red pepper, tubers, coconut, cacao, plantain and fruit trees. The annual and perennial crops are scattered on small parcels throughout the area. For the present study the semi-detailed land use map of the situation in 1987 (Overtoom *et al.* 1987) has been used (Figure 1.7).

A semi-detailed soil map of the settlement originates from De Bruin (1992). The soil map was generalized on the basis of soil fertility and drainage (Figure 1.5). The infertile, well drained soils are found on both flat and sloping positions. The group SIW is therefore subdivided in SIW_f with slopes of less than 6% and SIW_s with slopes over 6%. The swamps (P) are excluded from the calculations as it is not suitable for any of the land utilization types. There is a small discrepancy concerning the distribution of the swamps between the soil map and the land utilization types. Within the calculations the swamps of the land utilization map are used for the calculations. The parts that are classified as swamps on the soil map and not on the land utilization map are dealt with as soil type SFP. Table 4.1 presents the general characteristics of the different soils.

4.2.2 The calculation of the nutrient balance

Stoorvogel and Smaling (1991) developed a nutrient balance model (NUTBAL) for Sub-Saharan soils. This model is based on separate assessments for 5 input and 5 output factors. Input comprises the input of mineral (IN 1) and organic (IN 2) fertilizers, wet and dry deposition (IN 3), nitrogen fixation (IN 4) and sedimentation (IN 5). Output comprises the harvested product (OUT 1), removal of crop residues (OUT 2), leaching (OUT 3), gaseous losses (OUT 4) and erosion (OUT 5). The model

Table 4.1 Area and average soil properties for the generalized soil groups

Soil type	Area	Clay Content	Organic Matter	Total P	Total K	Exch. K	Bulk Density	Slope	K-factor
	km^2	%	%	%	%	meq 100g^{-1}	g cm^{-3}	%	
SFW	5.3	10	4.4	0.3	1.6	1.6	0.7	1	0.15
SFP	1.4	31	4.3	0.3	1.4	1.4	0.9	0	0.00
SIWf	16.6	55	6.3	0.1	0.3	0.6	0.9	3	0.11
SIWs	10.2	55	6.3	0.1	0.3	0.6	0.9	15	0.11
P	3.0	---------- not relevant ----------							

[De Bruin 1992]

was further elaborated by Smaling *et al.* (1992), who presented a more detailed version of NUTBAL (using the same input and output factors) for a high potential agricultural area in Kenya. The latter has been calibrated for the study area using data from Costa Rica. For the forest areas, a separate assessment is made based on the work carried out by Parker (1985).

For annual crops the use of mineral fertilizers (IN 1) is common in the Neguev area (Brink 1988, Van de Berg and Droog 1992). Although the fertilizer recommendations for the different soils vary considerably, no significant difference in actual fertilizer use between the soil types can be found. Therefore, fertilizer use is set at a fixed input per crop as presented in Table 4.2.

In the land use systems in the Neguev settlement no animal manure or any other organic fertilizer is applied. For the grasslands part of the nutrients are returned directly by the manure of the grazing animals. In general, 90% of the nutrients which are grazed return to the system in the form of excrements (IN 2) on pastures where the animals remain on the field (Stoorvogel and Smaling 1991). However, the nutrients in manure are subsequently exposed to an increased leaching and denitrification.

In humid regions with large quantities of rainfall, wet deposition (IN 3) normally provides more nutrients than dry deposition. Consequently, the latter is not considered in the present study. Although nutrient concentrations in rainfall are very low, the large amount of rainfall still provides a substantial amount of nutrients. Parker (1985) gives

Table 4.2 Harvest for the main crops in relation to soil type and the crop residue removal

LUT	Production kg $ha^{-1}yr^{-1}$			Fertilizer kg $ha^{-1}yr^{-1}$			Harvest index
	SFW	SFP	$SIW_{r,s}$	N	P	K	%
Maize	2300	800	1200	67	0	0	15
Red pepper	650	250	350	30	20	8	30
Tubers	5000	1000	3000	0	0	0	10
Coconut	3200	1500	2200	0	0	0	-
Cacao	800	---	300	0	0	0	-
Plantain	8000	4000	3000	12	0	0	-
Fruit trees	11300	4000	6000	0	0	0	-
Pasture	5000	4500	3000	0	0	0	-
Intercropping				0	0	0	-
Plantain	4500	2200	1700				
Cacao	400	---	150				

[Brink 1988, MAG 1991, Van de Berg and Droog 1992, Van Sluis *et al.* 1987]

an average annual nutrient input for La Selva (50 km North East of the study area with an average annual rainfall of 4007 mm) of 1.7 kg N ha^{-1}, 0.17 kg P ha^{-1} and 5.4 kg K ha^{-1}. These values have been corrected to 1.5, 0.15 and 4.9 kg $ha^{-1}yr^{-1}$ for N, P and K respectively according to the difference in rainfall.

Biological nitrogen fixation (IN 4) occurs in living fences, which are found on all fields in the Neguev, containing many leguminous trees like *Erythrina spp.* and *Gliricidia sepium*. The area under leguminous crops is negligible. In all the land use types a fixed contribution of 5 kg N $ha^{-1}yr^{-1}$ from the living fences and non-symbiotic N-fixers was assumed (Stoorvogel and Smaling 1991).

A separate inventory of the land units was made to assess the nutrient input by sedimentation (IN 5). All areas less than 5 meters above a major stream or 2 meters above a minor stream are supposed to have a nutrient input by sedimentation. The sediment input has been set at 500 kg $ha^{-1}yr^{-1}$. The concentrations of N, P and K in fresh sediment were respectively set at 0.1%, 0.15% and 0.8% (based on organic matter content and total P and K figures for 4 samples of fresh sediment in the region).

Harvest (**OUT 1**) estimates (Table 4.2) are based on interviews carried out in the study area (Brink 1988, Van de Berg and Droog 1992) and on literature (MAG 1991). For the pastures an estimate was based on the average cattle density (1.7 heads ha^{-1}, De Oñoro 1990). The major part of the nutrients removed by grazing return to the system as manure (IN 2). Nutrient concentrations (Table 4.3) in the harvested product are based on the review by Stoorvogel and Smaling (1991).

Crop residue removal (**OUT 2**) only occurs for cassava, red pepper and maize on a limited scale. Removal of wood from trees does not seem to form a substantial amount of nutrient removal and is therefore omitted in the present study. Nutrient concentrations in the crop residues (Table 4.3) are based on the review by Stoorvogel and Smaling (1991).

Especially for the fertile soils, leaching (**OUT 3**) may be an important export of nutrients. It is however a factor which is difficult to measure or to estimate. Leaching of N is determined as in the Kenyan case (Smaling *et al.* 1992) by calculating the amount of mineralized nitrogen as a function of the organic matter content. The fraction of mineralized nitrogen and fertilizer N which is leached is determined as a function of the clay content. The nitrogen mineralisation rate was set at 3%, which is relatively low due to the complexes which occur between organic compounds and short-range-order material in these volcanic soils. The fraction of mineralized soil N and total fertilizer N submitted to leaching was set at 40% for SFW, 30% for SIW_f, 25% for SIW_s and 20% for SFP. P-leaching was omitted since P is very immobile in the soil and most soils in the study area have a high P-fixation (70-100%) (De Bruin 1992). Leaching of soil K and fertilizer K has been determined as a function of exchangeable K and clay content as shown in Table 4.4. A reduction of the calculated leaching by 15% is performed for land units with a slope gradient of more than 6% due to a reduction of the infiltration by overland flow. Although the leaching sub-model is originally developed for Kenya, the results were in agreement with limited data for fertile soils presented by Rosales *et al.* (1992).

Denitrification (**OUT 4**) has been studied at La Selva experimental station (Keller *et al.* 1991; Keller pers. comm. 1992). Their results show that denitrification varies roughly between 4-10 kg ha^{-1} yr^{-1}, depending on drainage and fertilizer applications.

Table 4.3 Areas of the land utilization types and the nutrient concentrations in harvested product and crop residues

LUT	Area	Harvested product (kg/ton)			Crop residues (kg/ton)			C-factor
	(ha)	N	P	K	N	P	K	
Annuals								
Maize	66	16.8	9.4	5.7	9.7	4.4	25.7	0.5
Red pepper	23	3.3	0.7	3.4	7.1	4.8	4.3	0.5
Tubers	36	4.2	1.1	5.1	3.4	2.1	1.7	0.3
Perennials								
Coconut	10	61.0	16.5	11.8	-	-	-	0.1
Cacao	27	40.0	19.5	23.1	-	-	-	0.2
Plantain	30	0.7	0.1	3.8	1.2	0.3	7.1	0.2
Fruit trees	12	1.8	0.5	2.8	0.6	0.5	5.3	0.1
Intercropping								
Plantain/cocoa	14	--------------- see individual crops -----------------						0.2
Pasture	2336	15.0	2.3	15.0	-	-	-	0.01
Forest	737							
Swamps	250							
Village	80							

[Overtoom *et al.* 1987, Stoorvogel and Smaling 1991, ILACO 1981]

Table 4.4 The fraction of exchangeable K and fertilizer K submitted to leaching (in %) (based on Smaling *et al.* 1992)

Soil Type	Clay content	Fertilizer K	Exchangeable K
SFW	10	40	0.9
SFP	31	30	0.9
SIW_f	55	20	0.8
SIW_s	55	15	0.7

This corresponds with a yearly denitrification of 2% of all mineralized and fertilizer N for the land unit SFW and 3% for the land units SFP, SIW_f and SIW_s. For the calculation of the mineralized N again a yearly mineralization rate of soil organic matter of 3% has been used (See OUT 3).

Erosion (OUT 5) is not considered a serious problem and tolerable in the Atlantic Zone of Costa Rica with soil losses of less than 10 tons ha^{-1} yr^{-1} (Dercksen 1991). Nevertheless, Stoorvogel and Smaling (1991) indicated that even such low rates may contribute significantly to the nutrient balance. To assess the erosion in the area the universal soil loss equation (Wischmeier and Smith 1978) was used. The annual soil loss was estimated on the basis of the rainfall erosivity (R), the soil erodibility (K), the topography (LS), land cover (C) and management (P). The rainfall erosivity has been calculated for a number of weather stations around the study area by Vahrson (1991) and was on average 741. The soil erodibility factors are determined with the nomograph for estimation of soil erodibility (after Wischmeier *et al.* 1971) and given in Table 4.2. Slope gradient has been set at 1% for the flat areas and at 8% for land unit SIW_s with a slope length of 100 m resulting in values for LS of 0.02 and 0.15 respectively. Land cover factors are presented in Table 4.3. The management factor is set at 0.2 for SIW_f flat areas and at 0.4 for land unit SIW_s (according to FAO and Senacsa 1989 cited by Sanchez and Alvarez 1991). The nutrient concentrations in the eroded soil material are derived from organic matter content, total P and total K figures from the topsoil and making a correction with an enrichment factor for the sediment of 1.5 (Stocking 1984).

Forest covers 22% of the study area. Parker (1985) found for La Selva a negative nutrient balance for nitrogen and a nearly neutral one for phosphorus and potassium. The study by Parker did not include gaseous gains and losses of nitrogen resulting in an incomplete nitrogen balance. The nitrogen balance is therefore set at equilibrium like the balances for phosphorus and potassium. The extensive review on nutrient balances for forested areas by Bruijnzeel (1990) supports the use of a neutral nutrient balance for the forested area.

4.2.3 The linear programming model

LP models were originally developed for economic analysis (Hazell and Norton 1986). The models maximize or minimize a linear objective function subject to a set of constraints. Fu (1989) already indicated the possibility to use LP-models for the combination of two goals: (i) maximum agricultural production and (ii) minimal erosion. In the present study the distribution of land use was optimized to minimize nutrient depletion with the agricultural production for the area defined within the constraints. The variables in the constraint equations and in the objective function are the areas of the different land use systems. The 4 land units (swamps were excluded) under 9 different land utilization types result in a total of 36 land use systems and thus 36 variables. The first set of constraint equations is based on the total areas of the different land units and land utilization types. There are 9 constraints on the total production of the 9 land utilization types which are not allowed to decrease more than 5%. The objective function contains coefficients which are the quantities of nutrient depletion for each land use system. The simplex method (Hazell and Norton 1986) was used to calculate the solution of the model.

Two different scenarios were elaborated. The first one assumed that the land utilization type may be located on every site excluding the swamps and within the constraints on the total production of the land utilization types. The second scenario is based on the actual location of the forest areas, by excluding them from the relocation of the land utilization types over the area.

The link between NUTBAL and the LP-model

The maps of land units and land utilization types have been stored in PC Arc/Info. The nutrient depletion model and the LP-model were programmed in Turbo Pascal[4] with a direct link to Arc/Info through the simple macro language provided by PC Arc/Info. The integration of models and GIS results in an operational method.

4.2.4 The scenario results

The nutrient balance for 1984 land use is presented in Table 4.5 indicating a loss of 22 kg N $ha^{-1}yr^{-1}$, 5 kg P $ha^{-1}yr^{-1}$ and 13 kg K $ha^{-1}yr^{-1}$. Due to the large area of pasture a relatively important input of manure can be observed. However, part of these nutrients are subsequently removed by leaching and denitrification. Similar to the studies for Sub-Saharan Africa and Kenya (Stoorvogel *et al.* 1992, Smaling *et al.* 1992) erosion has a large impact on the nutrient balance.

[4] Turbo Pascal is a registered trademark of Borland International Inc. Scotts Valley, CA, USA.

Table 4.5 The nutrient balances for actual land use and optimized land use scenarios (in kg ha^{-1})

Nutrient	In-1	In-2	In-3	In-4	In-5	Out-1	Out-2	Out-3	Out-4	Out-5	Total
					Actual situation						
N	1.6	19.2	1.5	3.9	0.0	-22.6	-0.1	-7.1	-6.9	-11.8	-22.0
P	0.1	2.9	0.2	0.0	0.1	-3.8	0.0	0.0	0.0	-3.9	-4.5
K	0.1	19.2	4.9	0.0	0.1	-22.1	-0.1	-5.7	0.0	-9.6	-12.8
					First scenario						
N	1.6	17.5	1.5	3.9	0.0	-23.1	-0.1	-6.9	-6.9	-5.3	-17.5
P	0.1	2.6	0.2	0.0	0.1	-3.9	0.0	0.0	0.0	-1.2	-2.2
K	0.1	17.5	4.9	0.0	0.1	-22.8	-0.1	-5.6	0.0	-3.5	-9.0
					Second scenario						
N	1.6	18.4	1.5	3.9	0.0	-24.3	-0.1	-7.1	-6.9	-6.1	-18.8
P	0.1	2.7	0.2	0.0	0.1	-4.2	0.0	0.0	0.0	-1.5	-2.7
K	0.1	18.4	4.9	0.0	0.1	-23.8	-0.1	-5.8	0.0	-3.8	-9.6

The first scenario (redistributing all land utilization types) shows a decrease of nutrient depletion to a depletion of 18 kg N $ha^{-1}yr^{-1}$, 2 kg P $ha^{-1}yr^{-1}$ and 9 kg K $ha^{-1}yr^{-1}$ (Table 4.5). The annual and perennial crops are now concentrated on the fertile soils (SFW) along the river and the forest areas on the slopes (Figure 4.4). Therefore, most annual and perennial crops show an increase in production coinciding with this decrease in nutrient depletion mainly due to a decrease in the erosion.

