Spatial dynamics in allocation of scarce water

A study about inter-annual and spatial dynamics in agricultural land use and irrigation water use under changing water availabilities around the Orós reservoir in the Northeast of Brazil.

J. G. Leskens
Sept, 2006

University of Twente
Enschede - The Netherlands

ITC

Spatial dynamics in allocation of scarce water

A study about inter-annual and spatial dynamics in agricultural land use and irrigation water use under changing water availabilities around the Orós reservoir in the Northeast of Brazil.

By

Anne Leskens
(Sept 2006)

Master thesis submitted to the Water Engineering and Management department of the University of Twente in cooperation with the International Institute for Geo-information Science and Earth Observation (ITC) in fulfillment of the requirements for the degree of Ingenieur (Master of Science) in Civil Engineering, specialization in Water Engineering and Management.

Assessment board

Chairman: Dr. M.S. Krol
Daily supervisor: Ir. P.R. van Oel
Daily supervisor: Ing. R.J.J. Dost (M.Sc.)
Member: Dr. Ir. C.M.M. Mannaerts

University of Twente
Enschede - The Netherlands

ITC

Conflicts in distribution of available water resources are common in the semi-arid Northeast of Brazil, mostly because of the scarcity of availability of water resulting from climatic conditions. Lacking knowledge about spatial dynamics of land use and its irrigation water use under changing water availabilities limits the determination of suitable solutions in finding the best way to allocate the scarce and strongly varying amount of available water.

The goal of study is to get more knowledge about the spatial dynamics of agricultural land use and irrigation water use under different situations of water availability around strategic water reservoirs in the semi-arid northeast region of Brazil. This is done by analyzing these aspects a spatial way, using GIS and Remote Sensing-techniques, for a research area around the Orós reservoir located in the Northeast of Brazil, during the dry seasons in 2000 to 2005.

Firstly a downscaling is carried out to 6 'areas of interest', each mainly supplied by one source of water availability. Four aspects of water availability could be distinguished in the research area: rainfall, river discharges/reservoir releases, reservoir volumes and locally stored runoff. Each type of water availability is quantified for each area of interest. Secondly the agricultural land use, as largest water user, is determined by applying a land cover classification to satellite images of each year. Thirdly the irrigation water use of each area of interest in each dry season within the research period is estimated by using the Cropwat model, a model able to calculate crop irrigation requirements. The results of these three components are analyzed in a inter-annual way (how are the components evolving during the research period in a particular area of interest) and in a spatial way (how do the components of different areas of interest influence each other during the research period).

The analysis showed that the different types of water sources have different spatial and temporal ranges and different water availabilities. This determines the way in which agricultural water users, dependent on a certain water source, effect the water availability of other agricultural water users.
Rainfall and *river discharge* in the dry season are far too low to fulfill the irrigation water requirement of the areas that are connected to these water sources. *Locally stored water* is only local available; irrigation water use of this source does not noticeable effect downstream water availabilities. Water users around the *reservoirs* (upstream, edge, downstream) in the area influence each others water availability on different time scales: The relevant time scale around small reservoirs are estimated on 1-3 years and > 10 year around the largest reservoir. It is recommended to take these different time scales into account by defining water management actions.

The type of crops grown effects the course of the irrigation water use: paddy and beans require a shorter but a somewhat more intense water supply than banana. The reservoir edge is therefore suitable for temporal crops requiring intense water supplies. Areas downstream of reservoirs are more suitable for banana in case the intense water requirements of paddy and beans cannot be fulfilled. Banana is infeasible in areas where land availability can be temporally limited by overflow.

Conflicten in het verdelen van beschikbaar water komen veel voor in het Noord-Oosten van Brazilië, voornamelijk als gevolg van schaarste in het beschikbaar water. Ontbrekende kennis over de ruimtelijke dynamiek van agrarisch landgebruik en geïrrigeerd watergebruik, onder wisselende situaties van water beschikbaarheid, limiteren het bepalen van bruikbare oplossingen om tot een goede verdeling van het sterk variende beschikbare water te komen.

Het doel van dit onderzoek is om meer inzicht te krijgen te krijgen in de ruimtelijke dynamiek van agrarische landgebruik en geïrrigeerd water gebruik onder wisselende omstandigheden van waterbeschikbaarheid rond strategische reservoirs in het semi-aride Noord-Oosten van Brazilië. Dit is gedaan door de ontwikkelingen in agrarisch landgebruik, geïrrigeerd watergebruik en water beschikbaarheid op een ruimtelijke manier te analyseren, gebruikmakend van GIS- en Remote Sensing-technieken in het gebied rond het Orós reservoir voor de droge seizoenen van 2000 tot 2005.

Zes gebieden, elke voornamelijk afhankelijk van één soort water toevoer, zijn vastgesteld. De aanwezige soorten van waterbeschikbaarheid in het gebied zijn: regenval, rivierwater, reservoir water en plaatselijk opgeslagen runoff. Elk type waterbeschikbaarheid is gekwantificeerd voor elk van de 6 gebieden.
Vervolgens is het agrarsiche landgebruik bepaald met behulp van satteliet beelden en een landgebruik classificatie techniek. De oppervlakten van gewassen die niet bepaald konden worden met behulp van deze techniek zijn verkregen uit data bestanden van locale instanties.
Het geïrrigeerd watergebruik is geschat met behulp van het Cropwat model; een model dat in staat is de irrigatie water behoefte van een gewas te berekenen.
De drie bepaalde componenten zijn op een tijdelijke en een ruimtelijke schaal geanalyseerd.

Deze analyse liet zien dat de verschillende waterbronnen verschillende reikwijdtes en verschillende voorraden hebben en dat de relevante relaties tussen water gebruik en water beschikbaarheid zich op verschillend tijdschalen afspelen. Deze componenten bepalen voornamelijk het effect dat het watergebruik van watergebruikers van een bepaalde waterbron heeft op de waterbeschikbaarheid van andere gebruikers:
De waterbeschikbaarheid uit *regenval* en uit *river afvoer* in het droge seizoen kan niet aan de vraag van water voldoen. *Lokaal opgeslagen runoff* heeft alleen een lokale beschikbaarheid en beïnvloed daarom niet de waterbeschikbaarheid van benedenstroomse watergebruikers.
Water gebruikers rond de *Reservoirs* (bovenstrooms, aan de reservoirrand en benedenstrooms) in het gebied beïnvloeden elkaars water beschikbaarheid in verschillende tijdschalen: Rond de kleine reservoirs is de relevante tijdschaal geschat 1-3 jaar, rond het grootste reservoir is de relevante tijdschaal > 10 jaar. Het wordt aangeraden het water management aan te laten sluiten bij deze relevante tijdschalen.

Het type gewas bepaald het verloop van het benodigde irrigatie water. Rijst en bonen hebben een kortere maar intensere water toevoer nodig dan banaan. De reservoir rand is daarom geschikt voor gewassen met een intensere water behoefte, terwijl de gebieden bendenstrooms van een dam geschikter zijn voor gewassen met een gematigde constantere water behoefte, als banaan. Voor gebieden die tijdelijk onder water kunnen staan is het semi permanente gewas banaan niet geschikt.

First of all I want to express my thanks to my daily supervisor of the Water Engineering and Management department of the University of Twente: Pieter van Oel. During the whole research he helped me with al kind of comments and feedback. Together we explored the semi-arid 'sertão' of Brazil to retrieve as much 'ground truth' and information from local agencies as possible. Also outside the research work we had a very good time together, especially during the period in Brazil. In that aspect I also want to thank Marjella de Vries and Bertien Koopman for their good company in Brazil.

I would like to thank the chairman of my assessment board, Maarten Krol, for his rigorous and careful supervising.

I would like to express my gratitude to the water department of the International Institute for Geo-information Science and Earth Observation (ITC) in Enschede. They gave me the opportunity to get knowledge about GIS and Remote Sensing and they offered me a place to work during the GIS and Remote Sensing analysis. A special word of thank for my supervisors of the ITC: Remo Dost and Chris Mannaerts for helping me to solve all kind of problems during the GIS and Remote Sensing analysis and reviewing my report.

I am very grateful for the hospitality of the 'Departamento de Engenharia Hidráulica e Ambiental' at the Federal University of Ceará (UFC) in Fortaleza (Brazil) and especially for the help of professor José Carlos de Araújo and the daily company of Allexandre Cunha-Costa.

I am greatly indebted to the Brazilian family Cunha-Costa for the accommodations and the heartwarming hospitality and care they offered me during my stay in Brazil. They are justly 'minha familia Brasileira'. Muito obrigado!

Finally, I extend a special note of thanks for the members of the graduation room at the University Twente for their good company during the writing of my thesis.

Table of figures

1 Introduction

1.1 Introduction

The semi-arid region of Brazil is a dry area which is subjected to recurrent, severe droughts. Several problems emerge as a consequence of these droughts. An important issue is the allocation of the available water in the dry periods. The subject dealt in this thesis is the strain between the availability of water and the use of this water.

In this chapter the reader will be firstly introduced to some important background factors of the problem about water allocation in dry periods. These background factors are the natural and the socio-economic conditions related to water allocation in the semi-arid area of Brazil (paragraph 1.2 and paragraph 1.3). Hereafter, the research problem will be defined in paragraph 1.4. The research problem is owned by a large area in the semi-arid region of Brazil. Because the area is too large to investigate, a relatively small, but representative, area is selected to take part in this research. This area is the direct neighborhood of the Orós reservoir in the state of Ceará. (see figure 1.5). Paragraph 1.5 will describe the research approach. This research area will be introduced in paragraph 1.6.

1.2 Natural conditions

1.2.1 Hydrological properties

The northeast of Brazil is a semi-arid region for which precipitation is the most important climatologically variable. Recurrent droughts, "secas", with annual rainfall totals far below average, castigate the region (Werner and Gerstengabe, 2003). For this, the region is called ´drought polygon´. Some important hydrological properties of Ceará are shown in table 1.1.

Hydrological aspect	Quantity (mm/y)
Precipitation	900
Potential evaporation	2200
Actual evapotranspiration	700
Runoff	120
Percolation	90

table 1.1: Hydrological properties of Ceará (Frischkorn et al., 2003)

The very high potentical evaporation in Ceará is powered by 3000 hours of sunshine per year. The real evapotranspiration of 700 mm leaves only 120 mm for runoff and 90 mm for percolation. The annual temporal distribution is characterized by a rainy period during approximately from December to May and a dry period during approximately from June to November. The inter-annual course of the mean precipitation sum for the area shows a high variability. For example, the mean annual precipitation in the states Ceará and Piauí, both states are located within the drought polygon, was 400 mm in the year 1915 and 1500 mm in 1985 (see figure 1.1). Cyclical droughts occur every five years or less and can last more than 2-3 years (Johnsson, 2004). So there is a high probability that prolonged critical periods will occur (COGERH and ENGESOFT, 1999).

Also the moment of beginning and ending of the dry period, and due to that, the length of the dry period in each year is uncertain (see figure 1.2).

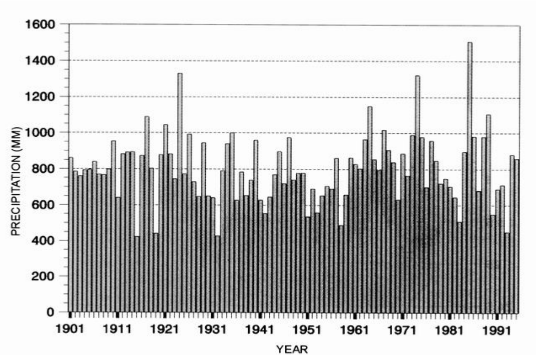


figure 1.1: Temporal course of the mean precipitation sum for the area of Ceará and Piauí (Werner and Gerstengarbe, 2003)

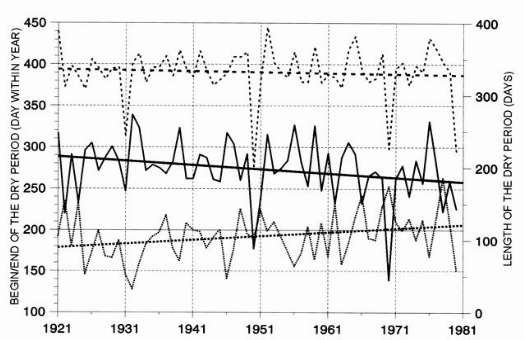


figure 1.2: Temporal course of the beginning (dotted), end (dashed) and length (solid) of the dry period at the station Cedro, period 1921-1980 (Werner and Gerstengarbe, 2003)

There is not only a temporal variation in precipitation, also a spatial distribution occur. Precipitation is distinctly decreasing from the coast to the interior of the country, with an average rainfall of approximately 2000 mm per year along the coastline and approximately 500 mm in the interior (Werner and Gerstengarbe, 2003). The conclusion of what is written above is that precipitation shows a high temporal irregularity in each particular year and a high inter-annual irregularity. Also a high spatial irregularity can be concluded.

1.2.2 Climate

The irregularity of the temporal precipitation and the recurrent droughts are caused by the climatologically properties of the area. The area is defined by a 'trade wind region with the summery maritime weather of the trade wind limit'. This climate is most importantly characterized by a long winter period with little precipitation. The annual cycle of rainy and drought periods is controlled by large-scale circulations. The two seasons, dry and wet, are determined by the position of the Inter Tropical Convergence Zone (ITCZ): "Between 30°N and 30°S latitude, the solar energy, which is higher around the equator than in surrounding areas, is transported by a relatively simple overturning circulation, with rising motion near the equator, poleward motion near the tropopause, sinking motion in the subtropics, and an equatorward return flow near the surface. The region in which the equatorward moving surface flows converge and rise is known as the ITCZ, a high-precipitation band of thunderstorms". The positions of high-pressure areas influence the position of the ITCZ. This high-pressure areas depend on the present sea surface temperature (SST) (see figure 1.3). If, for example, the ITCZ does not shift far enough to the south in one year due to these conditions, this can lead to a weakening of the rain period that can cause a drought (Werner and Gerstengabe, 2003)

The El Niño phenomenon is also called as a reason for the recurrent droughts. So far, a strong correlation with the "El Niño" phenomenon has been proved (Popelewski and Halpert, 1987 in: Frischkorn et al., 2003). El Niño is a climatic phenomenon which causes higher ocean temperatures around the equator in the *east side* of the Pacific Ocean. Due to this higher sea temperature, changes in wind en pressures areas occur, which are causing weather changes all over the world, as well as in the tropical Atlantic.

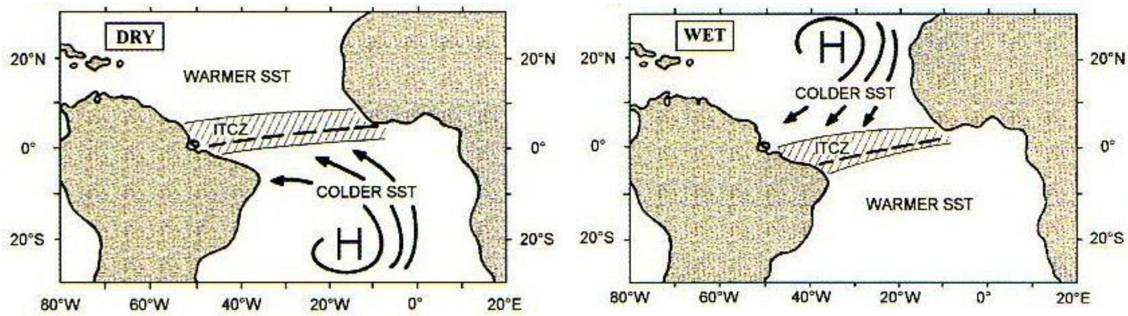


figure 1.3: Scheme of the large-scale circulation patterns over the tropical Atlantic (Werner and Gerstengarbe, 2003)

1.2.3 Water resources

Surface water

In the state of Ceará all rivers are naturally intermittent due to highly seasonally irregular precipitation rates, high evaporation rates and shallow soils on top of the crystalline basement (a basement of rock made up of minerals in a clearly crystalline state). Due to water scarcity in rivers in the dry season and to the low groundwater reserves, the only way to have water available with reasonable confidence for social and economical use is to store the excess of water during the rainy season. This policy, historically known as the 'hydraulic approach', directs all efforts towards the construction of artificial reservoirs (Frischkorn et al., 2003).

Groundwater

As mentioned above, the groundwater reserves in the northeast, and thus in Ceará, are low, due to the shallow soils on top of the crystalline basement. Studies about runoff reached the conclusion that, in the crystalline areas of the Semi-Arid zone an average of only 6 to 8% of the rainfall actually flows superficially or feeds aquifers by percolation. The remaining 92 to 94 % are retained by the soil or lost by evaporation or evapotranspiration (DNOCS, 1982). Around 75% of the area of Ceara contains a crystalline basement. The remaining 25% contains sedimentary water basins. These area's with groundwater resources are located at the borders of the state. (Frischkorn, Araújo and Santiago, 2003). Even though the groundwater reserves are small compared to surface water availability, they are crucial for many cities and irrigated areas, sometimes constituting the only water source. Despite this resource's importance, there is still little knowledge of the real availability and utilization of groundwater in the whole basin (Johnsson, 2004) (DNOCS, 1982). During the field visit in the Iguatú municipality in the state of Ceará, agricultural areas supplied by water form wells and natural lakes, were seen. This will be discussed later in chapter 3.

1.3 Socio-economic conditions

"Natural production risks in the northeast are extremely high. Due to climatic variability, agricultural production suffers extreme yield insecurity, up to consecutive years of total loss. This insecurity also affects the industry and trade sectors, since the processing of agrarian primary inputs and the production of food play a major role in the economy of northeastern Brazil" (Gaese, 2003). "A survey by SUDENE (1999) concerning the drought programs reveals that the drought-affected populations – small-scale farmers, farm workers, day laborers, settlers, etc. – are aware of the problems and have commonly experienced inequalities in receiving their assigned benefit. Undoubtedly, the Brazilian government has so far not succeeded in overcoming the evolving drought calamities" (Gaese, 2003). "Until the early 1990s, the water scarcity and recurrent droughts were treated as essentially a supply problem. This was resolved by massive construction of reservoirs and related water infrastructure. Water reform has introduced new practices aimed at complementing this supply-side approach with demand management" (Johnsson and Kemper, 2005). "This demand management is characterized by a combination of organized public organizations, private entities and civil society representatives which make the implementation of the water resources management instruments possible, in accordance with principles established in law. The river basin committees, for example, bring stakeholders together for debates and decisions regarding problems in the watershed. Among these stakeholders are public organizations such as state secretariats and local government representatives, as well as private and state-owned water users and NGO's. The composition of the board of directors of the Water Agencies is also a mixture of all the previous representatives" (Garrido, 2001).

1.4 Problem definition

Despite recent efforts by Ceara's government to promote more efficient water management, the relationship between water availability and water demand is still unbalanced (COGERH and Engesoft, 1999) (Johnsson and Kemper, 2005).

Conflicts about the distribution of the available water resources are common, mostly because of the strong variation of the availability resulting from climatic conditions. Most frequent conflicts arise between the users that depend directly on reservoir water and those located further downstream and among users in the valleys rendered perennial through regulations (Johnsson, 2004). A proposition is made that the conflicts, which are reemerging in every negotiation, are related to the geographical locations of different irrigation water users around the reservoirs (Van Oel, 2006). Knowledge about this spatial dynamics of land use and its irrigation water use under changing water availabilities is limited. It is expected that better insight in this topic will give more understanding in problems of water allocation and how they can be solved. Therefore, the problem that thesis is dealing with is:

The lacking knowledge about spatial dynamics of land use and its irrigation water use under changing water availabilities.

1.5 Goal of study, research questions and thesis outline

Goal of study

The goal of this study is to explain the inter annual spatial dynamics of land use and irrigation water use under different conditions of water availability around strategic water reservoirs in the semi-arid northeast region of Brazil, by analyzing changes in land use, irrigation water use and water availability in a spatial way for the period from 2000 until 2005 in a representative area.

Demarcation

Since it is not achievable within the available research time, given the chosen research method, to investigate the total semi-arid region of Brazil, a spatial demarcation has to be made. The 'representative area', mentioned in the goal of study, is the direct surrounding of the Orós reservoir, which includes the Trussu reservoir and the Lima-Campos reservoir. These reservoirs and their surrounding are located in the Jaguaribe River Basin. The surrounding of the Orós reservoir is selected as research area, because it is one of the most important reservoirs in the Upper basin of the Jaguaribe River Basin. This area has several spatial distributed water sources for agricultural land use and is therefore suitable for the goal of study. Another reason to select

this area is the availability of satellite images and hydrological data. A further description of this area will be given in paragraph 1.6.
The research period includes the dry season of 2000 to 2005. This temporal demarcation will be explained in paragraph 3.2. Only agricultural land uses supplied by irrigation water will be examined, because these land uses have the highest water demand in the basin during the dry periods.

Research questions

1. What is the annual amount of available water for irrigation in the dry season and which areas have access to it?
2. How much land do the different irrigated agricultural land uses cover in the dry season and what is their spatial distribution?
3. What are the amounts of irrigation water use, related to land use, in the dry season?
4. How are water availability, agricultural land use and irrigation water use related between 2000 and 2005 in different focus areas in the research area?
5. How are spatially explicit water availability, agricultural land use and irrigation water use of different focus areas mutually related in the research area between 2000 and 2005.

Structure of thesis
The methods and theories to answer the research questions will be discussed in chapter 2. The results of research questions 1, 2 and 3 will be dealt in chapter 3: ‘Data processing’. Research questions 4 and 5 will be dealt in chapter 4: ‘Results’. In chapter 5 the results will be discussed. The conclusions, as final answers to the research questions, will be given in chapter 6, together with the recommendations.

1.6 Research area

In figure 1.4 the location of the research area is shown. It is located in the Northeast of Brazil, in the south of the state Ceará. In figure 1.5 a more detailed map of the research area is shown. The area contains three major reservoirs: the Orós reservoir, the Trussu reservoir and the Lima Campos reservoir. The Jaguaribe river is the main river in the area and flows from southwest to the Orós reservoir. Before the Jaguaribe river meets the Orós reservoir, the Trussu river joins the Jaguaribe river. Between the Orós reservoir and the Lima Campos reservoir a water tunnel is located. The water through the tunnel flows from the Orós reservoir to the Lima Campos reservoir. The main cities in the area are: Orós, Iguatú, Quixelo and Icó. The elevation of the area is in between 145 and 480 m. In the plain areas around the river and the reservoirs much agricultural activity takes place.

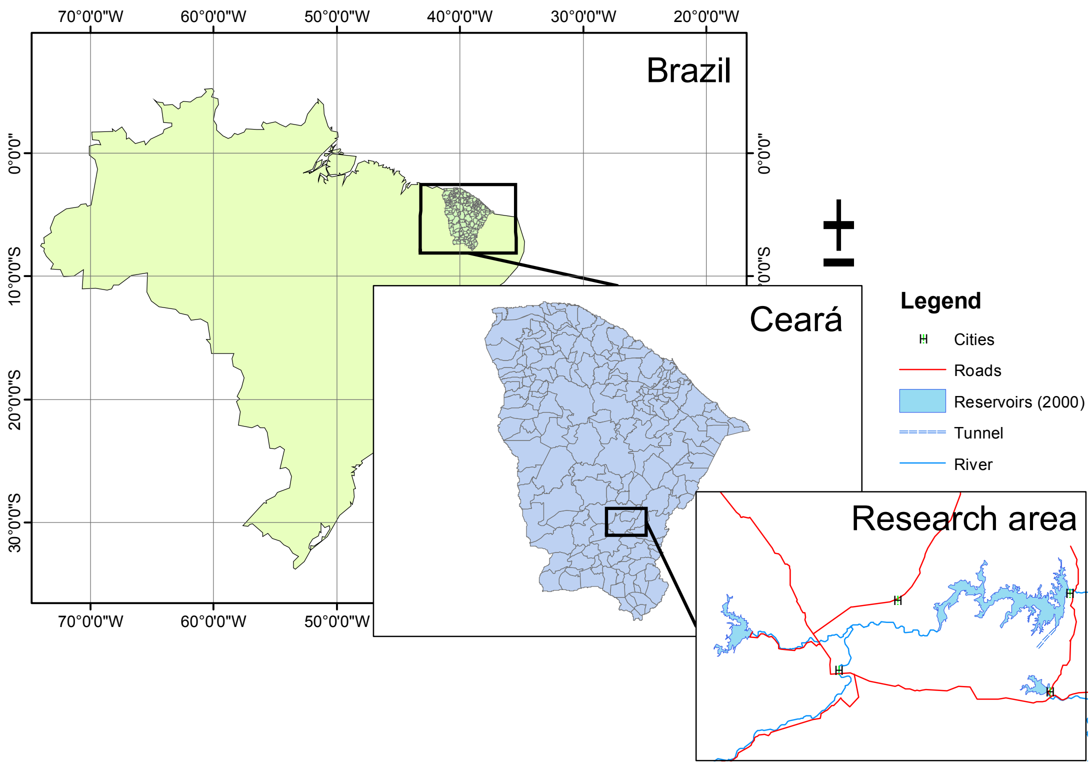


figure 1.4: Location of research area

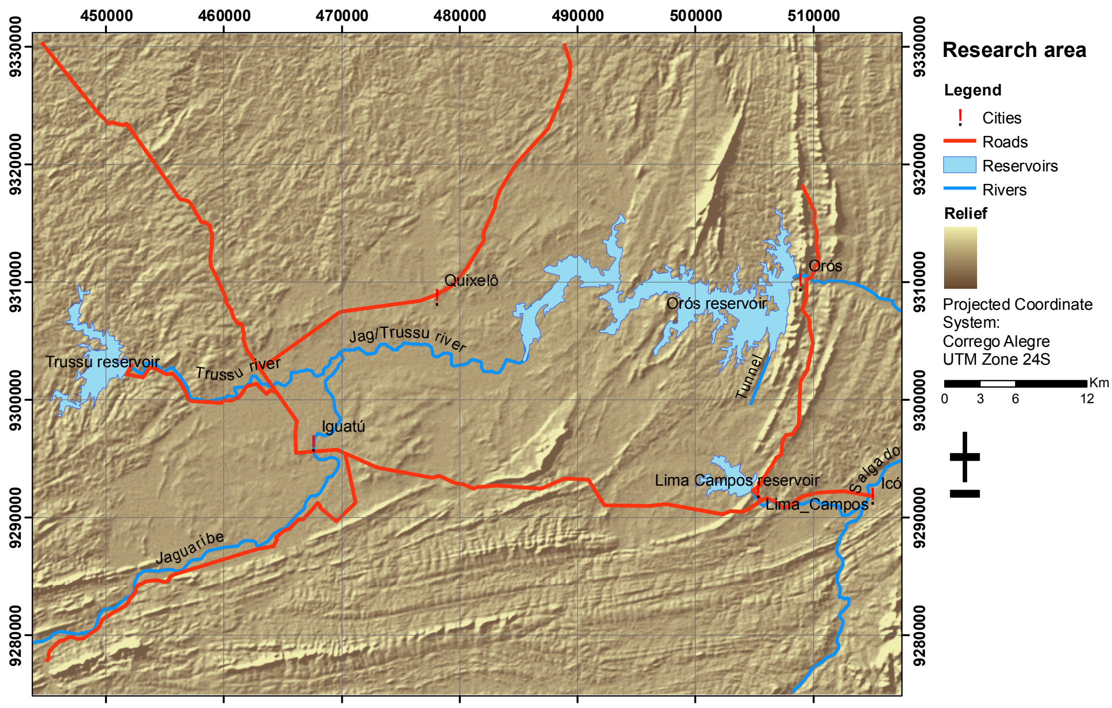


figure 1.5: Research area detailed

2 Theory and methods

2.1 *Introduction: Methodology*

This section introduces the methods that were chosen to be applied for the research. These choices will be explained in this paragraph. This report is build up in the same way as the described methodology, so it forms also the guideline through this report.

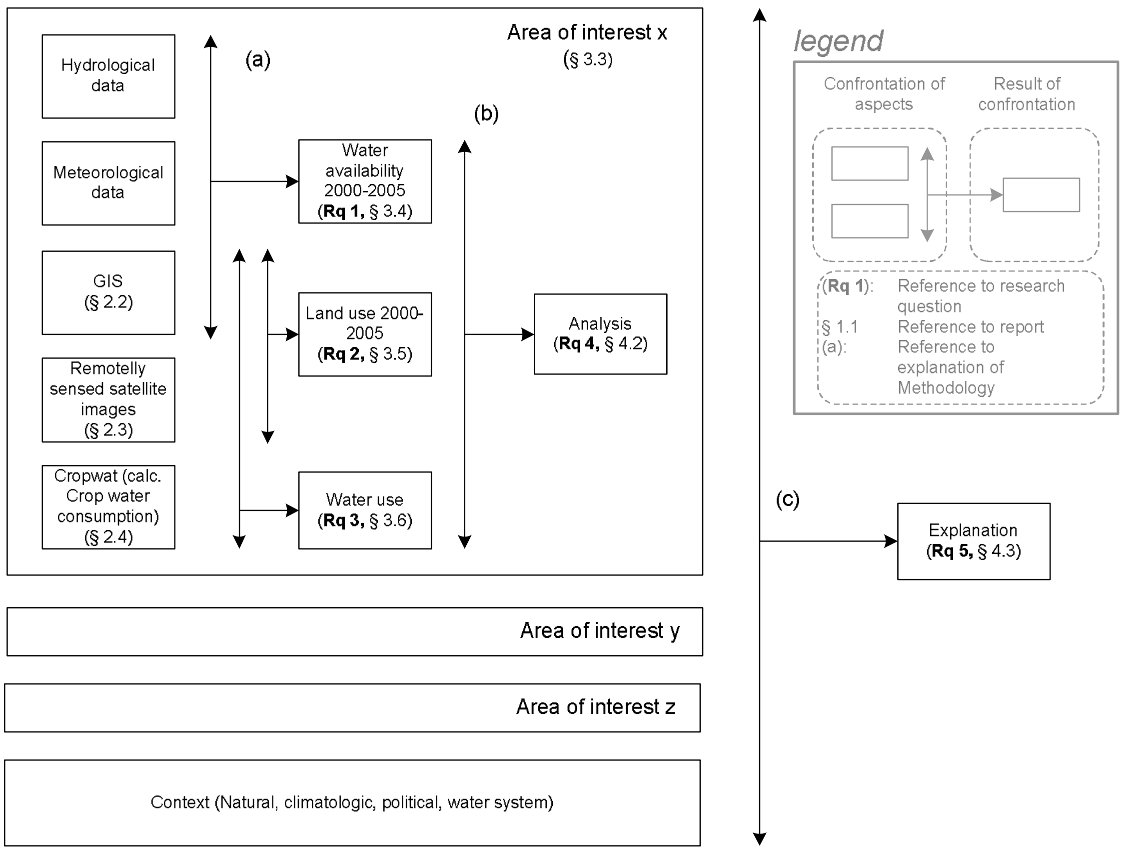


figure 2.1: Methodology

In figure 2.1 the used methodology is shown. The way this methodology is built up will be explained now:

(a) These arrows show how the different ingredients of the research are linked to answer the research questions 1 to 3.

Research question 1 could be answered by carry out a data analysis over the available data (meteorological and hydrological), which is described in paragraph 3.4. To satisfy the spatial aspect of research question 1, the water availability data is mapped in geographically referenced maps. These maps are stored in a Geographical Information System (GIS). Since different geographically determined aspects of this research area (water availability, water use, etc.) have to be combined to draw conclusions, the usage of a GIS is suitable for this research. The backgrounds of GIS's are explained in paragraph 2.2.

The land use in the research area is determined by an analysis of remotely sensed satellite images. A land cover classification is carried out on these images to determine the areas of the different agricultural land uses. Land over classification is a suitable method to answer research question 2, because it is able 'measure' the land cover in a direct way on a large scale and to make a downscaling to focus areas which are relevant for this research. Besides this, land use data of local agricultural agencies are doubtful; political decision about growing a crop in a certain quantity and the real planted area of the concerning crop are not clear distinguishable. The results of the land cover classification and its theoretical backgrounds are respectively described in paragraph 3.5 and 2.3.

For calculating the irrigation water use (research question 3), the results of the land use analysis is combined with crop water requirements calculations. The used calculation software is Cropwat. The results of the irrigation water use calculations and information about Cropwat is given in the paragraphs 3.6 and 2.4.

The described efforts in step (a) will be carried out for different focus areas in the research area to satisfy the spatial aspect of the research. In paragraph 3.3 the borders of these areas, called 'areas of interest', are clarified.

(b) To answer research question 4 ("how are spatial explicit water availability, agricultural land use and irrigation water use related between 2000 and 2005 in different focus areas in the research area?"), the findings of research question 1 to 3 are analyzed per area of interest. This analysis will be carried out by a comparison of the inter-annual trends of water availability, agricultural land use and irrigation water use. This analysis is described in paragraph 4.2.

(c) In the final phase of the research, the findings of research question 4 of the different areas of interest will be compared. It will be tried to "explain the inter annual spatial dynamics of land use and irrigation water use under different water availabilities" (goal of study). This comparison will be carried out by analyzing how the inter-annual trends of water availability, agricultural land use and irrigation water use of area of interest *x* correlates with the water availability, agricultural land use and irrigation water use of area of interest *y*. Also knowledge of the research area, called context in figure 2.1, will be used to give explain these correlations. This analysis is described in paragraph 4.3.

2.2 Geographical Information Systems (GIS)

This paragraph is derived from a course book about GIS (De By, 2004).

A geographical information system (GIS) can be defined as a computerized system that facilitates the phases of data entry, data analysis and data presentation especially in cases when we are dealing with georeferenced data. Data is georeferenced when coordinates from a geographic space have been associated with it. The georeferenced data tells us where the object represented by the data is (or was or will be). The main characteristics of a GIS software package are its analytical functions that provide means for deriving new geoinformation from existing spatial and 'attribute data' (see glossary). In using GIS software we obtain some computer representations of the geographic phenomena which exist in the real world. Continuing with manipulating the data with techniques like 'image classification' (see paragraph 2.3) or overlaying to maps (layers) to retrieve a new map (see figure 2.2), results in additional computer representations. Since the spatial referencing is an important issue in GIS, a further explanation of topics which have to do with spatial referencing (georeferencing and resampling) is given in appendix 1.

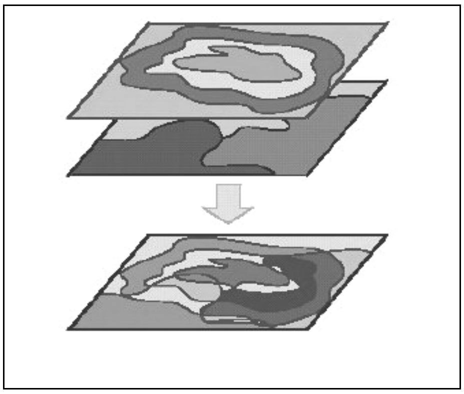


figure 2.2: Example of creating a new visualization. Two different layers can be overlaid to look for spatial correlations

These computer representations, existing of bits and bites, can be used for making visualizations, either on-screen, printed on paper, or otherwise. In this thesis the goal of making computer representations and visualizations of the real world, is to obtain additional data which are required for doing the analysis about the spatial relation between water availability and irrigation water use. The main technique used to obtain the research goal of this theses was 'image classification'. To understand the principles of image classification, the reader of this master thesis has to know some basic concepts about 'Remote sensing'. Backgrounds of remote sensing are explained in appendix 2 and 'image classification' will be explained in the following paragraph.

2.3 Image classification

2.3.1 *Introduction*

Sources for this paragraph are (Kerle et al., 2004), (Lillesand and Kiefer, 1994) and (ITC, 2001).

Image classification is a sub area of Remote sensing. It is based on the different spectral characteristics of different materials on the earth's surface and their reflection to a sensor. The overall objective of image classification procedures is to automatically categorize all pixels in an image into land cover classes or themes. Several image classification methods are available in literature. The most commonly used methods are: 'density slicing', 'unsupervised spectral image classification' and 'supervised spectral image classification'. These methods will be discussed in the next sub-paragraphs. The scheme in figure 2.3 shows the main line of these sub-paragraphs. Sub-paragraph 2.3.6 will discuss some shortcomings of the selected method.

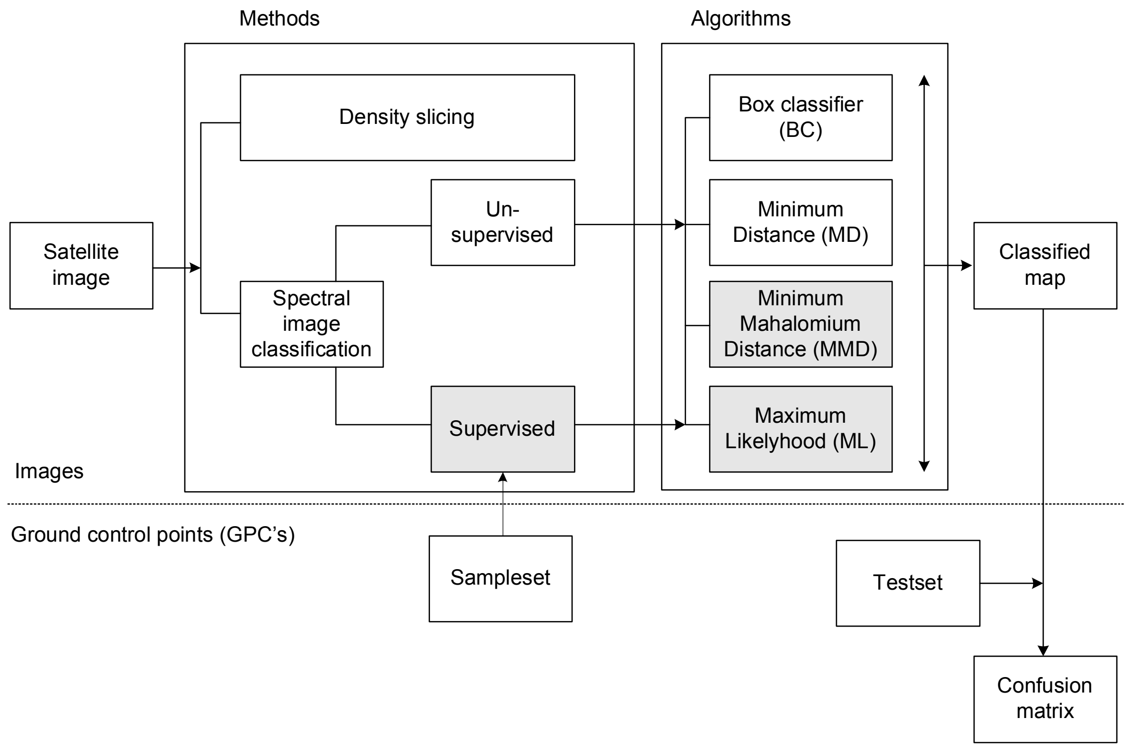


figure 2.3: Image classification methods and algorithms

2.3.2 *Methods*

Density slicing is a technique, whereby the Digital Numbers (DN-values, see appendix 2.7) distributed along the horizontal axis of an image histogram[1], are divided into a series of user specified intervals or slices. Density slicing will only give reasonable results, if the DN-values of the cover classes are not overlapping each other.