The second scenario with fixed location of the forest area results in a different distribution of land use (Figure 4.5), but leads to an average nutrient balance which almost equals the first scenario with a depletion of 19 kg N $ha^{-1}yr^{-1}$, 3 kg P $ha^{-1}yr^{-1}$ and 10 kg K $ha^{-1}yr^{-1}$ (Table 4.5). Again all the annual and perennial crops are replaced from the sloping areas (SIW_s) to the fertile areas along the river. In contrast with the first scenario the sloping areas are now under pasture.

The production of all annual and perennial crops is on average a 4% higher in the scenarios compared to the actual situation. The production of the pasture is 2% lower. This originates from the general redistribution of all the annual crops towards the fertile soils and, coinciding, of all the pastures to the soils with the low soil fertility.

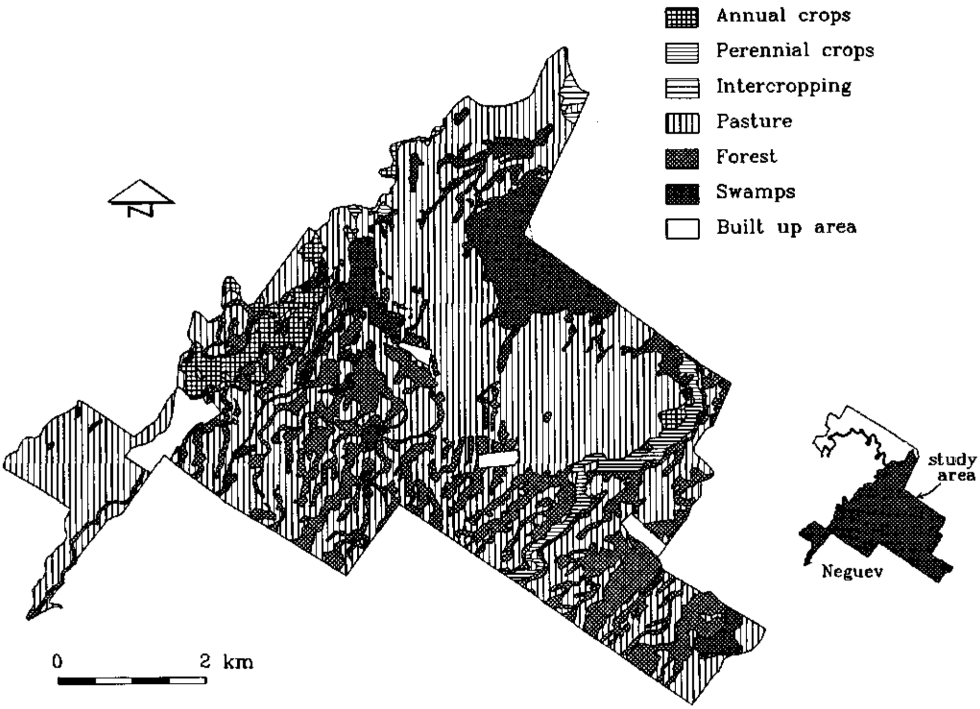


Figure 4.4 The first scenario: redistributing all the actual land utilization types

4.2.5 The possibilities of integrating nutrient depletion models in land use analysis

The estimated depletion rates are likely to result in a reduction of yields on a short term. Leaching (OUT 1) and erosion (OUT 5) are the most important factors determining the negative nutrient balance. The decrease of nutrient depletion may not only be established by changing the geographical distribution of the land utilization types over the different land units but also by adapting the management system. This study shows the decrease of nutrient depletion with a slight increase of agricultural production. The relocation of land utilization types results in an decrease of erosion by locating forest (in the first scenario) and pasture (in the second scenario) on the sloping areas. The lower depletion is mainly a result of the lower erosion rate. Additional measures reducing erosion will decrease the depletion even further. Leaching and

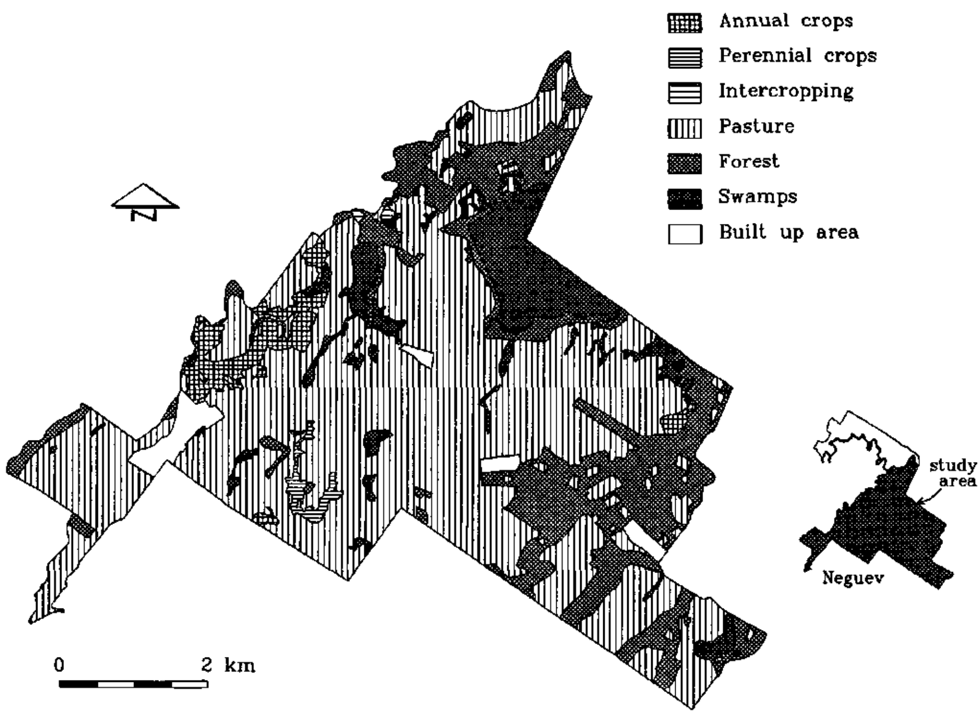


Figure 4.5 The second scenario: redistributing all actual land utilization types except the forest areas

denitrification on the other hand are almost not changed by the relocation of the land utilization types, but can be reduced by measures like split fertilizer applications.

Although it is clear that in reality land use can not be as easily manipulated as in a GIS, this case shows that techniques are available to incorporate nutrient depletion models in land use planning. The same technique can be used to include other types of models in planning as Fu (1989) showed in the case of an erosion model.

As shown by De Oñoro (1990) most farms do have annual crops. For successful land use planning land use should not be redistributed over a region but on farm level, being the level at which decisions are taken. This would provide a good option for the individual farmers.

For the integration of the different models into an operational method a GIS with good links to the different models is essential. The strength of the method is the linkage of existing techniques which will be necessary in a growing number of cases where one model does not provide the solution.

In this case study the distribution of the land utilization types has been optimized towards a minimum nutrient depletion. In practical situations it may be more realistic to maximize the production of the area under restrictions of nutrient depletion. The question now arises where one should put the threshold value for nutrient depletion. In general one can put the threshold as a percentage of the soil nutrient stock. In that case two aspects should be put forward: the replenishment of soil nutrients through weathering of soil material and ecological events like e.g. large scale inundations and volcanic eruptions, which are not taken into consideration in IN 5 or IN 3. The threshold value should not only been taken into consideration during land use planning but continue afterwards by monitoring of the soil nutrient balance (Smaling and Fresco 1993).

As indicated by Smaling and Fresco (1993) a good calibration of NUTBAL requires a serious investment. Nevertheless for a sustainable way of land use planning it will be necessary. Fertilizer trials and erosion measurement become, even in developing countries, more numerous and the combination of all existing information may already provide an acceptable level of detail in the model.

4.3 GIS for agricultural planning

4.3.1 Integration of computer-based models and tools

Agricultural science developed a large variety of computer-based models to be used for the characterization of agricultural systems, ranging from very simple expert systems to complex deterministic models (e.g. Bouma and Hoosbeek 1995). In addition computer-based tools, such as a GIS, are needed for data manipulation and georeferenced presentation. Each of these models and tools has its merits and drawbacks, with which users have to cope. In many cases, relatively complex and comprehensive models remain on the drawing board as data availability and field variability prohibit their practical application. However, integration of existing models and tools offers promising prospects for dealing with multifaceted and multi-disciplinary problems. By integrating models and tools the drawbacks of some can be compensated by the merits of others. In addition, extremely complex problems, which can not be handled by a single model, can be tackled by an integrated set of tools. Although integration *a-priori* may seem a logical development, problems are likely to arise during integration and operationalization. These problems originate in the incompatibility of databases, software and, sometimes, even hardware (e.g. Abel *et al.* 1994).

Typically, the analysis and planning of agricultural land use was based on techniques for land evaluation (FAO 1976) or farming systems analysis (Beets 1990). At the moment, methodologies are developed for the analysis of land use scenarios (e.g. WRR 1992; Van Keulen and Veeneklaas 1993). A land use scenario is defined as a set of hypothesized changes in the socio-economic and/or bio-physical environment. Land use analysis focuses on the possible effects of these changes on crop and technology choice (including the resulting consequences for the environment). Land use scenarios can be used for the analysis of "what-if" questions of agricultural policies and economic incentives (e.g. what will happen when the prices of fertilizers are lowered by subsidies). The scenarios are, typically, analyzed by an optimization model, which is, in most cases, a LP model. The optimization model maximizes an objective function by selecting LUSTs given a set of constraints which e.g. indicate resource availability. The different methodologies for the analysis of land use scenarios serve different objectives and work at different levels of detail (farm, region, nation). Therefore, they vary in the procedures used. Data requirements for the analysis of land use scenarios are similar to those of land evaluation and farming system analysis and mainly comprise input/output type data on agricultural activities, and data on resource availability (e.g. land, labor and capital). The activities, or LUSTs, are described by so-called input/output (I/O) coefficients defined in matrices. The available resources of land, labor and capital are inventoried in a specific way for each methodology, which mainly depends on the scale. Some methodologies include crop growth simulation and expert systems to define sets of crop-wise alternative LUSTs, others only include actual LUSTs which have been found in a region during farm surveys. Some of the drawbacks of LP models is that they can not describe potential LUSTs and are not georeferenced. To compensate for these drawbacks, the LP model can be combined with crop growth simulation models and GIS respectively.

The integration of different models and tools into the USTED methodology is mainly accomplished by establishing one common database in combination with a custom-designed software package. Ultimately, the land use scenarios are aimed at the evaluation of agricultural policies and economic incentives for a more sustainable agricultural production. The methodology has been developed for the Neguev settlement.

4.3.2 The USTED methodology

The general structure of the USTED methodology is presented in Fig. 4.6. For actual as well as alternative LUSTs, I/O matrices are described. The actual LUSTs are characterized mainly on the basis of farm survey data. Alternative LUSTs are identified with the help of crop growth simulation models and by expert knowledge. In the case of maize a calibrated simulation model (MACROS) was available (Jansen and Schipper 1995). For other crops, like cassava and palm heart, no calibrated crop growth simulation model is available and expert knowledge is used. Expert knowledge comprises data from experts but also data from farm surveys, where quantitative data was collected on different management practices and yields. Expert knowledge (including survey data) will probably remain an important data source as no comprehensive model is available to estimate outputs on the basis of crop management (including land preparation, fertilizer application and pest management).

Attribute data are based on surveys and literature and comprise e.g. data on prices and chemical compositions of inputs and outputs. On the basis of the I/O data for the different LUSTs, and attribute data, the coefficients for the LP model are calculated. In the USTED methodology, sustainability is included as a set of quantifiable, operational parameters. In the case of the Neguev settlement, biocide use and soil nutrient depletion are identified as the principal constraints on sustainability (Jansen *et al.* 1995). Biocide use is expressed as an index value on the basis of the concentration of active ingredients, their toxicity, and half life time, and is calculated for each operation and summed for the LUST:

$$BI = \sum_{Oper_1}^{Oper_n} (AI * TI * \sqrt{HLT})$$

Equation 4.1

In which:	BI	=	biocide index
	$Oper_i$	=	operation number i
	AI	=	amount of active ingredient (kg/ha)
	TI	=	toxicity index (based on WHO classification)
	HLT	=	half life time (days)

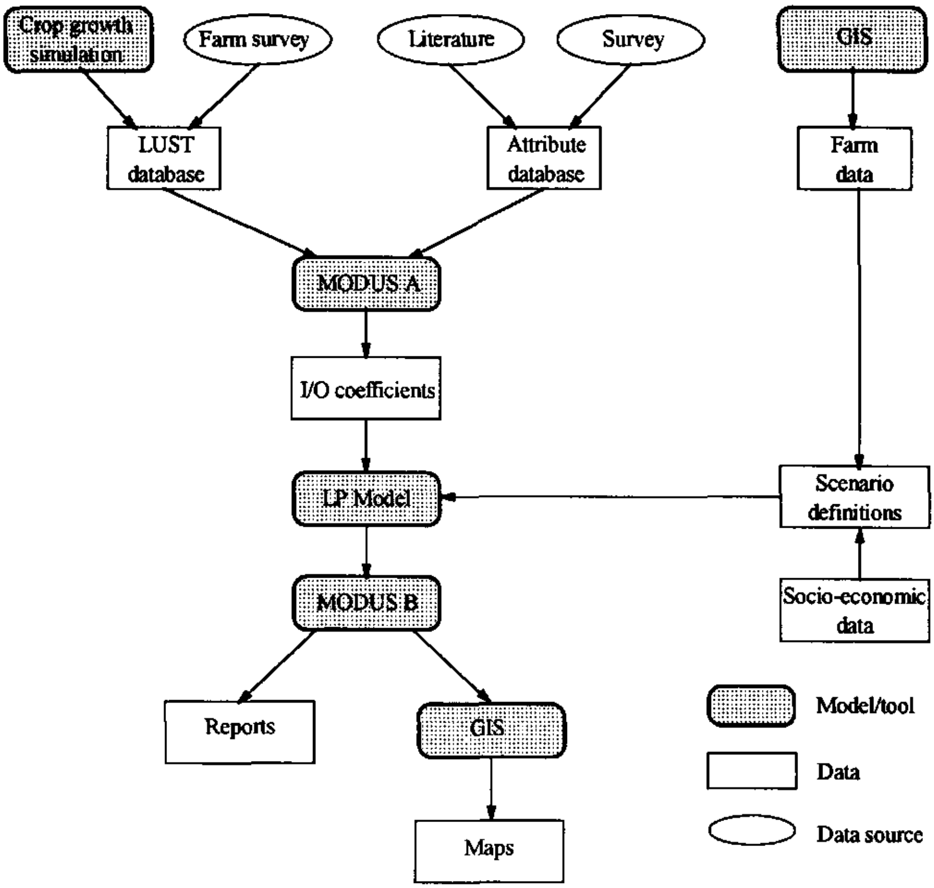


Figure 4.6 The structure of the USTED methodology

The nutrient balance is estimated on the basis of separate assessments of 4 nutrient inputs (mineral and organic fertilizers, wet deposition and N-fixation) and 4 nutrient outputs (product, stover, gaseouses losses and leaching) by an adapted version of the NUTBAL model (Section 4.2). Sedimentation and erosion are excluded from the NUTBAL model because the LUST descriptions do not include slope and sedimentation data. The scenario definition, i.e. the changes in the socio-economic and/or bio-physical environment, are translated into changes in attribute data, technologies (i.e. LUSTs), variables for the calculation of the I/O coefficients (e.g. the discount rate), and resource constraints. For specific scenarios, their definition may imply additional constraints or changes in the objective function for the LP model (e.g. constraint on biocide use). Although the methodology focuses on the regional level, the individual farms are included as the final decision makers. To deal with the relatively large number of farms in the settlement, a restricted number of farm types are defined as farms of a similar size and with a similar soil distribution. The geographical data for the availability of soil resources are generated by the GIS (Figure 1.6). LUSTs are combinations of soil types, crops and technologies. This implies that for each soil type, a set of LUSTs will be generated. The original soil map distinguishes 21 different soil series, although part of these soil series are classified on the basis of soil genesis and are similar in their management. To keep the number of LUSTs limited, three main soil groups are identified on the basis of drainage and inherent soil fertility.

The LP model maximizes the sum of the net farm incomes of the different farm types subject to resource constraints regarding soil, labor, and capital availability for the individual farm types, as well as constraints regarding labor availability within the settlement (hired labor must equal off-farm labor) and employment possibilities outside the settlement (Schipper *et al.* 1995). Net farm income equals the value of the production, plus income from labor on (banana) plantations outside the settlement, minus total production costs (including hired labor), and, therefore can be interpreted as the net return to land and family labor. The results of the LP model are visualized by the GIS.

4.3.3 Systems integration in USTED

The approach towards systems integration depends on its objectives. Procedures at different levels of detail can be integrated to reduce data requirements (Bouma *et al.* 1993). At one level of detail, procedures can be integrated to deal with multi-disciplinary problems, in which case different disciplinary models are integrated. Operationalization of systems integration can be achieved with appropriate database structures. The database structures can be jointly used by the different systems and form the link between them. In the USTED methodology, agronomic, edaphologic, and economic models are integrated together with database managements systems (including a GIS).

Typically, systems integration has to tackle problems related to incompatibilities between different applications. These incompatibilities may originate from differences in database requirements and structures, softwares or hardwares. The USTED methodology integrates computer applications by a common database and a special software package which adapts the various databases to the specific requirements of the different applications.

A common database

To avoid incompatibilities between the data used in the different procedures, one common database should be generated, from which the different procedures can derive the basic data required. For the USTED methodology the database is split in four groups. Geographical data are stored in the GIS. A second database comprises data on LUSTs, which is stored as an operation sequence and a reference to a certain soil type. The third database comprises the characteristics (e.g. prices and chemical characteristics) of the different input and outputs (attribute data). The last database is linked directly to the LP model and includes the availability of the non-geographical resources (labor and capital). Standard identifiers are included for the different inputs and outputs to allow referencing between the various data sets.

MODUS

A special software package denominated MODUS (MOdules for Data management in USted) is developed to derive for each module the appropriate input parameters and to read the output files from the different modules. MODUS reads the characteristics of the different farms from the GIS and translates them to a constraint file for the LP model. The data from the LUST files are combined with the attribute files to calculate the I/O coefficients for the LP model. For the calculation of the sustainability parameters, separate models are invoked. The scenario definition has to be translated by the user into changes in attribute files (e.g. prices), variables in MODUS (e.g. the discount rate) or into separate constraints which are added to the LP matrix (e.g. limitations on biocide use).

After the optimisation with the LP model, MODUS reads the results and generates reports and files for the GIS. The GIS enables a rapid interpretation by visualizing the results.

It is important to distinguish between a modular methodology, which is the result of the linkage of applications based on primarily data exchange, and a new overall procedure. Full integration of databases and software will lead to the development of a new, more complex procedure. Although integration is necessary for the development of a rapid interactive procedure, such an overall procedure will reduce the comprehension and overview for the users. It is therefore, likely to function as a black box. Different models and tools which are necessary for many multifaceted problems often originate from specific disciplinary problems. It is, therefore, relatively easy to develop one multi-disciplinary database on the basis of which each discipline will develop their own procedures. Such a multi-disciplinary database can therefore be used in an integrated procedure which, compared to a new overall procedure, is relatively transparent and user-friendly.

4.3.4 Alternative land use scenarios for the Neguev settlement using USTED

An infinitely large number of scenarios can be evaluated using the USTED methodology. In this paper, three different sets of scenarios will be presented, respectively dealing with biocide use, nutrient balances, and credit availability. The basic settings for the LP model are presented in Table 4.6. A total of 122 LUSTs is available for each of the farm types, comprising forest management systems for the forest areas, and tree plantation, palm heart, cassava, maize, and pineapple for the non-forest areas. The LP model maximizes the returns to land and family labor with no restrictions on the sustainability parameters. The LP model selects maize on the fertile, well drained soils, palm heart with some cassava on the infertile, well drained soils, and cassava on the poorly drained soils (Table 4.7, Figure 4.7 US$ 75 / ha).

Biocide use

Alternative scenarios involving certain specific policy measures may be preceded by an explorative study regarding the technical possibilities. In the case of biocides, one might study the negative effects on net farm income of a physical reduction in their application. Three alternative scenarios with respectively 75%, 50% and 25% of the maximum biocide use at farm level of the base scenario are carried out (Table 4.7). For each of the levels, net farm income is maximized and a new set of LUSTs is selected for the different farm types. Although palm heart remains one of the major land uses, changes in the cultivation of cassava and pastures will take place. Striking is the low decrease in net income. This means that even with an on-average reduction of biocide use (expressed by the biocide index) with 75%, net income only decreases 3%. Given the restrictions of the methodology and compared with the base scenario, this would indicate that farmers can decrease the use of biocides, without large losses in their net income. They, however, should change the technology with which they cultivate their crops.

Table 4.6 Constraints for the Neguev settlement

	Farm types					Region
	FT-1	FT-2	FT-3	FT-4	FT-5	
Number of farms	33	4	46	35	189	307
Fertile, well drained soil (ha/farm)	1.4	3.2	4.9	11.6	0.4	
Infertile, well drained soil (ha/farm)	3.5	21.6	4.3	0.6	9.4	
Fertile, poorly drained soil (ha/farm)	6.8	3.6	0.9	0.5	0.7	
Forest (ha/farm)	3.9	3.7	3.5	1.3	2.7	
Family labor units (1,485 hours/yr)	2	2	2	2	2	
Capital (US$/yr)	1,178	2,408	1,013	1,058	982.5	
Plantation employment[1] (days/yr)						5813

[1] Constraint per month, slight deviations during the year are included

Table 4.7 The effect of restrictions on biocide use (areas of LUSTs in ha for the total settlement)

Land utilization type	Soil type	Scenario			
		Base	Biocide constraint as % of use in base		
			75%	50%	25%
Forest		857	857	857	857
Tree plantation	SFW	404	340	273	304
Palm heart	SFW	329	389	458	428
	SIW	1774	1709	1671	1708
Cassava	SFW	47	47	47	47
	SIW	96	366	289	165
	SFP	413	91	39	14
Maize	SFW	5	5	0	0
Plantation labor (days per farm per year)		144	165	173	185
Net income (% of base)		100%	99%	98%	97%

Nutrient depletion

Nutrient depletion will, at least in the long run, have negative effects on soil resources and can therefore be regarded as a negative contribution to farm income (Solorzano *et al.* 1991). Even though simple models are available to estimate soil nutrient balances (Stoorvogel 1993[B]), it is still difficult to estimate the direct financial impact of nutrient depletion (partially as it depends on future land use). More difficult yet is the estimation of the indirect impact on soil and ground water quality. Most analyzes aimed at estimating such financial impacts are, therefore, based on arbitrary nutrient depletion penalties, often based on the cost to replace nutrient losses. Minimum values can be based on actual prices (US$ 0.45 / kg N, US$ 1.20 / P, and US$ 0.35 / kg K for 1992), and approximations for fertilizer efficiencies (40% for N, 95% for P and 60% for K). In this context, fertilizer efficiency is defined as the percentage which of the fertilizer application which positively contributes to the soil nutrient stock. The replacement costs for the Neguev settlement are estimated at US$ 1.36 / kg N, US$ 1.55 / P, and US$ 0.74 / kg K. With the minimum replacement costs for nutrient depletion the selected LUSTs only slightly vary with the base scenario (with a 4% reduction in farm income). However, if the costs to replace the lost nutrients are higher, larger changes in the selected LUSTs and net income take place (Table 4.8). In addition, large changes in the average nutrient balance take place, with, at 8 times the minimum prices, nutrient balances which are almost in equilibrium. The changes in nutrient balances partly originate in the selected crops, but also in the alternative technologies which are chosen.

Capital

One of the main farm resources, and often the target of policy interventions, is capital. Although large differences occur between farms, the average capital availability is estimated at US$ 75 per ha per year, yielding the total capital availabilities per farm as presented in Table 4.6. The effect on land use of lower (US$ 25 and US$ 50) and higher (US$ 100 and US$ 150) capital availability is depicted in Table 4.9 and Figure 4.7. Even though changes in the selected LUSTs do take place, increases in capital availability do not strongly influence net farm income. Different LUSTs are provided to the optimization model which vary in the investment costs but vary relatively little in the net return. Therefore the optimization model is not able to reach increase net income, with more capital available. This indicates the importance to describe alternative LUSTs.

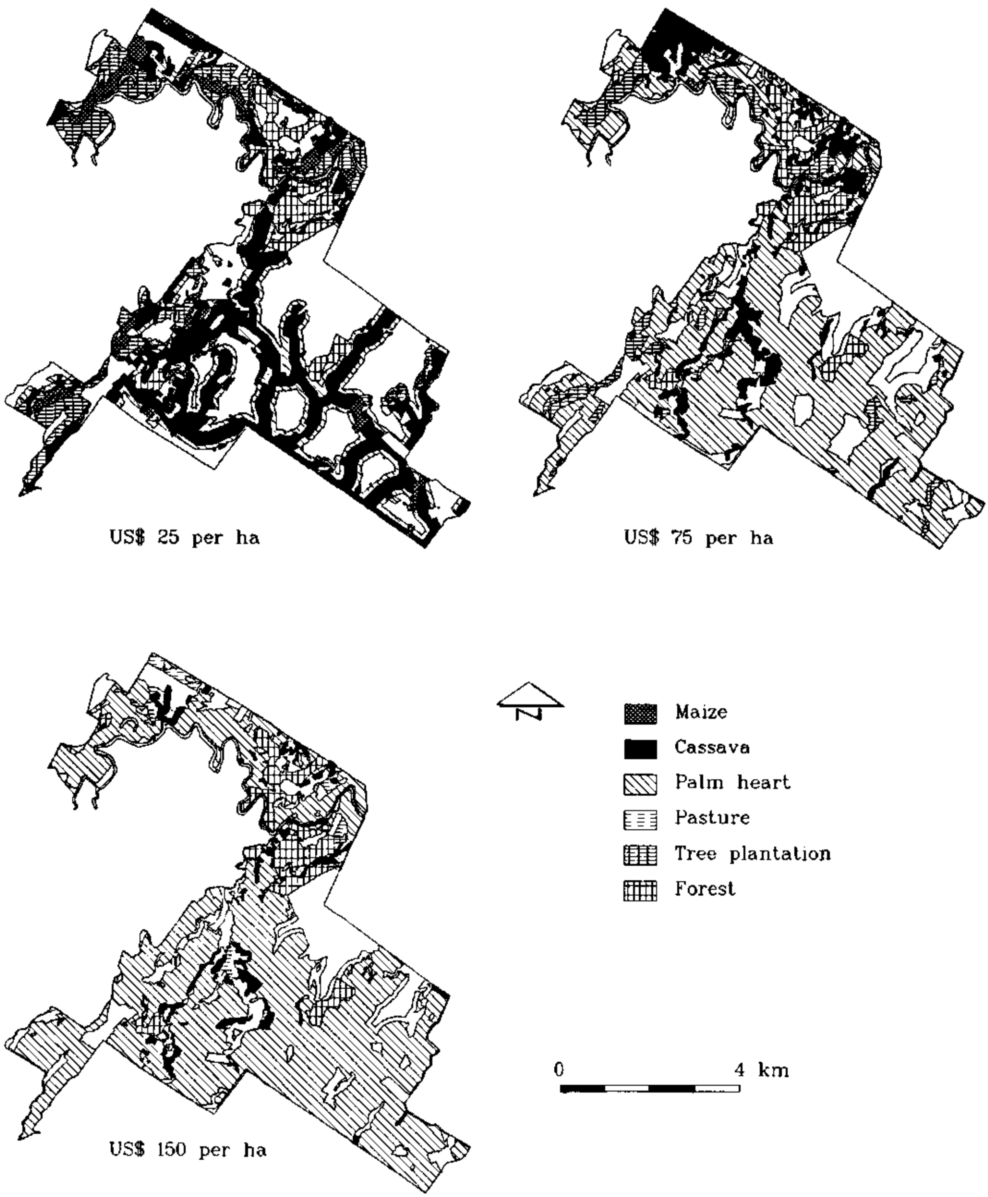


Figure 4.7 Land use distribution for different levels of capital availability

Table 4.8 The effect of incorporating nutrient depletion in net income on land use in the Neguev (areas of LUSTs in ha for the total settlement)

Land utilization type	Soil type	Scenarios			
		Number of times of actual prices for the replacement of nutrients			
		1	2	4	8
Forest		857	857	857	857
Tree plantation	SFW	404	404	717	449
Palm heart	SFW	329	329	0	328
	SIW	1775	1754	1698	271
Cassava	SFW	47	47	86	0
	SIW	246	368	98	288
	SFP	242	118	70	0
Maize	SFW	5	5	0	0
Plantation labor (days)		156	165	215	208
Average N balance (kg/ha,yr)		-15.4	-15.3	-10.9	-0.7
Average P balance (kg/ha,yr)		-0.2	-0.1	+0.9	-4.8
Average K balance (kg/ha,yr)		-11.2	-11.0	-8.1	-3.9
Net income		100%	96%	87%	85%

4.3.5 The linkage of models and tools in USTED

An integration of different models and tools as modules in a methodology requires an explicit description of the different assumptions of the individual modules to avoid that the overall methodology functions as a black box. This offers a unique possibility to integrate socio-economic factors with agro-ecological factors. Each of the modules within the USTED methodology has a clear role: the LP model for the optimization, crop growth simulation models and expert system for the description of LUSTs, a nutrient balance model for the quantification of the corresponding sustainability parameter, and the GIS for a rapid and quick interpretation of the results, as well as for data storage. The models and tools can be developed and calibrated independently,

Table 4.9 The effect of capital availability on the selection of LUSTs (areas of LUSTs in ha for the settlement)

Land utilization type	Soil type	Capital availability (US$ / ha) scenarios				
		25	50	75	100	150
Forest		857	857	857	857	857
Tree plantation	SFW	481	640	404	229	24
Pasture	SFP	0	0	0	0	108
Palm heart	SIW	476	1215	1774	2140	1971
	SFW	0	126	329	358	716
Cassava	SIW	1087	513	96	57	227
	SFP	263	411	413	250	51
	SFW	0	0	47	187	22
Maize	SFW	308	33	5	0	0
Plantation labor (days)		187	180	144	113	128
Net income		77%	89%	100%	110%	113%

and, therefore, a modular approach is proposed to systems integration. This will enable a clear and comprehensive development of the integrated methodology, avoiding an integration which results in one highly complex tool that is much more difficult to use. The operationalisation of the systems integration is thus mainly focused on data interchange between individual models and tools. In USTED, a separate system (MODUS) has been developed, to operationalise the overall integration. Typically, a modular approach deals very well with multi-disciplinary problems. Nevertheless, the development of one overall database is extremely important.