In an *unsupervised classification*, clustering algorithms are used to partition the feature space[2] into a number of clusters

In a *supervised classification*, the partitioning of the feature space is realized by an operator who defines the spectral characteristics of the classes by identifying sample areas (training areas). The operator needs to know where to find the classes of interest in the area covered by the image. This information can be derived from 'general area knowledge' or from dedicated field observations.

Important background information about classification methods is given in appendix 3.1.

Comparison of classification methods

Since the histograms of the individual bands and the histograms of the Normalized Defined Vegetation Indices do not show separate land cover classes, *Density slicing* is not an appropriate method to carry out an image classification (see appendix 3.1)

[1] For definition, see Glossary

[2] For an explanation of the meaning of a feature space, see appendix 2, paragraph 2.7. See also Glossary

Since there was no land cover data of the area available the result of an *unsupervised classification* could not be validated. Therefore, this method was scientifically untenable. In combination with validation data the unsupervised classification could be useful. But when a field visit will be made and the expectance is that many sample data will be collected, it is more logical to choose for the *supervised classification*, since the basis of this method is the 'real' field data. Recapitulating: in the supervised approach we define useful information categories and then examine their spectral reparability; in the unsupervised approach we determine spectrally separable classes and then define their informational utility. Since many 'useful information categories' is collected during the field visit, the supervised classification was the most suitable classification method in this case.

2.3.3 Algorithms

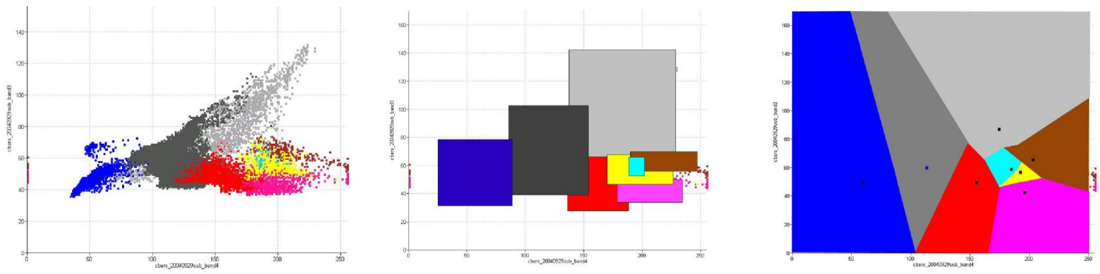
After the training sample sets have been defined, classification of the image can be carried out by applying a classification algorithm. Several classification algorithms exist. The following algorithms will be explained, because they are well-known in literature and the used software package, ILWIS, allows only these methods to use for image classification: Box classifier (BC), Minimum Distance (MD), Minimum Mahalomium Distance (MMD), Maximum Likelihood (ML). See appendix 3, paragraph 3.2 for a detailed description of these algorithms.

To demonstrate and evaluate the results of these four methods, they are applied to the sample sets of the (2004_right). See table 2.1 for the accuracies of bananas, beans, paddy and water. How these accuracies are calculated will be explained later in sub-paragraph 2.3.4. In paragraph 3.5.2, 'classification results', the four methods will be applied to all sample sets and the final result of the land cover classification will be given then.

	BC	MD	MMD	ML
Banana (yellow)	0.38	0.44	0.84	0.81
Beans (brown)	0.40	0.74	0.74	0.75
Paddy (red and pink)	0.65	0.68	0.96	0.96
Water (blue)	0.91	1.00	1.00	1.00
Overall	0.86	0.95	0.99	0.99

table 2.1: Accuracies of algorithm methods for 2004_right

The accuracy of bananas is very low when BC and MD classifiers are used. The density and the distribution of the banana point-cloud is not taken into consideration in these methods. The un-dens cloud of beans (brown) and coco (light blue) are prevailing the banana (yellow), see figure 2.4. The low score of the accuracy of beans using the BC method can be explained by overleap of the beans by coco. In the BC and the MD method, paddy scores low. In figure 2.4 it can be seen that bare land (grey) is overleaping the point-cloud of paddy (red and pink). MMD and ML classifiers have better scores for paddy, because they are bringing the dens and concentrated character of the sample set of paddy into account.
Conclusion: the methods which are using the variance-covariance matrix (MMD and ML), thus the methods that bring variability into account, have the best scores and are therefore the most useful for this research. Which of both methods has to be used for each individual image will be determined by the accuracy assessment of both methods (see sub-paragraph 2.3.4).



Feature space of sample set *Box classification* *Minimum distance classification*

figure 2.4: Comparison classification algorithms for sample set of 2004_right

Threshold distance

Variation in threshold distance, explained in appendix 3, gave no improvement of the accuracy of the ML and MMD classifiers. ML classifications with a threshold distance of 100, 80, 60, 40, 20, 10 are carried out with the sample set of 2004_right and gave all an overall accuracy of 98.64%. This is the same as the overall accuracy in case no threshold distance is specified (see table 3.6) A threshold distance of 5 gave 98.63%, just 0.01% lower (!). Apparently, almost all feature vectors of the image do not fall far away from the mean values of the clusters. Therefore, the rest of the images are classified with a non-specified threshold distance.

2.3.4 *Accuracy assessment of the classification result*

> "A classification is not complete until its accuracy is assessed"
> (Lillesand and Kiefer, 1994).

One of the most common means of expressing classification accuracy is the preparation of a classification confusion matrix (also called an error matrix or a contingency table). Confusion matrices compare, on a category-by-category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. Therefore, the total ground truth data is split up in a 'sample set' and a 'test set'. The sample set is used as training for the classifier (the left figure of figure 2.4 is an example of the feature space of a sample set), the test set is used as the reference data in a confusion matrix. An *example* of a confusion matrix is shown in table 2.2. This example is shown to evaluate the properties and the results of a confusion matrix, not to discuss the results it is showing. The final confusion matrices of all classification will be evaluated later in sub-paragraph 3.5.2

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	**295**	0	25	8	0	0	1	0.9
Beans	0	**0**	0	0	0	0	0	?
Coco	0	0	**0**	0	0	0	0	?
Grass	0	0	0	**0**	0	0	0	?
Paddy	37	0	8	0	**879**	0	2	0.95
Water	0	0	0	0	0	**3193**	64	0.98
Reliability	0.89	?	0	0	1	1		

Average Accuracy	94.33
Average Reliability	96.33
Overall Accuracy	96.79
KHAT-value	92.76

table 2.2: Example of a confusion matrix (2003_right)

A confusion matrix gives a value for the accuracy and the reliability for each class. The accuracy is defined as: the fraction of correctly classified test set pixels of a certain ground truth class. For each class of test set pixels, the number of correctly classified pixels is divided by the total number of test pixels in that class. The reliability is defined as the fraction of correctly classified test set pixels of a certain class in the classified image. For each class in the classified image (column), the number of correctly classified pixels is divided by the total number of pixels which were classified as this class. Also statistics to assess the accuracy and reliability of the whole classification are available: the average accuracy, the average reliability, the overall reliability and the KHAT-value. In appendix 3, paragraph 3.3 a more detailed explanation is given about the main concepts of a confusion matrix.

2.3.5 Sample set and test set

A part of the total retrieved ground truth data about land covers, in the case of this thesis agricultural field with a certain crop, has to form the test set. The other part will be the sample set. The first consideration is: "when can an agricultural field be part of the ground truth data?" The second question is: "how many sample areas are required?" The last point to discuss in this context is: "which part of the total amount of test areas should form the test set?

Ground truth data

Ground truth data in the case of this thesis are the DN-values of a group of pixels which form a homogenous agricultural field. The type, or class, of this field is retrieved in the research area from conversations with farmers. In short interviews, the class of land cover is determined in the research period (August to December in 2000 until 2005). With the use of a GPS-logger the exact positions of the fields are recorded. In the training stage (the assignment of pixels to the known land cover class), the homogeneous pixels around a GPS-point are assigned to the known land cover class from the interviews. To judge which pixels belong to a homogeneous field, the feature spaces, which are automatically be generated in the used software package ILWIS, are used. The points of one field should form a concentrated point-cloud in the feature spaces. How concentrated this should be is a part of the subjective judgment and skills of the operator.

Amount of sample areas
Different guidelines for the size of a sample set are given in literature.

- A minimum of 50 sample areas per class in the case of a few different classes and 75 to 100 per class in the case of 12 classes or more (Lillesand and Kiefer, 1994).
- A minimum of 30 sample areas per class (Kerle et al., 2004).

The conditions in each research area are different, so these quantities are just guidelines. Besides this, not any class has the same variability. 'Water' for example, requires less sample areas than banana, which can vary in grow stage.

Division sample set – test set
The ground truth data set has to be separated in a sample set and a test set. Which fields should be part of the test set and how many? Systematically sampling or random sampling are the two main options. In systematically sampling, only the field with a high level of certainty are selected for the test set. Random sampling selects random field for the test set. Because the amount of the ground truth data in this thesis is limited, the systematically way of sampling is applied. The relation sample set : test set in these thesis is 1:4. This division key is commonly used.

2.3.6 *Shortcoming of image classification*

Processing of the training set results in spectral classes. Sometimes these spectral classes are not equal to the real classes in consideration. For example: banana fields do not give the same spectral reflectance, due to differences in grow stage, banana type and ground type.
Another main problem is the existence of 'mixels'. Each pixel in a classified image can only represent one land cover class. In reality, more than one land cover can occur within one pixel. The reflectances of these different cover classes are mixed to one DN-value per band. The result of this phenomena is called a mixel. For example: when the pixel covers the border of an agricultural field and the road besides. In the case of this thesis, this is an important point of consideration, because the pixel sizes of the used images are 20 m and 30 m.

2.4 *Irrigation water use*

2.4.1 *Introduction*

Since no data are available about quantities of irrigation water used for growing crops in the research period, the irrigation water use has to be derived from other parameters. This paragraph will explain the used calculation methods to calculate the irrigation water use of crops. Firstly the physical phenomenon of transpiration and soil moisture will be explained. Thereafter the chosen method for calculating the irrigation water use of a crop will be clarified.

2.4.2 *Soil-plant-water relationships*

This sub-paragraph is derived from Bailey (1990).
One of the raw materials required by green plants to build the substances they need for growth and energy is water. Water carrying essential nutrients in solution passes into the roots, moves upward through the plant depositing the nutrients where they are required, evaporates within the leaf and passes though very tiny holes in the covering of the leaves. The water lost from the leaves is then replaced by more water via the roots. This continuous movement of water from the soil to the atmosphere is called the 'transpiration stream'. The rate at which this occurs is known

as the 'actual transpiration rate'. The driving force which determines this is the weather, but it is also regulated by the type of crop and the availability of moisture in the soil. It is increased by bright sunshine, high temperatures or strong winds.
The basic model, used in irrigation, is that the soil under constant conditions contains a constant amount of water; called the 'field capacity'. The force holding the water is called 'surface tension'. As a crop removes water from the soil, the soil is described as having a certain soil moisture deficit (SMD). The SMD at field capacity is zero. If a crop extracts 50 mm of water from a soil, and there has been no rainfall or irrigation to replenish it, the soil is said to have an SMD of 50 mm. When there is no restriction on soil moisture, and the rate for a given crop is determined solely by the weather, it is said to be proceeding at its full potential which is referred to as the potential transpiration rate. When the SMD increases, the actual transpiration of a crop may fall below the potential transpiration because the crop reacts on the lack of soil moisture by closing its tiny transpiration holes. The SMD at which this occurs is called the 'critical SMD'. In most situations the objective of irrigation is to keep the SMD below the critical level.

2.4.3 Calculating the irrigation water use by using Cropwat

Cropwat (FAO, 1992) is a computer program for irrigation planning and management. It is a commonly used method to give estimations about crop water requirement without requiring many measurements (many standard values of, for example, climate conditions are available as input for Cropwat). Since there is a little knowledge about each particular agricultural field in the research area, Cropwat is a suitable method for calculating the crop water requirements in the research area. It is not the aim of this study to evaluate the working of Cropwat. Therefore, the calculations methods of Cropwat will be briefly discussed. This brief description is mainly taken from the 'water report nr 22' of FAO (FAO, 2000)

Cropwat calculates the transpiration (T) of a crop and the evaporation (E) of the soil of an agricultural field, together called the evapotranspiration (ET). The calculation of a reference evapotranspiration (ET_o) is based on the FAO Penman-Monteith method (FAO, 1998). Input data include monthly and ten-daily for temperature (maximum and minimum), humidity, sunshine, and wind-speed. Crop water requirements (ET_{crop}) over the growing season are determined from ET_o and estimates of crop evaporation rates, expressed as crop coefficients (K_c), based on well-established procedures (FAO, 1977), according to the following equation:

$$ET_{crop} = K_c \times ET_0$$

Through estimates of effective rainfall (taking into account that a part of the rain runs off), crop irrigation requirements are calculated assuming optimal water supply. Inputs on the cropping pattern will allow estimates of scheme irrigation requirements. With inputs on soil water retention and infiltration characteristics and estimates of rooting depth, a daily soil water balance is calculated, predicting water content in the rooted soil by means of a water conservation equation, which takes into account the incoming and outgoing flow of water.
Stress conditions in the root zone are defined by the SMD. The SMD varies for different crops and different crop stages and is determined by the rooting density characteristics of the crop, evaporation rate and by the soil type.

It is assumed that the crops in the research area are irrigated in an efficient way, that is: refilling the soil water content to field capacity when the critical SMD is reached. By doing this assumption, Cropwat can be used to calculate the irrigation water use of the crops in the research area.
A further explanation and report of the input data and the output data of Cropwat is given in appendix 8.

3 Data processing

3.1 *Introduction*

This chapter deals with the processing of rough data to a format which is suitable for answering the research questions 1 to 3 about the evolving of water availability, agricultural land use and irrigation water use. Firstly the period of interest and the area of interest will be further specified to focus the efforts of translating the available data to the goal of study. Thereafter, the way the water availability, the land use and the irrigation water use is quantified will be explained in respectively paragraph 3.3, 3.4 and 3.5. These paragraphs will be the ingredients for the next chapter: Results.

3.2 *Period of interest*

Three factors are important to consider to make a choice for a suitable period of interest.
Firstly, it has to be a clear-cut period in which the phenomena of the problem this thesis is dealing with -explaining the inter-annual dynamics of irrigation water use and land use under different water availabilities- are clearly come up. From the field work it is known that the research area knows two of these clear-cut periods: the dry season and the wet season. In the dry season, the scarcity of water is larger than in the wet season. It is expected that in this period the expected relations between the behavior of land use, irrigation water use and water availability become more clear distinguishable than in the wet season.
Secondly the grow season is important, since the crops are the actual irrigation water users. The period of interest should start when the earliest crop in the dry season is planted and it should stop when the last crop in the dry season is harvested.
Finally, the availability of data is an important point of concern. Since this research uses a land cover classification, cloudless satellite images are required. Only in the dry season cloudless images were available.

Resuming the issues pointed above, the period of interest starts when the first crop in the dry season is planted and ends when the last crop is harvested. This is approximately from June up to December.

3.3 *Areas of interest*

Since the goal of this study is to analyze the dynamics of land use and irrigation water use under different water availabilities, these water availabilities have been categorized. It has been tried to select areas that are dependent on one certain type of water availability. In this way, the dynamics of land uses and irrigation water uses connected to a particular type of water availability can be compared and analyzed. The following types of water availability can be identified: locally stored runoff, river water, reservoir water and rainwater.

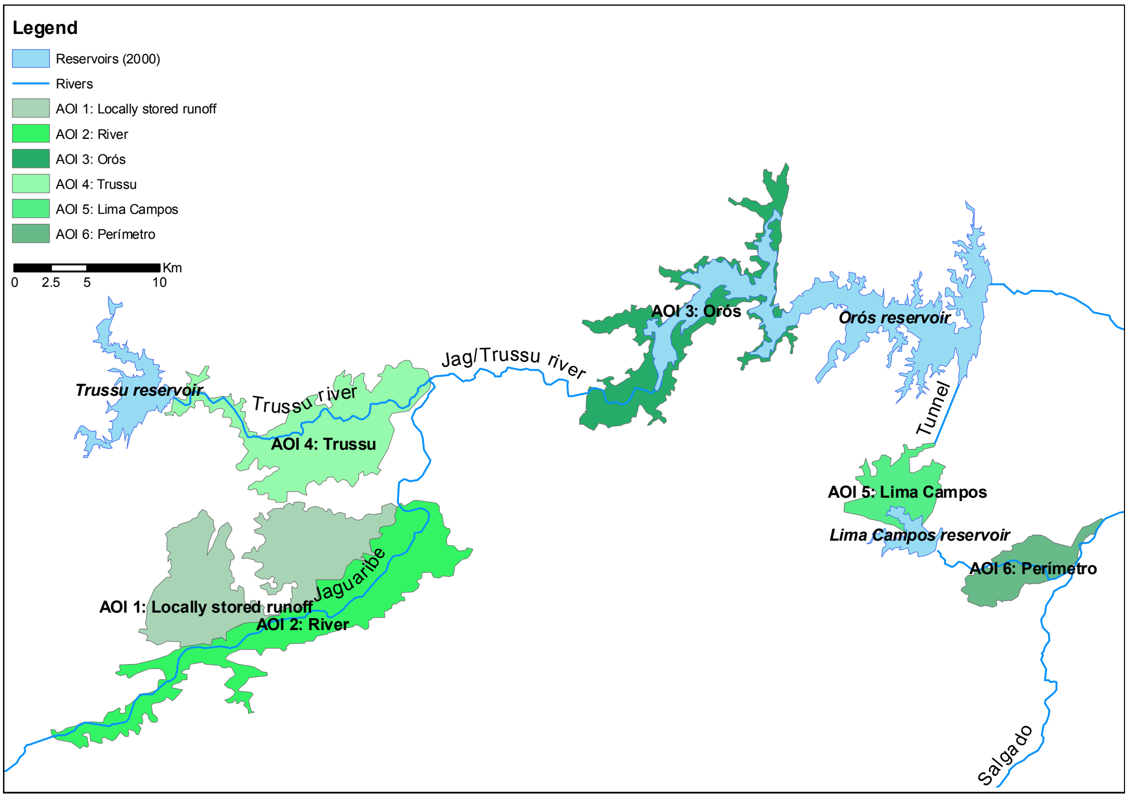


figure 3.1: AOI's

3.3.1 *AOI 1: Locally stored runoff*

It is mentioned in sub-paragraph 1.2.3 that studies about runoff reached the conclusion that, in the crystalline areas of the Semi-Arid zone of Brazil, an average of only 6 to 8% of the rainfall actually flows superficially or feeds aquifers by percolation. During the field visit, several plots irrigated by this kind of stored runoff have been visited (see figure 3.2).

These plots were concentrated in the area between the Trussu reservoir and the Jaguaribe river. A 'relative elevation'-map, as drawin in figure 3.3, gives more insight in the question why these area uses locally stored runoff. 'Relative elevation' is defined as the elevation of a location minus the elevation of nearest point of the river. How the relative elevation is determined is explained in appendix 5.

figure 3.2: Groundwater from wells

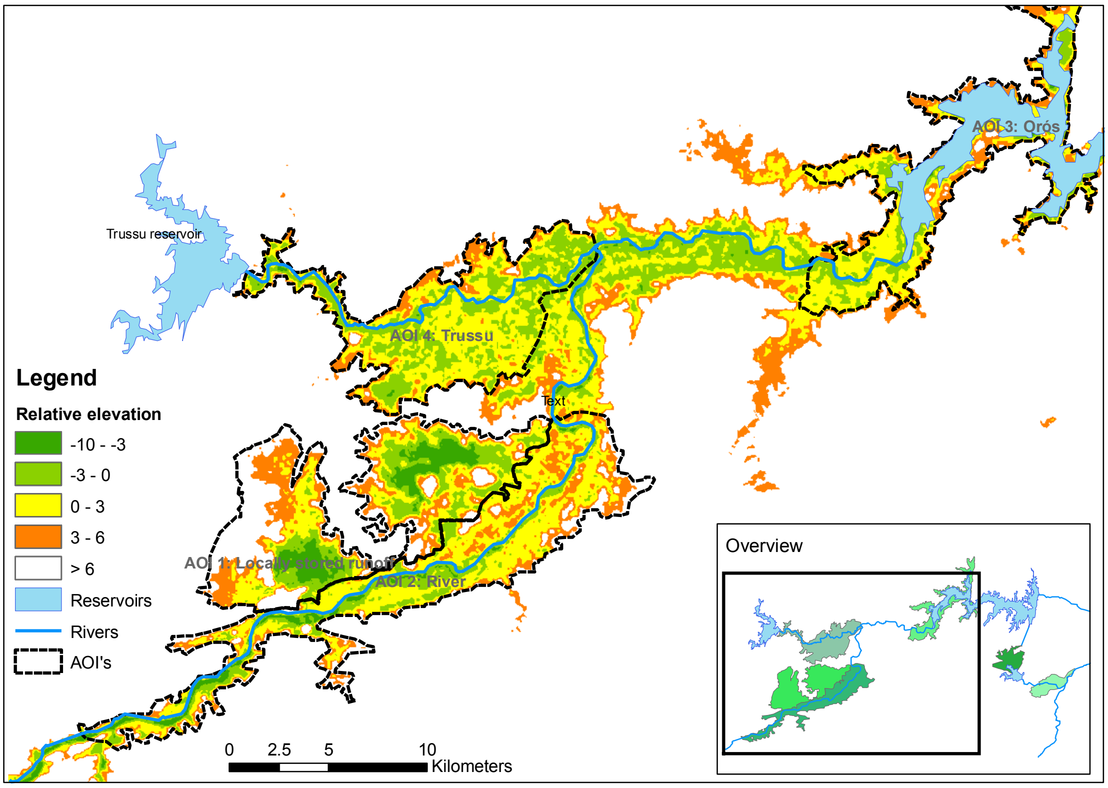


figure 3.3: Relative elevation perpendicular to river

It can be read in figure 3.3 that the relative elevation in the northern cross-direction of the Jaguaribe river first increases and subsequently decreases in the form of two 'bowls'. So two local depressions can be seen in which the relative elevation to the river decreases to -6 m.
Another property of this area is that the upper part of the soil has mainly an alluvial soil or a vertisoil (see the graph of AOI 1 figure 3.4 'soil types per AOI'). These soil types are called 'heavy', which means that it contains fine particles like clay and has a good water holding capacity.
The third point of attention is that water wells and natural lakes were observed in this area. The lakes originate in the center of these local depressions. Two reasons can be given for this phenomenon: either the surface intersects the groundwater table or the rainwater of the wet season is retained by the soil. Since lack of data about groundwater tables and the existence of aquifers, this will not be investigated further. Together with the knowledge that the distance to the Trussu river and the Jaguraribe river is to large to get water from there, it is assumed that this area gets water from locally stored runoff.
AOI 1 contains both local depressions. The borders of these bowls are determined by the rule: *relative elevation to the river = 6 m*.

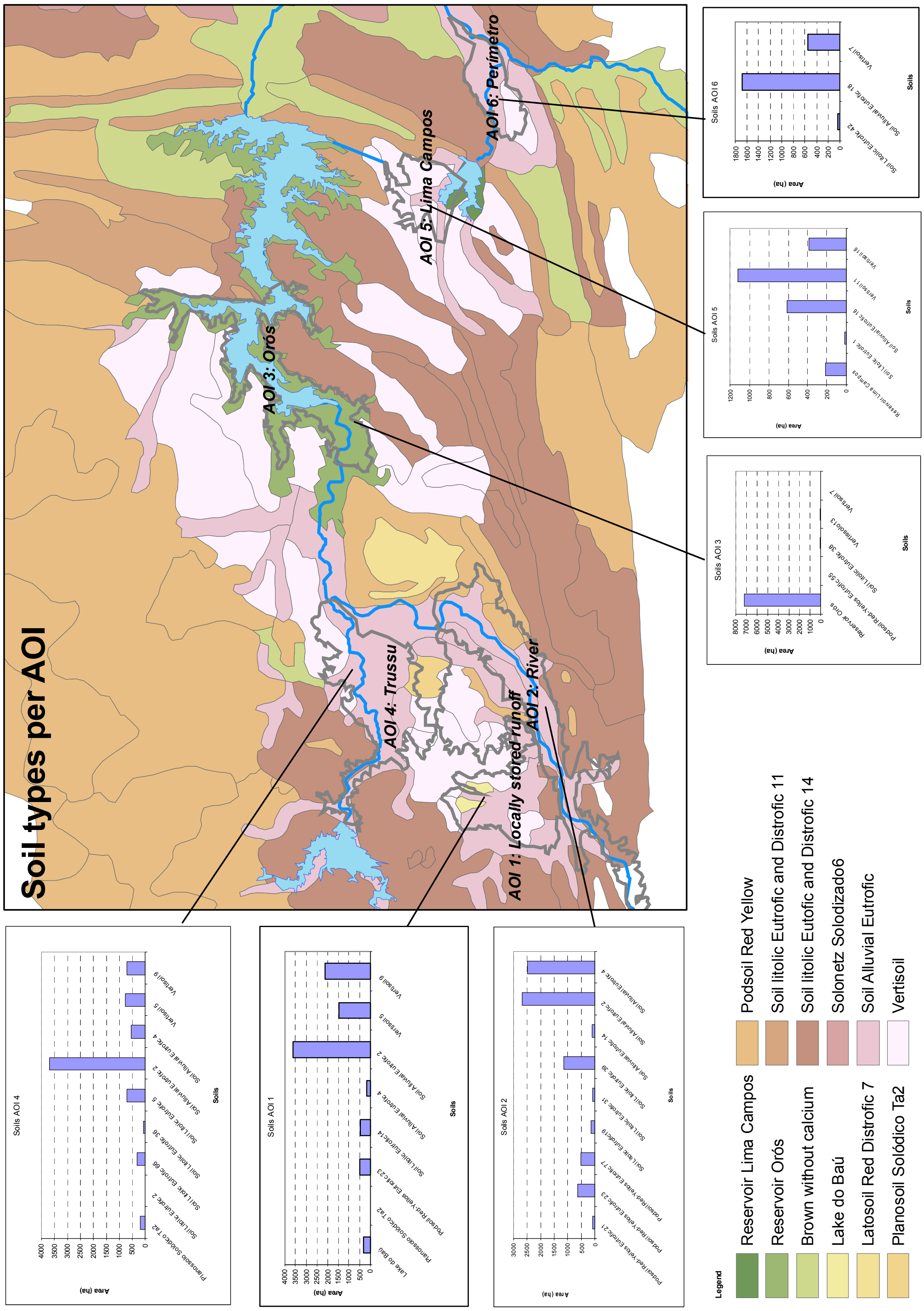


figure 3.4: Soil types per AOI

3.3.2 *AOI 2: Water from river (Jaguaribe)*

The map in figure 3.3 can also be used to determine the areas that have access to water from the river. During the field visit it became clear that the water from the river is pumped to the areas around the river. From information retrieved from local farmers the pumps used can pump the water up to around 6 meters, the area of AOI 2 is determined by this rule: *relative elevation to the river = 6 m.*

3.3.3 *AOI 3: Edge Orós reservoir*

Similar to AOI 2, farmers at the edge of the Orós reservoir also pump up the water to their agricultural fields. To obtain a suitable area in which the agricultural fields are supplied by only reservoir water, a same kind of decision rule to determine the border as in AOI 2 can be used. A problem comes up when a fixed relative elevation rule is applied: at the location where the Jaguaribe/Trussu river enters the Orós reservoir, the borders drawn by the fixed relative elevation rule are too far away from the reservoir to make it reasonably possible to pump the water from the Orós reservoir to these areas (see the box marked with a question mark in figure 3.5. How this land use map is made will be explained in 3.5).
This figure shows also that in other areas on the reservoir edge, the rule 'relative elevation = 6 m covers the suspected areas supplied by reservoir water.

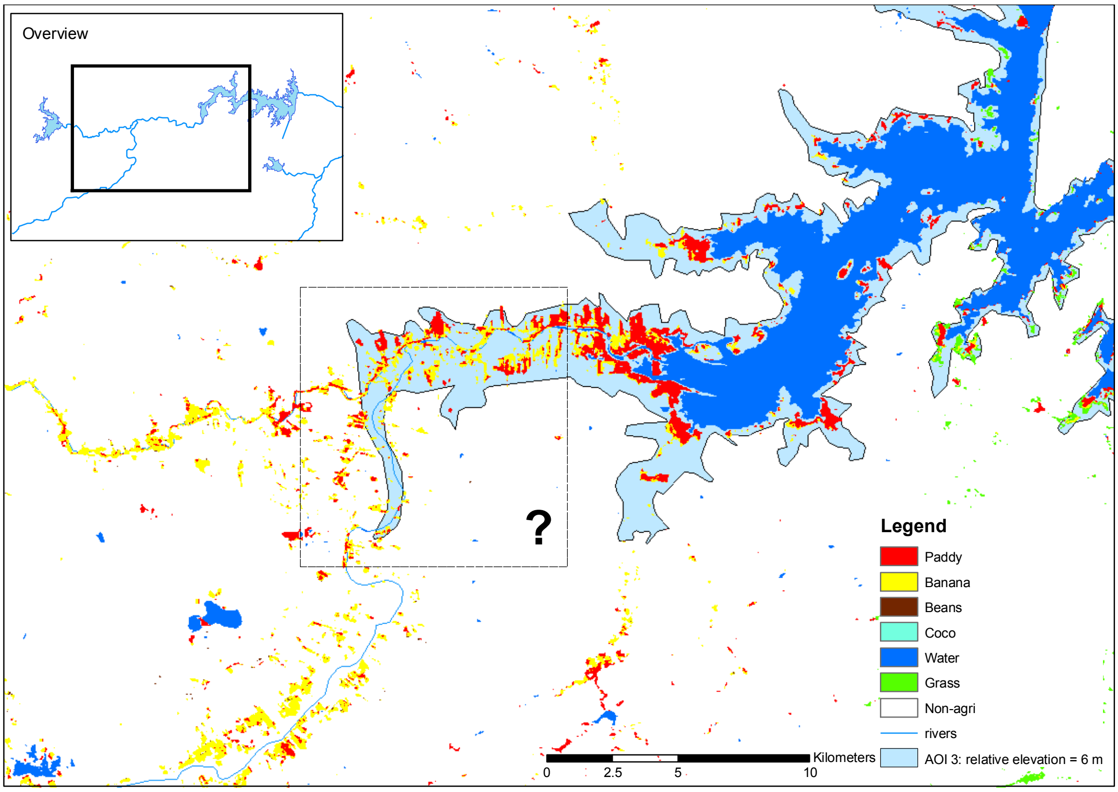


figure 3.5: Border AOI 3 in 2005 due to relative elevation from the reservoir level of 6 m

Since the decision rule 'relative elevation from reservoir level = 6 m' is ineffective at the area marked by a question mark, an extra rule has to determine of AOI 3 in this area.

With the available information (land use data, DEM, discharge data, reservoir data) it is hard to determine where the exact border between fields that are supplied by water from the river and water from the Orós reservoir has to be drawn. So an estimation has to be made. The following points are considered in this estimation:

1. The satellite images show agricultural activity in the surrounding of the river. Other types of water availability, besides the reservoir water, can therefore not be neglected in the dry season.
2. In a land use classification map as shown in figure 3.6 , it can be seen that water in the Jaguaribe/Trussu river during the dry season is dammed by little dams in the river bed. Upstream from these little dams, small reservoirs in the river bed emerge. It is assumed that agricultural fields around such small reservoirs are supplied by water from these small reservoirs and *not* by water form the Orós reservoir.
3. When the small reservoirs are close to the Orós reservoir and close to each other it is assumed that the water from the Orós reservoir is pumped up from reservoir to reservoir. This is a hypothesis, not an observation from the field visit.
4. When the shape of the fields are clearly rectangular and perpendicular to the river, it is assumed that they are supplied by water stored in the river bed.

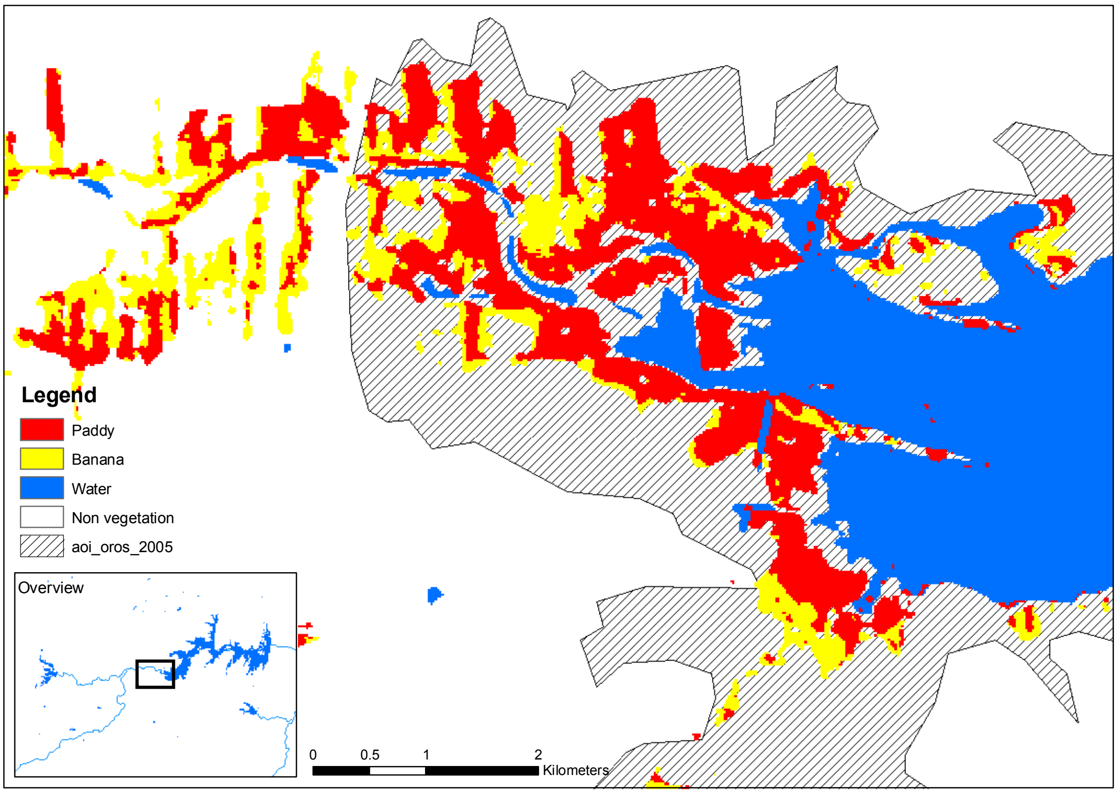


figure 3.6: Boundary of AOI 3 in 2005 at inlet point

In figure 3.6 the result of this estimation is shown for the year 2005. Since the reservoir level is changing every year, the boundaries of AOI 3 are also different in each year. In figure 3.7 the borders for each year are drawn, with the border of 2004 as reference. Also the areas of AOI 3 (included the water) in each year are shown in a table.