In many cases, GIS enables a quick presentation of model results and, therefore, in a rapid evaluation. The visualization of results may facilitate the communication between disciplines and result in an interactive procedure.

4.4 Considerations for the linkage of GIS and models

1. Integration of models and tools is able to deal with multifaceted (e.g. multidisciplinary) problems such as the analysis of land use scenarios, or to deal with sustainability. On the basis of a combination of a common database and an integrating software package, systems integration can be performed successfully.
2. The integration of applications is necessary for the development of an operational methodology, which can be used in land use planning. For the evaluation of agricultural policies and economic incentives, an interactive analysis of land use scenarios is necessary. The integration with a GIS enables a quick interpretation of the results by the user, as well as a quick definition of alternative scenarios.
3. A modular approach to systems integration is proposed to keep the middle between full integration and a loose set of procedures which is not operational. However, this approach implies the development of a procedure which performs this systems integration, combining the different modules.

5. Data requirements

Chapter 5 is based on the following publications:

- Stoorvogel, J.J., 1994. Data needs for sustainable land use scenarios for humid tropical Costa Rica. Invited paper at the XV[th] ISSS conference in Acapulco, Mexico.
- Stoorvogel, J.J., R.A. Schipper, and D.M. Jansen, 1995. USTED, a methodology for land use planning including the sustainability. *Netherlands Journal of Agricultural Science* 43: 5-18.

5.1 Land use analysis

Actual land use is the result of a long history of regional development. Driving forces influencing land use may be internal, like the available natural resources, or external, like prices on the world market. The development of land-use can be directed by incentives and regulations. These incentives and regulations can vary widely (Lutz and Daily 1991) but most of them have the same goal, to influence farmers' decisions concerning land allocation and management. As planners are interested in the effects of incentives and regulations on regional land use, they may define scenarios to study their possible effects. Although several techniques such as crop growth simulation models, GIS and LP models are available to support the analysis of scenarios, it is still very often the amount and quality of available data that limits the value of prognoses of these scenarios.

The amount and type of data needed for land use analysis depends mainly on the scale at which the study is carried out. But even at one specific scale level, the complexity of the problem and the required level of detail of the results may correspond with different data requirements (Figure 5.1). In this chapter the soil data requirements are studied. In studies where the complexity of the problem as well as the required level of detail of results are low, available soil survey data may fulfill the data need for the analysis. With increasing complexity and level of detail, additional data are necessary. These additional data may be based on e.g. a simple batch experiment. Complex problems and problems where a high level of detail is required need field trials, such as fertilizer experiments, for the calibration of models or as independent data sets. At a certain level of detail, problems become too complex or the level of required detail is so high that the data requirements become out of range. At that point, one has to change to a higher aggregation level or change the level of detail to answer the problem appropriately. The complexity of the problem is often inherent to the analysis. The required level of detail is often determined by the user, who defines the problem.

5.2 Four practical cases

Four different case studies are elaborated: (i) the possibilities for maize cultivation in the Atlantic Zone to be evaluated by a crop growth simulation model and a nutrient depletion model, (ii) the risk for ground and surface water pollution in the Atlantic Zone with a commonly used nematicide Ethoprop, (iii) the identification of

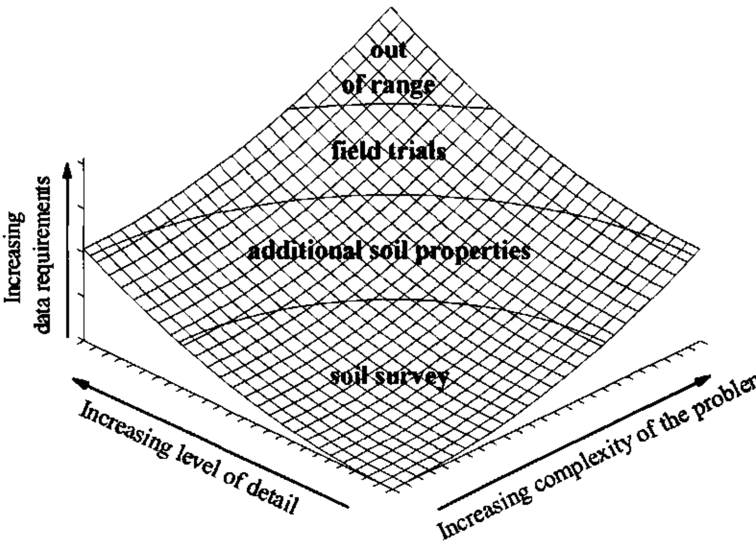


Figure 5.1 The amount of data required for a study on the basis of the complexity of the problem and the level of required detail

sustainability related problems of actual agricultural land use in the Atlantic Zone, and iv) alternative land use scenarios for the Neguev settlement.

The four case studies originate from different user groups. A policy maker might be working on import and export regulations for maize, and is interested in the regional production possibilities. The conservationist is worried about the contaminations of ground and surface waters with Ethoprop, causing the death of a large number of fish. Local government officials are interested in the sustainability of the agricultural sector. The Ministry of Planning wants to analyze the regional effects of incentives. The different scenarios all need soil data to express different land qualities but vary in quantity and level of detail, which is the object of this chapter.

5.2.1 Possibilities for maize cultivation

To study the possibilities for maize cultivation, a two step approach can be followed (Figure 5.2). In the first step, a qualitative land evaluation is used to screen for potential areas. In the second step, a more quantitative approach is followed to estimate yield levels. A rapid assessment of soil suitability can be obtained by a qualitative land evaluation. The accuracy of such a procedure is in most cases unsatisfactory to advise a farmer, but may be sufficient to decide whether and where detailed studies are worthwhile. This first step can be taken on the basis of a soil survey. With additional data, like specific fertility analysis, a more quantitative approach can be followed, which may include production assessments on the basis of expert knowledge. Reliable quantitative assessments of production and its sustainability can only be reached after field experiments and model calibration.

The qualitative evaluation to screen for potential production areas is based on altitude, land cover and land use and excludes high altitude areas and areas with natural vegetation or plantations. In the potential areas the nutrient limited production is estimated with QUEFTS (Quantitative Evaluation of the Fertility of Tropical Soils; Janssen *et al.* 1989). QUEFTS can estimate the nutrient limited production of maize for different land use systems and is calibrated on the basis of field trials for the local conditions (Guiking *et al.* 1994). Twenty one relatively fertile, well drained soil types are found in the potential areas. The topsoils in these soils were grouped into five functional A-horizons. From the soil database the input parameters for QUEFTS (pH-H_2O, C-Kurmies, exchangeable K and P-Olsen) for the functional A-horizons were determined. Production possibilities for three levels of fertilization were analyzed. Figure 5.3 illustrates the results for non-fertilized maize. QUEFTS evaluates the soils on their present nutritional status and does not include an analysis of the sustainability of the production. Models like NUTDEP (see Section 4.2, Stoorvogel 1993[B]) enable an additional evaluation of the nutrient balance of land use systems. Table 5.1 shows the estimated average nutrient balance for two possible levels of fertilization. Not only does fertilizer use reduce to a certain extent the loss of nutrients, it also increases the production. However, for an equilibrium of the nutrient balance an integrated management of the soil nutrient stock is necessary.

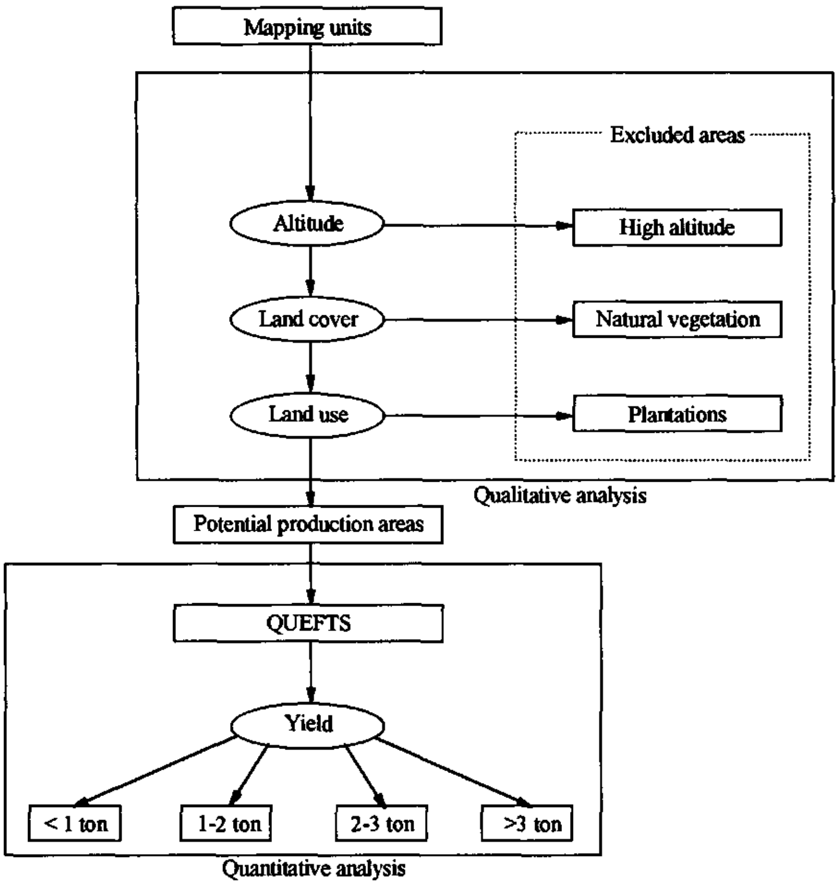


Figure 5.2 A procedure to evaluate the possibilities for maize cultivation in a region

The two levels of detail at which the problem is analyzed include a general inventory and a quantitative assessment of productivity, including a possible evaluation of the sustainability of the production on the basis of the nutrient balance. The data requirement for both levels of detail can be seen in Figure 5.4. The qualitative study (position 1) is based on the soil survey with accompanying data, and the quantitative study (position 2) on the basis of field trials, which were used for the calibration of QUEFTS. The complexity of the problem as treated by QUEFTS is relatively low.

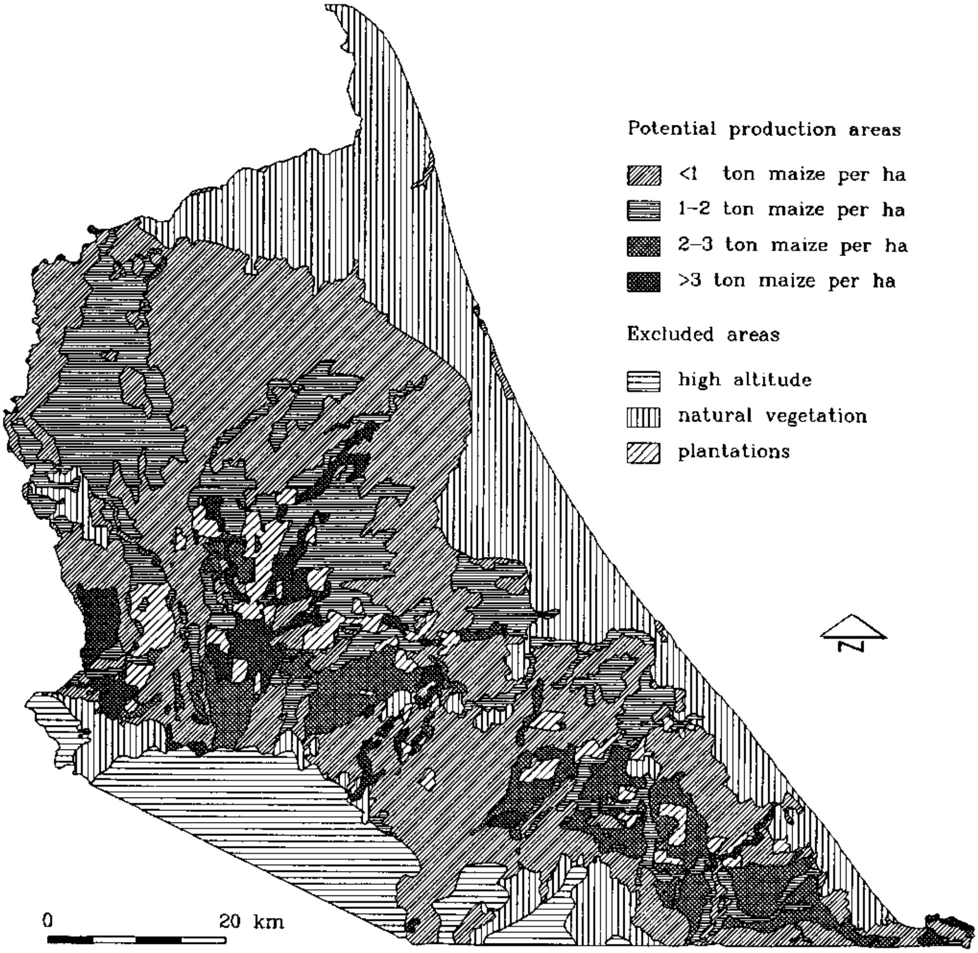


Figure 5.3 The productivity of non-fertilized maize in the Atlantic Zone of Costa Rica (average per mapping unit)

If the potential Ethoprop fixation (E_{fix}) exceeds twenty times a normal Ethoprop application (E_{app}) of 10 kg Ethoprop per ha, the soil is not considered to be prone to Ethoprop leaching. The factor twenty is chosen arbitrarily on the basis of two considerations. When Ethoprop is applied to the main crops in the Atlantic Zone the application normally takes place close to banana and palm heart plants and is thus not equally spread over the soil. Secondly it is known that water transport in most soils does not take place uniformly but along preferential patterns of water flow (Bouma 1991). This indicates that fixation of Ethoprop may occur around these preferential patterns of water flow and not in the complete soil matrix. For this exploratory study the factor is set at twenty to be sure that only soils with no Ethoprop leaching are selected. For soils where $E_{app} \geq E_{fix}$ Ethoprop leaching is almost certain to occur and the soils are classified to be extremely susceptible. In soils where $E_{fix} > E_{app} > 20\ E_{fix}$, the present study considers the risk of Ethoprop leaching to be intermediate. Figure 2.10 indicates the relative hazard of Ethoprop leaching in the Atlantic Zone of Costa Rica as based on the above procedure.

The contamination problem is a complex one, especially in a humid tropical environment where few data on biocide behaviour are available. In the present case study only a general overview of problem areas is required. The amount of necessary data is limited to data available from existing soil survey and some additional data i.e. Ethoprop fixation estimates for the major soil horizons or functional horizons (Position 3, Figure 5.4).

If the level of required detail is higher and, for example, accurate assessments of critical Ethoprop applications are required, quantitative simulation and further measurements are inevitable. The problem then moves to position 4 in Figure 5.4, where field trials are necessary to calibrate quantitative simulation on Ethoprop behaviour in the soil.

The integration of qualitative and quantitative procedures are increasingly propagated (Van Diepen *et al.* 1991, Reinds and Van Lanen 1992). If in the case of Ethoprop leaching a high level of detail is required, the problem may be similarly structured. Firstly, on the basis of existing data and few additional measurements an inventory of problem areas is made. Secondly, in problem areas where the qualitative procedure does not yield a clear answer on the extent of Ethoprop leaching, additional data collection and quantitative modelling can take place.

5.2.3 Sustainability indicators for actual land use

Agricultural land use analysis can, typically, only deal with a limited number of sustainability indicators (Jansen *et al.* 1995). It is, therefore, extremely important that the indicators are selected carefully. The sustainability of the agricultural sector of the Atlantic Zone may be assessed on the basis of three indicators which can be evaluated in sequence to obtain a general impression of sustainability problems at regional level: (i) nutrient depletion, (ii) degradation of soil physical characteristics due to compaction and, finally, (iii) contamination of ground water and surface waters with biocides. The basis for the analysis is the inventory of land use in land use zones (Huising 1993). Studying sustainability on a regional level will be restricted to a general evaluation of the different land use zones in terms of biocide leaching, the nutrient balance and a degradation of soil physical properties. The six generalized land use zones for the study area (Figure 1.4) will be discussed in terms of their sustainability.

- **Natural vegetation** is normally considered to be the most sustainable land cover. Nevertheless, it may be influenced by land use around it. An example would be the contamination of ground and surface water draining in the direction of the ecosystem. As a result the biodiversity of the natural vegetation may be threatened.
- A **colonization area** is defined as an area where both natural vegetation and agriculture are found and where the latter is gaining in importance. Agricultural land use in these areas is normally extensive. However, some nutrient depletion, biocide leaching and compaction may occur on the agricultural areas.
- Although in general grazing pressure in **pastures** of the Atlantic Zone is low, compaction, resulting in a decrease of the infiltration capacity, is found in almost all pastures (Spaans *et al.* 1989).
- **Mixed agricultural use** is found in the Neguev settlement for which the nutrient balance was calculated (Stoorvogel 1993[B]). The net annual loss per ha was estimated at 22 kg N, 5 kg P and 13 kg K. It is likely that other areas with comparable land use have similar depletion rates. Additionally, compaction will occur in the areas under pasture.
- **Annual crops** had, in general, a higher loss of nutrients than pastures. Although nutrient inputs and the technology level is higher in this land use zone, it is likely that mining of the soil nutrient stock takes place.
- **Plantations** in the area have very high levels of fertilization, compensating completely for nutrient losses (Stoorvogel 1993[A,B]). However, they also use large quantities of biocides, associated with a high risk of contamination.

Figure 5.5 illustrates the results. The complexity of this problem is moderately high due to the wide variety of aspects which are included. The inventory as presented above has a very low level of detail and general knowledge on the basis of a soil and land use inventory will be sufficient (Position 5, Figure 5.4) To come to relatively accurate assessments for nutrient depletion, compaction and biocide leaching field trials and case studies will be necessary.

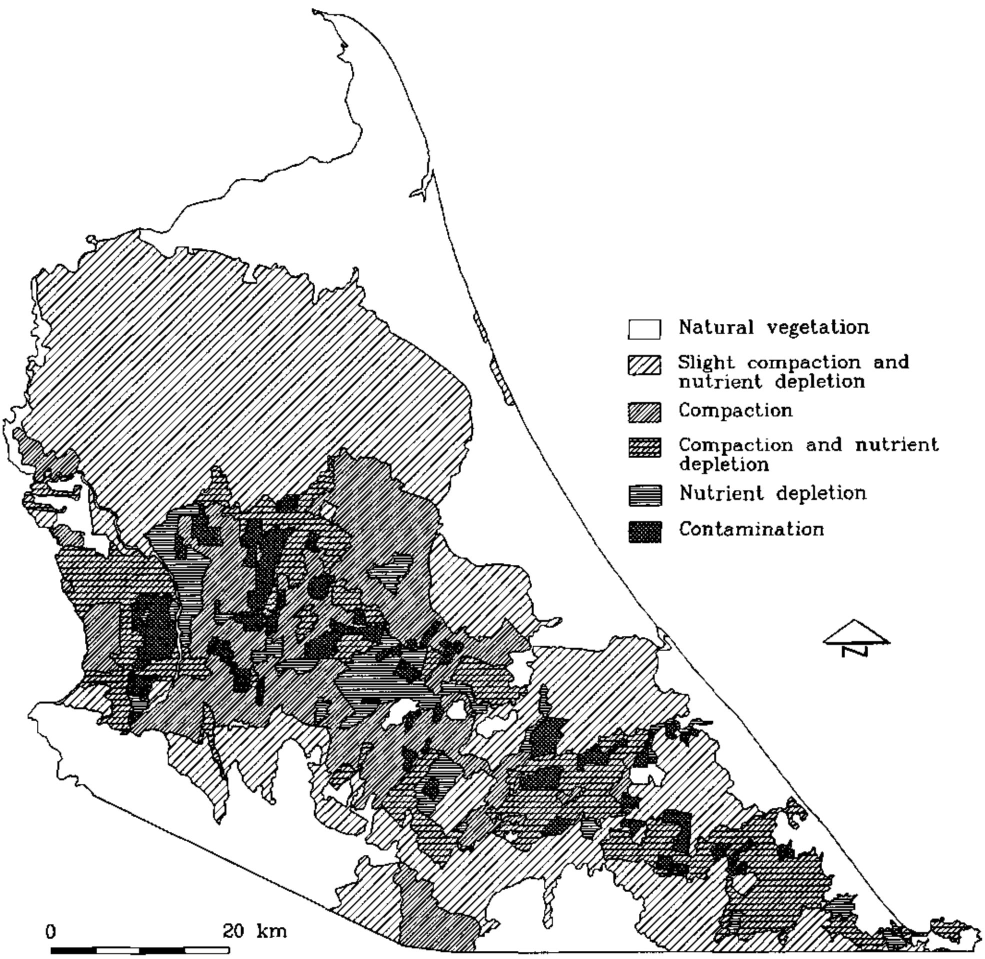


Figure 5.5 Principal sustainability indicators for 1984 land use in the Atlantic Zone of Costa Rica

5.2.4 The analysis of alternative land use scenarios

The recent literature on land evaluation includes several methodologies to evaluate alternative land use scenarios on the basis of a LP model and a GIS. In the case of Veeneklaas (1990) or Despotakis (1991), LP models are used for the analysis of regional agro-technical possibilities, whereas for WRR (1992) similar objectives played a role at the supra-national level. In both studies the farming systems were not taken into account and the analysis is focused on agro-technical possibilities. The USTED methodology presented in Section 4.3 is focused on the analysis of the effects of incentives and regulations including their effect on the natural resources.

USTED provides possibilities to evaluate land use scenarios at a sub-regional level, where the region is not treated as one farm, but differences between farms are included. However, it should be realized that only one overall goal function is maximized, namely the total net income for the sub-region, which is the sum of the net incomes of the individual farm types. Lower net incomes for one farm type may occur when the farm income of others increases. The farm level is included in the regional level with a number of regional constraints which include total labor availability and possibilities for off-farm work.

Objectives of the methodology are similar to those of many land use planners. Limitations, however, occur due to the vast data collection efforts which are necessary and the expert knowledge required to interpret the results. The methodology does not give one simple answer to its user, but rather provides for the possibility to evaluate scenarios of which the results have to be analyzed. The assumptions and limitations of the methodology should always be kept in mind during interpretation of the results. For example, in the case of extremely high productions of one product, its price can be expected to decrease. However, this would be in contrast with the assumption that prices are fixed. With increasing importance of a (sub)region in the market-share of agricultural products, it becomes more important to evaluate the effects of this assumption.

Although the USTED methodology was developed by a multi-disciplinary team of researchers without a demand from potential users, it is a potentially useful instrument for any organization dealing with land use planning. Even though USTED in itself is not a planning tool, it supports agricultural planning. Possible users within the Costa Rican context include the following:

(1) In the Atlantic Zone of Costa Rica, the Costa Rican Institute for Agricultural Development (IDA) manages and distributes land in agricultural settlements, including the Neguev settlement, which has figured as the case study for the USTED

methodology. Critics of IDA argue that, as most of the agricultural settlements are located on infertile soils with inadequate farm sizes, farming has a low potential. This is supposedly one of the main reasons why farmers sell their land or start working off-farm. IDA provides support to the farmers by training, extension, credit and legal advice (De Vries, 1992). The USTED methodology may help to determine more specifically how the extension and credit can be directed. In addition it may indicate the productive potential of a certain settlement.

(2) Disciplinary studies may yield data on alternative management practices to increase the sustainability of agricultural production. Such studies, however, typically lack data to evaluate the practices in a regional context and as a consequence, are unable to determine the potential for adoption of alternative practices. In the USTED methodology, alternative practices may be translated in alternative LUSTs and be evaluated for different farm types.

(3) Lutz and Daily (1991) identified a large number of possible incentives and regulations which might affect land use in Costa Rica. The effect on agricultural land use of these policy incentives may be evaluated with the USTED methodology. Schipper *et al.* (1995), for example, analyze the effect of increasing prices of biocides on their use in the Neguev settlement. To quantify and evaluate the effect of this kind of measures, USTED may be a useful tool.

Like any analysis of the agricultural sector, USTED requires a considerable amount of data at farm, sub-regional and regional level. Local institutes may not have the necessary resources to collect the data needed. Existing data, however, may provide at least part of the necessary database. Minimum data needs for the USTED methodology are:

- A general purpose soil survey. At the farm level most farmers have insight in the general geographic distribution of the main soil types. At the sub-regional level the required semi-detailed soil map may be available, for example in cases of agricultural settlements, and irrigation schemes.
- Numbers, sizes and soil types of farms.
- Quantitative descriptions of the principal LUSTs in the area and the corresponding attribute database.
- Insight in the farming systems, for example data on labor availability.
- Data on the region for detailed descriptions of the constraints and the alternative land use scenarios.
- Insight in a number of sustainability parameters to enable a quantification of these parameters for each of the LUSTs.

The flows of (disciplinary) data into common inter disciplinary databases is regulated by MODUS (see Section 4.3), thereby integrating different disciplinary models in the methodology.

For a full analysis of alternative land use scenarios, a large multi-disciplinary database is required, including the descriptions of actual and alternative land use systems and insight in the processes governing the sustainability indicators. Field experiments will therefore be necessary and the data requirements can be found in position 6 in Figure 5.4.

5.3 Data needs for land use analysis

Data needs for land use scenarios vary as a consequence of the complexity of the problem and the level of detail required for the results. The latter mainly depends on the objectives of the users as shown by the different examples. General inventory studies for government planning may require a relatively low level of detail. On the other hand, agricultural extensionists require a high level of detail to allow accurate advise to farmers. Tools like GIS may facilitate general inventory studies so that they become more valuable to studies where a high level of detail is required. Their main value is that they may indicate the best strategy for additional data collection or field trials (Van Lanen *et al.* 1992). The level of detail requested by users influences the scale at which the study takes place. A combination of a high level of detail and a complex problem may require a larger scale and, on the contrary, a very low level of detail with a relatively simple problem can be evaluated at a smaller scale with less data requirements. Next to the user, it is often the available data that determines the scale of studies. In the case of the Atlantic Zone of Costa Rica most data are available at a scale of 1:100,000 and 1:150,000, leading almost automatically to the generation of land use scenarios at the same scale.

6. Conclusions

6.1 Future challenges

GIS technology is nowadays widely applied in various types of applications (Bonham-Carter 1994, Worboys 1994, Fotheringham and Rogerson 1994, Michener *et al.* 1994, etc.). An inventory of these applications shows that this technology is, however, mostly applied in an *ad hoc* manner without a generally accepted theoretical basis. This theoretical basis and a standardisation of disciplinary procedures is clearly needed and can, in principle, be realized by applying GIS theory. Molenaar (1991) already identified the need for such a GIS theory, which comprises basic semantics and definitions to be used for quality control. Nevertheless, a standard theory still lacks and general concepts presented in the eighties (e.g. Burrough 1986) still function as a general theory. Does this mean that the development of GIS technology stagnated since 1986? Not really. The plethora of GIS related publications indicates that developments continue but that GIS related research is split into two separate groups. One group deals with theoretical concepts, and the efficiency of data storage (e.g. Egenhofer and Herring 1991, Kämpke 1994). The other applies GIS for their specific disciplinary problems (e.g. Liff *et al.* 1994). The gap between the two groups needs to be closed before standard disciplinary procedures can be developed that are supported by present theoretical insights into GIS technology. Only then can GIS fully support agricultural sciences in an efficient manner.

Soil science has to include GIS technology in survey procedures as defined by Soil Survey Staff (1951) or in more recent updates. Procedures as presented in Chapter 2 for the storage of soil survey data, but also work by e.g. Burrough (1991) and Bregt (1992) can form the basis for such adaptations. Standard procedures have to allow for variations due to scale and objectives. Quality indicators for soil survey such as the number of observations per square centimetre on the final soil map are useful, but GIS technology enables us to include new kinds of indicators. At the same time, variety in scales and applications most likely results in the need to develop several standards. At large scale applications, site-specific management and geo-statistics may play an important role, whereas at small scales satellite imagery and landscape stratification by applying principles of geomorphology are more important.