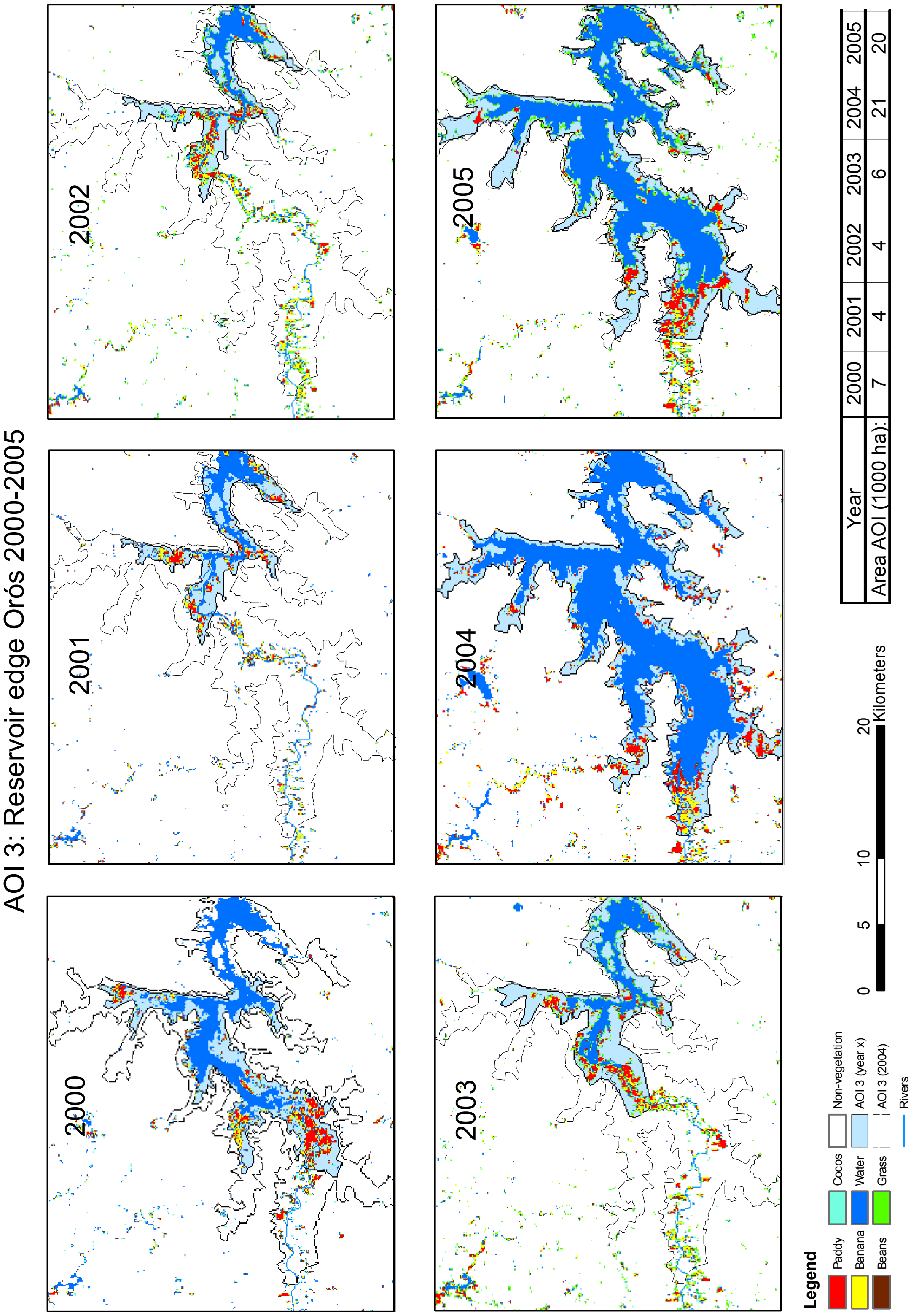


figure 3.7: AOI 3 in each year

Points of discussion:

- It is questionable how far the pumps can transport the water in horizontal way. In the drawn borders, this distance is around 2,5 km. Extra information about this distance should make it possible to draw a more certain border.
  Since the water level of the reservoir is changing during the dry season (it is decreasing, see figure 3.11), also the area of interest is changing during the dry season according to the fixed elevation difference determined by the pumps. Since the goal is to determine a fixed AOI per dry season, a reference water level had to be chosen. The acquisition date of the satellite images is chosen as date for the reference water level, since the land use maps, which are derived from these satellite images with a certain acquisition date, are also used to determine the borders (as described above). A fixed elevation difference form the water level (which is decreasing in the dry season) makes the water level at the end of the dry season determinative for the areas than can be supplied by water during the whole season. Since the acquisition dates of the satellite images are at the end of the dry season (21 October – 13 November) they are useful for this goal. In figure 3.7 it becomes clear that the fields around the reservoir edge are included in AOI 3 by using these reference dates, which makes these dates suitable for determining the borders of AOI 3.
- The water level at the acquisition dates could be determined in two ways: crossing the DEM with the satellite image and reading the elevation of the pixels at the water edge *or* converting the available daily volume data (COGERH, 2005) with a given 'volume-water level'-relationship. In each year the difference between both methods was around 6 m (see table 3.2). Probably the reference of one of both data sources is not correct.
- The difficulty to determine the border at the inlet of the Orós reservoir (area with question marked in
- figure 3.5) gives this area of interest the most uncertainty. In chapter 5, Discussion, this border will be varied to see the impact of this uncertainty on the conclusions of this research.

Year	Water level in DEM (m)	Water level in COGERH data (m)	Difference DEM – COGERH (m)
2000	196	191	5
2001	192	185	7
2002	192	185	7
2003	195	188	7
2004	204	198	6
2005	203	197	6
			Mean ≈ 6

table 3.1: Water level data of DEM and COGERH (2006)

3.3.4 AOI 4: Downstream of Trussu reservoir

The water from the Trussu reservoir is released for use of irrigation downstream from the Trussu dam (COGERH Iguatú, 2006). AOI 4 covers an area with agricultural fields that have mainly access to this water source by taking it from the Trussu river. Again it is assumed that a relative elevation, determined by the capacity of the pumps, is determinative for the boundary of the AOI since no canals have been seen during the field visit. The riverbed is also lower than its surrounding area, so no canals, perpendicular to the Trussu river, can be expected. The capacity of the pumps used in the area is estimated on 6 m (see former paragraphs).

The area from the joining point of the Jaguaribe river and the Trussu river to AOI 3 can be supplied by three sources: discharges from the Jaguaribe, releases from the Trussu river and water from the Orós reservoir (due to the uncertain border of AOI 3). Therefore, this area was not suitable for joining AOI 4 or to form an extra AOI (see also 3.3.7).

3.3.5 AOI 5: Edge Lima Campos reservoir

The Lima Campos reservoir is supplied by a tunnel/canal from the Orós reservoir. From the outlet of the tunnel, two main canals are lead to the Lima Campos reservoir. The agricultural fields north of the Lima Campos reservoir are supplied by water from these canals. The agricultural field west of the Lima Campos reservoir are not supplied by these canals. Since the AOI's are assumed to be areas with one main water source, the areas which are supplied by the canals are selected. The boundaries of AOI 5 are determined by the location of the outlet of the tunnel (recorded by GPS) in combination with the elevation map and observations during the field visit (see figure 5.2 in Appendix 5).

3.3.6 AOI 6: Downstream Lima Campos reservoir, irrigation scheme

The boundaries of AOI 6 are determined by the irrigation scheme which contains fixed locations of fields. The fixed network of cannels gets its water from the release from the Lima Campos reservoir by operations of the dam.

3.3.7 Remaining points of consideration

With the selection of the AOI's, the main part of the irrigated area is joined in the analysis. Except the area between AOI 3: Orós and AOI 4: Trussu. The reason for this is the uncertainty about which water source is used: the discharge of the Jaguaribe river, or the release from the Trussu reservoir. Since the goal of this study is to explain the inter-annual dynamics of land use and irrigation water use under different water availabilities, a complete fitting water balance with all irrigated areas is not required. As long as each field in a certain AOI gets its water from the same water source and all water sources are joining the analysis the analysis can be carried out.
The motive for choosing the goal of this research was the 'unbalanced relationship between irrigation water use and water availability' and the problem of distributing the available water along the different irrigation water users as described in paragraph 1.4. It is assumed that the actors in these problems, the farmers, can be categorized in the same way as the AOI's are categorized.

3.4 *Water and land availability data*

3.4.1 Precipitation

Source of data

The rainfall data is obtained from FUNCEME (FUNCEME, 2006). A selection had to be made since some rain stations gave empty-values for certain periods. Per year is examined which rain stations were operative; the operative rain stations were used. In figure 3.8 the operative rainfall stations of 2000 are shown.

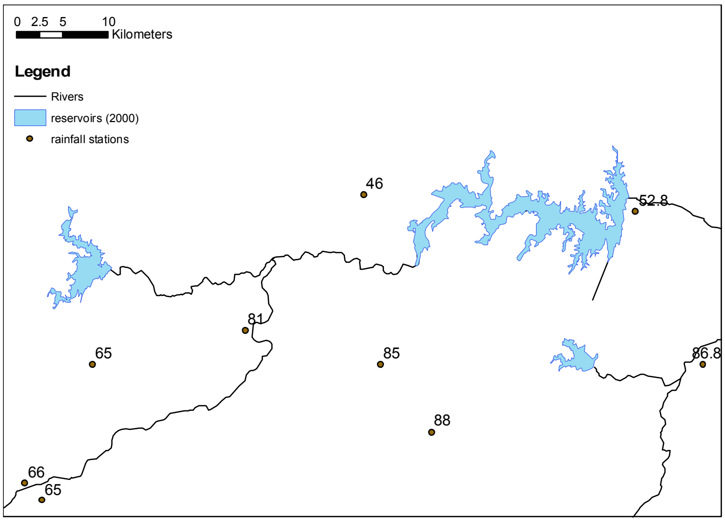


figure 3.8: Operative rainfall stations in 2000 (values are from December, in mm)

Conversion to mean monthly data per AOI

The daily rainfall data is converted to monthly rainfall data, since this scale fits better to the scale of the research and Cropwat, the software that is used for determining the irrigation water use, is using the same scale for the input of rainfall data. The monthly point data is interpolated to give each pixel in the research area a rainfall value. Several interpolations methods were available in the used software (Budde et al., 2005):

- Thiessen (each pixels is attributed the value of the closest rainfall stations)
- Moving average (each pixel-value is determined by the weighted average of the surrounding rainfall stations)
- Surface (One polynomial surface is calculated by a least squares fit through the values of all rainfalls stations. A value, related to this surface, is assigned to each pixel in the map)
- Kriging (similar to moving average but the weight factor is not determined by the distance to a rainfall stations, but by a user-specified semi-variogram)

Since the interpolation had to be done 72 times, a simple method was preferable. The Thiessen method and the Moving average method can be marked as simple method, since a few input variables are required. To decide which of both method to choose, the rainfall stations of a representative month are divided in a sample set and a test set. Both methods are carried out on the sample set and the results are compared with the test set. The representative month was December 2001. The year 2001 had the most operative rainfall stations (19) in the period 2000-2005, so a division in a test set and a sample set left enough points for the sample set to carry out the interpolations. Since the test set counts only 4 rainfall stations, the ratio of error can be remarkably vary by different test sets. Therefore, this calculations is carried out in 3 with different test sets. The results are shown in table 3.2. Since the Moving Average method scores better over these three test sets, this method is used for all interpolations.
The 'ratio of error' is calculated by:

$$ratio = \frac{\sum_{n=1}^{4}\left(\frac{S_n - T_n}{T_n} \times 100\%\right)}{4}$$

With:
n = rainfall stations of test set
S_n = Result of interpolation of sample set at location n
T_n = Measured rainfall at location n

Test set	Ratio of error Nearest Point (%)	ratio of error Moving Average (%)
1	24	18
2	35	36
3	30	11

table 3.2: results of error of interpolation methods rainfall

Figure 6.1 in appendix 6 shows, as example, the result of the interpolation of the monthly rainfall data of December 2000 using the Moving Average method, figure 6.2 of appendix 6 illustrates the mean rainfall per AOI. The graph of figure 3.9 shows the monthly rainfall per AOI. The results show in some years a high variation between different AOI's: the wet seasons of 2003 and 2004 show in some months a variability around 150 mm. This result confirms the temporal and spatial variation in the area, that as concluded in sub-paragraph 1.2.1.

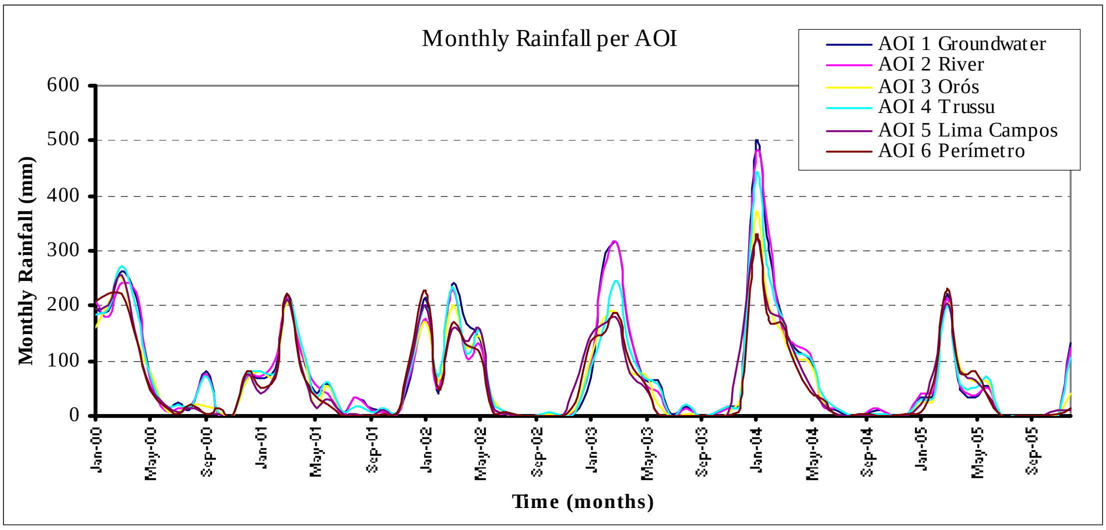


figure 3.9: Monthly rainfall per AOI

Reliability of rainfall data

The reliability of the rainfall data is hard to determine since no rainfall meter is looked at during the field visit and no information about the reliability of the data could be found. General used uncertainties in estimations of rainfall uncertainty are in the range of < 10% (Ungersböck et al., 2001). The accuracy analysis, from which the results are shown in table 3.2, are showing an accuracy of between 11% and 36%. This indicated the performance of the interpolation method, but not the accuracy of the rainfall data itself. Due to the many rainfall stations and the long period of measuring of COGERH, the reliability of the data is assumed to be < 10%.

3.4.2 *River discharge*

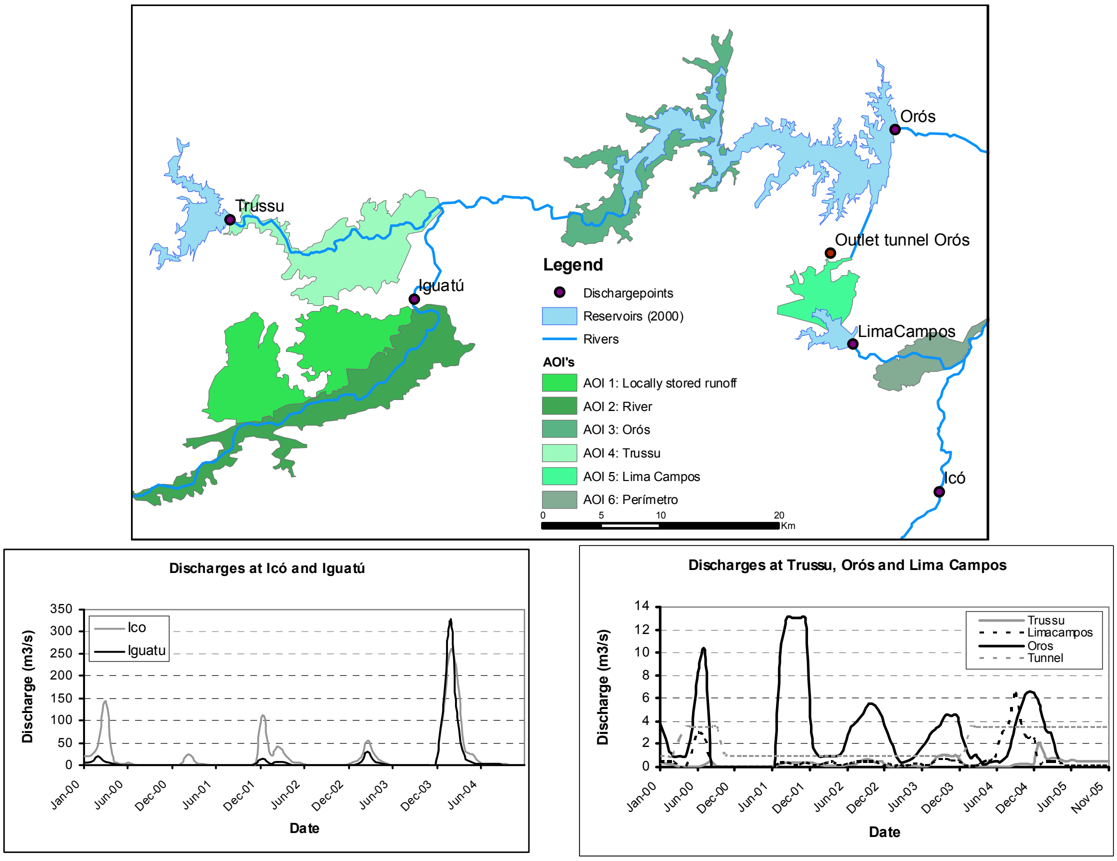


figure 3.10: Discharges

Source of data

Discharge data at the dams of Orós, Lima Campos and Trussu and discharge data of the tunnel between the Orós reservoir and the Lima Campos reservoir were recorded by COGERH (COGERH, 2006), the organization responsible for the management of the water recourses in the area. Discharge data at Iguatú, along the Jaguaribe river, and Icó, along the Salgado river, were available by CPRM, (CPRM, 2006), which is the geographical service in Brazil. The data of these 6 measurement stations were recorded with a temporal resolution of 1 month. In 2004, daily data was available at the dams.

Discharge data in the Jaguaribe and Salgado rivers of 2005 was not available at the moment of writing this thesis.

Reliability of river discharge data

The quantity of released water of the reservoirs are the result of a political decision. It is possible that this data shows only these decisions to release a certain amount of water and not the real amount of released water. Since there is no other data about reservoir releases available, this data will be used. The influence of uncertainties in this data on the conclusions will be discussed later.

The river discharge data at Icó and Iguatú are the result of a measurement with a stage-discharge relationship; the stage is constantly measured and the discharge is calculated using the stage-discharge relationship. This relationship is among other things dependent of the geometry in the cross-direction of the river. Since the riverbed of river is constantly changing, especially in a river that has a high variability in discharge like the Jaguaribe, this relationship has to be constantly updated (Shaw, 1994). The stage-discharge relationships at Icó and Iguatú are expected to be *regularly* updated. The frequency and the method of updating the stage-discharge relationship are unknown. Since this is the only discharge data available, it will be used for the analysis. The influence of possible uncertainties on the conclusions will be discussed later.

3.4.3 Reservoirs

Source of data

The graph in figure 3.11 shows the relative volumes of each reservoir in the research area. The data is obtained from COGERH (COGERH, 2006). Between 2001 and 2003, the reservoirs were relatively empty; in 2004 the reservoirs were full.

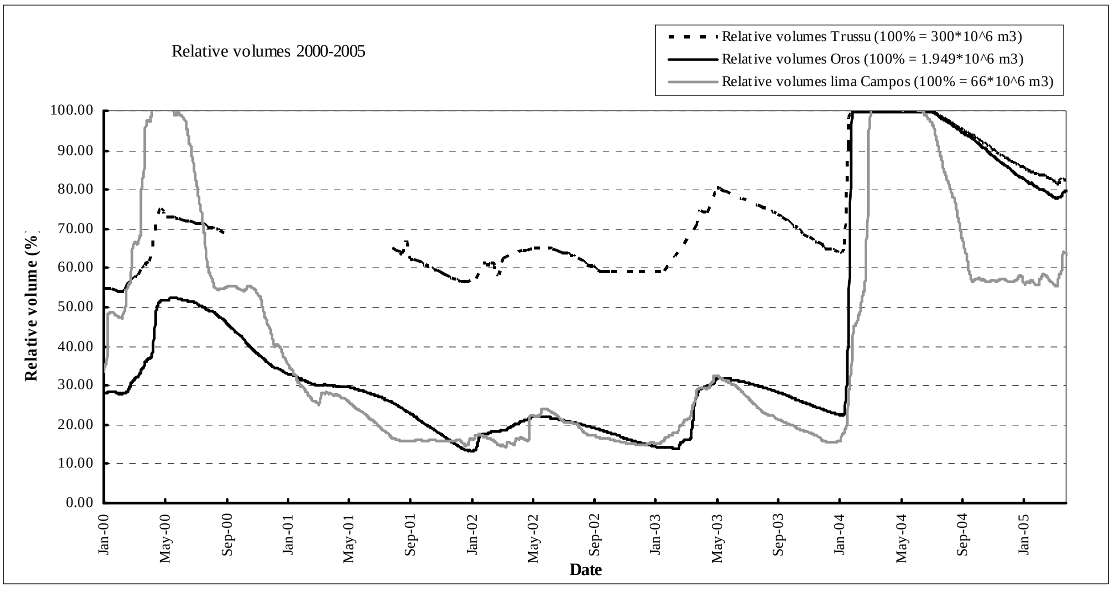


figure 3.11: Relative reservoir volumes (Cogerh, 2006)

As mentioned before, the availability of reservoir water is limited to a certain range around the reservoir, determined by the relative elevation from the water level.

According to information from farmers during the field visit the elevation difference which a pump can deal is 6 m. The land that is located in the area between the water level and 6 m above the water level is called 'the available land'

Land availability around the Orós reservoir
Since the range of the pumps is assumed to be 6 m, the availability per water level can be calculated using the following rule:

$$AL = \sum_{n=1}^{6} A_{h+n}$$

With:
AL = Available land
h = water level of reservoir
A_{h+n} = Area of land lying n meter higher than h

The river bed of the Jaguaribe/Trussu is lower than the surrounding area; by only applying the '6 m relative elevation'- rule, to much of this riverbed is counted under the available land.
This is shown in figure 3.12 in sub-figure 'water level = 203': the available land due to the 6 m relative elevation rule reaches to Iguatú (past the curve in the Jaguaribe). This is implausible since the distances from the water edge to this area is around 9 km, which is likely to be too long for horizontally pumping. This has to do with the same problem as discussed in the definition of AOI 3, sub-paragraph 3.3.3. In that situation the goal was to define an area in which all agricultural fields supplied by reservoir water are joined. In this case the goal is to determine the amount of available land around the Orós reservoir. In practice, both areas will show similarities.

To solve the problem of a too large distance from the reservoir edge, a contour line around the reservoir edge is drawn with a distance of 2,5 km from the reservoir edge. The reservoir edge at each elevation is determined by using the DEM. (The reservoir edge at elevations lower than 196 m was hard to determine, due to irregularities in the DEM, so these edges are less detailed). The 2,5 km is based on observation in figure 3.7: 'AOI 3 in each year'. The horizontal pumping distance to the agricultural fields which are likely to pump water form the reservoir is generally not longer than 2,5 km. Only the available land due to the 6 m relative elevation rule within this contour is counted as available land.
The available land is also limited by the suitability of the ground for agricultural land use. The soil map in figure 3.4 shows that the concerned area contains soils suitable for agricultural land use: vertisoils and alluvial soils. So it is assumed that the land available in the reach of 192 to 204 m water levels, is not limited by the soil.
In figure 3.12 shows the available land per water level. Water levels of 198 and 201 m show some local peaks in land availability, but the general trend is: 'the higher the water level, the more land becomes available'. This trend can mainly be explained by the increasing outline of the reservoir that gives more land access to the water.

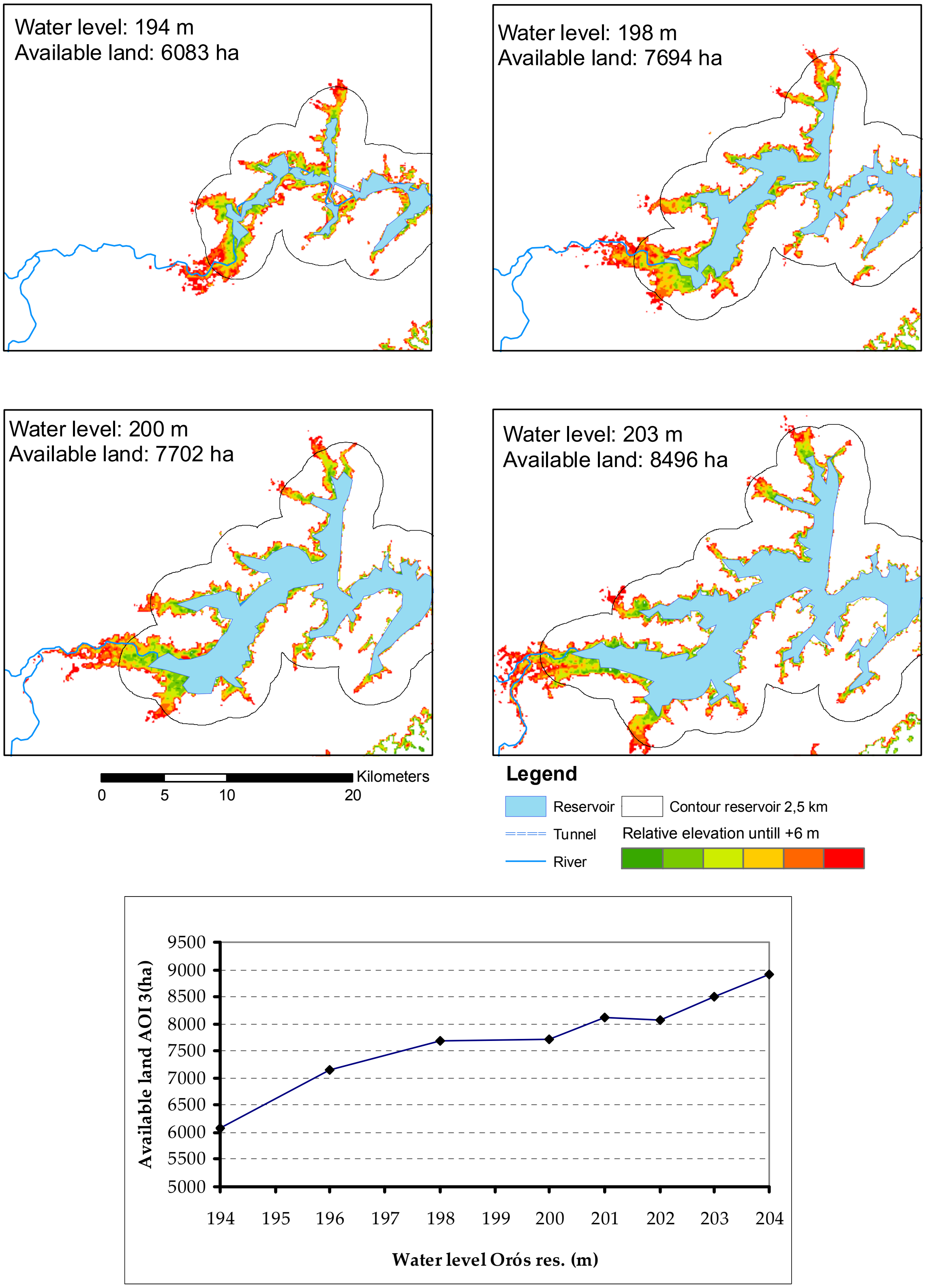


figure 3.12: Relationship 'water level (h)' and 'available land' at reservoir edge of Orós

Land availability around the Lima Campos reservoir.
Since the irrigated area of the AOI around the Lima Campos reservoir (AOI 5) gets its water from the tunnel from the Orós reservoir, the level of the Lima Campos reservoir is only influencing the land availability and not the water availability. As mentioned before, the water from the tunnel is divided over 2 sub canals and distributed over AOI 5. When the water level in the Lima Campos reservoir is high, a part of AOI 5 is inundated. Thus, the lower the level of the Lima Campos reservoir, the more land is available for agricultural land use. This phenomenon becomes clear in figure 3.13: the lower water level of 2003 provides more land available close to the reservoir.

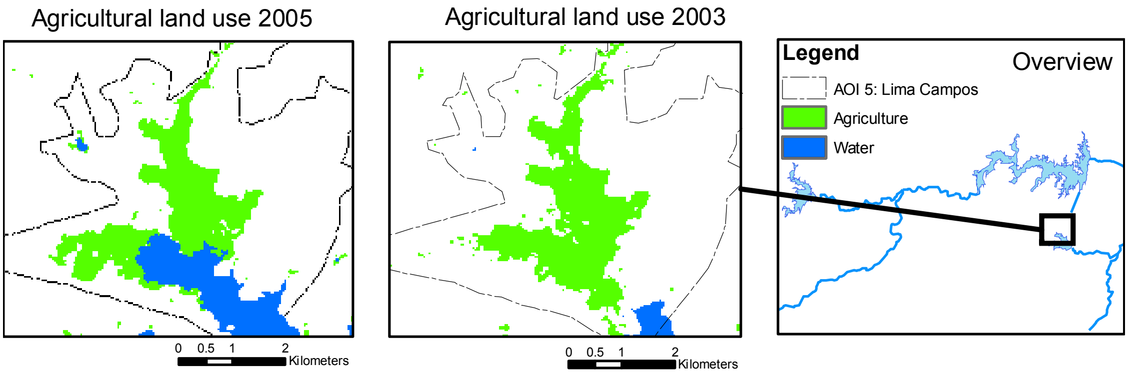


figure 3.13: Land availability due to water level Lima Campos reservoir

Reliability of reservoir and elevation data
The volumes in the reservoir are determined by COGERH using a known bathymetry and a measured water level. Since the bathymetry is changing due to erosion and sedimentation, the volumes have a unreliability. Since the research period is 6 years and sedimentation is estimated to be a long term process, it is assumed that the reservoir beds did not remarkably change during the research period. Also the measurements of the water level can contain an uncertainty. Since know knowledge was collected about how the water levels are measured, the uncertainty is estimated with the uncertainty of a standard staff gauge: +/- 3 mm (Shaw, 1994). On a depth of around 20 m (see figure 4.19: Waterlevel-volume relationship Orós reservoir (DNOCS, 1982)), this uncertainty is negligible. Another source of uncertainty is a different zero level of the staff gauge compared to the zero level of the water. Since the water depth during the research period was large (30 – 50 m) in comparison to the order in which this fault can be expected (+/- 1 m), the relative uncertainty is not expected to be larger than 5%.
For the calculation of the land availability around the Orós reservoir, the uncertainty of the crossing of the satellite images with the DEM (as explained before) are determinative. When the DEM is overlaid with the reservoir edge of an image, the reservoir edge should show similar elevation values. By doing this test, the variation in DEM values was around +/- 1 m. This variation can be caused by vegetation (which influences the measured elevation by the sensor) or by a spatial shift of the DEM compared to the image.

With a GIS-operation, the DEM can be transformed in a slope map; each pixel is labeled with a slope (see appendix 7). This slope map shows in the area concerned, a relatively constant slope between 0 and 9%. When the slope of the area concerned is assumed to be *constant*, this variation of 1 m on a elevation difference of 6 m causes an 1/6 uncertainty in the horizontal length over the elevation difference. (see figure 3.14). When the land within this reach of 6 m elevation is simplified as a streak (see figure 3.15), with a certain length, the area has an maximum uncertainty of 1/6. The chance that along the whole +6 m line, the real elevation is 1 meter higher than the elevation on the DEM is very small. It can rather be expected that this variation of +/- 1 m is leveled to 0 m along the line. Therefore, the 1/6 uncertainty is a rough overestimation and the real uncertainty is expected to be small. This will be further discussed in chapter 5: Discussion.

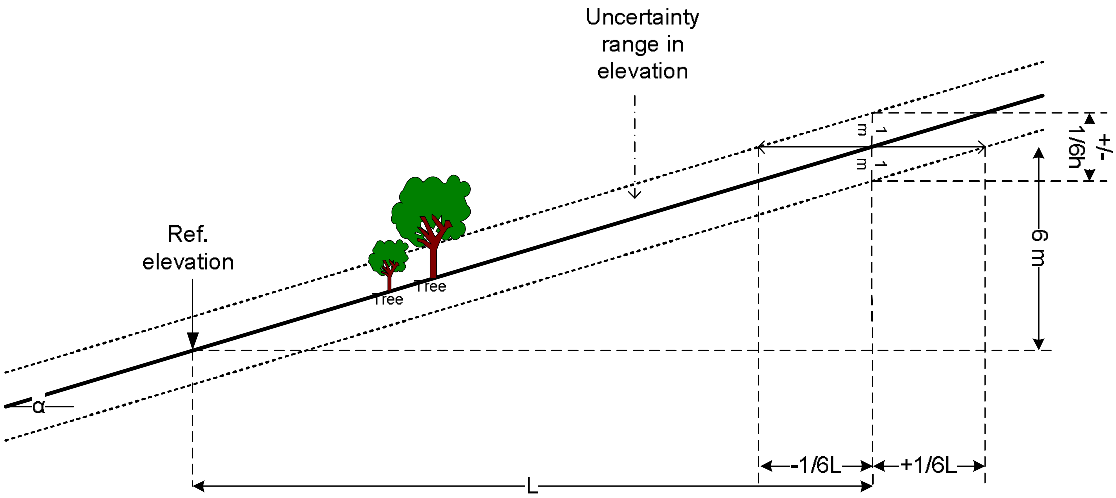


figure 3.14: Uncertainty of DEM; cross-section

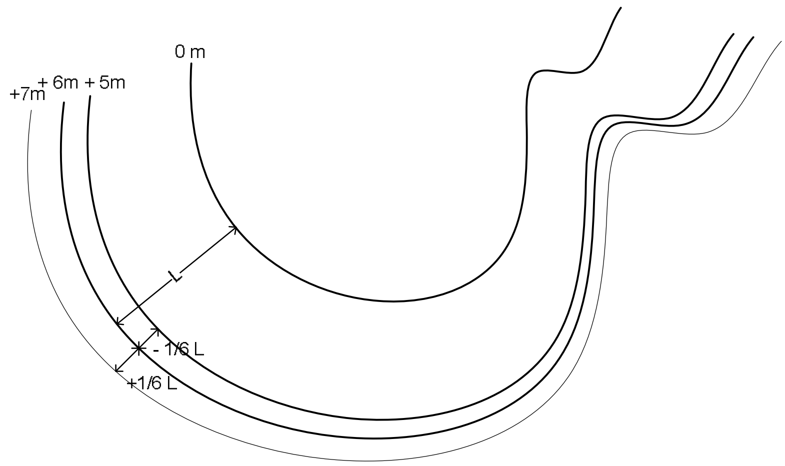


figure 3.15: Uncertainty DEM overview

3.4.4 *Locally stored runoff*

Paragraph 3.3.1 describes the area that is dependent of locally stored runoff. Runoff is influenced by many factors, like amount and intensity of rainfall (Shaw, 1994). Since most of these data is not available, only the amount of rainfall of foregoing periods will be considered to estimate the amount of locally stored runoff in a certain period. It is expected that the wet season (Jan – May) should cause locally stored runoff in the dry season (Jun – Dec); a wet season with much rainfall should give a high water availability in the dry season.

3.4.5 *Summary of water and land availability*

In table 3.3 a summary is given of this paragraph. It tries to summarize how each type of water and land availability is converted to a format usefull in this research. It also summarize which AOI's are expected to be supplied by which water availability. The described estimation of uncertainty is tried to express in a value in the last column.

Water availability	Concerned AOI's	Data used	Source	Estimation of maximum uncertainty
Precipitation	All	Daily rainfall data, converted to monthly rainfall per AOI	FUNCEME (2006)	10%
Discharges (from river and reservoir releases)	*River*: AOI 2: River *Releases*: AOI 4: Trussu, AOI 5: Lima Campos (tunnel) AOI 6: Perímetro	Monthly data of rivers discharge, daily data of release points	COGERH (2006), CPRM (2006)	Release data: 10% River data: unknown
Reservoir volume	AOI 3: Orós	Daily and 10-daily data of volumes of reservoirs.	COGERH (2006)	5%
Locally stored runoff	AOI 1: Locally stored runoff	Daily rainfall data, converted to monthly rainfall per AOI	COGERH (2006)	10%
Land availability	AOI 3: Orós AOI 5: Lima Campos	Digital Elevation Model (DEM)	CGIAR-CSI (2000)	17%

table 3.3: Summary water and land availability

3.5 *Land use data*

3.5.1 *Ground truth data*

As mentioned in appendix 2.6, CBERS-2 images and Landsat7 images were available for the land cover classification. The CBERS-2 images do not cover the total study area, so theses scenes are composed of two images. Before the both sides are resampled (the resample process is explained in appendix 1.3) and glued, the land cover classification, as described in paragraph 2.3, inclusive the accuracy assessment, is carried out on both images. The resampling is carried out after the land cover classification, since the DN-values are changing during a resampling process.
As mentioned in paragraph 2.3.5, a ground truth data set is required to carry out a multi-spectral image classification. The number of usable Ground Control Points (GCP's) and the division in sampleset – testset are listed in table 3.4 and table 3.5.

	2000			2001			2002		
	ss	ts	tot	ss	ts	tot	ss	ts	tot
Paddy	36	9	**45**	25	6	31	22	5	**27**
Banana	38	8	**46**	46	11	57	32	8	**40**
Beans	14	-	**14**	-	-	-	6	-	**-**
Grass	4	-	**4**	4	-	-	4	-	**-**
Coco	4	-	**4**	-	-	-	2	-	**-**

table 3.4: GPC's: division 'sample set (ss) – test set' (ts) and total (tot), Landsat7 images

	2003_Right			2003_Left			2004_Right			2004_Left			2005_Right			2005_Left		
	ss	ts	tot	ss	ts	tot	ss	ts	tot	Ss	ts	tot	ss	ts	tot	ss	ts	tot
Paddy	33	8	**41**	19	4	**23**	34	9	**43**	23	5	**28**	30	7	**37**	25	6	**31**
Banana	28	7	**35**	31	7	**38**	26	6	**32**	28	7	**35**	26	6	**32**	43	10	**53**
Beans	-	-	**-**	6	-	**-**	11	-	**11**	3	-	**3**	-	-	**-**	-	-	**-**
Grass	4	-	**-**	-	-	**-**	4	-	**-**	-	-	**-**	6	-	**-**	-	-	**-**
Coco	3	-	**-**	-	-	**-**	2	-	**-**	-	-	**-**	4	-	**-**	-	-	**-**

table 3.5: GPC's: division 'sample set (ss) – test set (ts)' and total (tot), CBERS-2 images

The data in table 3.4 and table 3.5 are the result of a fieldwork session, carried out in March and April 2006. Local organizations and farmers are questioned about the kind of crop that was grown during the dry periods of 2000 till 2005 on different fields. This information forms the dataset with GCP's.
Since the farmers and local organizations can make a mistake, the DN-values at the location of each GCP are compared with the DN-values of reliable GCP's of the same crop at other locations. When these values were corresponding, the GCP's was marked as usable GCP, when this was not the case, the GCP's was excluded for the dataset of that year (the same GCP can be a useable GCP in another year!). The rule of thumb to have a minimum of 30 usable GPC's per land cover class (paragraph 2.3.5) is not achieved in all years. The reasons for this are:

- There are only a few fields in the research area for some of the relevant crops (for example Coco)
- In some years a particular crop were scarcely cultivated (for example paddy in 2002)
- The total data set of GCP's had to be split for the CBERS-2 images, since an individual classification was carried out on each side.
- Some crops, growing in the research period, were harvested or very young at the acquisition date of the image (for example: beans), and therefore not recognizable on the image.

As a result of these points, beans, grass and coco were excluded from the accuracy assessments that are described in the next paragraph, since an accuracy assessment with a few usable GCP's has no statistical meaning. The reliability of the output of the classification of these crops will be discussed in paragraph 3.5.3. The dataset of crops with enough usable GPC's (more than 30) is split up in a test set and sample set; the test set to carry out the image classification, the test set to assess the accuracy of the image classification. For the contribution over both sets is chosen for the usual proportion 4:1 (sample set : test set)

Note: The class 'Grass' in the table is a kind of rest-class. 'Grass' is in the sample set assigned to fields that were indicated in the ground truth data as beans, but showed on the images so little vegetation reflectance (similar to area's outside the reach of water sources) that they could not be beans; when these fields were assigned to beans, the classification results showed many areas of beans which were in reality uncultivated land.

3.5.2 Classification results for classes with sufficient GPC's

As mentioned in sub-paragraph 2.3.3, a classification algorithm has to be used to classify each pixel of an image to a certain land cover. All images can be classified with the same algorithm, chosen by the best average accuracy in the accuracy assessment of all classifications, or each

image can be classified with the algorithm that gives the best accuracy on that particular image. In the last option, not all images are classified with the same algorithm.
Since the classification algorithms use the shape of the feature clouds expressed in parameters as 'mean value per class', 'variance' and 'co-variance' (see sub-paragraph 2.3.3), algorithm x can give a better accuracy than algorithm y due to this shape of feature clouds. Since all feature spaces have different shapes, due to different acquisition dates, different atmospheric conditions, different sample sets and different sensors, an optimal classification algorithm for each image can be found. Therefore, for each image the classification algorithm with the highest accuracy has been selected.
In table 3.6 an overview of the statistical results for all images is given. In appendix 4.2 all confusion matrices are given. In appendix 4.4, all maps are shown.

In the second column of table 3.6 the classification algorithm is listed; ML = Maximum Likelyhood, MML = Minimum Mahalonabis Distance. The accuracy and reliability of the most dominant crops, Paddy and Banana are listed in the last four columns.

Image	Method	Overall accuracy	Overall Reliability	KHAT	Paddy		Banana	
					Acc	Rel	Acc	Rel
2000	ML	0.93	0.98	0.84	0.99	0.96	0.81	0.99
2001	ML	0.94	0.86	0.84	**0.65**	0.86	0.89	0.72
2002	MMD	0.96	0.96	0.88	0.91	0.96	0.87	0.93
2003_L	ML	0.96	0.95	0.90	0.88	0.97	0.86	0.89
2003_R	MMD	0.98	0.97	0.96	0.97	1.00	0.97	0.92
2004_L	MMD	0.99	0.98	0.97	0.97	0.99	0.97	0.95
2004_R	ML	0.99	0.95	0.96	0.96	1.0	0.81	0.94
2005_L	MMD	0.99	0.96	0.97	0.97	0.93	0.91	0.96
2005_R	ML	1.00	0.99	0.98	0.99	0.97	0.88	0.99

table 3.6: Used classification methods per image and results of accuracy assessment

Note by table 3.6:

- The overall accuracy of the MML method for 2001 was higher than for the ML method: 94.19%. The reason to choose the ML method was the low accuracy of Paddy in the MML method: 55%. ML gives accuracy for Paddy of 65%.

The main reasons that a land cover map, as a result of a classification algorithm, is not 100% correlated to the test set are:

- The problem of 'mixels' (see sub-paragraph 2.3.6)
- Overlapping feature clouds (see feature spaces in appendix 4.1). A specific issue in the land cover classifications applied to this research area was that bananas in certain stages have almost an equal reflection in all available bands as coco and full grown paddy. This can be read in the feature spaces in appendix 4.1
- A field with a crop in an early growing stage also reflects radiation from the bare soil because of low crop coverage. Therefore, crops in an early growing stage are hard to assign to a certain class.
- The GPC's of the first years have more uncertainty than later years since the human mind can be forgetful. Besides this, people from Cogerh and Ematerce who guided our tour, did not work for these companies before 2003. Since the crop patterns of many fields did not change in the former 10 years (according to information of local farmers), the GPC's taken at these fields gave enough data for the classification.

Other considerations:

- The class 'paddy' is split up in the sample set, since two types of paddy could be distinguished: young paddy, where the water was good visible and almost full grown paddy, where the water was less visible in the reflectance values. After the classification both sub-classes are merged.
- Since the GPC's are only taken in irrigated areas with certain temporal growing patterns, the results for areas were crops are grown by using other methods and other temporal patterns have less accuracy, because they have other DN-values. (Lillesand and Kiefer, 1994). So the calculated accuracy values count only for the AOI's and the areas that have comparable growing conditions.

3.5.3 Accuracy of classes with a few GPC's

Coco

To get an idea about the accuracy of a the classification algorithm in assigning pixels to a certain class, without using a confusion matrix, the classified map can be 'crossed' with a classified map from other years. The cross operation performs an overlay of both maps: pixels on the same positions in both maps are compared; the occurring combinations of class names in the first input map and those of pixels in the second input map are stored. These combinations give an output cross table. The cross table includes the combinations of input classes, the number of pixels that occur for each combination and the area for each combination. When the combination '*coco from the first map* and *coco from the second map*' contains a large part of the total pixels of coco of the first map, then this indicates consistency in the classification between both years. Each year is classified with an own sample set, so consistent classification results are indicating a good accuracy of the particular land cover classifications.

Especially 'coco', as a *permanent* crop with a long life period, is suitable for this kind of accuracy assessment. Due to this permanency, the consistency should be high. Data of DNOCS (see table 3.13) also shows that between 2001 and 2004 the area of coco is not changing in AOI 6: Perímetro (this data will be discussed later in this report).

In the period 2001 – 2004, each year is crossed with all other years, so 3 cross-tables per year are made. In each cross-table the consistency of coco is calculated by dividing the amount of pixels in the combination 'coco_map$_x$-coco_map$_y$' by the total amount of coco pixels in map$_x$. The mean consistency per test year is shown in table 3.7. This value gives an answer to the question: "were the coco-pixels of year *a* also coco pixels in year *b,c* and *d*?" A value close to 1 indicates "always", an answer close to 0 answers "never". The total result of all cross-tables is shown in appendix 4.3, table 4.1.

Test year	2001	2002	2003	2004
Consistency with other years	13	11	27	16

table 3.7: Inter-annual consistency of coco


figure 3.16: Cocos in AOI 6

The results are very low, which indicates a low consistency of classified pixels to coco. As mentioned before, coco should have had a high consistency. So it can be concluded that the accuracy of the pixels in the class 'coco' is very low. Table 4.1 in appendix 4.3 shows also to which classes the 'coco-pixels' of another year are classified in the reference year. In all years, 'banana' and 'non-agriculture' score high on this aspect. So many pixels that are in a certain year coco, are in another year assigned to banana or non-agriculture.

A possible explanation for this is: Coco trees are planted relatively far away from each other (see table 3.13) as a result of which the ground between the coco trees has an important influence on the DN-values of a pixel in a coco-field. These DN-values can be very similar to banana fields that are just planted and with non-agricultural land in common. This overlap can also be seen in the feature spaces (see appendix 4.1).
A second point of consideration is the large amount of coco pixels in 2001 and 2002 (see first column of table in appendix 4.3). 2001 and 2002 were both dry years. (this can be concluded from the emptiness in the reservoirs in figure 3.11). In these years, there was much bare land as a consequence of the lack of water. It is assumed that the pixels that are assigned to coco, were in reality non-agricultural land.

Conclusion:
The result of the classification of coco is very uncertain. The pixels that are assigned to 'coco' have a high probability to be 'non-agricultural' or 'banana'. Since the irrigation water use of coco is relatively low ánd the coco-pixels in other years are mostly assigned to 'non-agricultural', the pixels assigned to coco in the classification are considered as 'non-agricultural'. This will have not great impact on the total irrigation water use of an AOI, since the total amount of coco-fields is small in comparison to other crops. The possibilities of using other data will be examined in the next parageraph.

Beans and grass in AOI 6
As explained in sub paragraph 3.5.1 only a few ground truth data points of beans could be identified on the images. The acquisition date of almost all images (between end of October and half of November) was too late in the season to see beans present on the fields. Because the images of 2004 were some week's earlier (end of September) some fields with beans could be identified. This is the reason why the number of GPC's in 2004 is higher than the surrounding years (see table 3.5). From interviews with local farmers in the Perímetro, it is known that the cropping pattern of beans is quasi fixed in the last 10 years. It is also known that after the beans are harvested the land becomes 'rough land' and is grazed by cattle, until a new crop is planted (mostly corn in the wet season). So the presence of grass in AOI 6: Perímetro at locations which were assigned to beans in 2004, can indicate a harvested beans field.

To check to which class the bean-pixels of 2004 are assigned in other years, the AOI: Perímetro is used again for crossings, since this area has 'fixed crop patterns'. Again cross-tables are made. A summary of these tables is given in table 3.8.

Crossing		Banana (%)	Beans (%)	Coco (%)	Grass (%)	Paddy (%)	Water (%)	Non-agri (%)	Grass + Non-agri (%)
Beans 2004	All crops 2000 (date: 10-25)	**14**	**13**	6	0	9	0	**58**	58
Beans 2004	All crops 2001 (date: 11-13)	10	0	9	0	3	0	**78**	78
Beans 2004	All crops 2002 (date: 10-31)	10	6	**18**	**13**	3	0	**50**	63
Beans 2004	All crops 2003 (date: 11-22)	**13**	0	1	**14**	9	0	**63**	77
Beans 2004	All crops 2005 (date: 10-21)	22	0	2	**50**	4	0	22	72

table 3.8: Bean-pixels of 2004 crossed with classified maps of other years (AOI 6: Perímetro)

Explanation table 3.5: The table shows to which class the bean-pixels of 2004 are assigned in other years. The bold percentages in table 3.8 indicate the main classes that cross with the bean-pixels of 2004. For example: the pixels that were assigned to beans in 2004, were mainly assigned to banana, beans and non-agricultural in 2000. In all years, the main part of the beans-pixels of 2004 is assigned to the non-agricultural class (see fore last column of table 3.8). The three years that contain grass GPC's (2002, 2003, 2005) show a relation with the bean-pixels of 2004. In 2002 and 2003 this relation is weak (13 and 14%); in 2005 it is relatively strong (50%). The strong relation with the non-agricultural can be explained by the fact that after the harvesting of beans, the land becomes a rough land which has much similarity with non-agricultural land. This process from just harvested land to rough land can go fast, since this period is very dry. Just after the harvesting, the land is still a little green and the soil has probably some water content which is enough for grass to grow. This can be the reason why the early images of 2000 and 2005, both end-October, show lower percentages in the non-agricultural column than 2001 and 2003, which show higher values in the non-agricultural column. Around 70% of the beans in 2004 cross with grass+non-agricultural land in other years (see last column)

An extra indication about the location of the harvested beans was found in an available CBERS-2 image of 2005 at the 2nd of September. At the locations that were indicated as beans in the ground truth data this image showed, in contradiction to the image of 2005 of the 24th of October, high vegetation activity in the DN-values. These locations were therefore sampled as beans. No accuracy assessment could be made because there were still to few beans GPC's, so the results can only give an indication. It is expected that this indication is quite good, since the feature space shows a separated point cloud of beans and the classified map shows a normal pattern (see appendix 4.3, figure 4.1 and 4.2). The classification results of beans in the image of the 2nd of September were crossed with the classification results of all crops in the image of the 24th of October in the same year. In table 3.9 the results are shown: the results are underlining the former observations: the fields that contain beans in the period August-September, become grass and bare land (79%) in the months afterwards. Banana counts 20%. Reasons for this are expected to be the same as called in 3.5.2

When the classification results of the image of the 2nd of September 2005 are crossed with the classification results of the image of 29th of September of 2004, consistency in the locations of the bean-fields can be seen: 53% of the pixels that are assigned to beans in the image of 2005 (09/02) are also assigned to beans in 2004. 22% of the bean-pixels of 2005 are non-agriculture in 2004.

Crossing crops	Crop 02/09/05	Crop 24/10/05	Area (ha)	Area (%)
Beans * Paddy	Beans	Paddy	0.04	0
Beans * Banana	Beans	Banana	14.68	20
Beans * Coco	Beans	Coco	0.80	1
Beans * Grass	Beans	Grass	32.12	44
Beans * Non-agri	Beans	Non-Agri	26.08	35
		Total:	73.72	100

table 3.9: Beans of 02/09/05 crossed with crops of 24/10/05, AOI 6: Perímetro

Crossing crops	Crop 02/09/05	Crop 29/09/04	Area (ha)	Area (%)
Beans * Paddy	Beans	Paddy	0.56	1
Beans * Banana	Beans	Banana	8.32	11
Beans * Beans	Beans	Beans	39.16	53
Beans * Coco	Beans	Coco	9.16	12
Beans * Non-agri	Beans	Non-agri	16.52	22
		Total:	73.72	100

table 3.10: Beans of 02/09/05 crossed with crops of 29/09/2004, AOI 6: Perímetro

Conclusion:
The locations of beans in images with an early acquisition date have a high correlation with the locations of non-agricultural land and grass in images with a late acquisition date. For this reason, the numerical data of COGERH about the areas of beans can be added to the results of the land cover classification; beans are replacing areas without irrigation water use (grass and non-agriculture). The areas of beans as a result of the land cover classifications of images with an early acquisition date (2004 and 2005) are neglected, to maintain consistency with other years. The numerical data is listed in table 3.11. The cells with the '-' symbol indicate that no sample points of beans were available, so it did not join the classification. From interviews during the fieldwork it is found that beans are mainly planted in the dry season. During this dry season, beans are one or two times grown on each field. It is assumed that the fields that contain beans in two cycles in the dry period are counted twice in the numerical data

	Area beans in AOI 6 (ha)	
year	COGERH	Classification
2000	350	-
2001	169	-
2002	18	-
2003	82	-
2004	301	205
2005	328	74

table 3.11: Numerical data (Cogher, 2006) of beans areas

Beans and grass in other AOI's

In table 3.12 the results of the land cover classification for AOI 1 to 5 are listed. The values indicate the percentages of the total agricultural land use. The share of beans and grass is quite small, except for the years 2000 and 2003 in AOI 2: river and AOI 4: Trussu. It is assumed that grass is former beans-land. The grass and beans area's are summed and assigned to one beans cycle in the dry season.

	Crop	2000 A (%)	2001 A (%)	2002 A (%)	2003 A (%)	2004 A (%)	2005 A (%)
AOI 1	Beans	9	0	2	5	0	3
	Grass	0	0	2	0	0	0
AOI 2	Beans	**19**	0	5	**18**	1	3
	Grass	0	0	5	0	0	0
AOI 3	Beans	2	0	1	0	0	0
	Grass	0	0	3	2	0	2
AOI 4	Beans	**28**	0	4	**22**	0	1
	Grass	0	0	5	0	0	0
AOI 5	Beans	3	0	2	0	6	0
	Grass	0	0	9	2	0	9

table 3.12: Results of classification in AOI 1-5 for beans and grass as percentage of total agricultural land use

Although the results of the land cover classification of beans and grass are very uncertain, using these outcomes is better than assigning these areas to non-agricultural land, since the DN-values of these areas indicate vegetation that is different from non-agricultural land. To give an indication for the final irrigation water use of AOI 4 in 2000, which has the highest percentage of beans; the difference in total irrigation water use for the case that these 28 percent were considered as beans or the case that these 28 percent were non-water using is 19% (after doing a irrigation water use calculations as explained in paragraph 2.4.3).

3.5.4 Comparison classification results with numerical data

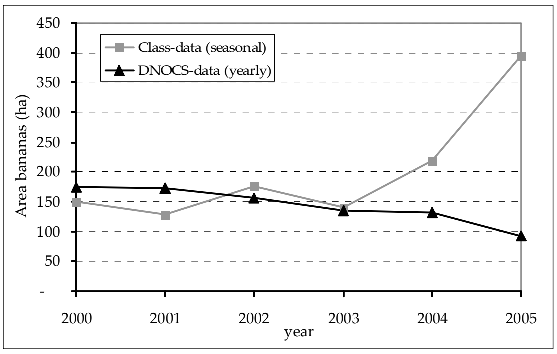
DNOCS-data of AOI 6: Perímetro

DNOCS, the governmental organization for preventive infrastructural works against droughts, had data about agricultural land use available of AOI 6: Perímetro. It is *annual* data. The data of the classification are only for the *dry season*.

	Area 2000 (ha)		Area 2001 (ha)		Area 2002 (ha)		Area 2003 (ha)		Area 2004 (ha)		Area 2005 (ha)	
	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)
Paddy	830	553	39	91	5	124	204	462	897	354	501	481
Banana	175	150	173	128	156	176	135	140	132	219	92	395
Beans	350	73	169	-	18	64	82	-	301	205	328	-
Maize	60	-	47	-	1	-	41	-	55	-	138	-
Capim	90	-	199	-	199	-	199	-	189	-	149	-
Grass	-	-	-	-	-	*117*	-	*165*	-	-	-	*500*
Coco	6	106	8	211	8	261	8	21	8	83	12	45
Goiaba	1	-	3	-	3	-	3	-	3	-	14	-

table 3.13: Comparison annual DNOCS data (DNOCS, 2006) (='Num') with classification results (='Class'), AOI 6: Perímetro

In table 3.13 the numerical data of DNOCS are listed together with the results of the classification. For permanent crops, like banana, the annual data of DNOCS and the seasonal data of the classification should be similar. The banana data of the classification shows a good similarity with the DNOCS data for the years 2000 – 2004 (see figure 3.17, left graph). 2005 shows a large difference. Also the paddy data shows similarity (figure 3.17, right graph). Since most paddy are grown in the dry season, according to information of local farmers during the field visit, the annual data of DNOCS is comparable with the seasonal data of the classification



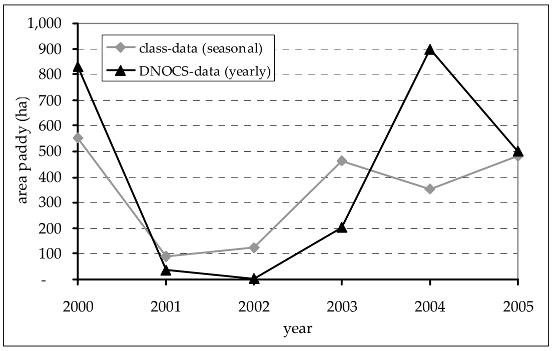


figure 3.17: Comparison DNOCS data and classification results (class) about bananas (left) and paddy (right) in AOI 6

The classification data should in any case not be higher, since it is only from the dry season. The classifications of 2001, 2002 and 2003 have however higher areas for paddy than the DNOCS data. After examining the classification maps of these years (see appendix 4.4, figure 4.8), it can be seen that 2001 and 2002, both dry years, have a few paddy fields that are located on locations that also contain paddy in other years. This is a reliable situation, so is decided not to doubt the results of the classification. 2003 can contain probably an overestimation of paddy, since some paddy fields are banana fields in other years. It is hard to say how large the supposed overestimation is. 2004 shows a deviant difference. According to the DNOCS data, the paddy area of 2004 should be twice as large. Looking at figure 4.8 of appendix 4.4 it has to be concluded that this is unrealistic; the classification result from 2004 on paddy seems more reliable.
The data about the results of the classification on beans and coco are underlining the conclusions of the former paragraph. Maize will not be discussed, since it is mainly grown in the wet season. Goiaba has a very small contribution to the total area of crops and will also be not discussed.
There is no field data collected on capim, since 'capim' is a kind of rest crop. There are several types of capim which are varying from pasture land to large reed-kind plants that are used as food for cattle. Remarkable is the large area of capim (between 90 and 200 ha). When these areas are compared to the classified areas of grass of 2002 and 2003, a similarity can be seen. Since at least a significant part of capim is grass land, it is assumed in this study that capim is not an irrigation water user.

IBGE-data of Iguatú municipality

IBGE, the Brazilian institution for geographical and statistical data, supplies *annual* agricultural data about the planted and the harvested areas per municipality. In the concerning years, the planted and the harvested areas were similar. The research area contains one complete municipality and parts of other municipalities. This complete municipality is called Iguatú. Overlaying the available borders of this municipality that are used in the WAVES-project over the classification results provided the land use in the Iguatú municipality. The borders of the Iguatú municipality are shown in figure 4.9 in appendix 4.5. It can be seen that this area also contains areas outside AOI's. Although the classification is based on ground truth data inside the AOI's, the results just outside the AOI's, inside the municipality Iguatú, are expected to be reliable, since properties as 'soil', 'growing season' and 'type of crops' are equal. An extra note has to be made by the location of the borders of the Iguatú municipality: It is expected that the

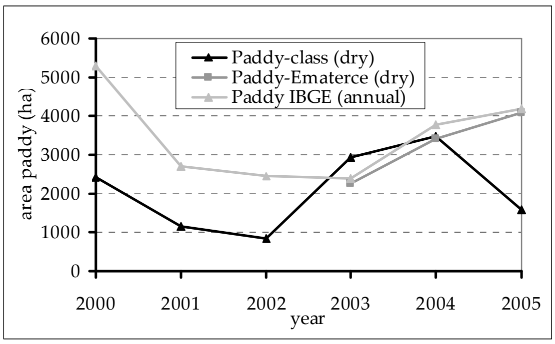
river between AOI 3 and AOI 4 (Jaguaribe), being a natural border, is similar to the municipality border. This is not the case in the border data retrieved from the WAVES-project. After checking this in another map with municipality borders (IPECE, 1998) (see figure 4.10 in appendix 4.5) the borders are located more or less on the same locations as the WAVES-data. Probably because of the changing course of the river, the border is not similar to the river anymore. Still the borders of both sources are not similar in the northern part. This will cause an uncertainty in the computation of the total areas of the crops in the classified maps. Since most agricultural activity is around the river, this uncertainty will be small for the purpose it is used for in this research.
The land use according to the IBGE data and the classification data are listed in table 3.14: Only banana is shown, because annual and seasonal data can be compared because of its (semi) permanency.

	Area 2000 (ha)		Area 2001 (ha)		Area 2002 (ha)		Area 2003 (ha)		Area 2004 (ha)		Aera 2005 (ha)	
	Num (annual	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)	Num (annual)	Class (dry)
Banana	430	1331	500	1860	500	1707	475	2085	373	3061	?	2502

table 3.14:Comparison of IBGE data (IBGE, 2006) (='Num') and classification results (='Class') in the Iguatú municipality

The table shows a large difference between the IBGE-data and the classification data for banana in all years (factor 3-10 higher values in the classification) and for paddy in 2000 till 2002. An explanation for this difference is hard to give. It is possible that IBGE counts the area for each harvesting. This can be three times a year, so that would make the difference even larger. Beans and paddy data of IBGE are shown in figure 3.18 and will be discussed below.

Ematerce-data of Iguatú municipality
Ematerce is a local institute that supervises the distribution of seeds for agricultural use and assists in the production process. With the amount of seeds as input, they can calculate a value for the production area of a crop. Data in 2003, 2004 and 2005 for Paddy and Beans in the dry season was available. It is drawn, together with the classification results and the IBGE-data, in figure 3.18



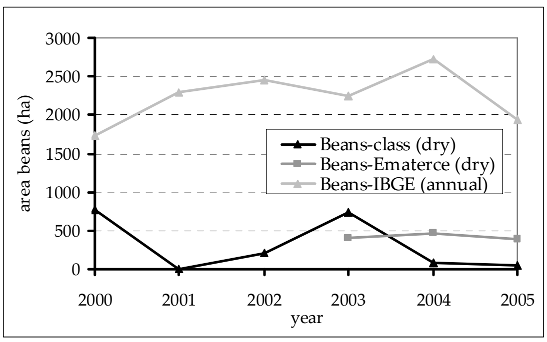


figure 3.18: Comparison seasonal classification results, seasonal Ematerce-data (2006) and annual IBGE-data (2006) of Paddy (left graph) and Beans (right graph)

Paddy
The *annual* IBGE data and the *seasonal* Ematerce data have the same quantities for 2003 – 2005. It is assumed that the main part of yearly paddy area is planted in the dry season. This is also in line with the field observations; during the field visit in March and April 2006 (wet season),

almost all paddy-fields ware 'bare'. Annual data of paddy and data of the dry season can therefore be compared in these years. In the period 2000 – 2002 the same trend in the classification results and the IBGE-data is visible, but the IBGE-values are higher (factor 2-3). A possible reason for this is that in these years two growing cycles for paddy were common. The classification-value of 2005 seems to be an outlier. A possible reason for this outlier is that the sample set of the left image of 2005 shows an unusual feature space for banana (see appendix 4.1). But still the accuracy assessment, as an indirect result of this feature space, shows a very high accuracy value for the classification of the left image of 2005 (0.97), so no conclusion can be given on this point.

Beans

For beans, the difference between the seasonal Ematerce data and the results of the classification is a factor 2 - 8. Possibly, areas with beans can be assigned to other crops, like banana. Banana lies close to beans in the feature spaces, so this can be a possibility. The fact that banana pixels from the test set are sometimes crossed with beans pixels can under scribe this hypothesis (appendix 4.2, 2004L and 2005R). Still this remains a weak basis to explain the differences in figure 3.18 between the classification results and the Ematerce-data.

The differences between the classification results and the IBGE-data even much larger. Since the differences between the *seasonal* Ematerce data and the *annual* IBGE data are large as well, the conclusion can be drawn that a vast part of the beans are grown in the wet season. Therefore, the annual IBGE data is not useful for a comparison with the seasonal classification results.

Summary

The numerical data of DNOCS of bananas and paddy in AOI 6: Perímetro show a similar trend to the classification results. The comparison on beans are underlining the conclusions of the former paragraph; the results of the classification on beans are unreliable.

The Ematerce and IBGE data of paddy in the municipality Iguatú show the same trend as the classification data (with exception of 2005) and therefore give a weak validation of the classification results.

The comparison between banana in the classification results and the IBGE data shows large differences. A suitable reason for this cannot be given.

It was already concluded that the classification results of beans are very uncertain. This conclusion is supported by the seasonal numerical data of Ematerce. From a comparison between the seasonal data of Ematerce and the annual data of IBGE the conclusion can be drawn that beans are mainly grown in the wet season. The areas of beans in the dry season are small. Therefore, the uncertainty will have a small impact on the uncertainty of the total irrigation water use.

3.6 Irrigation water use

For each AOI in each year the irrigation water use in the dry season is calculated according to the method described in paragraph 2.4. The areas as calculated in the land cover classification and derived from numerical data (AOI 6; coco and beans) are used as input for this irrigation water use calculations. Appendix 4.6 shows an overview of this input data. The results are shown in table 3.15 and figure 3.19. The results will be discussed in the next paragraph.

	AOI 1 ($10^6 m^3$)	AOI 2 ($10^6 m^3$)	AOI 3 ($10^6 m^3$)	AOI 4 ($10^6 m^3$)	AOI 5 ($10^6 m^3$)	AOI 6 ($10^6 m^3$)
2000	10.01	14.40	16.33	4.20	5.98	9.37
2001	5.77	17.45	9.16	6.23	4.89	4.86
2002	4.03	12.97	9.11	6.29	3.87	9.46
2003	17.00	23.09	7.66	10.22	6.05	7.95
2004	15.94	23.05	22.01	12.77	2.56	10.03
2005	3.45	18.75	18.52	11.32	5.21	11.58

table 3.15: Results Water use calculations

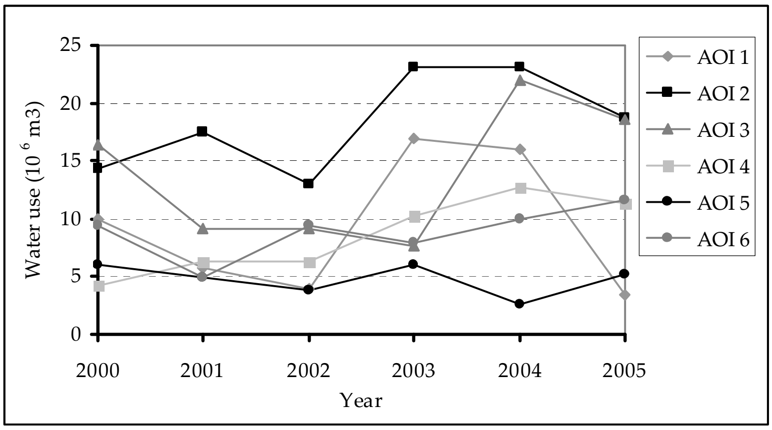


figure 3.19: Amounts of water use per AOI per year

To get a better insight in the ways the different crops use water, the courses of irrigation water use for the three main crops are calculated for a test field of 1000 ha in the research area (see figure 3.20): Growing rice demands the highest intensity of water supply, banana is the most constant and beans demands in a short period a high supply of water.

The total irrigation water use in the dry season per crop is listed in table 3.16; the proportion in irrigation water use between banana, rice and beans is as 8:6:3

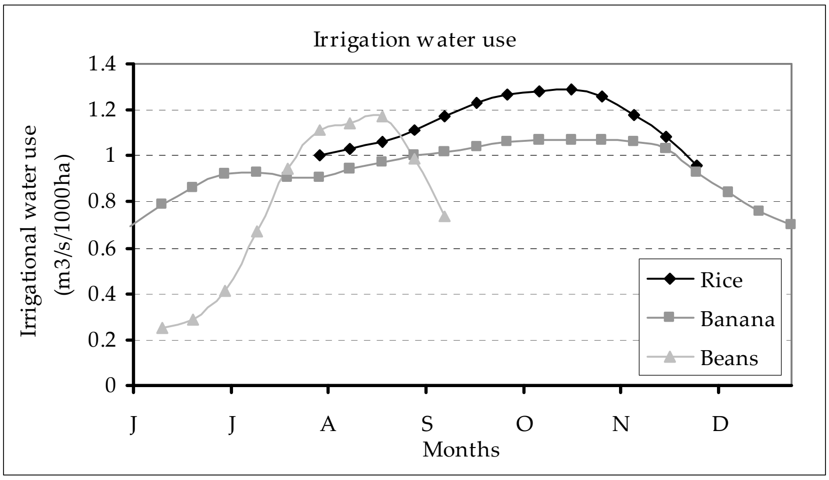


figure 3.20: irrigation water use banana, rice and beans

Crop	Total irrigation water use dry season (m3/1000 ha)	Irrigation depth (mm)	Proportion
Banana	$18*10^6$	180	8
Rice	$13*10^6$	130	6
Beans	$7*10^6$	70	3

table 3.16: Irrigation water use banana, rice and beans

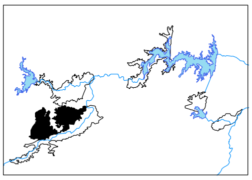
4 Results

4.1 Introduction

This chapter deals with research questions 4 and 5. In the definition of the AOI's (paragraph 3.3), hypothesis have been made about the dependency of certain areas on certain types of water availability. In this chapter the real data of the former chapter will be used to check these hypothesizes. This will be done in paragraph 4.2 by analyzing the evolving of water availability, land use and irrigation water use *per AOI*. In paragraph 4.3, the results of paragraph 4.2 will be used to identify trends in mutually conflicting, inter-annual, water uses between *different AOI's*. In both paragraphs, the AOI's will be analyzed one by one.

4.2 Inter-annual analysis per AOI

4.2.1 AOI 1: Locally stored runoff



Available water

AOI 1 is drawn with the goal to select an area which is mainly dependent of locally stored runoff. Since no data of locally stored runoff is available, the origin of locally stored runoff; rainfall will be considered as water source. The discharge of the Jaguaribe will not be excluded in this analysis, since it is in the direct neighborhood of AOI 1. Both water sources and the irrigation water use are drawn in figure 4.1.

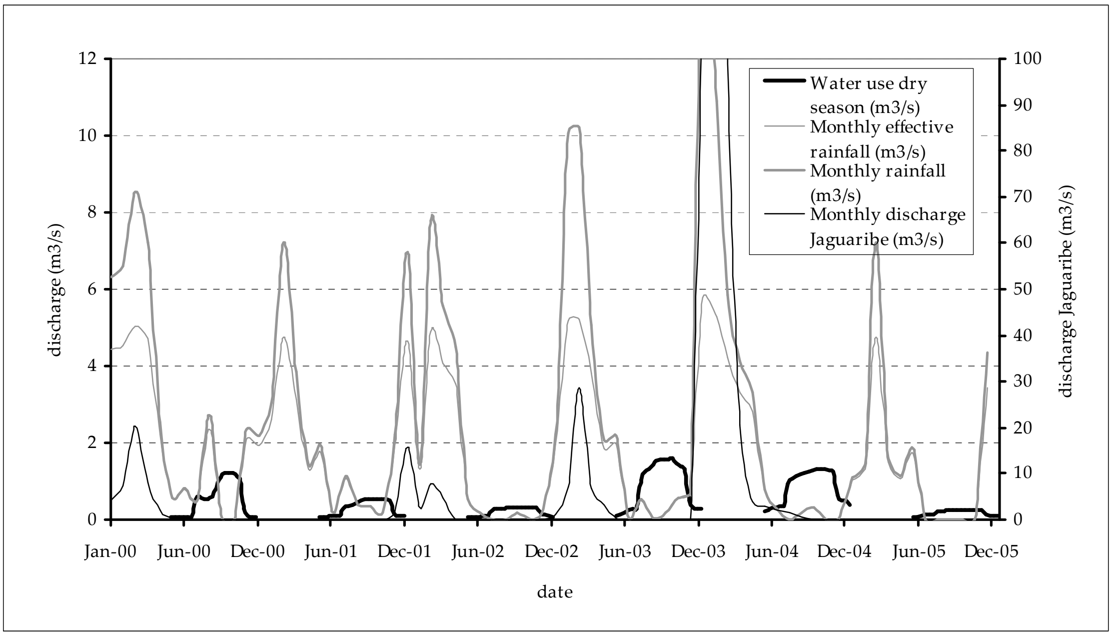


figure 4.1: Irrigation water use and water availability AOI 1

Observations on irrigation water use:

1. The discharge in the Jaguaribe river shows a high correlation with the rainfall (discharge data of 2001 and 2005 not available) (see figure 4.1).
2. The river discharge is negligible in comparison to the irrigation water use.
3. Effective rainfall in the dry season (especially 2000-2002) has a remarkable addition to the water availability (figure 4.1). The total effective rainfall in the dry season shows no correlation with the irrigation water use (figure 4.2)
4. Annual rainfall and wet season rainfall are positively correlated to the irrigation water use in the dry season of the same year (see figure 4.1 and figure 4.2): the three wet years 2000, 2003 and 2004 have higher irrigation water use than the dry years 2001, 2002 and 2005.
5. The correlation of point 4 (above) is not visible in 2003 and 2004: 2004 has a higher rainfall, but a lower total land use and irrigation water use than 2003 (see figure 4.1 and figure 4.2).

Observations on land use

6. The three wettest years (2000, 2003, 2004) have also the three largest paddy areas (figure 4.3
7. 2004 is wetter than 2003 but has less land use, mainly caused by less paddy
8. The area of banana is constant from 2000 to 2002 and increases in 2003 and 2004. 2005 has a again a lower area bananas.

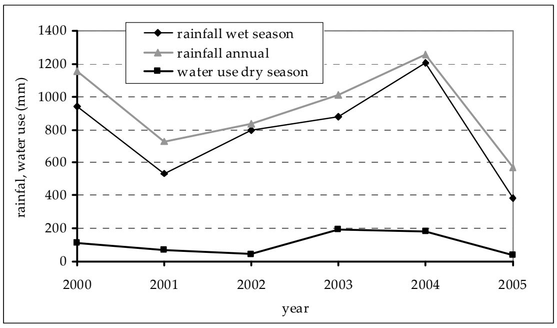


figure 4.2: total effective rainfall (wet season and annual) and total irrigation water use AOI 1

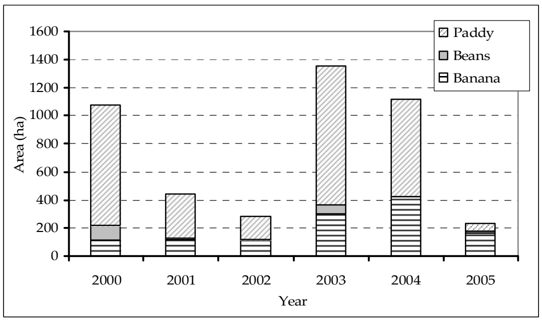


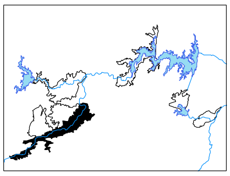
figure 4.3: Land use AOI 1

Interpretation of observations

- The decision rule to draw the border of this AOI is suitable: observation 2 shows that the only water source for AOI 1 in the dry season is locally stored runoff.
- The (Upper-) Jaguaribe river is dependent on runoff. This is confirmed by observation 1.
- Observation 3 is based on the total effective rainfall in the whole area of the AOI. Since only some fields in the AOI contain agricultural land use in the dry season (see appendix 4.4, figure 4.3 for maps of each year), the total effective rainfall for these fields is much smaller.
- Observation 5 and 7 can be explained by the filling of the natural lakes in the center of the bowls, which have an opposite effect on the irrigation water use, due to less land availability. For 2004, as wettest year, this can be the explanation of the unexpected lower irrigation water use than 2003 (see appendix 4.4, figure 4.3 for maps of each year)

- The expected reason for observation 6 is: Paddy is a flexible crop: it is grown for one season, so in each season the choice can be made to grow it. Only in a wet year, farmers choose to grow it.
- Observation 8 is hard to explain with water availability data. A *hypothesis* is that after the dry years of 2001 and 2002, a (political) shift is made to a crop that require a less intense water supply; banana (see figure 3.20: banana requires more m^3/s than paddy). The dip of 2005 can then be explained by the fact that this year was very dry, through which the planted banana's dried out.