Standard procedures for the inventory of land cover and land use may include the development of a generally accepted classification system for land cover, farming systems, and land use. On the basis of such a classification system, database structures can be developed and inventories can be structured. Case studies as presented in Chapter 3 and procedures such as the ones proposed by Huising (1993) (who combines techniques like remote sensing and GIS are combined with field surveys), present a starting point for the development of new methodology for land use inventories.

Increasingly, agricultural sciences will be confronted with problems for which multi-disciplinary research teams will have to be formed. An example is the analysis of alternative land use scenarios which requires the integration of different disciplinary models and databases (Section 4.3). Common concepts and definitions derived from GIS theory will facilitate interchange of data among different disciplines. An application-oriented approach in the development of geographic information theory is needed to achieve a relevant and successful theory. Tools for systems integration where GIS is used in combination with different simulation models or other non-spatial database systems will be increasingly important in the future. Only by development of these tools, the development of huge, highly complex models can be avoided. These complex models can only be understood, if that, by their original developers and, therefore, often they are of limited applicability.

The procedures for applications need, for all practical purposes, to be based on standard commercial software packages. At the same time, procedures should not stop at creating a digital database, but should continuously reflect results of the ongoing interaction with users. Thus, GIS is extended from computer-aided-mapping to a true scientific tool, opening new and creative opportunities for research.

Quantitative parameters are needed for spatial analysis to improve the often qualitative and subjective procedures to compare different maps. Spatial statistics to check whether two maps are significantly different can only successfully be applied when quantitative quality indicators are available for the individual maps.

6.2 General conclusions

1. The efficiency of attribute database structures largely depends on its future use. Therefore, an inventory of queries and applications has to precede the selection of the database structure. Quantitative indicators like the number of files, number of algorithmic steps for specific queries, and size of the database may help to objectively select the most appropriate database structure. Typically, the original surveyor possesses the expert knowledge for generalizations which are often needed when studying small-scale problems. However, users of the database who are unfamiliar with the data do not possess this expertise, even though they often have to make these generalizations. As a consequence, general purpose databases require additional rules for generalizations to be created by the original surveyor.

2. Few tools are available to describe temporal changes that have occurred when mapping areas at different times. Markov chains and especially those stratified for relevant driving forces may support the comparison, by being used as a quantitative parameter.

3. For the analysis of spatial data, GIS software can be linked to disciplinary models, as is demonstrated in this thesis. This enables the development of general-purpose GIS-software for disciplinary applications. Most problems to be studied have, however, an interdisciplinary character. A modular approach, linking different disciplinary GIS models, is preferred above a full integration of GIS and one comprehensive model as it increases the transparency and flexibility of the model structure and of procedures to be followed.

4. The level of detail and scale of applications varies widely when reviewing current land use problems, and so will the resulting data requirements. The required level of detail of the objectives of a study and the corresponding data needs have to match.

5. Available data sets, like soil survey data, can be used to screen for potentially interesting or hazard areas where more specific research can be focused on. As a result, data needs for a specific scale and level of detail can decrease significantly.

Abstract

Stoorvogel, J.J., 1995. Geographical information systems as a tool to explore land characteristics and land use, with reference to Costa Rica. Doctoral thesis. Wageningen Agricultural University, Wageningen, The Netherlands, 151 pp.

An adequate inventory of land characteristics and land use is increasingly necessary to support agricultural land use planning, especially in view of the conflicting demands on scarce land resources. Fortunately new tools like GIS are being developed and adapted to support these inventories. Although GIS may be a useful tool for the storage and management of spatial data, its development is often "technology driven" and not directly focused on the applications. This thesis presents approaches to use GIS in the inventory and analysis of land characteristics and land use. The approaches are explored and illustrated for the perhumid tropical lowlands in the northeast of Costa Rica. More detailed studies are focused on the Neguev settlement located in these lowlands.

In Chapter 2, a procedure is formulated to develop and select database structures for soil survey data. The procedure is based on a five step approach in which i) a data model is developed for the soil survey data, ii) alternative database structures are created, iii) possible queries are analyzed, iv) the efficiency of the database structures is evaluated on the basis of quantitative indicators, and v) the most appropriate database structure is selected. Applications are not always based on queries only. Therefore, the structure of the soil survey database is tested on the basis of a practical application: possible modelling approaches to deal with biocide leaching on the basis of the soil survey data. Biocide leaching is one of the main environmental problems in the Northern Atlantic Zone. A best possible assessment of the severity of the problem, from different perspectives, is needed for the various stakeholders. For the Atlantic Zone, one of the few readily available data sets comprises the soil survey data. To deal efficiently with the different approaches of users, the proposed soil information system needs to provide data at different levels of detail. Therefore, a rule base is developed for each of the hierarchical levels (mapping units, pedons, and soil horizons). The rule base includes decision rules for the generalization of data at a specific hierarchical level. Although aside from soil survey data, additional measurements may be necessary for many applications, soil survey data are useful to screen for potential risk areas and to select sites where additional measurements are most effective.

Typically, the analysis of land use at a regional scale should be focused on its changes over time, but this is rarely done in a systematic way. In Chapter 3, the use of GIS to quantitatively describe land use dynamics is explored. Three different

indicators for land use dynamics have been developed. The indicators include a single-time approach based on qualitative knowledge of the colonization history, Markov chains with soil type as a probability modifier, and Markov chains with a geographical analysis to stratify for polygon size, shape and neighbouring land covers.

Often, users of GIS require very specific, disciplinary operations on geo-information that are not supported by GIS. These operations can be made available to the GIS through links with external models. In Chapter 4, structures and examples are given to link GIS with models dealing with the sustainability of agricultural production. A general structure for the GIS-model interface is presented and identifies six consecutive steps: i) geometry operations, ii) attribute operations, iii) data export from the GIS to the external model, iv) model run, v) data import from the model into the GIS, and, vi) visualisation or spatial analysis of the model results with the GIS. This structure is illustrated for a case study where a GIS is linked with a LP model for the analysis of alternative land use scenarios. The structure can be operationalized, using the abilities of many commercial software packages to develop user oriented applications.

To explore the possibilities to reduce soil nutrient depletion in a settlement area, a GIS was linked with a model estimating soil nutrient depletion for land use systems and a LP model. The distribution of land use over different land units can be optimized with the LP model to minimize soil nutrient depletion in the settlement. This technique explores the geographical distribution of land utilization types to create a more sustainable basis for agriculture in the area. In contrast with traditional land use planning where land utilization types are matched with land units on the basis of maximizing present agricultural production, this approach focuses on long-term effects of land use and sustainability.

To explore the trade offs between sustainability and economic objectives, different models and tools were integrated for the analysis of different land use scenarios for the Neguev settlement. Crop growth simulation and expert systems were used to describe alternative land use systems. A GIS was used for data storage, and the analysis and presentation of results. The optimization of land use was carried out by a LP model. Using a series of relevant land use scenarios, effects are studied of: (i) restrictions on biocide use; (ii) nutrient depletion as a negative contribution to farm income, and (iii) changes in capital availability. For the integration of models and tools, a modular approach is proposed, which is based on separate software packages and appropriate database structures. The methodology is particularly appropriate for interdisciplinary research, integrating socio-economic and agro-ecological data.

In Chapter 5, the use of GIS databases and data needs for the analysis of land use and its sustainability is studied. Externally, land use can be affected by incentives and

regulations. Data needs are studied and discussed for the analysis of regional production possibilities of maize, an analysis of sustainability indicators, and the possible contamination of ground and surface water with the commonly used nematicide Ethoprop. The different cases vary in their complexity and the level of detail required for the results. Data requirements change correspondingly. General inventories may already indicate which type of data collection is useful. Studies with a low level of detail must precede more detailed studies, while complex detailed studies could benefit from a change of scale, associated with a more generalized representation of data.

Future challenges to incorporate the use of GIS in both disciplinary and interdisciplinary methodologies are recognized. This will require an integrated development of both GIS technology and applications. The ultimate challenge remains applying the proposed techniques to support the increasing demand for agricultural products and at the same time safeguarding the sustainability of the production and natural resources.

Additional index words: agricultural systems, database structures, GIS, land cover, land use dynamics, land use inventory, modelling, biocide leaching, soil nutrient depletion, soil survey.

Samenvatting (summary in Dutch)

Stoorvogel, J.J., 1995. Het gebruik van geografische informatie systemen als een instrument bij de verkenning van landkarakteristieken en landgebruik, met aandacht voor Costa Rica (in engels). Dissertatie. Landbouw Universiteit Wageningen, Wageningen, Nederland, 151 pp.

Agrarische landgebruiksplanning vereist in toenemende mate een goede inventarisatie van landkarakteristieken en landgebruik, met name in verband met de toenemende landschaarste. Nieuwe instrumenten als geografische informatie systemen (GIS) worden ontwikkeld en aangepast om deze inventarisatie te ondersteunen. Ondanks het feit dat GIS een nuttig instrument kan zijn voor de opslag en het beheer van ruimtelijke gegevens, wordt GIS vaak ontwikkeld vanuit een technologisch standpunt en is de ontwikkeling veelal niet direkt gericht op toepassingen. Deze dissertatie beschrijft de mogelijkheden voor het gebruik van GIS bij de inventarisatie en analyse van landkarakteristieken en landgebruik. De verschillende benaderingen zijn bestudeerd en geïllustreerd voor de humide tropische laaglanden in de Atlantische Zone van Costa Rica. Meer gedetailleerde studies zijn gericht op het landhervormingsproject Neguev, gelegen in deze laaglanden.

Hoofdstuk 2 beschrijft een procedure voor de ontwikkeling en selectie van structuren voor gegevensbestanden van bodemkarteringen. De procedure bestaat uit vijf opeenvolgende stappen: i) een gegevensmodel voor bodemkartingsgegevens wordt beschreven, ii) verschilllende alternatieve structuren voor gegevensbestanden worden gecreëerd, iii) informatie die mogelijkerwijze in de toekomst veelvuldig wordt opgevraagd wordt geïnventariseerd, iv) de efficiëntie van de structuren voor gegevensbestanden wordt geëvalueerd op basis van kwantitatieve indicatoren, en v) de meest geschikte structuur wordt geselecteerd. In veel gevallen worden de bodemkundige gegevens niet alleen geselecteerd of opgevraagd, maar vinden er analyses met behulp van modellen plaats. De structuur voor bodemkundige gegevensbestanden is daarom ook getest op basis van een praktische toepassing: het schatten van het risico van biocidenuitspoeling met behulp van verschillende modelaanpakken op basis van bodemkundige gegevens. De uitspoeling van biociden is één van de belangrijkste milieuproblemen in de Atlantische Zone. Een zo goed mogelijke inschatting van de ernst van het probleem is vanuit verschillende standpunten noodzakelijk voor diverse belangengroepen. De bodemkartering levert één van de weinige direkt beschikbare gegevenssets voor de Atlantische Zone. Om efficiënt met de verschillende aanpakken van gebruikers om te gaan moet het voorgestelde bodemkundig informatiesysteem gegevens in verschillende mate van detail kunnen aanleveren. Er is daarom op ieder hiërarchisch niveau (kaarteenheid, bodem en bodemhorizont) een set beslissingsregels voor generalisaties toegevoegd. Alhoewel er naast de bodemkartering aanvullende metingen noodzakelijk kunnen zijn, speelt de bodemkartering nog steeds een belangrijke rol bij het selecteren van potentiële risicogebieden en locaties voor monstername.

De analyse van landgebruik op een regionaal niveau is traditiegetrouw gericht op de analyse van veranderingen in de tijd. Dit wordt echter zelden systematisch uitgevoerd. In hoofdstuk 3 is het gebruik van GIS bij de kwantitatieve beschrijving van landgebruiksdynamiek onderzocht. Hiervoor zijn drie verschillende indicatoren ontwikkeld: i) een analyse op basis van een eenmalige inventarisatie van landgebruik in combinatie met kwalitatieve gegevens over de kolonisatiegeschiedenis, ii) Markov-ketens met een stratificatie per bodemtype, en iii) Markov ketens met een geografische analyse voor een stratificatie van polygonen op basis van hun grootte, vorm en de landbedekking in omliggende polygonen.

Gebruikers van GIS vragen vaak specifieke, disciplinaire operaties van de ruimtelijke gegevens die niet worden ondersteund door het GIS. Deze operaties kunnen toch worden uitgevoerd door het GIS aan externe modellen te koppelen. In hoofdstuk 4 wordt een structuur gegeven voor de koppeling van GIS aan modellen. De koppeling wordt geïllustreerd met een aantal voorbeelden uit het landhervormingsproject Neguev die gerelateerd zijn aan duurzaamheidsaspecten

De structuur voor de koppeling omvat zes opeenvolgende stappen: i) geometrische operaties, ii) attribuut operaties, iii) de uitvoer van gegevens van het GIS naar de externe modellen, iv) de modelberekeningen, v) de invoer van modeluitkomsten in het GIS, en vi) de visualisatie en ruimtelijke analyse van de resultaten met behulp van het GIS. De structuur wordt geïllustreerd met een voorbeeld waar een GIS wordt gekoppeld aan een lineair programmeringsmodel voor de analyse van alternatieve landgebruiksscenario's. De structuur kan worden geoperationaliseerd met behulp van de mogelijkheid binnen commerciële GIS-pakketten om gebruikersspecifieke toepassingen te ontwikkelen.

Om de mogelijkheden tot het terugbrengen van verliezen van bodemnutriënten te onderzoeken is een GIS gekoppeld aan een model dat een schatting maakt van de nutriëntenbalans onder verschillende vormen van landgebruik, en aan een lineair programmeringsmodel. De regionale verspreiding van landgebruik kan worden geoptimaliseerd met het lineair programmeringsmodel om aldus het nutriëntenverlies te minimaliseren. De techniek onderzoekt de geografische landgebruiksverdeling om te komen tot een meer duurzame landbouw in de regio. In tegenstelling tot traditionele landgebruiksplanning, waarbij de eisen van landgebruiksvormen gekoppeld worden aan de eigenschappen van de verschillende landschappelijke eenheden om de huidige landbouwkundige productie te maximaliseren, richt deze methode zich op de langdurige effecten van landgebruik en duurzaamheid.

Om de wisselwerking tussen duurzaamheid en economische doelstellingen verder te onderzoeken, zijn verschillende modellen en instrumenten geïntegreerd voor de analyse van alternatieve landgebruiksscenario's voor Neguev. Gewasgroeimodellen en expertsystemen zijn gebruikt om alternatieve landgebruikssystemen te beschrijven. Een GIS is gebruikt voor de opslag en analyse van ruimtelijke basisgegevens en de visuele presentatie van de resultaten. De optimalisatie van landgebruik is uitgevoerd met behulp van een lineair programmeringsmodel. Door middel van een reeks alternatieve scenario's zijn de effecten bestudeerd van i) beperkingen van het gebruik van biociden, ii) nutriëntenverlies als een negatieve bijdrage aan het inkomen, en iii) veranderingen in kapitaalbeschikbaarheid. Voor de integratie van de verschillende modellen en gereedschappen wordt een modulaire aanpak voorgesteld. De methodologie is geschikt voor interdisciplinair onderzoek waarbij sociaal-economische en agro-ecologische gegevens worden gekoppeld.