4.2.2 AOI 2: River



Available water

The border definition of AOI 2 is drawn in the supposition that this area is dependent on river discharge, so this water availability will be considered in the analysis of AOI 2. Since AOI 2 is close to AOI 1: groundwater, also the locally stored runoff will be considered as available water source by looking to the rainfall. Both water sources and the irrigation water use are drawn in figure 4.4.

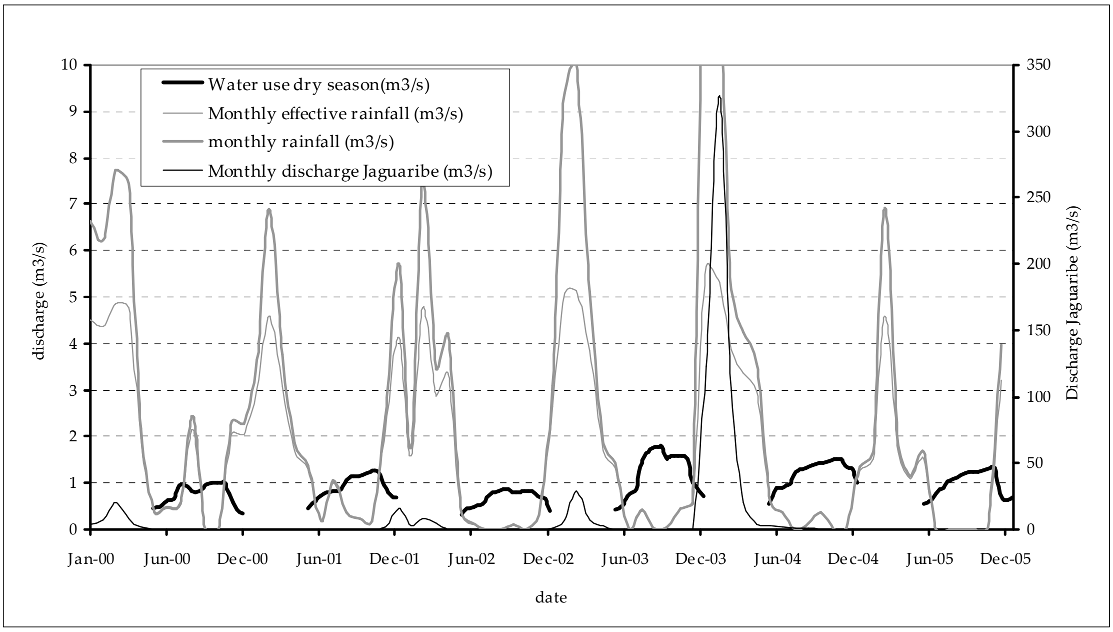


figure 4.4: Irrigation water use and water availability AOI 2

Observations on irrigation water use:

1. A correlation between river discharge and rainfall can be seen again (figure 4.4): the peaks of the river discharge are corresponding with the peaks of the rainfall (see also AOI 1).
2. The values of the effective rain in the dry season in figure 4.5 count for the whole AOI, the *real* total effective rain counts only for the total area of agricultural fields, which is maximal 40% of the total area of the AOI (see appendix 4.4, figure 4.4). Thus, the effective rainfall in the dry season is much lower than the total irrigation water use in the dry season.

3. The water availability of the river, as measured in Iguatú (after AOI 2) in the dry season is far below the irrigated water in the dry seasons.
4. From 2001 to 2005 a correlation between total annual rainfall and total irrigation water use in the dry season can be seen (figure 4.5). The variation in irrigation water use is much weaker as the variation in the rainfall.

Observation on land use

5. No remarkable trends in the distribution of crops can be concluded from the areas of land use as shown in figure 4.6

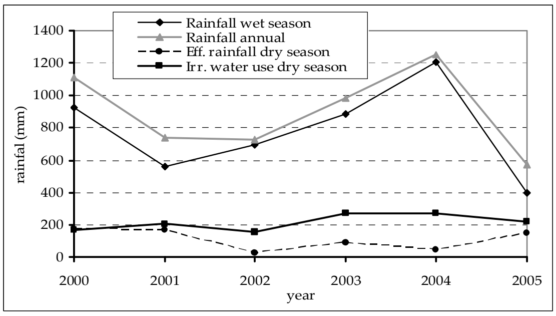


figure 4.5: total effective rainfall (wet season and annual) (left y-axis) and total irrigation water use (right y-axis) AOI2

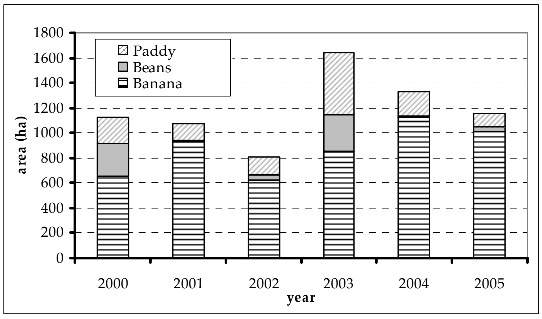
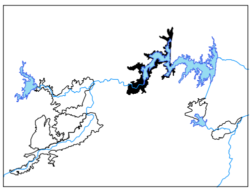


figure 4.6: Land use AOI 2

Interpretation of observations

- Resulting from observation 2 to 4 the conclusion can be drawn that this AOI is mainly dependent on locally stored runoff during the dry season. The hypothesis that this area is mainly supplied by river flow is falsified.
- Since the discharge correlates with the rainfall (observation 1) and the rainfall is low in the dry season; the measured low discharges *after* AOI 2 point out very low discharges in the whole river (not a high extraction of river water by AOI 2).
- No clear conclusion about changing crop patterns can be drawn from the observations.

4.2.3 AOI 3: Orós



Available water

The borders of AOI 3 are based on the assumption that this area is mainly supplied by pumped up water from the Orós reservoir. Since the Jaguaribe/Trussu river flows though the AOI *and* the separation between fields supplied by the Orós reservoir and fields supplied by other water sources is not clear (see sub-paragraph 3.3.3), also the discharge of the Jaguaribe/Trussu river will be analyzed. This discharge is composed as the discharge measured in Iguatú (after AOI 2) plus the discharge at the Trussu dam minus the water use of AOI 4. Since only the water use of AOI 4 in the dry season is known, only the discharges of the Jaguaribe/Truss river in the dry season are considered. In 2001 and 2005 the discharge at Iguatú was not measured. Since the discharges at Iguatú in the dry season in the other years are zero (see figure 3.10), 2001 and 2005 are also assumed to have a zero discharge in the dry season.

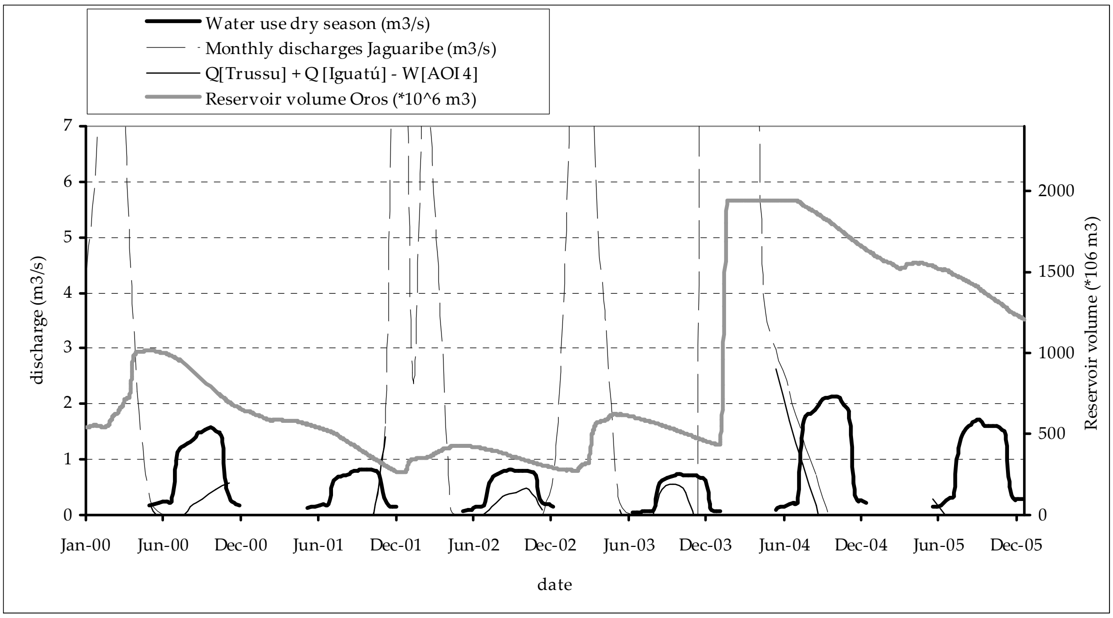


figure 4.7: Irrigation water use and water availability AOI 3

Observations on irrigation water use:

1. The discharge of '$Q_{Iguatú}$ + Q_{Trussu} – $W_{AOI\ 4}$' (W = 'water use') is always lower than the irrigation water use. In 2002 and 2003 this discharge is relatively high.
2. A high correlation between reservoir volume and irrigation water use can be concluded from figure 4.7: the more volume, the more irrigation water use.

Observation on land use

3. As figure 4.8 shows, the variation of total area of land use is dominated by the variation in paddy area
4. In comparison to Paddy, a constant small, area of banana is grown.

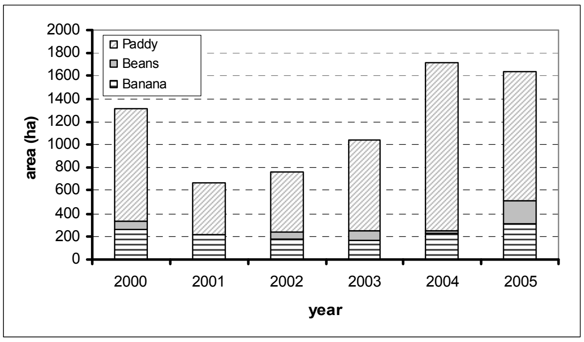


figure 4.8: Land use AOI 3

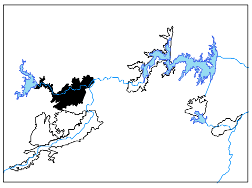
Interpretation of observations

- Observation 2 and the lower discharge of the Jaguaribe/Trussu river than the water use in AOI 3 (observation 1), support the hypothesis that this AOI is supplied by water from the reservoir.
- Taking figure 3.7 in consideration, it can be seen that upstream from AOI 3 also much agricultural land use occur. Especially in 2001-2003, due to the small Orós reservoir, much area is available between the point were the Jaguaribe/Trussu begins and AOI 3. Discharges from the Trussu dam to supply these areas are therefore a suitable explanation for observation 1.
- The correlation, as described in observation 2, can be explained by the 'water level – available land'-relationship from sub-paragraph 3.4.3: Higher volumes cause higher

water levels -> higher water levels cause more land availability -> more land availability causes more agricultural land use and irrigation water use.

- The reservoir edge is constantly changing, and therefore the location on which crops can be grown. So together with the large water availability from the reservoir, this makes the use of a temporal and high water intensity demanding crop as paddy logical (according to observation 3).

4.2.4 AOI 4: Trussu



Available water

Since AOI 4 is defined as an area supplied by released water from the Trussu reservoir, this source will be considered as available water. Also the effective rainfall in the dry season and the total annual rainfall (as derivative of locally stored runoff) will be analyzed. In figure 4.9 these three possible water sources are drawn, together with the irrigation water use. The total yearly values are drawn in figure 4.10.

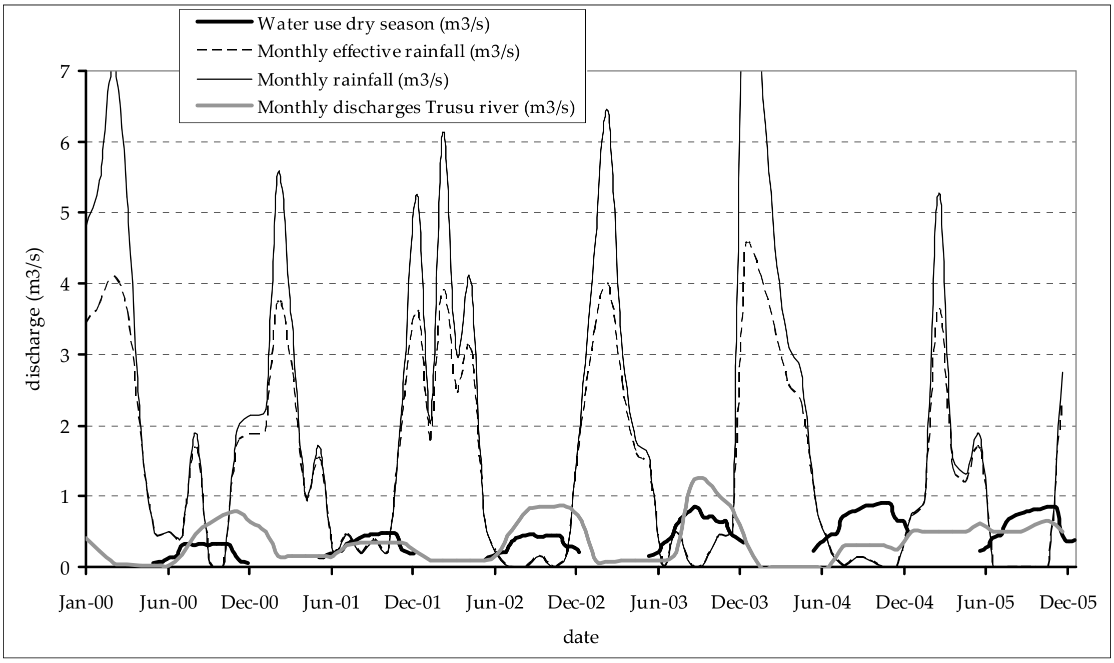


figure 4.9: Irrigation water use and water availability AOI 4

Observations on irrigation water use (based on figure 4.9 *and* figure 4.10*):*

1. From 2000 to 2005, the released water is higher or almost equal (in 2001 and 2005) to the water use, except for 2004.
2. In 2004 a very low release from the Trussu reservoir is combined with a very high annual rainfall.
3. No correlation between effective rainfall in the dry season and irrigation water use in the dry season can be concluded.
4. No correlation between rainfall in the wet season and irrigation water use in the dry season can be concluded.

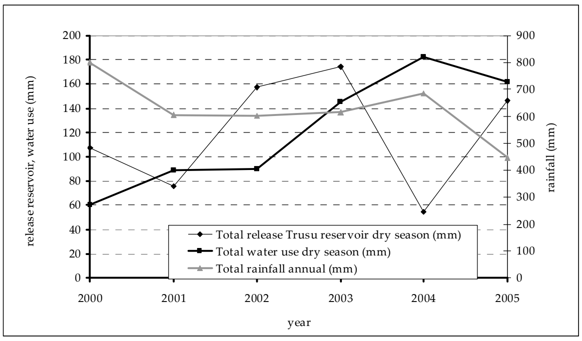


figure 4.10: Water sources and irr. water use AOI 4

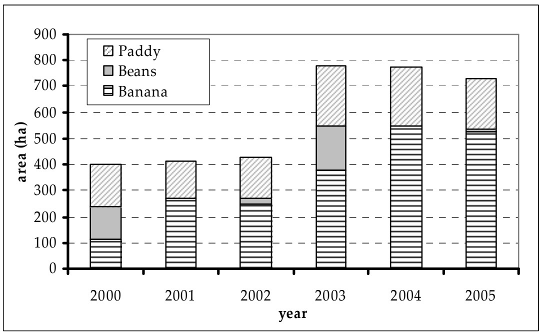


figure 4.11: Land use AOI 4

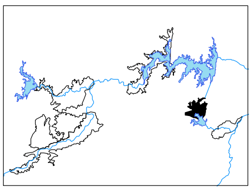
Observation on land use (based on figure 4.11*)*

5. The area of banana is increasing from 2000 to 2004
6. The area of paddy is constant.
7. From 2002 to 2003 an increase in the total agricultural area is visible.
8. 2000 and 2003 show a large area of beans in comparison to other years.

Interpretation of observations

- The hypothesis that AOI 4 is mainly supplied by release water from the Trussu reservoir is supported by the data (observation 1 and 2). The reason for the outlier of 2004 can be either that much water was locally stored in the soil due to the high rainfall in 2004 or the Trussu dam had a overflow which was not counted as release.
- The higher release than water use in some years (2002 and 2003) can be linked to the agricultural land use in the area between AOI 4 and AOI 3. In these years the Orós reservoir was small and much agricultural land use can be seen in this area (figure 3.7). 2001, also a year with a small Orós reservoir shows less agricultural land use in this area and also a lower discharge from the Trussu dam (figure 4.10).
- The higher irrigation water use in 2003-2005 is caused by more agricultural area and more banana area, which has the highest seasonal water use of the three considered crops (observation 6).
- Observation 6 to 9 cannot be explained in terms of water availability.

4.2.5 AOI 5: Lima Campos



Available water/land

In figure 4.12 the hypothesized sources of water availability are drawn; tunnel release, together with the irrigation water use. No storage for locally stored runoff (wells, lakes) have been seen during the field visit, so annual rainfall will not be further considered. The reservoir volume is an indication for the land availability as described in paragraph 3.4.3: the higher the water level of the Lima Campos reservoir, the less land available for agricultural land use. The water supply from the tunnel is influenced in the other way around by the Orós reservoir: The higher the water level of the Orós reservoir, the more water supplied.

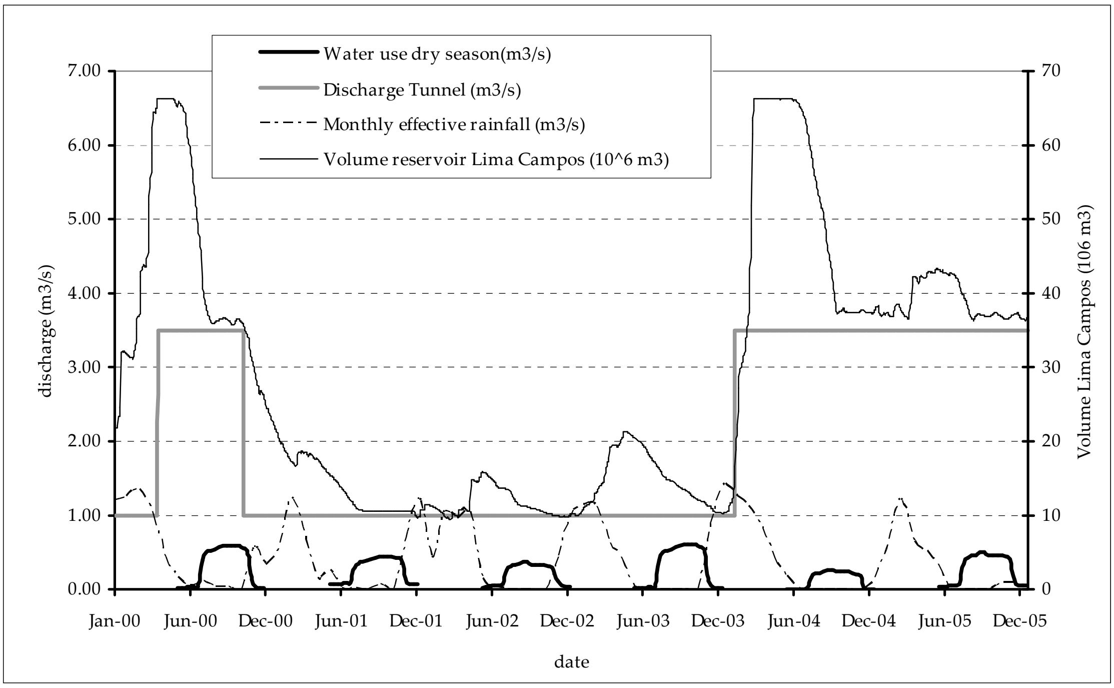


figure 4.12: Irrigation water use and water availability AOI 5

Observations on irrigation water use (based on figure 4.12*)*

1. Water supply by the tunnel is always much higher than the irrigation water use.
2. No correlation between tunnel release and irrigation water use can be concluded: years with the same tunnel release show different water uses and visa versa.
3. No correlation between reservoir volume and irrigation water use can be concluded: years with the same irrigation water use (2000 and 2003) show completely different reservoir volumes and years with the same reservoir volumes (2001 to 2003) show different irrigation water uses.

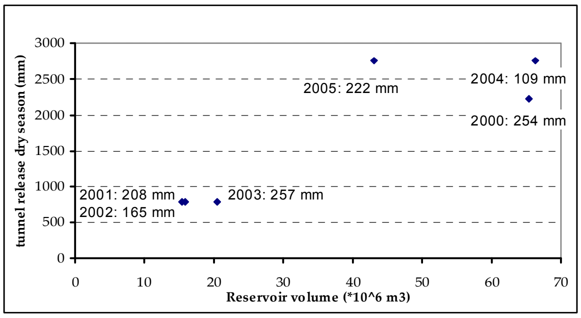


figure 4.13: Irr. water use at combinations of reservoir volume and tunnel release

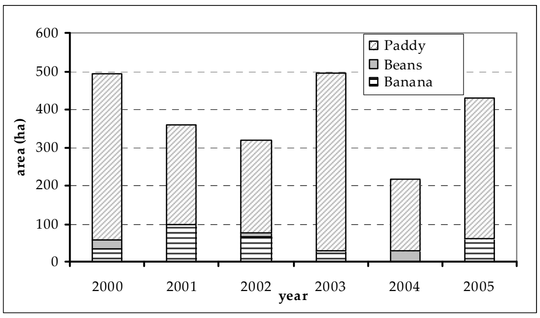


figure 4.14: Land use AOI 5

4. To get more insight in how the reservoir volume of Lima Campos and the tunnel release are correlated to the irrigation water use, figure 4.13 is drawn. For the occurring combinations of runnel release (in mm over the total AOI) and reservoir volume, the irrigation water use is shown. It can be concluded from this figure that the reservoir volume and the tunnel release are correlated for the years 2000 to 2005: High reservoir

volumes (low land availability) are combined with high tunnel releases (much water availability) and low reservoir volumes (high land availability) are combined with low tunnel release (low water availability). No relation to water use can be linked to this correlation (see labels of points)

Observation on land use (based on figure 4.14*)*

5. Paddy is the main crop in AOI 5. The variation in land use is also mainly determined by the variation in areas of paddy

Interpretation of observations

- Since the water supply by the tunnel is always much more than the required irrigation water use (observation 1), the water availability for AOI 5 seems not to constrain water use. Observation 1 can therefore be the reason that no correlation between irrigation water use and tunnel release can be seen (observation 2). Another possibility is that in cases of low tunnel release, the use of tunnel water is restricted by government in favor of downstream water users. To check this, figure 4.15 is drawn. The supposed restricted water availability of the tunnel is drawn by substracting the release from the Lima Campos dam from the Tunnel discharge. By considering figure 4.15, the governmental restriction can not be proved.

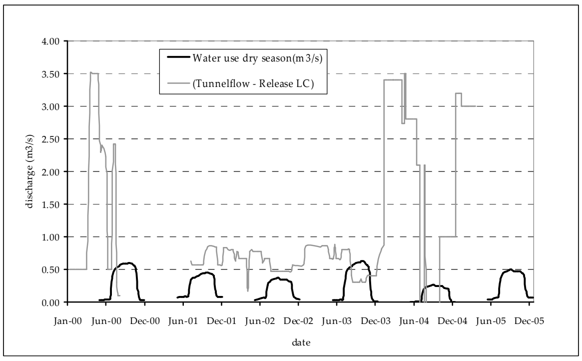
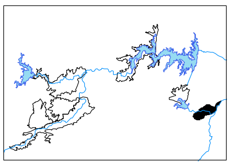


figure 4.15: Water use AOI 5 vs 'tunnel flow *minus* release Lima Campos'

- The hypothesis that lower reservoir volume (more available land) should cause higher land use and water use is not confirmed by the data. It is expected that the relation in observation 4; high land availability is combined with low water availability (possibly limited by political decisions as mentions above), is the cause of this.
- An optimal situation is expected in which the levels of the Orós reservoir and the Lima Campos reservoir are in a way that a high tunnel release is assured at a minimum water level of the Lima Campos. The year 2005 (see figure 4.13) seems to satisfy these requirements, but has a mediate land use and water use. So concluding from the available data, this expectation can not be confirmed.

4.2.6 AOI 6: Perímetro



Available water

The main types of water availability for AOI 6 (Lima Campos reservoir release and volume and effective rainfall) are drawn in figure 4.16, together with the irrigation water use. Data of Lima Campos reservoir release was not available between Jul-2000 and Jun-2001 and after Apr-2005.

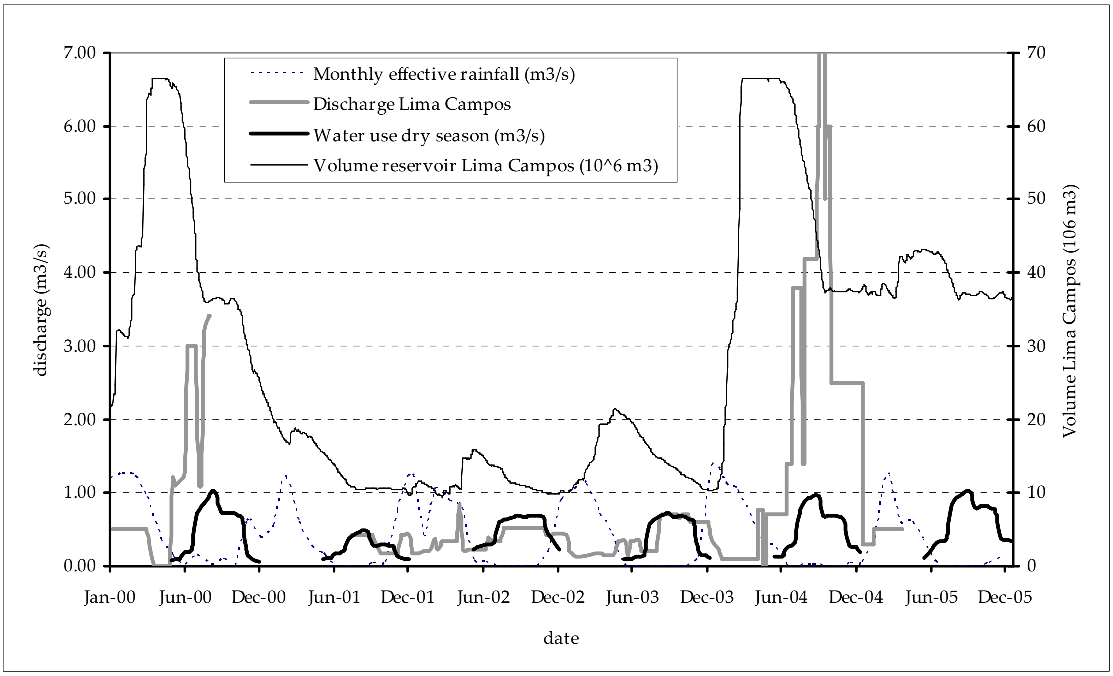


figure 4.16: Irrigation water use and water availability AOI 6

Observations on irrigation water use (based on figure 4.16*)*

1. The irrigation water use and the water availability from discharge from the Lima Campos reservoir are correlated in 2001 – 2003.
2. The years with a high reservoir volumes (2000 and 2004) show higher discharges from the Lima Campos reservoir than years with low volumes (2001 – 2003)
3. Years with a high discharge (2000 and 2004) show a higher irrigation water use than years with low discharges.

Observation on land use

4. In figure 4.17 a relatively constant area with banana between 2000 – 2004 is shown.
5. The area with paddy is very low in 2001 and 2002, both years with low reservoir volumes.
6. years with high irrigation water uses (2000, 2004, 2005) show high areas with beans.

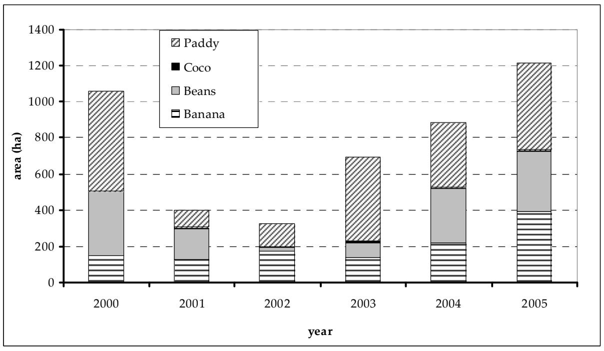


figure 4.17: Land use AOI 6

Interpretation of observations

- The hypothesis that AOI 6 is supplied by released water from the Lima Campos reservoir is confirmed by the data (observations 1-3).
- Since paddy and beans require the high intensities of supplied water (see paragraph 3.6) the low areas with these crops (observation 5) seems to be connected with the low volume and release of water from the Lima Campos reservoir in these years

4.2.7 *Summarizing table*

AOI	*Hypothesis:* AOI is mainly dependent for its water availability of:	*Data interpretation:* Correlation water/land availability – irrigation water use	Hypothesis confirmed by data interpretation?	*Data interpretation:* Correlation water/land availability – land use
1	Locally stored runoff	R_{wet} ↑ → W ↑ R_{wet} ↑↑→lakes↑ → W ↓	Yes	R_{wet} ↑ → Pa ↑ in same year Hypothesis: 2 years R_{wet} ↓→Ba↑
2	$Q_{Jaguaribe}$	R_{wet} ↑ → W ↑	No	-
3	$V_{Orós}$	$V_{Orós}$ ↑ → L ↑ L ↑ → W ↑	Yes	$V_{Orós}$ ↑ → L ↑→ Pa ↑
4	Q_{Trussu}	Q_{Trussu} ↑→ W ↑	Yes	-
5	Q_{Tunnel}	-	No	-
6	Q_{LC}	V_{LC}↑ → Q_{LC} ↑ Q_{LC} ↑ → W ↑	Yes	Q_{LC} ↑ → Pa,Be ↑

table 4.1: Summary correlations between water availability and land use and agricultural water use

Legend of table 4.1

R_{wet} = Rainfall in wet season (mm)
W = Irrigation water use (mm)
$V_{Orós}$ = Volume in Orós reservoir (m^3)
V_{LC} = Volume in Lima Campos reservoir (m^3)
L = Total agricultural land use (m^2)
$Q_{Jaguaribe}$ = Discharge in Jaguaribe river (m^3/s)
Q_{Trussu} = Release from Trussu reservoir (m^3/s)
Q_{LC} = Release from Lima Campos reservoir (m^3/s)
Pa = Area of paddy (m^2)
Ba = Area of banana (m^2)
Be = Area of beans (m^2)
↑ = High
↓ = Low
→ = 'Is correlated to'

4.3 *Inter-AOI analysis*

4.3.1 AOI 1

As mentioned before, AOI 1 is mainly supplied by locally stored runoff, coming from rainfall in the wet season. According to field observations (wells and natural lakes) it is assumed that this water is temporally stored in the upper layer of the soil and has a very slow downwards movement, due to gravitational forces. Since the discharges in the rivers are very low in the dry seasons (see figure 3.10), it is assumed that this water does not give a noticeable increase in discharges in the dry season. Therefore, the water use of this locally stored water does not influence the water availability of downstream users during the dry season: no influence on the water availability of users of other AOI's can be expected.
In case AOI 1 would not use the locally stored runoff, it is unknown were the water would go to. It can feed an aquifer by percolation and flow to the Jaguaribe river, (if the Jaguaribe intersect this aquifer). In case this will happen, the water will become available for other AOI's. This will be in any case later than in the dry season after the wet season which provides the considered water since the discharge in the Jaguaribe is very low in the dry season. In case it reaches the river in the next wet season, it can contribute to the total discharge in the Jaguaribe Since knowledge about aquifers is not available no final conclusions can be given on how water use in the dry season of year x influences the water availability in year x+n.

4.3.2 AOI 2

As concluded in sub-paragraph 4.2.2, AOI 2 is also dependent of locally stored runoff. Therefore, the same story as AOI 1, as described in sub-paragraph 4.3.1, can be applied on AOI 2. This similarity in behavior is confirmed by figure 4.18.

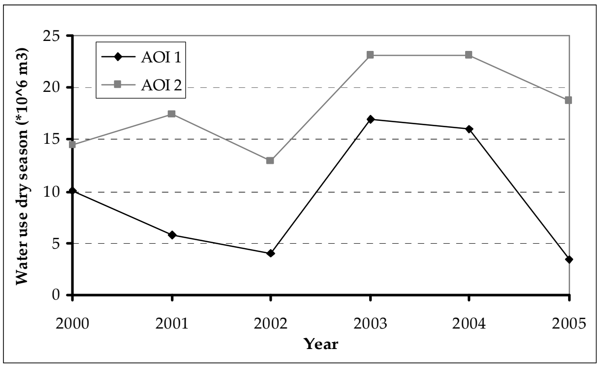


figure 4.18: Water use AOI 1 and AOI 2

4.3.3 AOI 4

[The AOI's will be discussed one by one, in downstream direction. Therefore, AOI 4 will be discussed before AOI 3].
In the former paragraph it is concluded that almost all released water from the Trussu reservoir is used by the downstream water users until AOI 3. The amount of released water seems to be controlled by the demand of these farmers. This observation is confirmed by conversations with local water agencies (Cogerh Iguatú, 2006). Only in case the reservoir is full (as in 2004, see figure 3.11), more water than the downstream farmers from the Trussu dam require is released (overflow). This causes a higher level in the Orós reservoir. A higher level in the Orós reservoir means an advantage for the land availability of AOI 3. It also means a higher level in the Lima Campos (when the level of the Orós reservoir is above the tunnel inlet), which means less land availability for farmers in AOI 5.
But, in cases the Trussu reservoir released more water than used by the farmers downstream from the Trussu dam, due to overflow, the Orós reservoir was also at its maximum during the research period (see figure 3.11). In that case, the water level of the Orós reservoir does not change, only the overflow at Orós increases. So in the observed situations during the research period no influences of AOI 4 on downstream AOI's occur.

This is true in case the existence of the Trussu reservoir is taken as a starting point. In case the Trussu reservoir did not exist, all runoff that is stored in the Trussu reservoir should be flow into the Orós reservoir. The existence of the Trussu reservoir is taken as starting-point in this research, so this option will not be further discussed.

4.3.4 AOI 3

The water use of AOI 3 is subtracted from the Orós reservoir and is therefore not available for downstream AOI's. To get more insight in the proportion of water subtraction from the reservoir AOI 3, one has to consider the values in table 4.2. The water use of AOI 3 is very small in comparison with the volume in the Orós reservoir. In a dry year, when the reservoir contains 15% of its capacity (see figure 3.11), the water use in that year causes a decrease in water level of only 0.2 m. This value is calculated by using the 'water level-volume'-relationship at the moment of the construction of the dam, as showed in figure 4.19 (DNOCS, 1982).

Capacity of Orór reservoir	$1949 \cdot 10^6$ m^3
Water use of AOI 3 in dry year (2003)	$7{,}66 \cdot 10^6$ m^3 (0.4% of capacity Orós)
Water use of AOI 3 in wet year (2004)	$22{,}0 \cdot 10^6$ m^3 (1.1% of capacity Orós)
Decrease water level Orós at 0.15*capacity due to extraction of $7.66 \cdot 10^6$ m^3 (as in 2003)	0,2 m
Decrease water level Orós at 1.0*capacity due to extraction of $22{,}0 \cdot 10^6$ m^3 (as in 2004)	0,1 m

table 4.2: Characteristics of water extraction from Orós reservoir by AOI 3

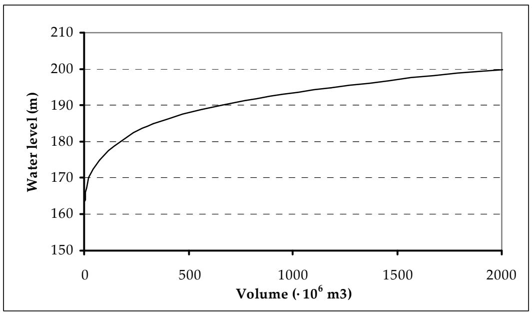


figure 4.19: Waterlevel-volume relationship Orós reservoir (DNOCS, 1982)

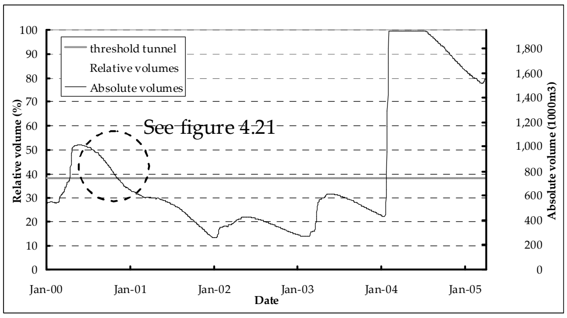


figure 4.20: Threshold tunnel

As mentioned before, the water supply from the Orós reservoir to the Lima Campos reservoir is performed by a tunnel. This supply is dependent on the water level of the Orós reservoir; as the level is under the tunnel inlet, the discharge is much lower (3,5 m^3/s -> 1,0 m^3/s). This level is called the threshold level. In figure 4.20 the threshold level is shown. In the past 5 years, the volume was three times for a short moment at a level close to this threshold level. In figure 4.21 one of these moments is zoomed out. The difference between a reservoir level with and without the total water use of 2000 is drawn. As a result of this difference, the threshold is only 4 days earlier intersected. The influence on the total water availability on the whole dry season is therefore very small.

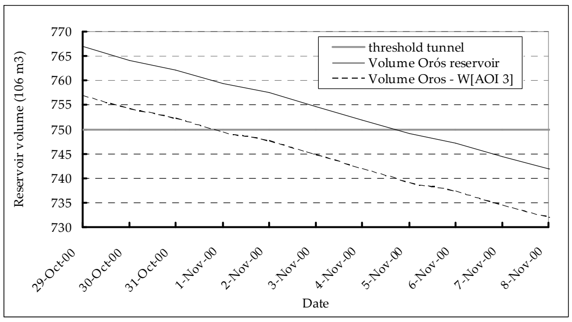


figure 4.21: Influence of water use AOI 3 on water level around tunnel inlet

An additional argument is that the calculated decrease in water level is calculated with the water use of the whole dry season: this is the total decrease of water level which occurs at the end of the dry season. Before this moment, the water level decrease is even smaller.
As a result of these arguments, it is concluded that the water use of AOI 3 does not influence other AOI's.