In hoofdstuk 5 wordt de behoefte aan ruimtelijke gegevens voor de analyse van landgebruik en duurzaamheid onderzocht. Databehoeftes zijn bestudeerd en bediscussiëerd voor de analyse van de regionale productiemogelijkheden van maïs, de analyse van duurzaamheidsindicatoren, een kwantitatieve procedure om de mogelijke vervuiling van bodem- en oppervlaktewater te bepalen, en de analyse van alternatieve landgebruiksscenario's. De verschillende voorbeeldstudies variëren in complexiteit en de mate van detail vereist in de resultaten en hebben daarom een verschillende databehoefte.

Het opnemen van GIS-technologie in zowel disciplinaire als interdisciplinaire procedures is een grote uitdaging voor de toekomst. De integratie vereist een gelijktijdige ontwikkeling van zowel GIS-technologie als GIS-toepassingen. De grootste uitdaging ligt in de toepassing van de voorgestelde technieken ter ondersteuning van de toenemende vraag naar landbouwprodukten, bij het bevorderen van duurzaamheid van de productie en het beschermen van de natuurlijke hulpbronnen.

Aanvullende index woorden: Landbouwkundige systemen, structuren voor gegevensbestanden, GIS, landgebruiksdynamiek, modelering, biocide-uitspoeling, bodemnutriënten balans, bodemkartering.

References

ABEL, D.J., P.J. KILBY, and R.J. DAVIS, 1994. The systems integration problem. *International Journal on Geographical Information Systems* 8:1-12.

ADDISCOT, T.M., and J. WAGENET, 1985. Concepts of solute leaching in soils: a review of modelling approaches. *Journal of Soil Science* 36: 411-424.

ALVARADO, A., E.E. GUTIÉRREZ, N. BALDARES, and L.G. BRENES, 1993. Indicadores de sostenibilidad para el sector agrícola y de recursos naturales. IX Congreso Nacional. La agricultura de hoy, para la Costa Rica del mañana. Colegio de Ingenieros Agronomos, San José, Costa Rica.

ANONYMOUS, 1981. Plano mosaico del proyecto Neguev. Instituto de Tierras y Colonización, San Pedro, 1 map.

ANONYMOUS, 1994. World Development Indicators. Socio-economic time-series access and retrieval system. International Bank for Reconstruction and Development, The World Bank, Washington, D.C., USA.

BAILEY, T.C., 1994. A review of statistical spatial analysis in geographical information systems. In: Fotheringham, S., and P. ROGERSON (Eds), 1994. Spatial Analysis and GIS. Taylor and Prancis, London, United Kingdom, 13-44.

BAUMGARDNER, M.F., and R.F. VAN DE WEG, 1989. Space and time dimensions of a world soils and terrain database. In: J. Bouma and A.K. Bregt (Eds). Land qualities in space and time. Pudoc, Wageningen, The Netherlands, 35-43.

BEETS, W.C., 1990. Raising and sustaining productivity of smallholder farming systems in the tropics. AgBé Publishing, Alkmaar, The Netherlands.

BELDER, M., 1994. Land use and land use dynamics in the Atlantic Zone of Costa Rica. Atlantic Zone Programme, Guápiles, Costa Rica.

BILSBORROW, R.W., and H.W.O. OKOTH-OGENDO, 1992. Population-driven changes in land use in developing countries. *AMBIO* 21:37-45.

BONHAM-CARTER, G.F., 1994. Geographic Information Systems for geoscientists: modelling with GIS. Computer Methods in the Geosciences, Vol. 13. Pergamon, Elsevier Sciences Ltd., Kidlington, United Kingdom.

BOUMA, J., 1991. Influence of soil macroporosity on environmental quality. *Advances in Agronomy* 46: 1-37.

BOUMA, J., W. DE VRIES, and P.A. FINKE, 1993. Models for predicting environmental impacts. In: R.C. Wood and J. Dumanski (Eds). Proceedings of the International workshop on sustainable land management for the 21st century (June 20-26, 1993). University of Lethbridge, Lethbridge, Canada, 239-249.

BOUMA, J., and M.R. HOOSBEEK, 1995. The contribution and importance of soil scientists in interdisciplinary studies dealing with land. In: R.J. Wagenet, J. Bouma, and J. Hutson (Eds). The role of soil science in interdisciplinary research. Soil Science Society of America. Madison WI, USA, (In press).

BOUMA, J., and H.A.J. VAN LANEN, 1987. Transfer functions and threshold values: from soil characteristics to land quelities. In: Quantified land evaluation. Proc. Workshop ISSS/SSSA. ITC Publication No. 6. Enschede, The Netherlands, 106-111.

BOUMA, J., H.A.J. VAN LANEN, A. BREEUWSMA, H.J.M. WÖSTEN, and M.J. KOOISTRA, 1986. Soil survey data needs when studying modern land use problems. *Soil use and management* 2: 125-130.

BOUMA, J., M.C.S. WOPEREIS, J.H.M. WÖSTEN, and A. STEIN, 1993. Soil data for crop-soil models. In: F.W.T. Penning de Vries, P. Teng, and K. Metselaar (Eds). Systems approaches for agricultural development. Kluwer, Dordrecht, The Netherlands, 207-220.

BRACKEN, I., and C. WEBSTER, 1989. Towards a typology of geographical information systems. *International Journal of Geographical Information Systems* 3: 137-152.

BREEUWSMA, A., J.H.M. WÖSTEN, J.J. VLEESHOUWER, A.M. VAN SLOBBE, and J. BOUMA, 1986. Derivation of land qualities to assess environmental problems from soil surveys. *Soil Science Society of America Journal* 50: 186-190.

BREGT, A.K., 1992. Processing of Soil Survey Data. PhD Thesis. Agricultural University, Wageningen, The Netherlands.

BRINK, M., 1988. Doblar o quitar. Sistemas de producción de maíz en la parte norte de la Zona Atlántica de Costa Rica. Field Report No. 16. Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

BROUWER, F.M., and M.J. CHADWICK (Eds.), 1991. Land use changes in Europe. Kluwer Academic publishers, Dordrecht, The Netherlands.

BRUIJNZEEL, L.A., 1990. Hydrology of moist tropical forests and effects of conversion: a state of knowledge review. Free University, Amsterdam, The Netherlands.

BURROUGH, P.A., 1986. Principles of geographical information systems for land resources assessment. Clarendon Press, Oxford, United Kingdom.

BURROUGH, P.A., 1989. Modelling land qualities in space and time: the role of geographical information systems. In: J. Bouma and A.K. Bregt (Eds). Land qualities in space and time. Pudoc, Wageningen, 317-320.

BURROUGH, P.A., 1991. Soil Information Systems. In: D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Eds). Geographical information systems. Volume 2: Applications. Longman Scientific & Technical, Harlow, United Kingdom, 153-169.

CAMPBELL, W.G., M.R. CHURCH, G.D. BISHOP, D.C. MORTENSON, and S.M. PIERSON, 1989. The role for a geographical information system in a large environmental project. *International Journal on Geographical Information Systems* 3: 349-362.

CASTILLO, L.E., C. RUEPERT, and E. SOLIS, 1993. Pesticide residues in aquatic ecosystems of lowland areas in Costa Rica. Poster presented at the 4th Workshop on Modern Pesticides in Prague.

CHEN, P.P., 1976. The entity-relationship model: toward a unified view of data. *ACM Transactions on Database Systems* 1: 9-36.

CHUVIECO, E., 1993. Integration of linear programming and GIS for land-use modelling. *International Journal on Geographical Information Systems* 7: 71-83.

COOKE, G.W., 1982. Fertilizing for maximum yield. English Language Book Society/Collins, London, United Kingdom.

COX, D.R., and H.D. MILLER, 1965. The theory of stochastic processes. Chapman and Hall, London, United Kingdom.

DE BRUIN, S., 1992. Estudio detallado de los suelos del asentamiento Neguev. Report No. 25 (Phase 2). Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

DE OÑORO, M.T. (ed.), 1990. El asentamiento Neguev. Interacción de campesinos y estado en el aprovechamiento de los recursos naturales. Informe Técnico No. 162, Programme paper No. 7. Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

DE VRIES, P., 1992. Unruly clients. PhD thesis. Agricultural University, Wageningen, The Netherlands.

DENT, D., and A., YOUNG, 1981. Soil survey and land evaluation. E&FN Sponman, London, United Kingdom.

DERCKSEN, P.M., 1991. A soil erosion mapping exercise in Costa Rica: purposes, methodology and results. In: W.-G. Vahrson, M. Alfaro, and G. Palacios (Eds). Taller de erosion de suelos. FAO-PRODAF-IPGH, Heredia, Costa Rica, 164-170.

DESPOTAKIS, V.K., 1991. Sustainable development planning using geographical information systems. PhD thesis, Free University, Amsterdam, The Netherlands.

EGENHOFER, M.J., and J.R. HERRING, 1991. High-level spatial data structures for GIS. In: D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Eds). Geographical information systems. Volume 1: Principles. Longman Scientific & Technical, Harlow, United Kingdom, 227-237.

FABER, D.C., 1986. The use of agronomic data in the socio-economic models of the Centre for World Food Studies. In: Van Keulen, H., and J. Wolf, (Eds). Modelling of agricultural production: weather, soils and crops. Pudoc, Wageningen, The Netherlands, 329-340.

FAO, 1972. Production Yearbook 1971, Vol 25. Food and Agricultural Organization, Rome, Italy.

FAO, 1976. A framework for land evaluation. Soils bulletin 32. FAO, Rome, Italy.

FAO, 1992. Production Yearbook 1991, Vol 45. Food and Agricultural Organization, Rome, Italy.

FAO, 1993. FESLM: an international framework for evaluating sustainable land management. World Resources Report 73, Land and Water development Division, FAO, Rome, Italy.

FARSHAD, A., and J.A. ZINCK, 1993. Seeking agricultural sustainability. *Agriculture, Ecosystems and Environment* 47: 1-12.

FOTHERINGHAM, S., and P. ROGERSON (Eds), 1994. Spatial Analysis and GIS. Taylor and Prancis, London, United Kingdom.

FOURNIER, F., 1989. The effect of human activity on soil quality. In: J. Bouma and A.K. Bregt (Eds). Land qualities in space and time. Pudoc, Wageningen, The Netherlands, 25-31.

FOURNIER, L.A., 1993. Recursos naturales. Editorial UNED, San José, Costa Rica.

FRANK, A.U., and D.M. MARK, 1991. Language issues in GIS. In: D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Eds). Geographical information systems. Volume 1: Principles. Longman Scientific & Technical, Harlow, United Kingdom, 147-163.

FRESCO, L., H. HUIZING, H. VAN KEULEN, H. LUNING, and R. SCHIPPER, 1990. Land evaluation and farming systems analysis for land use planning. FAO guidelines: Working document. FAO, Rome, Italy.

FU, B., 1989. An optimum model for land use in Xizhuanggou basin, China. In: J. Bouma and A.K. Bregt (Eds). Land qualities in space and time. Pudoc, Wageningen, The Netherlands, 317- 320.

GARRITY, D.P., and P.C. AGUSTIN, 1995. Historical land use evolution in a tropical acid upland agroecosystem. *Agriculture, Ecosystems & Environment* 53: 83-95.

GUIKING, F.C.T., D.M. JANSEN, and L.O. FRESCO, 1994. The use of simplified nutrient balances at farm level to determine boundary conditions for sustainable production. In: J.K. Syers and D.L. Rimmer (Eds). Soil science and sustainable land management in the tropics. CAB international, Wallingford, United Kingdom, 248-257.

HARTSHORN, G., L. HARTSHORN, A. ATMELLA, L.D. GÓMEZ, A. MATA, L. MATA, R. MORALES, R. OCAMPO, D. POOL, C. POOL, C. QUESADA, C. SOLERA, R. SOLÓRZANO, G. STILES, J. TOSHI, A. UMAÑA, C. VILLALOBOS, and R. WELLS, 1982. Costa Rica country and environmental profile. A field study. Tropical Science Center, San José, Costa Rica.

HASSAN, M.H., and C. HUTCHINSON (Eds), 1992. Natural resource and environmental information for decisionmaking. The World Bank, Washington, D.C., USA.

HAZELL, P.B.R., and R.D. NORTON, 1986. Mathematical programming for economic analysis in agriculture. Macmillan Publishing Company, New York, USA.

HERRERA, W., 1985. Clima de Costa Rica. EUNED, San José, Costa Rica.

HEUVELINK, G.B.M., 1993. Error propagation in quantitative spatial modelling, applications in Geographical Information Systems. PhD Thesis. University of Utrecht, The Netherlands.

HILJE, L., L.E. CASTILLO, L.A. THRUPP, I. WESSELING, 1987. El uso de los plaguicidas en Costa Rica. EUNED, San José, Costa Rica.

HOOSBEEK, M.R., and R.B. BRYANT, 1992. Towards the quantitative modeling of pedogenesis - a review. *Geoderma* 55: 183-210.

HOUGHTON, R.A., D.S. LEFKOWITZ, and D.L. SKOLE, 1991. Changes in the landscape of Latin America between 1850 and 1985. I. Progressive loss of forests. *Forest Ecology and management* 38:143-172.

HUISING, J., 1993. Land use zones and land use patterns in the Atlantic Zone of Costa Rica. PhD thesis. Agricultural University, Wageningen, The Netherlands.

HUTSON, J.L., 1993. Applying one-dimensional deterministic chemical fate models on a regional scale. *Geoderma* 60: 201-212

ILACO, 1981. Agricultural compendium for rural development in the tropics and subtropics. Elsevier Scientific Publishing Company, Amsterdam, The Netherlands.

JANSEN, D.M., and R.A. SCHIPPER, 1995. A static, descriptive approach to quantify land use systems. *Netherlands Journal of Agricultural Science* 43: 31-46.

JANSEN, D.M., J.J. STOORVOGEL, and R.A. SCHIPPER. 1995. Using sustainability indicators in agricultural land use analysis: an example from Costa Rica. *Netherlands Journal of Agricultural Science* 43: 61-82.

JANSEN, L., 1994. Methodology for updating terrain object data from remote sensing data: the application of Landsat TM data with respect to agricultural fields. PhD thesis, Agricultural University, Wageningen, The Netherlands.

JANSSEN, B.H., F.C.T. GUIKING, D. VAN DER EIJK, E.M.A. SMALING, J. WOLF, and H. VAN REULER, 1989. A system for quantitative evaluation of soil fertility and the response to fertilizers. In: J. Bouma and A.K. Bregt (Eds). Land qualities in space and time. Pudoc, Wageningen, The Netherlands, 185-188.