4.3.5 AOI 5

It is already explained how the water use of AOI 5 is influenced by land availability and water availability (see sub-paragraph 3.4.3). Unfortunately, no exact data is available about the operation of the tunnel. It is known that the discharge can be regulated only in a downstream direction; from the Orós reservoir to the Lima Campos reservoir. When water is subtracted from the Orós reservoir, the water level will decrease and less land will be available for agricultural land use in AOI 3. The water use of AOI 6 intensifies this. The impact of water use of AOI 5 and 6 on the land availability of AOI 3 is shown in table 4.3. By using the 'waterlevel-volume' relationship, the water levels of the different years could be determined. By using the 'waterlevel-available land' relationship (see figure 3.12) the available land could be determined. Since the DEM, with which the land availability is determined, is only usable for water elevations above 190 m (probably this was the water elevation at the acquisition date of the DEM), not all years are shown in table 4.3

year	Water use AOI 5+6 (*10^6 m3)	Volume Orós: Perc. of capacity (%)	Decrease water level Orós reservoir due to Water use AOI 5+6 (m)	Difference in land avail AOI 3. (ha)	Relative difference in land avail. AOI 3. (%)
2000	15.35	52	0.13	-130	-3.10
2004	12.58	100	0.07	-28	-0.39
2005	16.79	80	0.10	-45	-0.68

table 4.3: Impact water use of AOI 5+6 on land availability AOI 3

Concluding from table 4.3, the impact of the water use of AOI 5 and 6 on the land availability in the same year of AOI 3 is small. When it is considered that only a part of the available land is used for agricultural land use, the impact on the land use will even be smaller.

The impact of water use of AOI 4, 5 and 6 in the *wet* season on the land availability of AOI 3 in the *dry* season can have also impact. Since this falls outside this study, this will not be discussed.

The water used by AOI 5 cannot be used by the downstream AOI 6. In the observed period from 2000 to 2005, the Lima Campos reservoir never reached its dead volume (at 4% of its capacity (COGERH, 2006)), but 2001 till 2003 show low volumes (0,1 to 0,3 * capacity, see figure 3.11). In these dry years, the volume is a little increased due to the rainfall in the wet season and the supply by the tunnel is small in these years. An opposite behavior of water use of AOI 5 and 6 is shown in figure 4.22 in this period. The scarce water from the tunnel seems to be mainly used by AOI 5; this causes a conflict with the water requirement of AOI 6, which cannot irrigate the whole irrigation scheme (see appendix 4.4, figure 4.8). In case AOI 5 should use no water from the tunnel, AOI 6 would have enough; 1 m^3/s is enough to irrigate the whole irrigation scheme, as shown in figure 4.16. A regulation to protect the interests of AOI 6 is recommended.

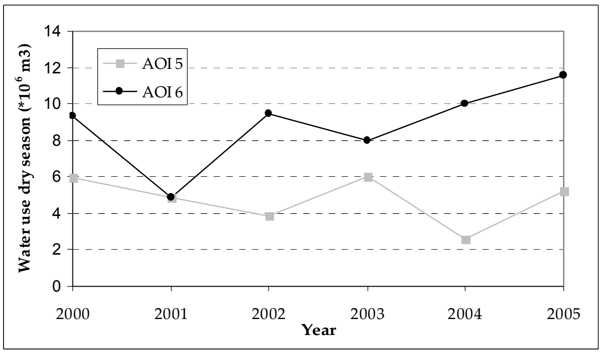


figure 4.22: Water uses dry season of AOI 5 and 6

4.3.6 AOI 6

The more AOI 6 consumes, the lower the Lima Campos reservoir level, the higher the land availability for AOI 5. So a high release from the Lima Campos reservoir destined for the water use of AOI 6 is advantageous for AOI 5. Unfortunately, it is not possible to make a water level-land availability'-relationship for AOI 5, since the DEM shows strange behavior in this area: peaks, which should be above the water level, are shown in the DEM, while these peaks were not observed during the field visit. It was already concluded in the former sub-paragraph that the water use of AOI 6 and AOI 5 has a little influence on the water availability of AOI 3.

4.3.7 AOI 4, 5 and 6 cumulative

As a result of the analysis of the former sub-paragraphs, it can be concluded that the Orós reservoir level is the most important variable for the water availability of AOI 3. It is also concluded that AOI 1 and 2 do not have a noticeable effect on downstream water availability and that AOI 4, 5 and 6 have a very small effect on the Orós reservoir volume in the same year. In this sub-paragraph the effect of AOI 4, 5 and 6 *together* over 6 years on the volume of the Orós reservoir will be considered This is called the 'cumulative effect'.

In the hypothetical situation that AOI 4, 5 and 6 would have a water use of zero, a cumulative effect over more than one year should be noticeable in the volume of the Orós reservoir: the water that would be used by these AOI's would have flow into the Orós reservoir. In figure 4.23 the water uses of AOI 4, 5 and 6 are summed per year; the cumulative effect on the reservoir volume over 6 years is add to the measured reservoir volume. To assess the differences, it is assumed that the release strategy remained the same. The effect on

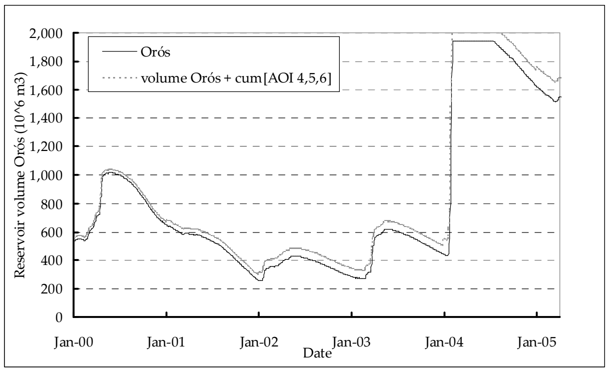


figure 4.23: Cumulative effect of water use AOI 4,5 and 6 on reservoir volume Orós

land availability in 2005 (in march with a volume of 1500 * 10^6 m^3) is shown in table 4.4: + 7.0 % (the values in this table are determined by using the 'volume-waterlevel'-relationship and the 'waterlevel-availableland'-relationship). When the volume in march 2005 would have been as low as in 2002 (300 * 10^6 m3) the relative effect on land availability would have been much larger. Unfortunately, the available land cannot be estimated by water levels under the 194 m (because at the acquisition data of the DEM the reservoir had a level of around 194 m). 194 m corresponds with a volume of 1050 * 10^6 m^3. This is the minimum volume at which the effect on available land can be determined. The cumulative effect of 6 years on the available land for this volume is: + 8.7% (see table 4.4).

In case AOI 4, 5 and 6 should use twice as much water, the volume of the Orós and the available land would be less. This will be relatively in the same order as in the case AOI 4, 5 and 6 do not use any water (as discussed above), since the volume-waterlevel'-relationship and the 'waterlevel-availableland'-relationship can be considered as linear between a reservoir volume difference of more or less 150 10^6 m^3.

	Volume Orós (10^6 m^3)	Water level (m)	Available land (ha)
March 2005	1500	197.1	7450
incl. cumulative effect of 6 years	1632	197.9	7670
difference (%)	+132 (9.6)	+0.8 (0.4)	+220 (7.0)
Minimum level at which land availability can be determined	1050	194	6090
incl. cumulative effect of 6 years	1183	195	6620
difference (%)	+132 (13)	+1.0 (0.5)	+530 (8.7)

table 4.4: Cumulative effect

Although the impact of this cumulative effect have only less than 10% effect over 6 year, on a long term (>10 years) it can have large impact, especially in dry years. When the reservoir reaches, due to a wet year, its capacity, the cumulative effect returns to zero. So when the time between two succeeding points in which the capacity of the reservoir was reached is usually small (around 4 years), the cumulative effect can not 'grow'. How long this time usually is, has been checked in the available data. Only reservoir data between 1986 and 2005 was available (COGERH, 2006). In this period, the maximum time between two succeeding points in which the capacity of the reservoir was reached, was 13 year (from 1991 to 2004). So long term effects of water use around the reservoir can have a large impact on the volume of the Orós reservoir and therefore on the water users, given the release strategy of the last 20 years. In reality the evaporation can play an important role as well. The amount of evaporation water is, among other factors, dependent on the geometry of the reservoir which influences the evaporation surface and the sensibility on temperature changes, both important

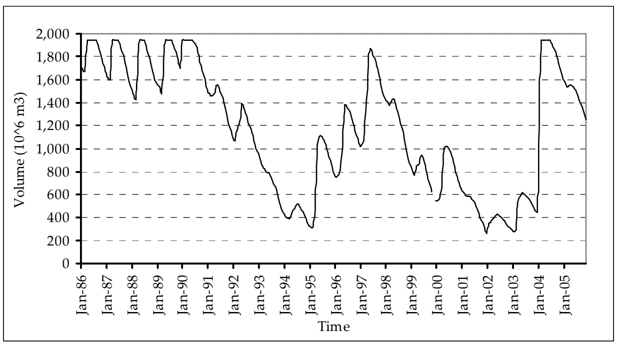


figure 4.24: Volume Orós reservoir 1986-2005 (COGERH, 2006)

variables for the computation of the evaporation (Shaw, 1994). A detailed analysis of this factors is left outside this research. It is expected that this factors will not change the raw relations as discussed above.

4.3.8 Summary of inter-AOI analysis

- The water uses of AOI 1 and AOI 2 are not influencing the water availabilities of other AOI's during the dry season. On longer term, the influence is unknown, due to lack of knowledge about groundwater and aquifers.
- The influences of AOI 4, 5 and 6 on the reservoir volume of the Orós reservoir, and therefore on the land availability of AOI 3 in the same dry season, are very small.
- There is a small noticeable cumulative effect of water use of AOI 4, 5 and 6 on the reservoir volume of the Orós reservoir, and therefore the land availability of AOI 3 *within the research period*. Considering a period of 20 year, under the given release strategy and climate and the same irrigation water use as in the research period, a large cumulative effect can be seen.
- The influence of the water use of AOI 3 on the discharge through the tunnel is very small. Therefore, water use of AOI 3 does not influence downstream water availability.
- The water use of AOI 5 is within the same year in conflict with the water availability of AOI 6.
- Water use in AOI 6 has a positive influence on the land availability of AOI 5 and water use in AOI 5 itself too.

5 Discussion

5.1 *Introduction*

This chapter deals with the uncertainties of the results presented in chapter 4. During the description of the different components of this results, the main uncertainties are already discussed. These uncertainties will be listed in paragraph 5.2. Also the mutually coherency of the different factors of uncertainty will be discussed.

5.2 *Uncertainty and impact analysis*

In this paragraph the main uncertainties in each important step in the research will be discussed. In figure 5.1 these steps and uncertainties are listed. Each uncertainty has consequences for the next step, and finally for the inter-annual analysis and the inter-AOI analysis.

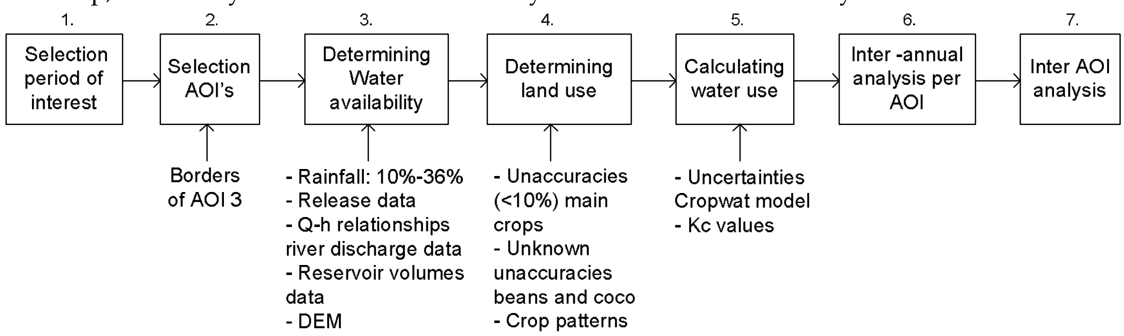


figure 5.1: Research steps and accompanying uncertainties

5.2.1 *Ad step 2: Selection AOI's*

As mentioned in sub-paragraph 3.3.3, the separation between agricultural fields supplied by reservoir water and other water sources is questionable in AOI 3 at the point where the Jaguaribe river enters the reservoir. It is hard to determine the border which separates the fields irrigated by reservoir water and water from the river. To determine the impact of this uncertain border on the area of land use and calculated irrigation water use in AOI 3, the border in this area is displaced for three scenarios (see figure 5.2). One scenario determined by the rules of sub-paragraph 3.3.3, one scenario as a overestimation of these rules and one scenario as an underestimation of these rules. The scenario's are respectively called 'medium', 'large' and 'small'. In appendix 4.4, figure 4.5, it can be seen that determining the border in a year with a high reservoir volume is harder than in a year with a low reservoir volume; the uncertainty for these years is the highest. The chosen year to calculate a representative variation in land use and irrigation water use as a consequence of the uncertain border for all years is a therefore a year with a high reservoir volume: 2005..

The variation around the medium scenario is +/- 250 ha (= +/-17%) total agricultural area. The calculated variation of the water use around the medium alternative is +16% , -17% (Large: $25.4 \cdot 10^6$ m^2, Medium: $21.9 \cdot 10^6$ m^2, Small: $18.3 \cdot 10^6$ m^2.

The difference in water use between years with much volume and little volume in the Orós reservoir is more than 100% (see figure 4.7). So the impact of the uncertain border of AOI 3 at the inlet point of the reservoir on the analysis of dependency between reservoir volume and water does not changes the drawn conclusions in the former paragraphs.

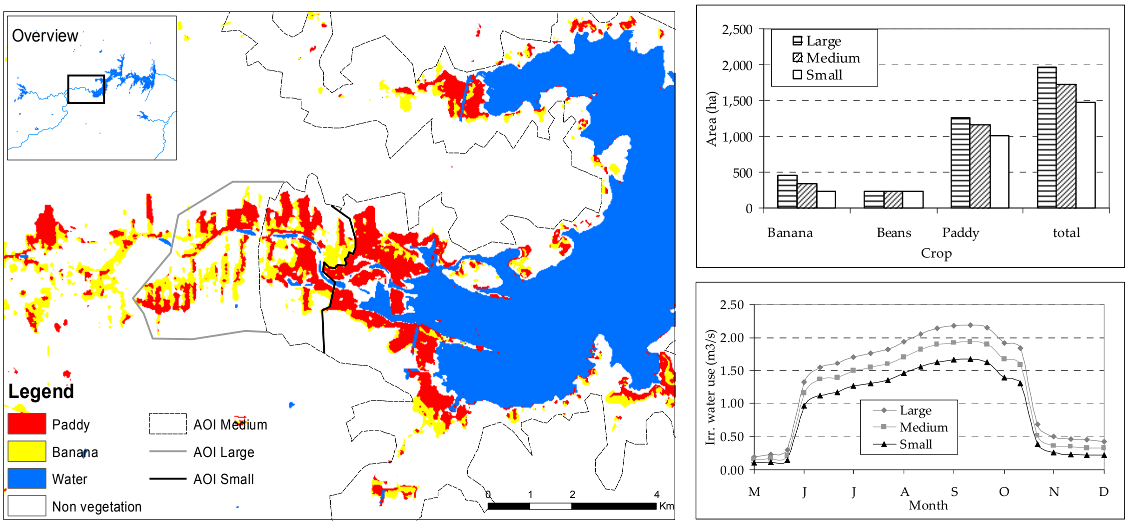


figure 5.2: Varying border of AOI 3 in 2005 (left), land use (upper right) and irr. water use (lower right)

5.2.2 *Ad step 3: Determining water availability*

Rainfall data

The unreliability of the rainfall data is assumed to be less than 10% (see sub-paragraph 3.4.1). The water availability data was important for proving the relation between rainfall in the wet season and locally stored runoff. Since only a trend is analyzed, the exact values of rainfall were not important. It is therefore assumed that the maximum unreliability of the rainfall data of 10% will not influence the conclusions about the relation between rainfall and locally stored runoff in AOI 1 and 2.

Discharge data

The release data has a reliability of <10%, see sub-paragraph 3.4.2. It is expected that this unreliability will not influence the conclusions based on this data: the water use of AOI 4 and AOI 6 is correlated to the reservoir releases.

Since the natural rivers are usually dried out in the dry season and none of the considered AOI's is supposed to use water from the natural rivers, the unreliability of the Q-h relations has no influence on the reliability of the results.

Reservoir volumes

Due to expected sedimentation the data about reservoir volumes, which are likely determined by using a 'water level-volume'-relationship, have uncertainties. Since the research period is 6 years and sedimentation is a long term process, it is assumed that the reservoir beds did not remarkably change during the research period. Since the reservoir data is only used for the analysis of trends, the sedimentation has no influence on the conclusions.

Land availability AOI 3:Orós

The water level of the Orós reservoir used for determining the available land, was determined by reading the elevation value at the reservoir edge in a crossed satellite image – DEM map. Since the elevation values at the water edge had a variation of +/- 1 meter, this indicates an uncertainty. In sub-paragraph 3.4.3 it is already discussed that this uncertainty of 16% is an upper estimate.

When the mean water level of each image is compared to the water level according to the COGERH (2006) data at the same date a structural difference of around 6 m can be seen. (see table 3.1). This structural difference causes no uncertainty, since only elevation differences are considered. It affirms the variation of +/- 1 meter in elevation values of the DEM.

To get an indication for the reliability of the trend between water level and available land, a trendline is drawn, using a polynomial regression (see figure 5.3), which satisfies the rising trend that was concluded in sub-paragraph 3.4.3. This type of regression gave a very high R^2: 0.94 (1.00 indicates a 100% correlation between a trend line and the considered points). At a water level of 202 m the deviation from the trendline is the largest: 300 ha lower than the trendline (-3,5 %). This deviation is assumed as maximum variation on each point. The effect on the trendline of a deviation is higher when the slope of the succeeding points is low. Therefore, the points at the water levels of 201 – 204, the range with the lowest slope, are decreased with 3,5 % (see figure 5.4) to examine the effect on the trendline. Still the trendline has a rising trend, although this trend is less strong at higher water levels.

Conclusion: For the interpretation of the dynamics in AOI 3 this nuance has no great impacts; still the lower irrigation water uses in years with low reservoir volumes and the high irrigation water uses in years with high reservoir volumes can be explained by this trend (see figure 4.7 in sub-paragraph 4.2.3).

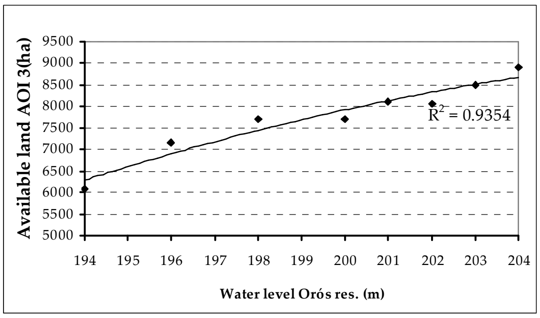


figure 5.3: Trendline 'waterlevel-available land'-relations

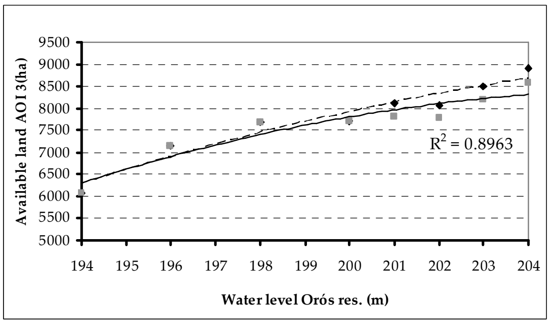


figure 5.4: Trenline after changing values at water level = 201 to 204

5.2.3 Ad step 4: Determining land use

Classes with accuracy assessment

In spite of the fact that the accuracy on the classification of banana and paddy areas is very high, some remarks have to be made. Only clear distinguishable fields joined the ground truth data set. So the high accuracy indicates a good ability of the classification algorithm to assign these clear fields to the right class. It says nothing about the ability of the classification algorithm to assign the unclear fields to the right class. 'Unclear fields' are, for example: banana fields in young stages and with various distances between the trees, very small fields with relatively many mixels and banana fields in a certain grow stage which have overleap in the feature space with paddy in a certain grow stages.

The question is how large this uncertainty is en which influence is has on the results of the calculated irrigation water use and the drawn conclusion after the comparison with the water availability. The uncertainty is anyhow higher than the accuracy values indicates. When the uncertainty in the areas of paddy and banana are estimated on 10%, this causes only 2% uncertainty in the total irrigation water use (see table 5.1). The influence of this uncertainty on the conclusions of this study will be discussed below.

Description	Area Banana (ha)	Area Paddy (ha)	Irr. Water use dry season (m3)
Reference situation	500	500	$15.2 \cdot 10^6$
10% less banana	450 (-10%)	550 (+10%)	$14.9 \cdot 10^6$ (-2%)
10% more banana	550 (+10%)	450(-10%)	$15.5 \cdot 10^6$ (+2%)

table 5.1: Variation in proportion banana and paddy of +/-10% in an area of 1000 ha

Classes without accuracy assessment

The uncertainties of the classes beans and coco are unknown (see sub-paragraph 3.5.3). The question is again: "How much influence has an estimated accuracy on these results of the inter-annual and inter-AOI analysis?". The percentage of area with coco of the total grown area is small (mostly < 5%, see appendix 4.6), so this uncertainty will be neglected.

Beans covers in some cases 30% of the total area. In this cases an uncertainty on this value means that a part of this area can be in reality another crop. According to the feature spaces, this should be banana. The impact on the irrigation water use in the dry season under a large variation of beans of 50% is shown in table 5.2. It is estimated that the uncertainty on the beans area is not larger than this 50%, so the impact on the uncertainty in the total irrigation water use will not be larger than 14%.

Description	Area Banana (ha)	Area Paddy (ha)	Area Beans (ha)	Irr. Water use dry season (m3)
Reference situation	350	350	300	$12.6 \cdot 10^6$
50% less beans	500	350	150 (-50%)	$14.3 \cdot 10^6$ (+14%)
50% more beans	200	350	450 (+50)	$10.9 \cdot 10^6$ (-13%)

table 5.2: Variation in proportion beans and banana under variation of +/-50% beans area. Total area: 1000 ha

In the worst case, the uncertainties of table 5.1 and table 5.2 have to be summed. The uncertainty on the irrigation water use in the dry season is than maximal 16%. The results in paragraph 4.2 and 4.3 are based on a much larger variation in irrigation water use (>50%) between different years.

Conclusion: The uncertainties in land use do not influence the results about relation between trends in water availabilities and irrigation water use.

Drawing conclusions as a result of changing proportions of different crops has an estimated uncertainty of 10% if only banana and paddy are concerned and 50% if beans are concerned.

5.2.4 Ad step 5: Calculating water use

Cropwat gives estimations about the irrigation water requirement with limited input data. This simplification of the reality introduced uncertainties in the outcomes. Also the input data about climate and crops parameters introduce uncertainties. Is has to be mentioned that it was not the goal of this study to calculate the exact values of water use per AOI, but to investigate the inter-annual trends as a result of changing land uses. Since the same system and data uncertainties of Cropwat are applied in all calculations, the estimations of Cropwat about irrigation water use as a result of different land uses are useful in this research.

6 Conclusions and recommendations

6.1 *Introduction*

The conclusions are defined as the answers to the research questions. Considering the findings of the former chapter, these answers are given in paragraph 6.2. In this chapter it has been tried to draw general conclusions, which can also be valuable for other areas with similar climate and spatial conditions. Recommendations for the water allocating policy in the area and for further research will are given in paragraph 6.3.

6.2 *Conclusions*

Research question 1:
What is the annual amount of available water for irrigation in the dry season and which areas have access to it?

Answer:
The amount of available water can be divided in four sources: rainfall, locally stored runoff, river discharges and water from reservoirs. The quantitative results will not be repeated here, they can be read in the different tables and graphs in paragraph 3.4

- Rainfall is negligible in the dry season.
- Availability of *locally stored runoff* at a certain moment is dependent on the relief of the area, the soil type and the amount of rainfall in the preceding period (wet season). This rainfall is the only variable factor.
- River *discharges* can be subdivided in two types: discharge as a consequence of water release from reservoirs and discharge as a consequence of natural rivers, fed by runoff. Discharges from natural rivers give negligible water availability in the dry season. Releases from reservoirs are important for the supply of water for areas downstream of the reservoir dams. In dry periods (e.g. 2001-2003) these water releases are not adequate to satisfy the water requirement for all downstream agricultural fields. In normal and wet years, the water supply is found to be sufficient.
- The amount of stored water in the *reservoirs* is large in comparison to the yearly water use of the agricultural fields which are dependent of this source.

The accessibility to the four sources of water are spatially determined.

- (Effective) rainfall is available for all areas.
- Locally stored runoff in soils is available for areas with soils with high water holding capacity as vertisoils and alluvial soils. When these areas have a relief that is characterized by local depressions, locally stored runoff in natural lakes is an important water source.
- River discharges from natural rivers are not available, since these discharges are negligible in the dry season.
- Accessibility to released reservoir water is connected to the capacity of the pumps (how high can the water pumped up) and in case there is a fixed irrigation scheme by canals.

- Accessibility to reservoir water is also determined by the capacity of the pumps. The area that falls within the range of these pumps is defined as 'available land'. The available land is variable under a changing water level of the reservoirs.

The spatial distribution of water availability is downscaled by selecting 6 areas of interest (AOI's). The borders of the AOI's are based on hypothesis's about water availability: each AOI is mainly dependent of one of the water availability sources. This hypothesis's are validated or invalidated by in research question 4.

Research question 2:
How much land do the different irrigated agricultural land uses cover in the dry season and what is their spatial distribution?

Answer:
The areas of different crops are determined per AOI by applying a land cover classification of satellite images of each year. The main crops grown in the area are paddy, banana, beans and coco. The accuracy of the classification of banana and paddy (the two main crops in the area) was very high. The accuracy of beans could not be determined, since not enough sample points were available. The main reason for this were the late acquisition dates of the images, which were to late to see beans present on the fields. To determine the area of beans, the correlation with grass areas (supposed to be old beans fields) is used to estimate the area of beans. When numerical data was available (in case of AOI 6) this data was used. The accuracy of coco could not be determined since very few coco fields are present in the area. Coco was neglected in case no numerical data was available, since the area of coco is very small in comparison to other crops. The final areas are listed in appendix 4.6. The differences in crops between different AOI's are describing the spatial distributions in each year. Along the Jaguaribe river and downstream from the Trussu reservoir until the joining point with the Jaguaribe river. mainly bananas are grown. On the reservoir edges of the Orós and Lima Campos reservoir mainly paddy is grown. In the irrigation scheme downstream from the Lima Campos reservoir, a mixture of crops is grown with as the three most important crops: paddy, banana and beans.

Research question 3:
What are the amounts of irrigation water use, related to land use, in the dry season?

Answer:
The amount of irrigation water use for each AOI in each year is calculated by using the Cropwat model. The irrigation water use per AOI in each year is shown in figure 3.19 and table 3.15. The water uses show a high variation between different AOI's ($3 - 23 \cdot 10^6$ m^3) and high yearly variation within each particular AOI (AOI 3: $8 - 22 \cdot 10^6$ m^3).

Research question 4:
How are water availability, agricultural land use and irrigation water use related between 2000 and 2005 in different focus areas in the research area?

Answer:
(These conclusions are also represented in table 4.1)
The results of research question 1 to 3 for the years 2000 to 2005 are drawn per AOI to find out what the inter-annual dynamics of each AOI were (paragraph 4.2). The conclusions are:

4.1. Areas that are supplied by locally stored runoff (AOI 1 and 2) are mainly dependent on the rainfall in the directly preceding wet season. The amount of locally stored runoff is determined by the quantity and the intensity of this rainfall. In very wet years, the rising water level of the natural lakes in the center of the bowls in AOI 1 causes decreasing land availability and therefore lower water use than in wet years without filling of the natural lakes. The crop distribution changed after the dry period of 2001 and 2002. More banana was planted and less rice. Within individual years, a good wet season seems to cause farmers to plant more rice.

4.2. Areas that are supplied by water pumped from a reservoir (in this case the water users at the upper edge of the Orós reservoir; AOI 3) are mainly influenced by the relation between water level and land availability. The relation is determined by the relief of the surrounding area and limited by the suitability of the ground around the reservoir. In case of the water users at the upper edge of the Orós reservoir during the research period, the relation is: ‘the higher the water level, the more land available’. At higher water levels, this effects becomes lower. The amount of available water was not an issue in AOI 3, since this was, even in dry years, much larger than the water used in the dry season. Due to mostly the changing reservoir edge and also the large amount of available water, paddy is grown most.

4.3. The water use of areas downstream from a reservoir dam are strongly correlated to the water that is released from the reservoir dam (AOI 4, 5 and 6). At the beginning of the dry season, the quantity of water to be released is a political decision. In case of AOI 5, the land availability, determined by the level of the Lima Campos reservoir, also plays an important role for the land use and the linked water use: the higher the water level of the Lima Campos reservoir, the less land is available for agricultural land use. The water release from the Orós reservoir to AOI 5 (through the tunnel) is limited when the level of the Orós reservoir is low.

There seems to be a tendency to increase the area with banana. Since this is a crop which requires a somewhat less intense and more constant water supply than paddy (the other main crop in the area) This seems to be a good decision for areas dependent on releases from reservoirs with fixed areas of available land.

Research question 5:

How are spatially explicit water availability, agricultural land use and irrigation water use of different focus areas mutually related in the research area between 2000 and 2005?

Answer

(This research question is mainly dealt in paragraph 4.3)

5.1. The water uses in the dry season of areas supplied by locally stored runoff are not influencing the water availabilities of other AOI’s in the dry season. On a long term, no conclusions can be drawn given the available data in this research.

5.2. The proportion of water use downstream from reservoirs (AOI 4, 5 and 6) to the decrease of the water level of the reservoir is important to assess the influence to the land availability of water users at the upper reservoir edge (AOI 3). In case of the Orós reservoir, this influence is very small within the same year. On a longer term (>10 years), a cumulative effect of the water extracting areas around the Orós reservoir on the volume of the reservoir becomes significant and influences therefore the water availability of the same AOI’s.

5.3. The water level of the Orós reservoir effects the water supply through the tunnel to AOI 5 and 6 because the water has to get over a threshold level. The influence of the irrigation water use of AOI 3 to the decrease in water level of the Orós reservoir (even in years with low reservoir volumes) is that small that AOI 3 does not influence the water supply to AOI 5 and 6. On a longer term (>10 years) it is expected that the cumulative water use of all the AOI's which surround the Orós reservoir influence the water level, and therefore also the tunnel supply.

5.4. When the water level of the reservoir is important for the water supply to downstream users (as in the case of AOI 5 by the tunnel) again the proportion of water use of upper reservoir water users to the change of water level of the reservoir has to be considered. In case of the Orós reservoir, this influence is very small.

5.5. In case of a small lake of the size of the Lima Campos reservoir, water use upstream of the reservoir (AOI 5) strongly limits the volume of the reservoir and limits therefore the release from this reservoir in dry periods, which disadvantage the farmers dependent on the water release (AOI 6).

5.6. In case the water use at a reservoir edge is only influenced by the water level of the considered reservoir for its land availability (the water availability comes from an upstream supply as in AOI 5), the downstream water use (AOI 6) is important for them. A low reservoir level can cause more land availability, so these users try to keep the level as low as possible by preventing the upstream water to flow into the reservoir by using it. Also water use downstream from the reservoir dam has a positive influence on the land availability.

Resulting from the points mentioned above, a categorization can be made. The main water sources can be categorized in a spatial and a temporal scale. In table 6.1 this categorization is made. The properties of this table will be explained below. In the conclusion drawn to this table, the conclusions from paragraph 4.3, as listed above, will be included.

Water source	Spatial scale		Temporal extent	Linked water user groups to water source	
	Extent	Resolution		Position to water source	Members of group (AOI)
Locally stored runoff	"Local scale"	1 AOI	Short ≈ 0.5-1.5 year	*Same*	1, 2
Trussu	"Regional scale"	1 AOI	Mediate ≈ 1-3 year	Upstream Edge Downstream	- - 4
Lima Campos	"Regional scale"	2 AOI's	Mediate ≈ 1-3 year	Upstream Edge Downstream	- 5 6
Orós reservoir	"Large scale"	4 AOI's	Long < 10 year	Upstream Edge Downstream	4 3 5, 6

table 6.1: Water sources categorized in spatial and temporal scales

Explanation table 6.1:

- In the first column the main water sources in the research area are listed. Water releases are linked to the reservoirs and are therefore not separately treated. From the analysis in chapter 4 it appeared that discharges from natural rivers do not play a significant role as water source in the dry season, so these water sources are not recorded in this column.

- The second and the third column are dealing the spatial scale of the water sources. This is an indication for the spatial extent on which the water source is available for or influenced by water users. A subdivision in ‘local scale’, ‘regional scale’ and ‘large scale’ is made in the column ‘extent’. In case of the locally stored runoff, this extent is small, and therefore labeled as ‘local’. In case of the Orós reservoir, many areas around the reservoir are linked to it and are influencing the volume. This scale is therefore categorized as ‘large’. The Lima Campos and the Trussu are in between local and large scale. This scale is called ‘regional’.
- The resolution indicates how many ‘spatial units’ have to be considered separately to describe the behavior water availability and water use related to the water source. Examples: Since AOI 1 and 2, both dependent on locally stored runoff, show the same behavior, without influencing each other, they can be threaten as one AOI; the resolution is therefore 1 AOI. In case of the Orós reservoir, 4 different AOI’s (3, 4, 5 and 6), with each a different behavior (see research question 4), are linked to it. Therefore, the resolution is 4 AOI’s.
- In the column ‘temporal extent’, an indication for the timescale on which the main relations between water availability and water use take place is given. Since locally stored water is dependent on the rainfall of the preceding wet season of the same year or (after very wet years) the former year, the scale is between a half year and one and a half year. As mentioned in the conclusions of paragraph 4.3, which are listed above, AOI’s linked to the Orós reservoir only influences each others water availability on a long term, due to the ‘cumulative effect’. The estimation of the time scale is therefore longer than 10 years. Since the Trussu and the Lima Campos reservoirs are both relatively small reservoirs, the water use of a year x can remarkable reduce the water availability of year x+1 or year x+2. In case of the Lima Campos the water use of AOI 5 restricts the water availability of AOI 6 even in the same year. The exact time scale is hard to conclude from the analysis. The conclusion that can be drawn is that the general time scale is longer than the time scale of locally stored runoff and shorter than the time scale in which the processes around the Orós reservoir take place. It is therefore labeled as ‘mediate’ and estimated on 1-3 year.
- AOI’s linked to the same water source, with a specified spatial and temporal scale, are forming ‘groups’. When more than one AOI’s are included in a group they can subdivided to the spatial location in relation to the location of the water source.

Conclusions of table 6.1:

- The spatial and temporal scale in which the main correlations between water availability and water use of a particular water source take place are roughly linked: a small spatial scale is linked to a small temporal scale and a large spatial scale is linked to a large temporal scale.
- When AOI’s are member of the same group and have different locations in relation to the water source, they are influencing each others water availability. Groups do not mutually influence each other, since they have different spatial and temporal scales. Examples:
    - AOI 1 and 2 are in the same group and do not have different locations, so no influence takes place within the group or with other AOI’s: see conclusion 5.1

  - The second 'group', related to the Trussu reservoir on a mediate time scale, does not have influences within the group, since AOI 4 is the only AOI in the group. No conflicts with other groups arise; see conclusion 5.2
  - The members of group 3 (AOI 5 and 6), linked to the Lima Campos reservoir on a mediate time scale, influence each others water/land availability (see conclusion 5.5 and 5.6); they have a different location in relation to the water source. They do not have influences with AOI's in other groups (see conclusion 5.3)
  - The members of group 4 (AOI 3, 4, 5 and 6), linked to the Orós reservoir on a large time scale, influence each others water availability due to the cumulative effect (conclusion 5.2 and 5.3)
- AOI 4, 5 and 6 are involved in more than one group and therefore involved in processes on different time scales.

6.3 Recommendations

6.3.1 *Policy recommendations for the research area*

- As a result of the findings in table 6.1, the distinguished groups with their specific temporal and spatial scales should be taken into account by the definition of water policies. When the 'demand management'-approach is applied by creating river basin committees as described in paragraph 1.3, the members of a committee should be linked to the defined groups in table 6.1.
- More knowledge about groundwater and aquifer is required to conclude the influence of AOI 1 and 2 on other AOI's on a long term.
- From the aspect of water availability and water use, banana is advised when the water supply can not give peaks (as in case of releases from reservoirs) and the land cannot be overflow, since banana is a semi-permanent crop. (areas not on the reservoir edge)
- Paddy is suitable when plenty water is available and can also be grown at areas which are overflow in some seasons, since paddy is a seasonal crop. The reservoir edge seems to be the most suitable place for the growing of paddy.

6.3.2 *Recommendations for further research*

- In this research the water availability in reservoirs is considered as a given value. This availability is determined by the hydrological processes in the wet season and the water use in the wet season. Since this last aspect is adjustable by policy statements, a better insight in processes in the wet season can lead to better policy recommendations that influence the water availability in the dry season.
- Since the climate in the area is very variable, it is hard to draw general conclusions and recommendations for 'the best' water allocation given a certain hydrological situation. With hydrological situation are, for example, meant: dry/wet years, dry/wet period, intense/spread wet season, early/late wet season, etc. A classification of a few characteristic hydrologic situations can lead to a better generalization of conclusions and recommendations.
- The chosen acquisition date (around the 1st of November) of the images used for the land cover classification does not give enough insight in the cropping pattern, especially for beans. Earlier images are better, also more images per season would give more insight in the cropping pattern.

- The research period included the dry seasons of 2000 to 2005, while the field visit took place in the wet season. For retrieving ground truth data, a visit in the dry season should have been better, since the cropping patterns in the dry and the wet season of some areas were very different.

Glossary

Most of the definitions are related to GIS and remote sensing and derived from the course books of the ITC about 'Principles of Geographic Information Systems' (De By, 2004) and 'Principles of Remote Sensing' (Kerle et al., 2004).

Active sensor:	Sensor with a built in source of energy. The sensor both emits and receives energy.
Attribute data:	An 'attribute' is the name of a column in a database table; it should suggest what the values in the column stand for. Theses values are known as attribute values.
Electromagnetic energy:	Energy with both electric and magnetic components. Both the wave model and photon model are used to explain this phenomenon. The measurement of reflected and emitted electromagnetic energy is an essential aspect in remote sensing
Electromagnetic spectrum:	The complete range of all wavelengths, from gamma rays (10^{-12} m) up to very long radio waves (10^{12} m)
Feature space:	The mathematical space describing the combinations of observations (in this case: DN-values in the different bands) of a multi spectral or multi-band image.
Geographic Information System (GIS):	(GIS) can be defined as a computerized system that facilitates the phases of data entry, data analysis and data presentation especially in cases when we are dealing with georeferenced data
Georeferenced data:	Data is georeferenced when coordinates from a geographic space have been associated with it. The georeferenced (spatial reference) tells us where the object represented by the data is (or was or will be).
Histogram:	Tabular or graphical representation showing the (absolute and/or relative) frequency. In the context of image data it relates to the distribution of the (DN) values of a set of pixels.
Map projection:	A map projection is a mathematically described technique of how to represent the curved planet's surface on a flat map.
Spatial data:	In the precise sense, spatial data is any data with which position is associated. 'Geospatial data' means that geographic positions data is part of it.
Spectral reflectance curve:	Spectral reflectance curves show the fraction of the incident radiations that is reflected as a function of wavelength

References

Bailey R. (1990), *Irrigated crops and their management*, Ipswich: Farming Press Books

By, R.A. de (2004), *Principles of Geographic Informations Systems*, Enschede, The Netherlands: ITC.

Budde, P., Hendrikse, J., Koolhoven, W., Nieuwenhuis, W., Nijmeijer, R., Retsios, B., Wang, L., *Ilwis 3.3 Academic (2005) Geographical Information System software package,* Enschede: ITC RSG/GSD

CGIAR-CSI The CGIAR Consortium for Spatial Information (2000), *Shuttle Rader Topography Mission (SRTM),* http://srtm.csi.cgiar.org/SELECTION/listImages.asp, date of visit: 02/2006, NGA/NASA

COGERH Companía de Gestão dos Recursos Hídricos (Management of water resources), *COGERH reservoir volume and release data of Trussu reservoir, Lima Campos reservoir and Orós reservoir between 1986 and 2005*, www.cogerh.com.br/versao3 -> monitoramento or: http://www.dnocs.gov.br -> recursos hídricos/açudagem pública, date of visit: 02/2006, COGERH: Fortaleza (Brazil)

COGERH and ENGESOFT (1999), *Plano de gerenciamento das águas da Bacia do rio Jaguaribe. Fase 1 (Diagnóstico), Volume 3 (Estudos de balanç oferta X demanda)*. Fortaleza

COGERH Iguatú Companía de Gestão dos Recursos Hídricos (2006), Conversation with Deborah at 04-2006, COGERH: Iguatú (Brazil

Colares D.S. (2004), *Análise Técnico-económica do Cultivo do Arroz no Perímetro Irrigado Morada Nova, Ceará*, Fortaleza, Brazil: Universidade Federal do Ceará, Centro de Ciências Agrárias, Departamento de engenharia agrícola curso de mestrado em irrigação e drenagem

CPRM Serviço Geológico do Brazil (geological agency of Brazil) (2006), *CPRM discharge data of Jaguaribe and Salgado river,* http://hidroweb.ana.gov.br, date of visit: 05/2006 CPRM: Brazil

DNOCS Departemento Nacional de obras contra as secas (National department of works against droughts) (1982), *Dams in the northeast of Brazil DNOCS experience in dams in semi-arid region*, DNOCS: Fortaleza (Brazil)

DNOCS Departemento Nacional de obras contra as secas (National department of works against droughts) (2006), *DNOCS agricultural land use data of perímetro,* www.dnocs.gov.br -> Irrigação/perímetros irrigados, date of visit: 05/2006, DNOCS: Fortaleza (Brazil)

Ematerce (Empresa de assistência Técnica e Extensão Rural) (2006), *Ematerce governmental data*, Ematerce: Iguatú (Brazil)

FAO (1977), *Guidelines for predicting crop water requirements* by J. Doorenbos & W.O. Pruitt, FAO Irrigation and Drainage Paper No. 24, Rome: FAO

FAO (1992), *CROPWAT - A computer program for irrigation planning and management* by M. Smith, FAO Irrigation and Drainage Paper No. 26, Rome: FAO

FAO (1998), *Crop evapotranspiration* by R. Allen, LA. Pereira, D. Raes & M. Smith, FAO Irrigation and Drainage Paper No. 56, Rome: FAO

FAO (2000), *Use of the FAO CROPWAT model in deficit irrigation studies Deficit irrigation practices* by M. Smith and D. Kivumbi, FAO Irrigation and Drainage Paper No. 22, Rome: FAO

Frischkorn, H., Araújo, J.C., Santiago, M.M.F. (2003), Water Resources of Piauí and Ceará, in: T. Gaiser, M. Krol, H. Frischkorn, J.C. de Araújo (eds.) (2003), *Global Change and Regional Impacts*, Berlin: Springer-Verlag, pp. 87-94.

FUNCEME (2006) (Fundaçãu Cearense de Meteorologia e Recursos Hídricos), *governmental rainfall data,* http://www.funceme.br/DEPAM/index.htm, date of visit: 02/2006 Fortaleza (Brazil)

Gaese, H. (2003), *Degradation Factors in a Risk-Prone Area: the Semiarid Northeast of Brazil*, in: T. Gaiser, M. Krol, H. Frischkorn, J.C. de Araújo (eds.) (2003), *Global Change and Regional Impacts*, Berlin: Springer-Verlag, pp 125-133

Garrido, R. (2001), *Water Resources National Policy in Brazil*, http://www.dams.org : The World Commission on Dams

IBGE (O Instituto Brasileiro de Geografia e Estatística) (2006), *IBGE governmental data, www.ibge.gov.br,* IBGE: Brazil

ITC Unit Geo software Development Sector Remote Sensing & GIS IT Department International Institute for Aerospace Survey and Earth Sciences (ITC) (2001), *ILWIS 3.0 Academic User's Guide*, Enschede: ITC

Johnsson, R.M.F.(2004), *Jaguaribe and Alto Tietê Riverbasins Brazil Background Paper*, World Bank.

Johnsson, R.M.F. and Kemper, K.R. (2005), *The Jaguaribe River Basin, Ceará, Brazil*, in *Institutional and policy analysis of river basin management*, World Bank.

Kerle K., Janssen L.F.J, Huurneman G.C. (eds.) (ITC Educatinal Textbook Series; 2) (2004 3rd edition), *Principles of Remote Sensing*, Enschede: ITC

Lillesand T.M., Kiefer R.W. (1994, 3rd edition), *Remote sensing and image interpretation*, New York: John Wiley & Sons, Inc.

Oel, P.R.van (2006), *Unpublished Work*, WEM, Twente University

Shaw, E.M. (1994), *Hydrology in Practice,* Cheltenham: Helson Thornes Ltd.

SUDENE (1999), *A Seca Nordestina de 1998-1999: Da Econômica à Calamidade Social,* Recife: SUDENE

Ungersböock, M., Rubel, F., Fuchs, T., Rudolf, B. (2001), Bias correction of global daily rain gauge measurements, *Physics and Chemistry of the Earth 26,* pp. 410-414

Werner P.C., Gerstengarbe F.W. (2003), The Climate of Piauí and Ceará, in: T. Gaiser, M. Krol, H. Frischkorn, J.C. de Araújo (eds.) (2003), *Global Change and Regional Impacts*, Berlin: Springer-Verlag, pp. 81-85.

Appendices

Appendices by report 'Spatial dynamics in allocation of scarce water'

University of Twente
Enschede - The Netherlands

ITC

1 Background concepts of GIS

1.1 Geometric transformation

For linking an image to a new map projection a geometric transformation can be carried out. The coordinates of two 'systems' are linked together with a geometric transformation. The first system is a dataset with Ground Control Points (GPC's). These GPC's include coordinates of recognizable landmarks in the research area. The second system is the satellite image itself. The geometric transformation assigns every pixel of the image to the map projection of the GPC dataset. Several transformation methods (= the formula that is used by transforming one system into another) are available in literature: conformal, affine, polynomial and projective. In figure 1.1 shows these four transformations methods.

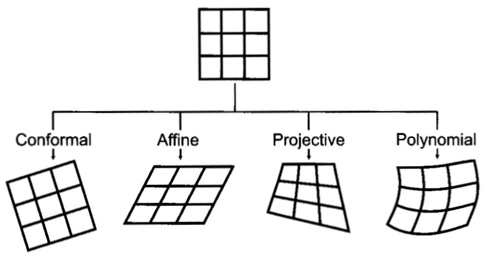


figure 1.1: Different image transformation methods

The types shown increase in complexity and parameter requirement from left to right. table 1.1 shows the results of the geo-referencing with these methods. The column called 'number of ground control points' shows the number of available GPC's in a particular image. Not all GPC's were useful in every image, because of clouds or vagueness. The value for 'sigma' is calculated with the following formula:

$$\Sigma = \sqrt{\frac{\sum\left(D_{row}{}^{2}\right)+\sum\left(D_{col}{}^{2}\right)}{(n-df)\cdot 2}} \quad \text{(Eq. 1.1)}$$

Where D_{row} and D_{col} are the differences between calculated row and column values and actual row and column values in pixels, n are the number of GPC's and $((n\text{-}df)\cdot 2)$ are the degrees of freedom, that is the number of required GPC's for a certain transformation. Sigma gives a measure for the overall accountability or credibility of the GPC's.

Image	Number of Ground Control Points (GPC's)	Sigma			
		Conformal	Affine	Polynomial (2^{nd} order)	Projective
2000	16	0.913	0.839	0.613	0.805
2001	18	0.635	0.592	0.640	0.610
2002	10	0.987	0.864	0.737	0.919
2003_left	6	0.825	0.664	-[1]	-
2003_right	10	1.853	0.941	0.616	0.985
2004_left	10	2.037	0.660	0.547	0.517
2004_right	11	2.435	1.038	1.106	1.106
2005_left	7	1.780	0.534	0.666	0.605
2005_right	10	2.166	0.802	0.452	0.535

table 1.1: Results geo-referencing

The shaded cells in table 1.1 indicate the best transformation method as a result of the lowest score for sigma.

1.2 Resampling

After geo-referencing only the pixels are referenced to the right coordinate in the new projection. The actual pixel shape of the old projection has not changed yet. To get a real presentation of the new projection, the image itself has to change. This changing is called 'resampling'. In this procedure, the values of the new pixels are determined by using the pixels of the original image. Each pixel has a value in a certain domain. In the case of a land use map, a certain land use (Paddy, Banana or Water etc.) is assigned to each pixel. In the case of a satellite image, each pixel has a value which indicates the reflection in a particular band. These values are called 'digital numbers', or DN's. More about 'pixels' and DN's' can be read in paragraph 2.7.
In the third sub figure in figure 1.2 (clockwise) it becomes clear that for some new pixels it is not clear which original pixel has to use to assign a value to it. Several methods can be used to assign a value to the new pixels: The *nearest neighbour* method assigns the value of the nearest original pixel to the new pixel. Using the *bilinear interpolation* the weighted mean is calculated for the four nearest pixels in the original image. *Cubic convolution* applies a polynomial approach based on the values of 16 surrounding pixels. Resampling methods assign DN's to the pixels in the new image by using the DN's of the original image. In case of bilinear interpolation and cubic convolution the new DN's are the result of a calculation, so the original DN's are lost. Since the original DN's are important for an appropriate classification, the resampling has to be carried out *after* the classification process. After this classification process the images in this research are resampled by using the nearest neighbour method (because bilinear interpolation and cubic convolution use values, these methods could not be used for a class-domain).

In figure 1.2 the process of transformation and resampling is schematized. Note that the 'image after transformation' is a conceptual illustration, since the actual pixel shape dos not change. The showed grid indicates how the new geo-reference system assigns new coordinates to the original image.

[1] Not enough GPC's

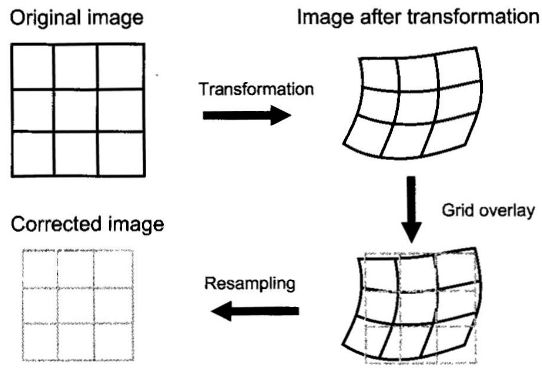


figure 1.2: Illustration of the transformation and resampling process

2 Background concepts of Remote sensing

2.1 *Introduction*

This paragraph is mainly taken from an educational textbooks about Remote Sensing (Kerle, Janssen and Huurneman, 2004) and from the user's guide of the GIS software package ILWIS (ITC, 2001), both published by the International Institute for Aeorspace survey and Earth Sciences (ITC)

2.2 *Electromagnetic energy*

Remotely sensed data, such as satellite images, are measurements of reflected solar radiation, energy emitted by the earth itself or energy emitted by Radar systems that is reflected by the earth.
The reflected solar radiation is called electromagnetic energy (EM). All matter with a temperature above absolute zero (0 K) radiates EM energy due to molecular agitations (agitation is the movement of the molecules).
Electromagnetic energy (EM) can be modeled in a wave model. In the wave model, electromagnetic energy is considered to propagate through space in the form of sine waves (see figure 2.1). These waves are characterized by electrical (E) and magnetic (M) fields, which are perpendicular to each other. Both fields propagate through space at the speed of light *c*.

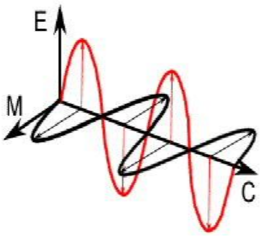


figure 2.1: Wave model

The most important characteristic of the wave model is the wavelength, λ [m], which is defined as the distance between successive wave crests. The frequency, v, is the number of cycles of a wave passing a fixed point over a specific period of time. Since the speed of light is constant, wavelength and frequency are inversely related to each other:

$$c = l \cdot v \qquad \text{(Eq. 2.1)}$$

With c as the speed of light ($3 \cdot 10^8$ m/s), λ as the wavelength (m) and v as the frequency (cycles per second, Hz)
The shorter the wavelength, the higher the frequency. Conversely, the longer the wavelength, the lower the frequency.

Another important characteristic is the EM energy of a photon. This characteristic can be found by the particle theory, in which EM energy is composed of discrete units called 'photons'. The amount of energy held by a photon of a specific wavelength is then given by:

$$Q = h \cdot v \qquad \text{(eq. 2.2)}$$

With Q as the energy of a photon (J), h as the Planck's constant ($6.6262 \cdot 10^{-34}$ Js) and v as the frequency (Hz)
Combining equation 2.1 and 2.2 results in:

$$Q = h \cdot \frac{c}{l} \qquad \text{(eq. 2.3)}$$

From equation 2.3 follows that the longer the wavelength, the lower is the energy content (see figure 2.2). An important consequence for remote sensing is that it is more difficult to measure the energy emitted in longer wavelengths than in shorter wavelengths.

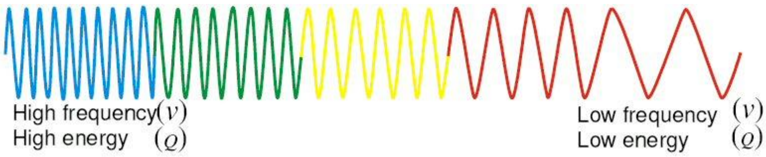


figure 2.2: Relation between wavelength, frequency and energy

2.3 *Electromagnetic spectrum*

The total range of wavelengths is commonly referred to as the electromagnetic spectrum (see figure 2.3). Remote sensing operates in several regions of the electromagnetic spectrum. The optical part of the EM spectrum refers to that part of the EM spectrum in which optical phenomena of reflection can be used to focus the radiations. The optical range extents from X-rays (0.02 μm) through the visible part of the EM spectrum up to and including far-infrared (1000 μm). The visible region of the spectrum is commonly called 'light'. Only this portion of the EM spectrum ca be associated with the concept of color.

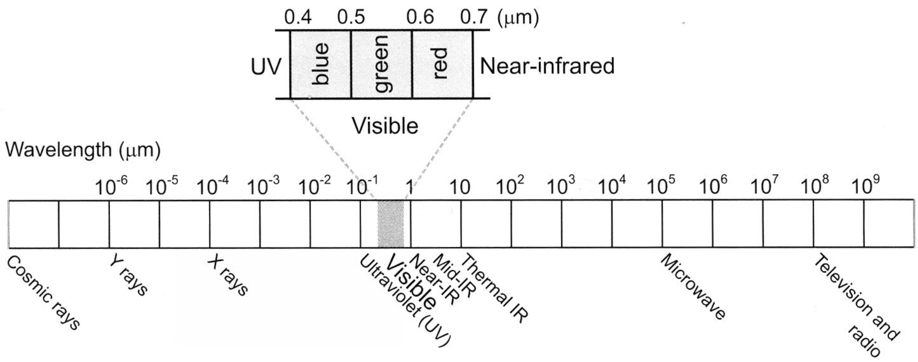


figure 2.3: Electromagnetic spectrum

2.4 *Energy interaction in the atmosphere*

An overview of interaction in the atmosphere (and at the surface) is given in figure 2.4. The most important aspects of this overview will be discussed now.

Electromagnetic energy traveling trough the atmosphere is partly absorbed by various molecules. The most efficient absorbers of solar radiations in the atmosphere are ozone (O_3), water vapour (H_2O) and carbon dioxide (CO_2). Only the regions outside the main absorption bands of the atmospheric gases can be used for remote sensing. These regions include a 'window' in the visible and reflected infrared regions, between 0.4-2.0 μm and three windows in the thermal infrared region. The window in the visible and reflected infrared regions is important for land cover classifications, since plants are reflecting EM energy in these particular regions.
Besides atmospheric absorption, atmospheric scattering takes place. Atmospheric scattering occurs when the particles or gaseous molecules present in the atmosphere cause the EM waves to be redirected from their original path. Three types of scattering take place: Rayleigh scattering, Mie scattering and Non-selective scattering. Rayleigh scattering predominates where EM radiation interacts with particles that are smaller than the wavelength of the incoming light. For satellite remote sensing, Rayleigh scattering is the most important type of scattering. Due to the Rayleigh effect, the shorter wavelengths are overestimated because larger wavelengths are more bothered by particles as tiny specks of dust, nitrogen and oxygen molecules. Mie scattering and Non-selective scattering occur when the size of atmospheric particles are respectively the similar size or of much larger size as the wavelength of the incoming radiation.

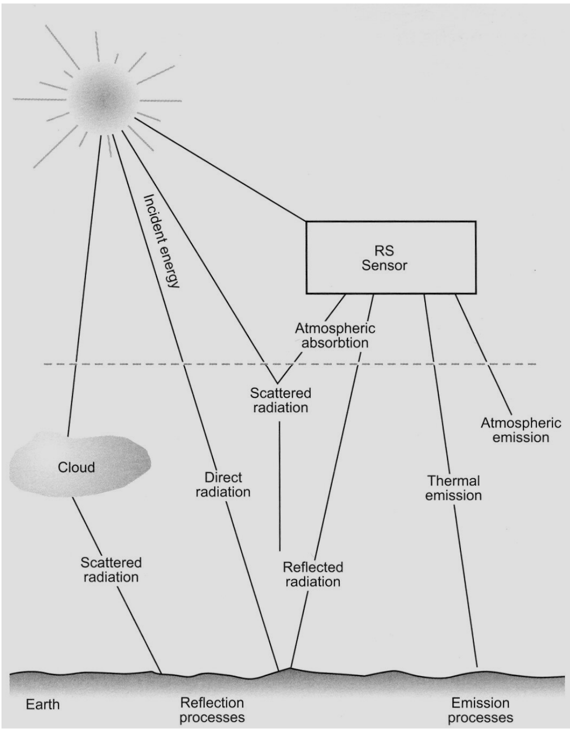


figure 2.4: Energy interactions in the atmosphere and at the surface

2.5 *Reflectance*

The energy reaching a certain surface is called irradiance. The energy reflected from the surface is called radiance. In the context of this thesis, we are most interested in the reflected radiation because this tells us something about surface characteristics. For each material, a specific reflectance curve can be established. Such curves show the fraction of the incident radiations that is reflected as a function of wavelength. In figure 2.5 a reflectance curve of healthy vegetation is shown. Most remote sensing sensors are sensitive to broader bands, for example from 0.4-0.8 μm. In these cases, the reflectance curve can give an estimation of the overall reflectance in such a band. The reflectance curves of some the most important land cover classes for this thesis will be discussed now: vegetation and water.

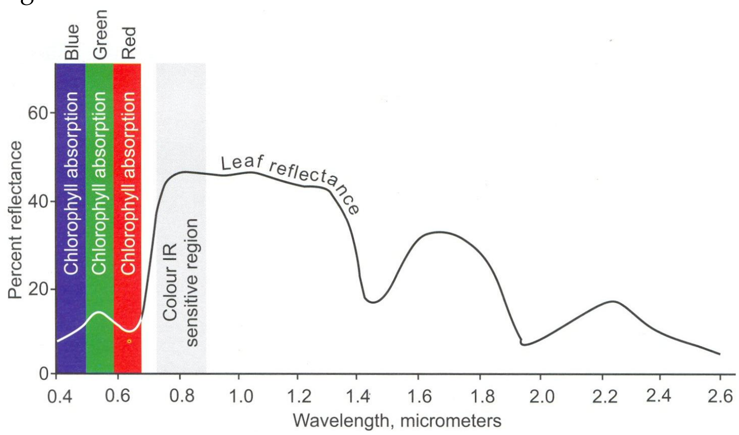


figure 2.5: Idealized spectral reflectance curve of a healthy vegetation

The reflectance characteristics of vegetation depend on the properties of the leafs, including the orientations and the structure of the leaf canopy. The proportion of the radiations reflected in the different parts of the spectrum depends on leaf pigmentation, leaf thickness and compositions (cell structure), and on the amount of water in the leaf tissue. The reflectance in the near-infrared is highest, but the amount depends on leaf development and cell structure. In the middle infrared, the reflectance is mainly determined by the free water in the leaf tissue; more water results in less reflectance.

Water has a low reflectance, compared to vegetations and soils; at most 10% of the incoming radiation. Water reflects EM energy in the visible up to the near-infrared. The reflectance curves of water and vegetation can easily be distinguished because of this difference in intensity and distribution over wavelengths. The distinction between different vegetation types is harder. The near-infrared region of the spectral reflectance curve of vegetation is important for distinguishing different vegetation types, since this region gives the most variations among different types of vegetation.

2.6 Sensors

The EM energy, which is discussed in the previous sections, can be measured and recorded by sensors. The resulting data can be used to derive information about surface characteristics. The measurements of electromagnetic energy are made by sensors that are attached to a static or moving platform. Different types of sensors and platforms have been developed for different applications. figure 2.6 shows an overview of different available sensor-platform combinations (for the definitions of active and passive sensors see Glossary).

Different criteria are considered for making the decision which sensor to use. The most important were in the situations of this thesis: 'spatio-temporal characteristics' and 'availability of image data'.

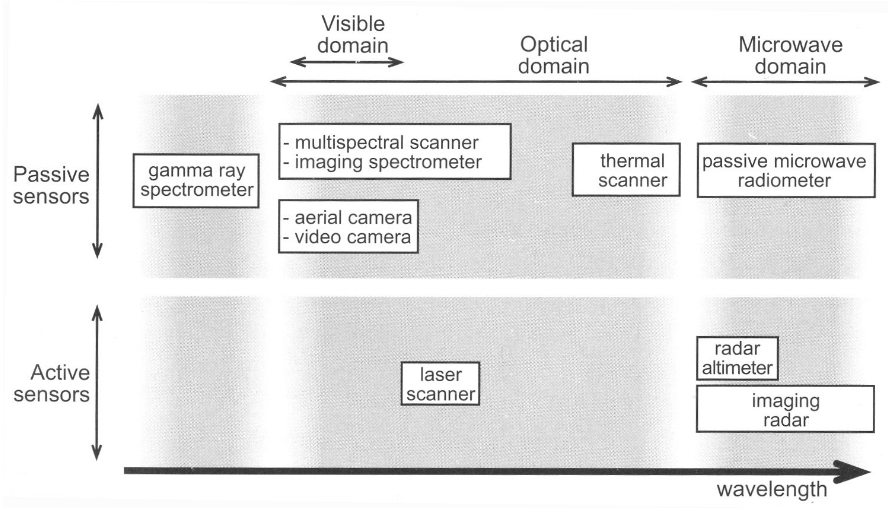


figure 2.6: Overview of sensors

Spatio-temporal characteristics:
A good understanding of the data requirements of the research is required to know which data type and sensor is appropriate. Since this thesis deals with a temporal comparison of a fixed moment in each particular year, the dry period between August and December, a condition was to get images of this moment of the year. Because the main concern was to classify agricultural field with a size of several hectares and large water bodies, a pixel size of around 30 m was appropriate. For image classification a multi spectral sensor was required.

Availability of image data:
For the year 2000 a free cloudless image of Landsat7 was available at the end of October. In the years 2003 up to 2005 appropriate CBERS-2 images in the same period were free available. The borders of these CBERS-2 images are crossing the research area. Therefore, 2 CBERS-2 images per year are used with a time difference of 3 days. These images are glued after the image classification process. In this way no radiometric data, required for the image classification, is lost due to resample techniques. For missing years, 2001 and 2002, Landsat7 images are bought. See table 2.1 for properties of both types of images.

System	Landsat 7	CBERS-2
Orbit	Altitude: 705 km Inclination angle: 98.2° Type: sun-synchronous Crosses equator at: 10:00 AM Revisit time: 16 days	Altitude: 778 km Type: sun-synchronous Revisit time: 26 days
Sensor	ETM+ (Enhanced Thematic Mapper)	CCD (Charge Couple Device)
Swath width	185 km	113 km
Spectral bands (μm)	Band 1: 0.45-0.52 Band 2: 0.52-0.60 Band 3: 0.63-0.69 Band 4: 0.76-0.90 Band 5: 1.55-1.75 Band 6: 10.4-12.5 Band 7: 2.08-2.34 PAN: 0.50-0.90	Band 1: 0.45-0.52 Band 2: 0.52-0.59 Band 3: 0.63-0.69 Band 4: 0.77-0.89 Band 5: 0.51-0.73 (PAN)
Spatial resolution	15 m (PAN), 30 m (bands 1-5,7), 60 m (band 6)	20 m (band 1-5)

table 2.1: Properties of Landsat/ETM+ images and CBERS-2 images

Explanation of important concepts in table 2.1:

Orbit:	A circular path described by the satellite in its revolution around the earth.
Inclination angle:	The angle (in degrees) between the orbital plane and the equatorial plane. The inclination angle of the orbit determines, together with the field of view of the sensor, the latitudes up to which the earth can be observed.
Sun-synchronous orbit:	This is a near-polar orbit chosen in such a way that the satellite always passes overhead at the same time. To achieve this, the inclination angle must be carefully chosen between 98° to 99°.
Swath width:	Swath width refers to the strip of the Earth's surface from which data are collected.
Band:	A specific rang of the EM spectrum to which the sensor is sensitive.
Panchromatic band (PAN):	A spectral band in the visible *and* near-infrared part of the EM spectrum.
Spatial resolution:	Indicates the smallest observable (measurable) difference at which objects can still be distinguished.

2.7 *Image characteristics*

Some important properties of images used for this thesis will be discussed in this paragraph. A remote sensing image is a measurement of EM energy. Image data are stored in a regular grid format (rows and columns). A single image element is called a pixel. For each pixel, the measurements are stored as Digital numbers, or DN-values. Typically, for each measured wavelength range a separate data set is stored, which is called a band.

Multi-band image: Depending on the sensor, data are recorded in *n* bands. In one pixel of this multi-band image, the values of the used bands can be regarded as an n-dimensional vector, the feature vector. For example, the CBERS-2 images have 5 bands; when these 5 bands are combined in one multi-band image, a vector of 5 DN-values per individual pixel emerge. The vector could be (240, 210, 10, 50, 60), which tells that the DN in band 1 is 240, the DN in band 2 is 210, the DN in band 3 is 10, etc. See figure 2.7 for a representation of a multi-band images consisting 3 bands.

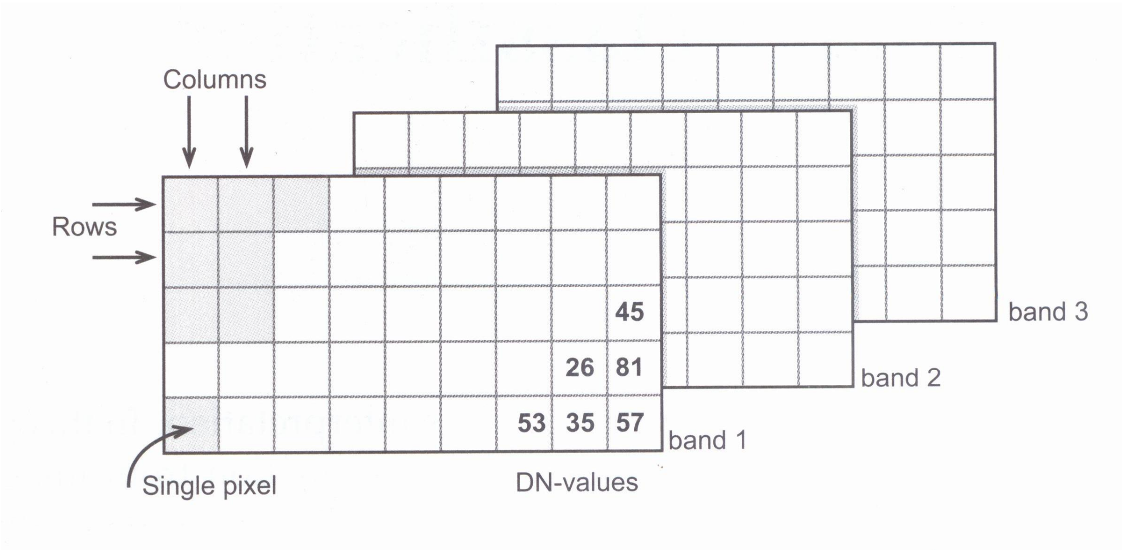


figure 2.7: Mulit-band image

Feature space:

A feature space is the mathematical space describing the combinations of 2 or more bands of a multispectral or multi-band image. In this thesis the 2 dimensional feature space is much used. The two axis are defined by the spectral ranges of the two bands (DN = 0- 255). Each pixel in the image has a value for both bands and can be drawn in the feature space. In figure 2.8 the feature space of band 3 and 4 of the left image of 2005 is shown. It is obviously that the pixels of the same land cover class, indicated by the color, are clustered.

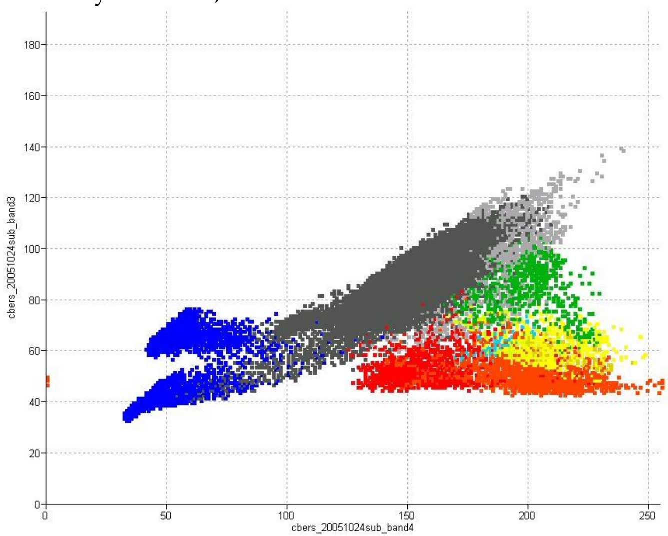


figure 2.8: Feature space of 2005_right: band 3 and band 4

Spatial characteristics refer to the area measured. Spatial resolution (see 'explanation of important concepts in table 2.1') and coverage are important issues. Spatial coverage refers to the total area covered by one image.

Spectral characteristics refer to the spectral range where the sensor is sensitive for. The spectral resolution is related to the widths of the spectral wavelength bands that the sensor is sensitive to. In case of the CBERS-2 sensor, the radiometric resolution for band 1 is 0.07 μm (see table 2.1)

Radiometric characteristics refer to the energy levels that are measured by the sensor. In case of the sensors of Landsat7 and CBERS-2, the DN-value is in between 0 and 255. When the value is zero, no EM energy is received. A value of 255 means a maximum possible response. The EM energy values between the minimum and the maximum response are proportional distributed over the 254 resting numbers.

2.8 Radiometric corrections

The presence of a heterogeneous, dens and layered terrestrial atmosphere composed of water vapor, aerosols and gases disturbs the signal reaching the senor in many ways. Many methods are available to make corrections to the disturbed signal. By buying an image, some of these methods are already carried out. In the case that a researcher is interested in the exact amount of EM energy in a certain pixel, the radiometric correction is of great importance. In the case of this thesis, where the comparison of several spectral bands in one image will result in a land use classification, the exact ground value is not required, so the operations that are carried out in the pre-processing of the image are sufficient.

2.9 Geometric corrections

The topic 'spatial referencing' is described in appendix paragraphs 1.1 and 1.2. The errors that are emerging due to these topics have to be solved to make it possible to determine the size of an area in the images. Since an important part of this thesis is to determine the area's of different types of land use, the geometric correction is an important issue.

3 Background concepts of Image classification

3.1 Classification methods

Slicing

Density slicing is a technique, whereby the DN-values distributed along the horizontal axis of an image histogram[2], are divided into a series of user specified intervals or slices. Density slicing will only give reasonable results, if the DN-values of the cover classes are not overlapping each other.

In figure 3.1 the histograms of band 2, 3 and 4 of the left image of 2003 are shown. It is hard to distinguish different cover classes from the histograms, because there are no separated ranges which indicate a certain cover class: the cover classes are overlapping each other. The histograms of the other images used in this these were showing the same phenomena of overlapping.

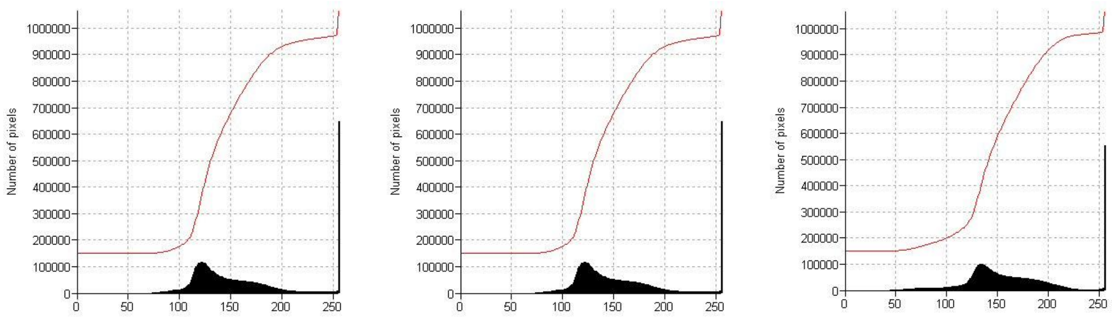


figure 3.1: Histograms of CBERS 2003 (left part), respectively band 2, 3 and 4

To make a better distinction between vegetation and no-vegetation a vegetation index can be used.

$$VI = NIR - VIS \qquad \text{(Eq 3.1)}$$

Or:

$$NDVI = \frac{NIR - VIS}{NIR + VIS} \qquad \text{(Eq. 3.2)}$$

With:

VI = Vegetation Index
NDVI = Normalized Difference Vegetation Index
NIR = Near-infrared band
VIS = Visible band (yellow, green or red)

[2] For definition, see Glossary

Vegetated areas will generally yield high values for a vegetation index because of their relatively high near-infrared reflectance and low visible reflectance (see appendix paragraph 2.5). In contrast, clouds, water, and snow have larger visible reflectance than near-infrared reflectance. Thus, these features yield negative index values. Rock and bare soil areas have similar reflectances in the two bands and result in vegetation indices near zero. Because the separation between vegetation and non-vegetation becomes better distinguishable using a VI or NDVI, the histogram of a VI or NDVI can give a better distinction between land cover classes. In figure 3.2 the histogram of a NDVI is shown. Unfortunately, the distinction between different vegetation classes is still hard to see.

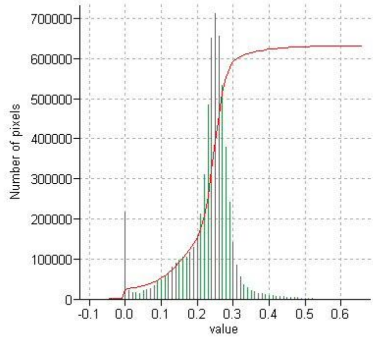


figure 3.2: Histogram the NDVI of the of 2003 (left part)

Unsupervised classification

In an unsupervised classification, clustering algorithms are used to partition the feature space into a number of clusters. For an explanation of the meaning of a feature space, see appendix paragraph 2.7. Unsupervised classification is at option when little knowledge of the area is available. Several methods of unsupervised classification exist, their main purpose being to produce spectral groupings based on certain spectral similarities. Important is that unsupervised classifiers do not utilize training data as the basis for classification.