JURY, W.A., and H. FLÜHLER, 1992. Transport of chemicals through soil: mechanisms, models, and field applications. *Advances in Agronomy* 47: 141-201.

KÄMPKE, T., 1994. Storing and retrieving changes in a sequence of polygons. *International Journal on Geographical Information Systems* 8: 493-513.

KELLER, M., D.J. JACOB, S.C. WOFSY, and R.C. HARRISS, 1991. Effects of tropical deforestation on global and regional atmospheric chemistry. *Climatic Change* 19: 139-158.

LANDON, J.R. (Ed.), 1991. Booker tropical soil manual: a handbook for soil survey and agricultural land evaluation in the tropics and subtropics. Booker Tate Ltd., Thame, United Kingdom.

LÉLÉ, S. M., 1991. Sustainable development: a critical review. *World Development* 19: 607-621.

LIFF, C.I., K.H. RIITERS and K.A. HERMANN, 1994. Forest health monitoring case study. In: W.K. Michener, J.W. Brunt, and S.G. Stafford (Eds). Environmental information management and Analysis: ecosystem to global scales. Taylor & Francis Ltd., London, United Kingdom, 333-355.

LIZANO, J.R., 1993. Desempeño del sector agropecuario durante 1992. MAG/SEPSA, San José, Costa Rica

LUTZ, E., and H. DAILY, 1991. Incentives, regulations and sustainable land use in Costa Rica. *Environmental and Resource Economics* 1: 179-194.

MAG, 1991. Aspéctos tecnicos sobre cuarenta y cinco cultivos agrícolas de Costa Rica. Ministerio de Agricultura y Ganadería, Dirección General de Investigación y Extensión Agrícola, San José, Costa Rica.

MAGUIRE D.J., M.F. GOODCHILD, and D.W. RHIND (Eds), 1991. Geographical information systems. Longman Scientific & Technical, Harlow, United Kingdom.

MEIJERINK, A.M.J., 1989. Modelling in the land and water domain with a versatile GIS, ILWIS; experiences from a large tropical catchment. In: J. Bouma and A.K. Bregt (Eds). Land qualities in space and time. Pudoc, Wageningen, The Netherlands, 73-87.

MICHENER, W.K., J.W. BRUNT, and S.G. STAFFORD (Eds), 1994. Environmental information management and analysis: ecosystem to global scales. Taylor & Francis Ltd., London, United Kingdom.

MOLENAAR, M., 1991. Status and problems of geographical information systems. The necessity of a geoinformation theory. *ISPRS Journal of Photogrammetry and Remote Sensing* 46: 85-103.

MOLENAAR, M., 1993. Object-hierarchies, why is data standardisation so difficult. *Geo Information Systems* 6: 22-28.

NIJKAMP, P., and H.J. SCHOLTEN, 1993. Spatial information systems: design, modelling, and use in planning. *International Journal on Geographical Information Systems* 7: 85-96.

O'KELLY, M.E., 1994. Spatial analysis and GIS. In: S. Fortheringham and P. Rogerson (Eds). Spatial analysis and GIS. Taylor & Francis, London, United Kingdom, 65-79.

OVERTOOM, T, H. MUDDE, and I. KOFFERMAN, 1987. Land use map of the Neguev settlement (1:20,000). unpublished. Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

PARKER, G.G., 1985. The effect of disturbance on water and solute budgets of hillslope tropical rainforest in northeastern Costa Rica. PhD thesis, University of Georgia, Athens, GA., USA.

RAMIREZ, A., and T. MALDONADO (Eds), 1988. Desarrollo socioeconónomico y el ambiente natural de Costa Rica. Situación actual y perspectivas. Series informes sobre el estado del ambiente, Fundación Neotrópica. Editorial Heliconia, San José, Costa Rica.

REINDS, G.J., and H.A.J. VAN LANEN, 1992. Crop production potential of rural areas within the European Communities. II: A physical land evaluation procedure for annual crops and grass. Working Document 66, Netherlands Scientific Council for Government Policy, The Hague, The Netherlands.

REINERS, W.A., A.F. BOUWMAN, W.F.J. PERSONS, and M. KELLER, 1994. Tropical rain forest conversion to pasture: changes in vegetation and soil properties. *Ecological Applications* 4: 363-377.

ROSALES, A., P. MAEBE, and R. SEVENHUYSEN, 1992. Determination of losses of nutrients and nematicides on a banana plantation in the Atlantic Zone. Report No. 33 (Phase 2). Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

SADER, S.A., and A.T. JOYCE, 1988. Deforestation rates and trends in Costa Rica, 1940 to 1983. *Biotropica* 20(1): 11-19.

SANCHEZ, M.A., and G.P. ALVAREZ, 1991. Aplicación de la EUPS a nivel de una microcuenca, el caso de la quebrada Pital, Puriscal. In: W.-G. Vahrson, M. Alfaro, and G. Palacios (Eds). Taller de erosión de suelos. FAO-PRODAF-IPGH, Heredia, Costa Rica, 144-163.

SCHIPPER, R.A., D.M. JANSEN and J.J. STOORVOGEL, 1995. Sub-regional linear programming models in land use analysis: a case study of the Neguev settlement, Costa Rica. *Netherlands Journal of Agricultural Science* 43: 83-109.

SEPSA, 1992. Metodología para la determinación de la capacidad de uso de las tierras de Costa Rica. SEPSA, San José, Costa Rica.

SHARIFI, M.A., 1992. Development of an appropriate resource information system to support agricultural management at farm enterprise level (ARIS). PhD Thesis, Agricultural University. Wageningen, The Netherlands.

SKOLE, D.L., and C. TUCKER, 1993. Tropical deforestation and habitat fragmentation in the Amazon: Satelite data from 1978 to 1988. *Science* 260: 1905-1910.

SMALING, E.M.A., and L.O. FRESCO, 1993. A decision support model for monitoring nutrient balances under agricultural land use (NUTMON). *Geoderma* 60: 235-256.

SMALING, E.M.A., J.J. STOORVOGEL, and P.N. WINDMEYER, 1992. Calculating soil nutrient balances in Africa at different scales. II. District scale. *Fertilizer Research* 35: 237-250.

SOIL SURVEY STAFF, 1951. Soil Survey Manual. Agric. Handb. 18. US Department of Agriculture. Washington, D.C., USA.

SOIL SURVEY SATFF, 1993. State soil geographical data base (STATSGO): Data Users Guide. U.S. Department of Agriculture, Soil Conservation Service, Miscellaneous Publication Number 1492. Washington, D.C., USA.

SOIL SURVEY STAFF, 1994. Keys to the Soil Taxonomy, Sixth Edition. US Department of Agriculture, Washington, D.C., USA.

SOLORZANO, R., R. DE CAMINO, R. WOODWARD, J. TOSI, V. WATSON, A. Vásquez, C. VILLALOBOS, J. JIMÉNEZ, R. REPETTOM, and W. CRUZ, 1991. Accounts overdue: natural resource depreciation in Costa Rica. Tropical Science Centre, San José, Costa Rica & World Resource Institute, Washington, D.C., USA.

SPAANS, E.J.A., G.A.M. BALTISSEN, J. BOUMA, R. MIEDEMA, A.L.E. LANSU, D. SCHOONDERBEEK, and W.G. WIELEMAKER, 1989. Changes in physical properties of young and old volcanic surface soils in Costa Rica after clearing of tropical rain forest. *Hydrological processes* 3: 383-392.

SSSA, 1971, Glossary of Soil Science Terms. SSSA, Madison, WI, USA.

STEYAERT, L.T., and M.F. GOODCHILD, 1994. Integrating geographical information systems and environmental simulation models: a status review. In: W.K. Michener, J.W. Brunt, and S.G. Stafford (Eds). Environmental information management and Analysis: ecosystem to global scales. Taylor & Francis Ltd., London, United Kingdom, 333-355.

STOCKING, M., 1984. Erosion and soil productivity: a review. Consultants' Working Paper No. 1. AGLS, FAO, Rome, Italy.

STOORVOGEL, J.J., 1993[A]. Experiencias de desarrollo sostenible en la Zona Atlántica, el caso de flores de corta. Proceedings of the 'IX Congreso Nacional Agropecuario y de Recursos Naturales', 17-21 October, San José, Costa Rica.

STOORVOGEL, J.J., 1993[B]. Optimizing land use distribution to minimize nutrient depletion: a case study for the Atlantic Zone of Costa Rica. *Geoderma* 60: 277-292.

STOORVOGEL, J.J., and G.L. EPPINK, 1995. Atlas de la Zona Atlántica. Atlantic Zone Programme (CATIE-WAU-MAG), Guápiles, Costa Rica.

STOORVOGEL, J.J., and E.M.A. SMALING, 1991. Assessment of soil nutrient depletion in Sub-Saharan Africa: 1983-2000. Report 28, The Winand Staring Centre for Integrated Land, Soil and Water Research, Wageningen, The Netherlands.

STOORVOGEL, J.J., E.M.A. SMALING, and B.H. JANSSEN, 1992. Calculating soil nutrient balances in Africa at different scales. I. Supra national scale. *Fertilizer Research*:

TAYLOR, D.R.F., 1991. GIS and developing nations. In: D.J. Maguire, M.F. Goodchild, and D.W. Rhind (Eds). Geographical information systems. Volume 2: Applications. Longman Scientific & Technical, Harlow, United Kingdom, 71-84.

TURNER, B.L. II, W.B. MEYER, and D.L. SKOLE, 1994. Global land-use/land-cover change: towards an integrated study. *AMBIO* 23:91-95.

ULLMAN, J.D., 1982, Principles of database systems. Computer Science Press Inc. Rockville, Maryland, USA.

VAHRSON, W.-G., 1991. Aspectos climáticos de la erosión hídrica en Costa Rica, America Central. In: W.-G. Vahrson, M. Alfaro, and G. Palacios (Eds). Taller de erosion de suelos. FAO-PRODAF-IPGH, Heredia, Costa Rica, 33-47.

VAN DE BERG, R., and R. DROOG, 1992. Quantification of farming systems in the Neguev settlement. Report No. 26 (Phase 2). Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

VAN DIEPEN, C.A., H. VAN KEULEN, J. WOLF, and J.A.A. BERKHOUT, 1991. Land evaluation: from intuition to quantification. *Advances in Soil Science* 15: 139-204.

VAN ENGELEN, V.W.O., and T.T. WEN (Eds), 1995. Global and national soils and terrain digital databases (SOTER): Procedures manual. ISRIC, Wageningen, The Netherlands.

VAN KEULEN, H., and F.R. VEENEKLAAS, 1993. Options for agricultural development: a case study for Mali's fifth region. In: F.W.T. Penning de Vries, P. Teng, and K. Metselaar (Eds). Systems approaches for agricultural development. Kluwer, Dordrecht, The Netherlands, 367-380.

VAN KEULEN, H., and J. WOLF (Eds), 1986. Modelling of agricultural production: weather, soils and crops. Pudoc, Wageningen, The Netherlands.

VAN LANEN, H.A.J., M.J.D. HACK TEN BROEKE, J. BOUMA, and W.J.M. DE GROOT, 1992. A mixed qualitative/quantitative physical land evaluation methodology. *Geoderma* 55: 37-54.

VAN OOSTEROM, P.J.M., 1990. Reactive data structures for geographic information systems. PhD thesis, Leiden University, Leiden, The Netherlands.

VAN SLUYS, F.R., H. WAAIJENBERG, W.G. WIELEMAKER, and J.F. WIENK, 1987. Agriculture in the Atlantic Zone of Costa Rica. Summarizing report of an exploratory survey. Informe Técnico No. 123, Programme paper No. 1, Atlantic Zone Programme (CATIE-WAU-MAG), Turrialba, Costa Rica.

VEENEKLAAS, F.R., 1990. Competition pour des ressources limitées: le cas de la cinquième region du Mali. Rapport 3. Description formelle du modèle d'optimisation MALI5. CABO, Wageningen, The Netherlands & ESPR, Mopti, Mali.

VELDKAMP, A., and L.O. FRESCO, 1995. Modelling land use changes and their temporal and spatial variability with CLUE. Department of Agronomy, Agricultural University, Wageningen, The Netherlands.

VELDKAMP, E., A.M. WEITZ, I.G. STARITSKY, and E.J. HUISING, 1992. Deforestation trends in the Atlantic Zone of Costa Rica: a case-study. *Land Degradation & Rehabilitation* 3: 71-84.

WIELEMAKER, W.G., and A.W. VOGEL (Eds), 1993. Un sistema de información de suelos y tierras para la Zona Atlántica de Costa Rica. Report No. 22 (Phase 2), Atlantic Zone Programme, CATIE-UAW-MAG, Turrialba, Costa Rica.

WISCHMEIER, W.H., C.B. JOHNSON, and B.V. CROSS, 1971. A soil erodibility nomograph for farmland and construction sites. *Journal on Soil Water Conservation* 26: 189-192.

WISCHMEIER, W.H. and D.D. SMITH, 1978. Predicting rainfall erosion losses. A guide to conservation planning. Agric. Handbook No 537. USDA, Washington, DC, USA

WORBOYS, M.F., 1994, Innovations in GIS 1. Taylor & Francis Ltd., London, United Kingdom.

WÖSTEN, J.H.M., J. BOUMA, and G.H. STOFFELSEN, 1985. Use of soil survey data for resional soil water simulation models. *Soil Science Society of America Journal* 49: 1238-1244.

WRR (Netherlands Scientific Council for Government Policy), 1992. Ground for choices. Four perspectives for the rural areas in the European Community. Report to the government no. 42, Sdu Uitgeverij, The Hague, The Netherlands.

Curriculum Vitae

Jetse Jacob Stoorvogel was born on July 24, 1965 in Hengelo (O), the Netherlands. In 1983, he started his study regional soil science at Wageningen Agricultural University. For his practicals in soil science, he carried out part of the soil survey for the Northern Atlantic Zone in Costa Rica. Additionally, he did thesis research on the classification and field recognition of Andisols in Costa Rica. After returning to the Netherlands he did thesis research on spatial structures of land characteristics and the effect on land evaluation. Before graduating in 1989, he did a third thesis research for the department of Agronomy where agro-ecological zones for cassava in sub-Saharan Africa were identified.

From August 1989 till April 1990, he worked at the Winand Staring Centre for Integrated Soil and Water Management on a study initiated by FAO to estimate soil nutrient balances in sub-Saharan Africa for both the 1990 and the projected 2000 situation.

From April 1990 till October 1991, he conducted a study financed by the Tropenbos Foundation for the Wageningen Agricultural University. The project involved field measurements of gross inputs and outputs of nutrients for a forested watershed in Côte d'Ivoire.

Since November 1991, he is employed as a staff member at the multi-disciplinary research project of the Wageningen Agricultural University in the Northern Atlantic Zone of Costa Rica, named the Atlantic Zone Programme. At the Atlantic Zone Programme, he deals with the GIS and soil science related aspects of the research which focuses on the development of a methodology for the analysis of alternative land use scenarios. Part of this work resulted in his PhD thesis.