In one of the most common approaches, the user has to define the maximum number of clusters in a data set. Based on this, the computer locates arbitrary mean vectors as the centre points of the clusters. Each pixel is then assigned to a cluster by the 'minimum distance to cluster centroid' decision rule. Once all the pixels have been labeled, recalculation of the cluster centre takes place and the process is repeated until the proper cluster centers are found and the pixels are labeled accordingly. The iteration stops when the cluster centers do not change any more. The results of an unsupervised classification are *spectral classes.* The analyst must compare the classified data with some form of reference data to determine the identity and informational value of the spectral classes.

The derived cluster statistics are then used to classify the complete image using a selected classification algorithm. A background of these classification algorithms will be given in sub-paragraph 3.2 of this appendix.

Supervised classification

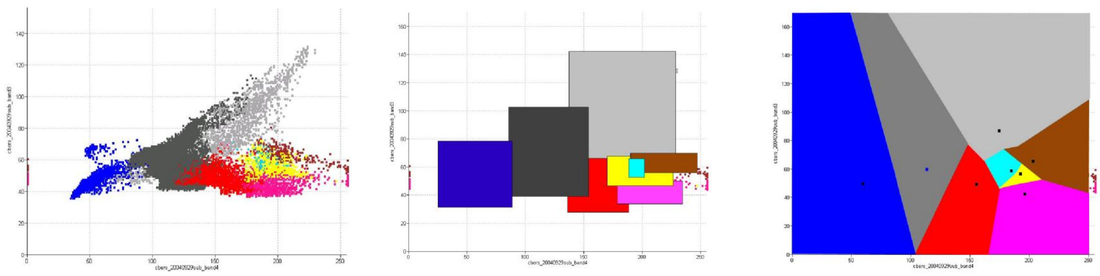
In supervised classification, the partitioning of the feature space is realized by an operator who defines the spectral characteristics of the classes by identifying sample areas (training areas). The operator needs to know where to find the classes of interest in the area covered by the image. This information can be derived from 'general area knowledge' or from dedicated field

observations. A sample of a specific class, comprising of a number of training pixels, forms a cluster in the feature space. The clusters as selected by the operator should form a representative data set for a given class. This means the variability within the image should be taken into account. The clusters should also not or only partially overlap with the other clusters, as otherwise a reliable separation is not possible.
Identifying sample areas in the area, called the training stage, is very important for a supervised classification since the sample set is used to classify each pixels in the image to a certain land cover class.

3.2 *Algorithms*

Box classifier

The box classifier is the simplest classification method. For this purpose, upper and lower limits are defined for each class. The limits may be based on the minimum and maximum values, or on the mean and standard deviation per class. When the lower and the upper limits are used, they define a box-like area in the feature space, which is why it is called box classifier. Bos classification is also known as parallelepiped classification. The disadvantage of the box classifier is the overlap between the classes. In such a case, a pixel is arbitrarily assigned the label of the first box it encounters. In the second figure of figure 3.3 this is demonstrated for the sample set of the image of 2004_right; it is obviously that the light grey area (bare land) and the light blue (coco) are too large and the yellow area (banana) is too small.



Feature space of sample set *Box classification* *Minimum distance classification*

figure 3.3: Comparison classification algorithms for sample set of 2004_right

Minimum Distance to Mean classifier (MD)

The basis for the Minimum Distance to Mean (MD) classifier is the cluster centers. During classification the Euclidean distances from an unknown pixel to various cluster centers are calculated. The unknown pixel is assigned to that class to which the distance to the mean DN-value of that class is least. The right figure of figure 3.3 illustrates this principle for the sample set of the right image of 2004.
Disadvantages of this method are that pixels which are far away from the center of a cluster may be assigned to this cluster when it is the nearest centre. To overcome this problem a threshold distance can be used. When a pixel falls outside the threshold distances of all means, it will be left unclassified. Another disadvantage of the Minimum Distance to Mean classifier is that it does not take the class variability into account: some clusters are small and dense, while other are large and dispersed. This becomes obviously when the feature space of the sample set and the right figure of figure 3.3 are compared. The dispersed brown 'points-cloud' (beans) retrieves a larger area in the classification than the bananas, whereas the point-cloud of bananas is larger and

denser than the point cloud of beans. Thus, also ‘density’ and ‘distribution’ should be taken into account.

Minimum Mahalanobis Distance classifier (MMD)
The same principle is used as in the ML classification method, but the MMD classifier uses the Mahalanobis distance instead of the ‘normal’ distance. The Mahalanobis distance depends on the distances towards class means and the variance-covariance matrix of each class.

Maximum likelihood classifier (ML)
The Maximum Likelihood (ML) classifier considers not only the cluster centre but also its shape, size and orientation. This is achieved by calculation a statistical distance based on the mean values and covariance matrix of the clusters. The statistical distance is a probability value: the probability that observation x belongs to a specific cluster. The pixel is assigned to the class (cluster) to which it has the highest probability. The assumption of most ML classifiers is that the statistics of the clusters have a ‘normal’ (Gaussian) distribution. Also in this method a threshold distance can be used.

3.3 *Accuracy assessment*

Confusion matrices compare, on a category-by-category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. Therefore, the total ground truth data is split up in a ‘sample set’ and a ‘test set’. The sample set is used as training for the classifier (the left figure of figure 3.3 is an example of a the feature space of a sample set), the test set is used as the reference data in a confusion matrix. A confusion matrix is shown in table 3.1.

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	**295**	0	25	8	0	0	1	0.9
Beans	0	**0**	0	0	0	0	0	?
Coco	0	0	**0**	0	0	0	0	?
Grass	0	0	0	**0**	0	0	0	?
Paddy	37	0	8	0	**879**	0	2	0.95
Water	0	0	0	0	0	**3193**	64	0.98
Reliability	0.89	?	0	0	1	1		
Average Accuracy		94.33						
Average Reliability		96.33						
Overall Accuracy		96.79						
KHAT-value		92.76						

table 3.1: Confusion matrix of 2003_right

Explanation of a confusion matrix as shown in table 3.1:

- Rows correspond to classes in the test set.
- Columns correspond to classes in the classification result.
- The *diagonal elements* in the matrix represent the number of correctly classified pixels of each class, i.e. the number of test set pixels with a certain class name that obtained this same class name during classification.

- The *off-diagonal elements* represent misclassified pixels or the classification errors, i.e. the number of ground truth pixels that ended up in another class during classification.
- The figures in column *Unclass* represent the ground truth pixels that were found not classified in the classified image.

Accuracy: The figures in column *Accuracy* present the accuracy of classification: it is the fraction of correctly classified test set pixels of a certain ground truth class. For each class of test set pixels, the number of correctly classified pixels is divided by the total number of test pixels in that class; for example for the 'Banana' class, the accuracy is 295/329 = 0.90 meaning that approximately 90% of the 'Banana' test set pixels also appear as 'Banana' pixels in the classified image.

Reliability: The figures in row *Reliability* present the reliability of classes in the classified image: it is the fraction of correctly classified test set pixels of a certain class in the classified image. For each class in the classified image (column), the number of correctly classified pixels is divided by the total number of pixels which were classified as this class. For example for the 'Banana' class, the reliability is 295/332 = 0.89 meaning that probably 89% if the 'Banana' pixels in the classified image are correct compared to the ground truth pixels.

The *average accuracy* is calculated as the sum of the accuracy figures in column Accuracy divided by the number of classes in the test set (question marks and zero-values are not included).
The *average reliability* is calculated as the sum of the reliability figures in column Reliability divided by the number of classes in the test set (question marks and zero-values are not included).
The *overall accuracy* is calculated as the sum of all correctly classified pixels (diagonal elements) divided by the total number of test pixels.

In the case of the accuracy assessment of table 3.1, only the classes 'Paddy', 'Banana' and 'Water' are assessed. Therefore only these classes have a value greater than zero in the column accuracy and the row reliability. Because of the limited amount of ground truth data of the classes 'Beans', 'Coco' and 'Grass' an accuracy assessment of these classes would have no meaning. In some other years, beans and grass could join the accuracy assessment.

Another indicator to assess the classification is known in literature: the KHAT-value (Lillesand and Kiefer, 1994). The KHAT statistic is a measure of the difference between the actual agreement between reference data and an automated classifier and the chance agreement between the reference data and a random classifier. In some cases a random assignment of pixels to classes can give a good result. The meaning of high accuracy values of an automated classifier is in such cases very low. A KHAT-value of 0 suggests that a given classification is no better than a random assignment of pixels, a KHAT-value of 0.90 indicates a 90% better result of the automated classifier than the random assignment.

The KHAT statistic is computed as

$$KHAT = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}\left(x_{i+} \cdot x_{+i}\right)}{N^2 - \sum_{i=1}^{r}\left(x_{i+} \cdot x_{+i}\right)}$$

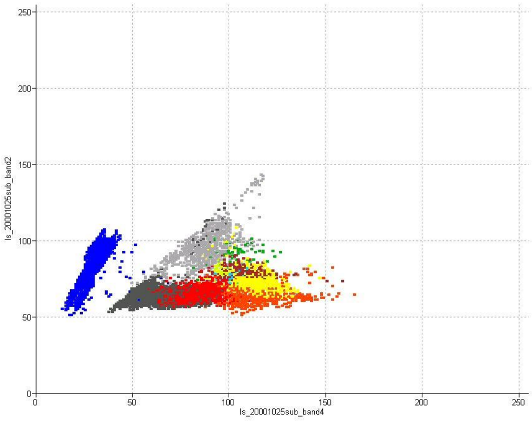
In which:

r	= number of rows in the confusion matrix
x_{ii}	= the number of observations in row i and column i (on the major diagonal)
x_{i+}	= total of observations in row i
x_{+i}	= total of observations in column i
N	= total number of observations included in matrix

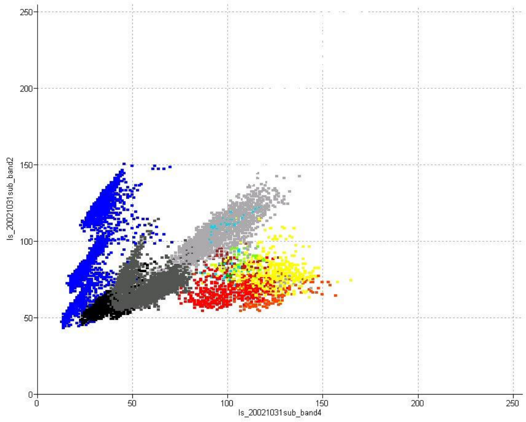
See the last value of table 3.1 for the result of the computation of the KHAT-value.

4 Image classification

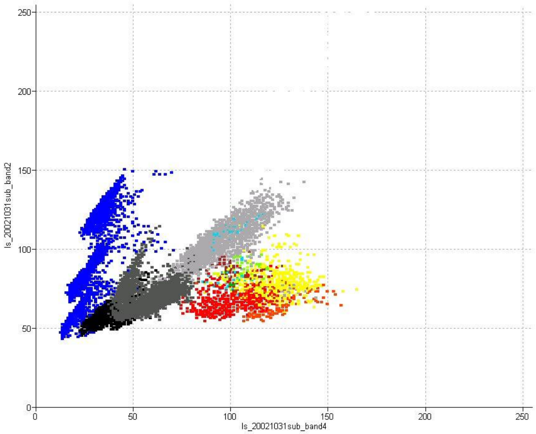
4.1 Feature spaces



2000



2001

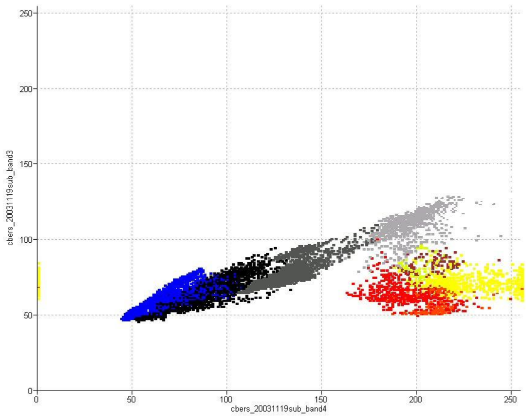


2002

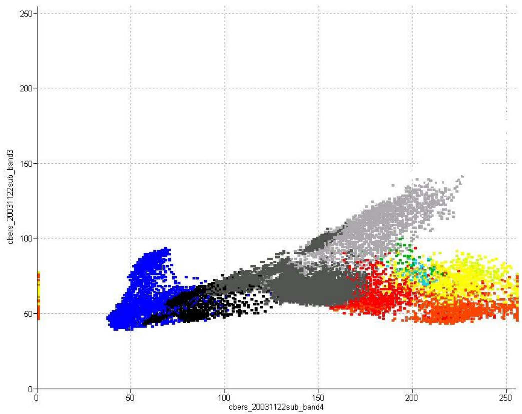
Legend:

- Banana
- Beans
- Coco
- Grass
- Paddy
- Water

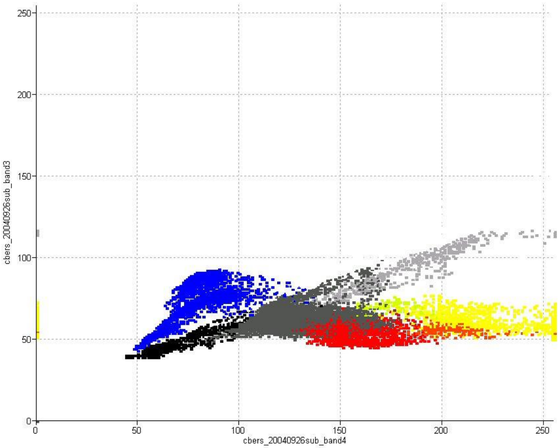
(Light grey and dark grey are non-agricultural lands, black is cloud shadow and white areas between 'water' are clouds)



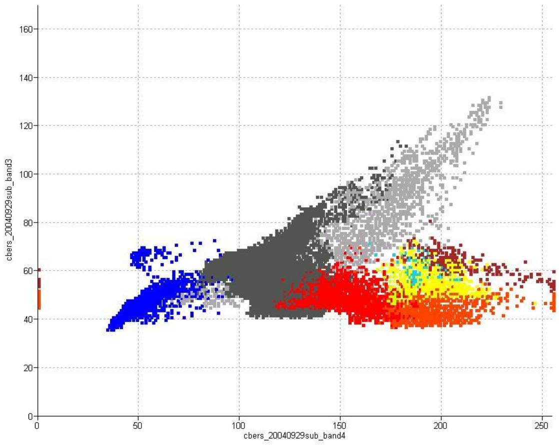
2003_L



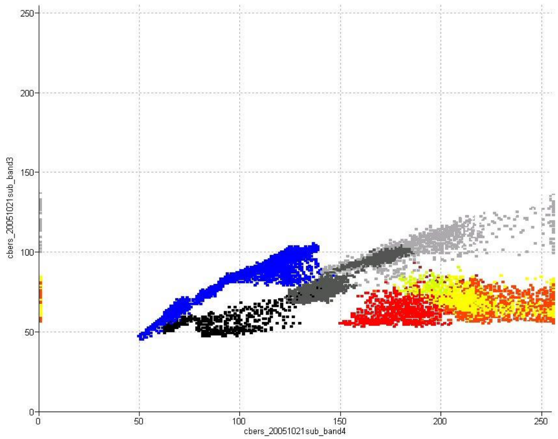
2003_R



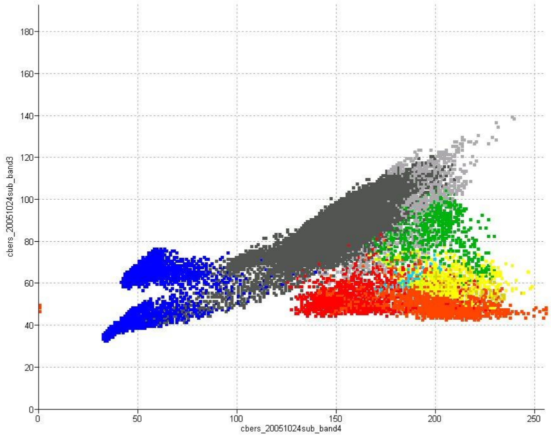
2004_L



2004_R



2005_L



2005_R

4.2 Confusion matrices

2000

	Banana	Beans	Coco	Grass	Paddy	Unclass	Accuracy
Banana	126	1	8	3	17	1	0.81
Beans	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	?
Paddy	1	0	2	0	366	2	0.99
Reliability	0.99	0	0	0	0.96		
Average Accuracy		0.90					
Average Reliability		0.98					
Overall Accuracy		0.93					
Khat		0.84					

2001

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	201	0	0	0	24	0	1	0.89
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	77	0	2	0	149	0	0	0.65
Water	0	0	0	0	0	1456	14	0.99
Reliability	0.72	?	0	?	0.86	1		
Average Accuracy		0.84						
Average Reliability		0.86						
Overall Accuracy		0.94						
Khat		0.84						

2002

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	138	0	1	7	4	0	8	0.87
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	10	0	0	0	105	0	0	0.91
Water	0	0	0	0	0	1461	30	0.98
Reliability	0.93	?	0	0	0.96	1		
Average Accuracy		0.92						
Average Reliability		0.96						
Overall Accuracy		0.97						
Khat		0.88						

2003_L

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	186	26	0	0	5	0	0	0.86
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	22	0	0	0	166	0	0	0.88
Water	0	0	0	0	0	1141	16	0.99
Reliability	0.89	0	?	?	0.97	1		
Average Accuracy		0.91						
Average Reliability		0.95						
Overall Accuracy		0.96						
Khat		0.90						

2003_R

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	319	0	6	1	0	0	3	0.97
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	27	0	0	0	895	0	4	0.97
Water	0	0	0	0	0	3215	42	0.99
Reliability	0.92	?	0	0	1	1		

Average Accuracy	0.98
Average Reliability	0.97
Overall Accuracy	0.98
Khat	0.96

2004_L

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	241	0	0	0	7	0	0	0.97
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	14	0	0	0	485	0	0	0.97
Water	0	0	0	0	0	3043	20	0.99
Reliability	0.95	?	?	?	0.99	1		

Average Accuracy	0.98
Average Reliability	0.98
Overall Accuracy	0.99
Khat	0.97

2004_R

	Banana	Beans	Paddy	Water	Unclass	Accuracy	total
Banana	166	15	0	0	5	0.81	186
Beans	9	103	0	0	14	0.75	126
Paddy	1	1	826	0	33	0.96	861
Water	0	0	0	7275	7	1.00	7282
Reliability	0.94	0.87	1.00	1.00			

Average Accuracy	0.88
Average Reliability	0.95
Overall Accuracy	0.99
Khat	0.96

2005_L

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	273	1	0	0	23	0	4	0.91
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	10	0	0	0	324	0	0	0.97
Water	0	0	0	0	0	3932	0	1
Reliability	0.96	0	?	?	0.93	1		

Average Accuracy	0.96
Average Reliability	0.96
Overall Accuracy	0.99
Khat	0.97

2005_R

	Banana	Beans	Coco	Grass	Paddy	Water	Unclass	Accuracy
Banana	141	0	0	0	19	0	0	0.88
Beans	0	0	0	0	0	0	0	?
Coco	0	0	0	0	0	0	0	?
Grass	0	0	0	0	0	0	0	?
Paddy	1	0	0	1	672	0	2	0.99
Water	0	0	0	0	0	6958	0	1
Reliability	0.99	?	?	0	0.97	1		

Average Accuracy	0.96
Average Reliability	0.99
Overall Accuracy	1.0
Khat	0.98

4.3 *Accuracy assessment of coco, beans and grass*

set	# pixels coco in 'set'	Test set	Banana (%)	Beans (%)	Coco (%)	Grass (%)	Paddy (%)	Water (%)	Non-agri (%)
2001	5354	2002	22	5	28	9	8	0	30
2001	5351	2003	11	0	4	25	27	0	33
2001	5335	2004	22	9	8	0	12	0	49
				Mean:	**13**				
2002	6556	2001	9	0	22	0	4	0	64
2002	6567	2003	11	0	2	16	18	0	53
2002	6567	2004	14	14	8	0	11	0	53
				Mean:	**11**				
2003	595	2001	25	0	40	0	6	0	29
2003	602	2002	28	5	22	16	6	0	22
2003	602	2004	33	13	19	0	12	0	22
				Mean:	**27**				
2004	2118	2001	14	0	20	0	6	0	60
2004	2143	2002	18	7	24	12	4	0	35
2004	2135	2003	18	0	5	20	15	0	43
				Mean:	**16**				

table 4.1: Consistency table of coco (AOI 6: Perímetro)

Notes by
table 4.1: The column 'coco' should contain most of the pixels from column '# pixels in testset'. This is obviously not the case. Most of the classified coco pixels of the 'set' are in other years assigned to banana, coco and 'non-agri'

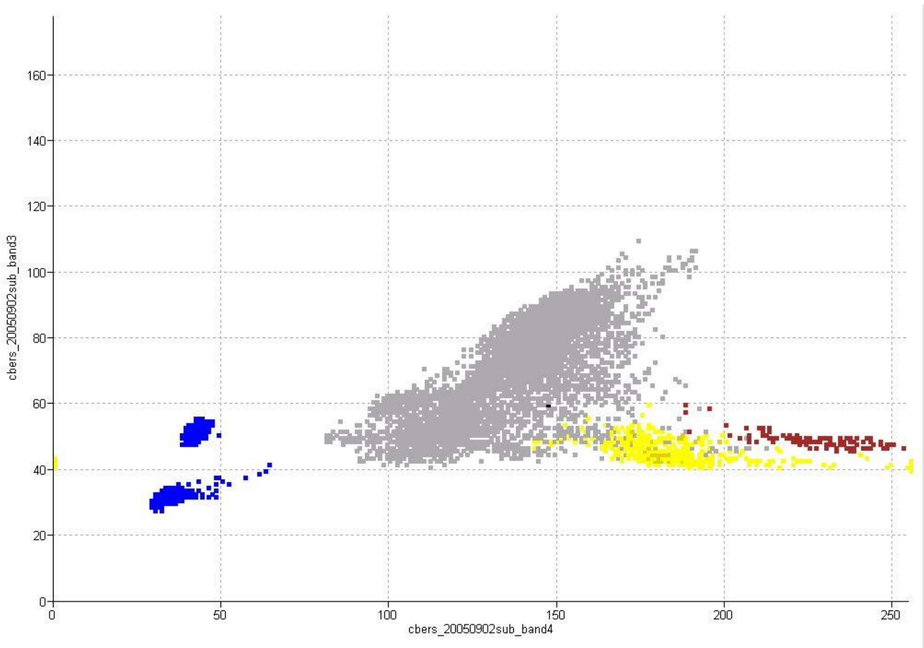


figure 4.1: Feature space of classification of CBERS-2 image of 2005-09-02. Blue: water, Yellow: banana, Brown: beans, Grey: rest

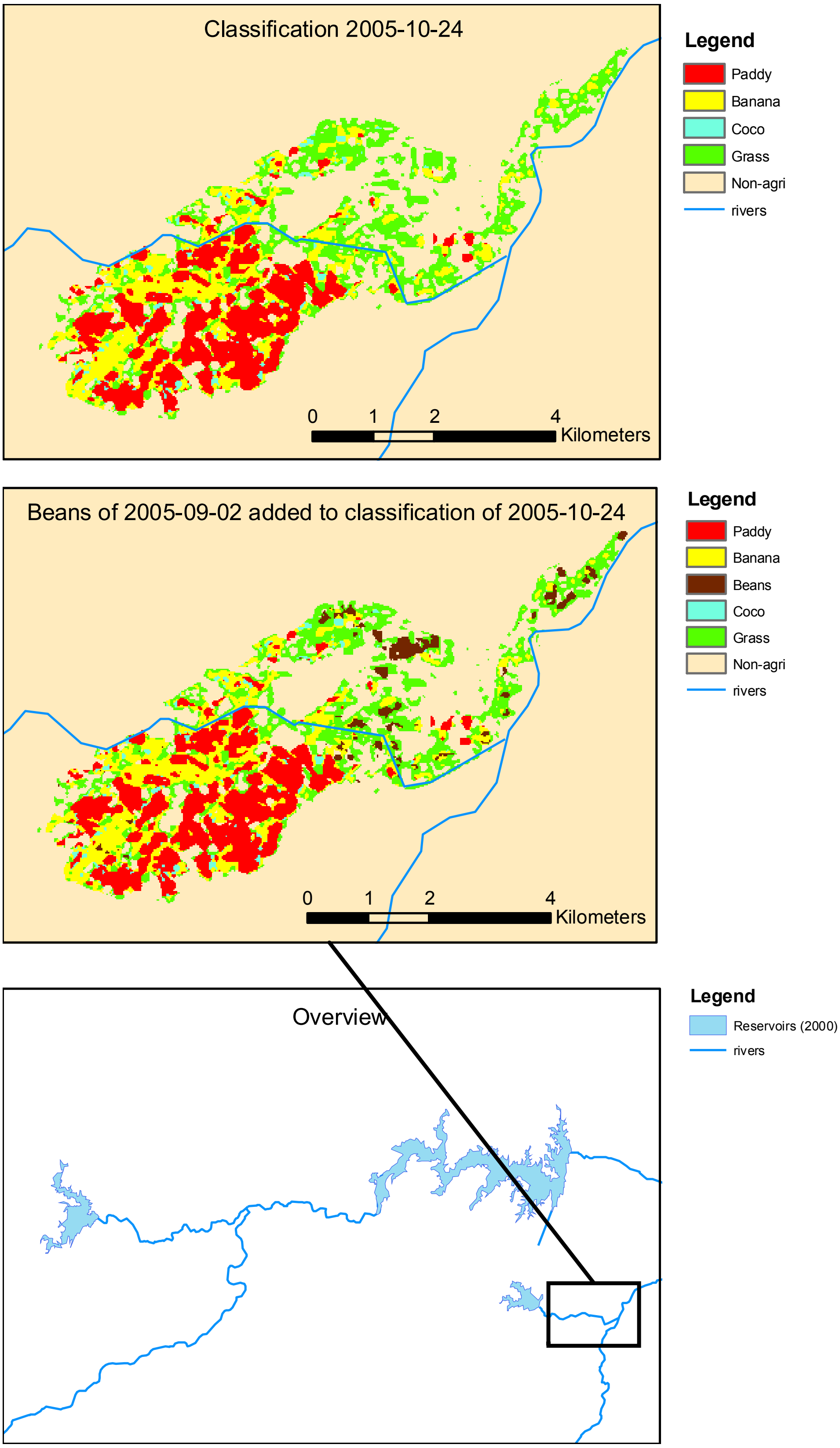


figure 4.2:Beans of 2005-09-02 added to classification of 2005-10-24, AOI 6: Perímetro

4.4 *Result land use classification*

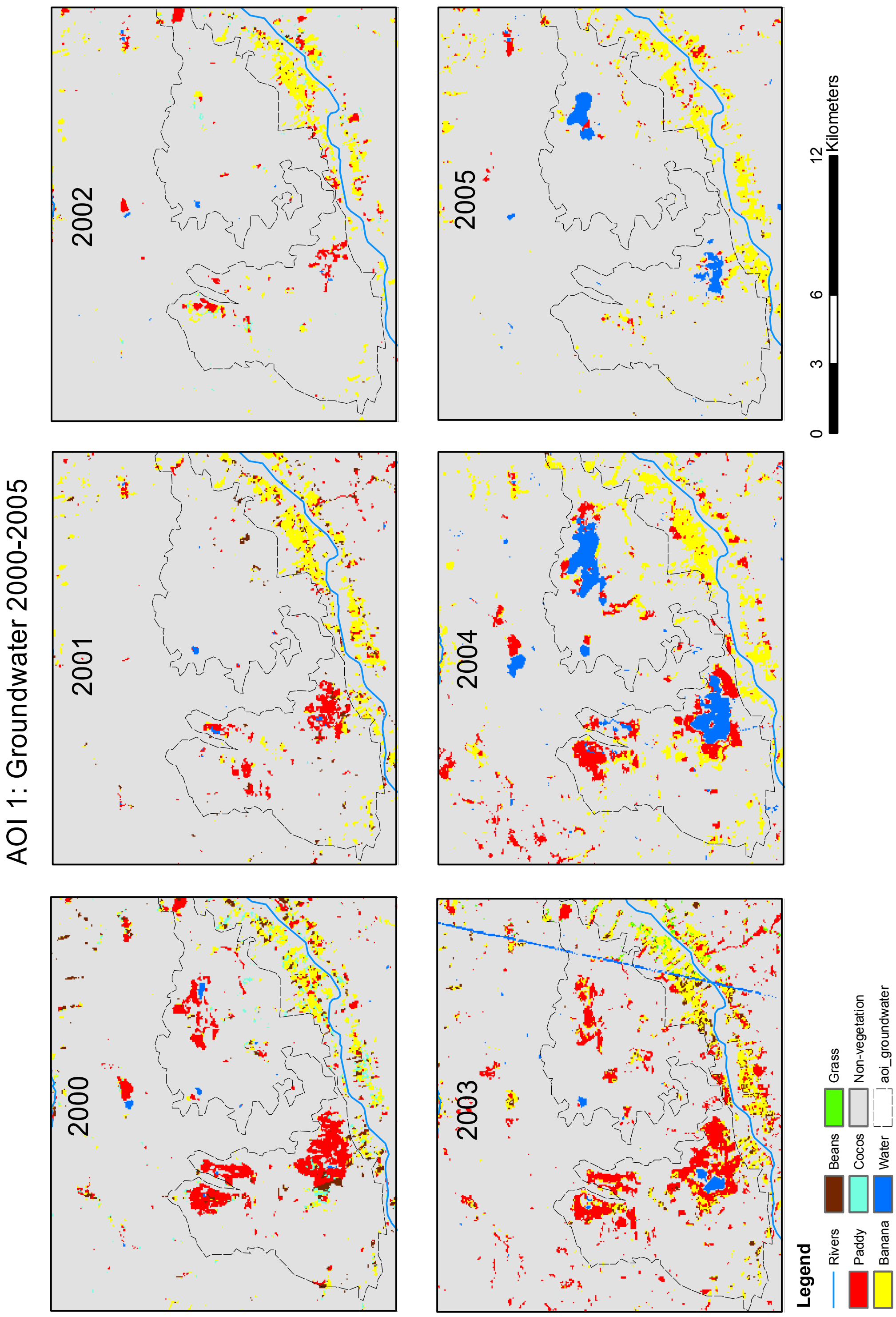


figure 4.3: Results classification AOI 1

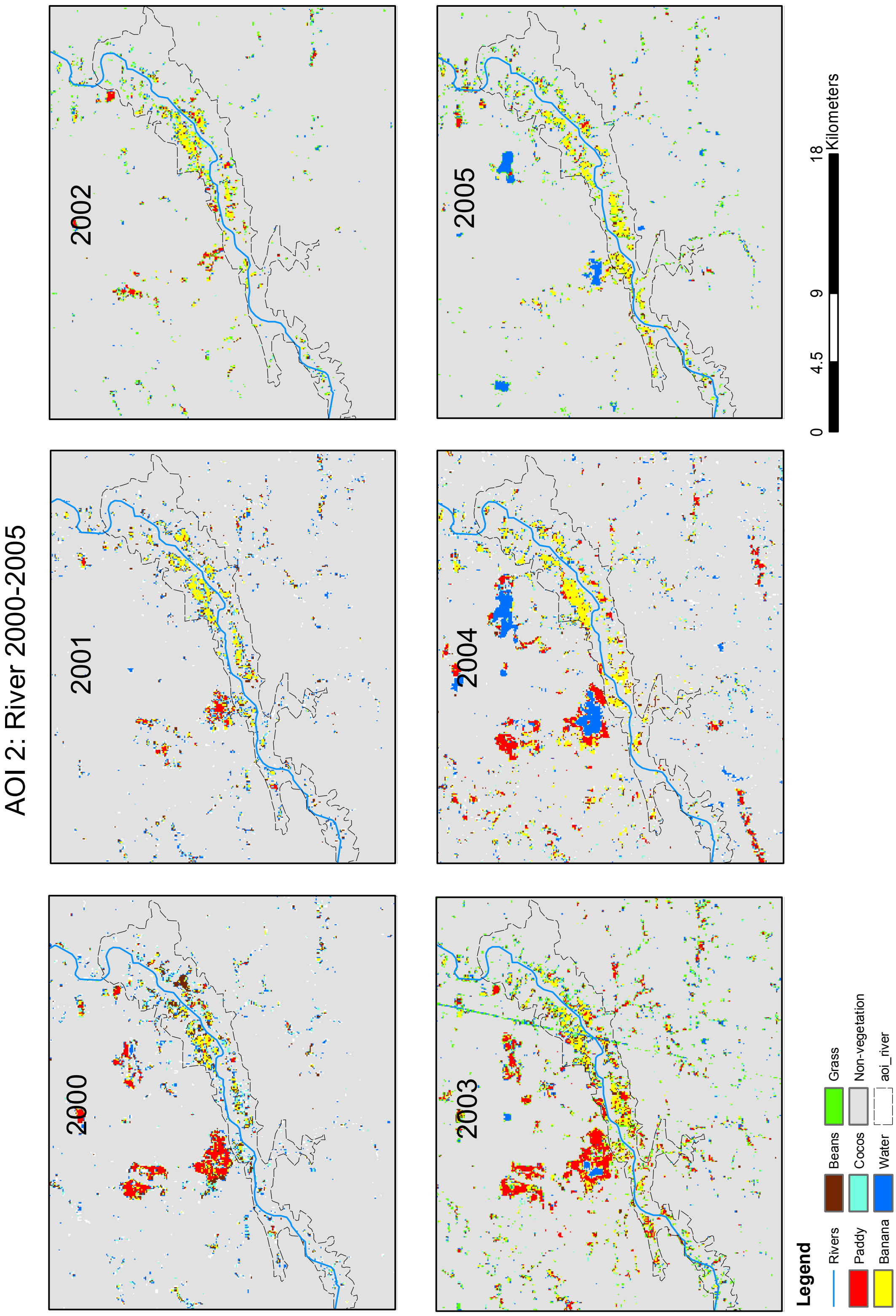


figure 4.4: Classification AOI 2

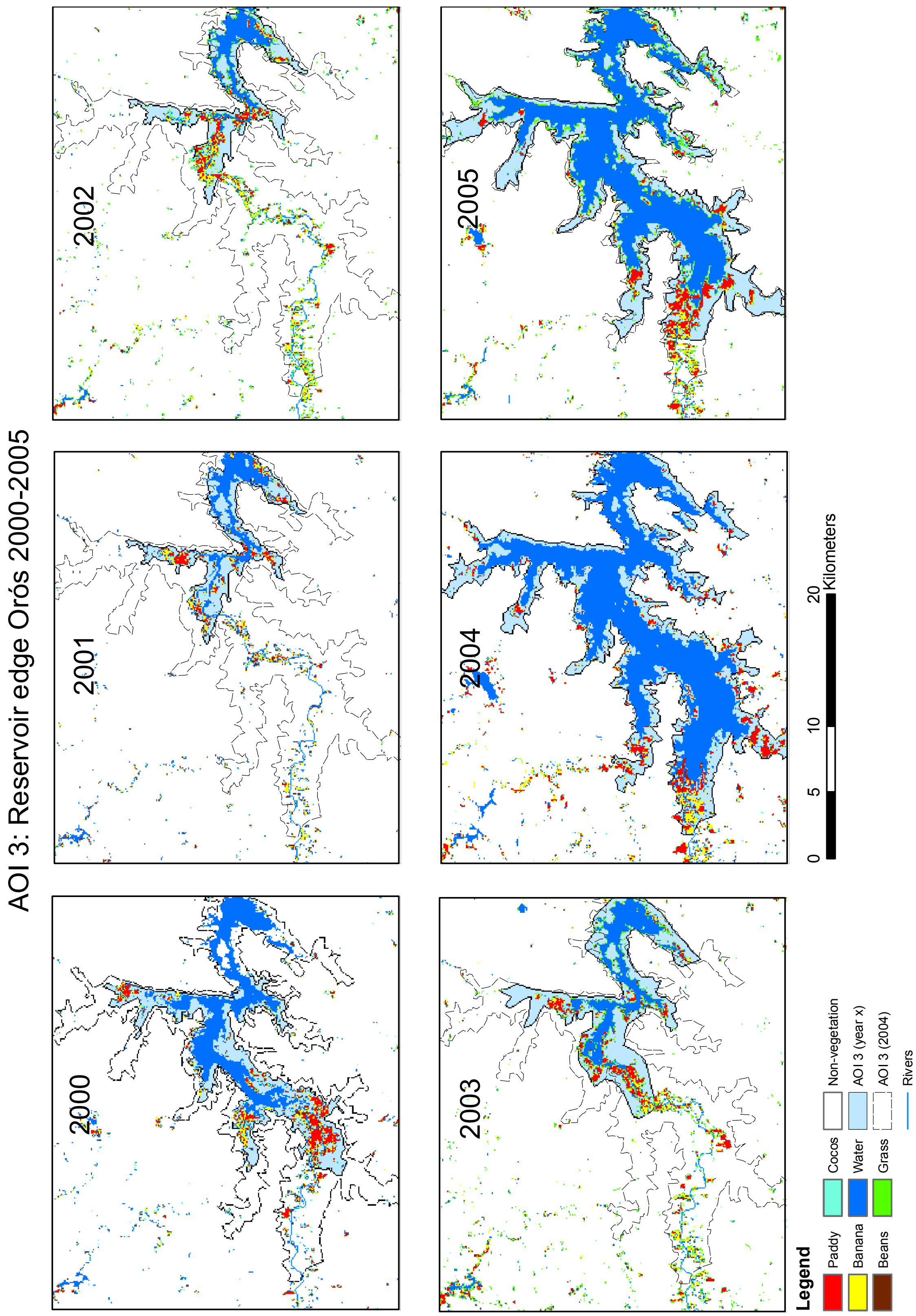


figure 4.5: Result classification AOI 3

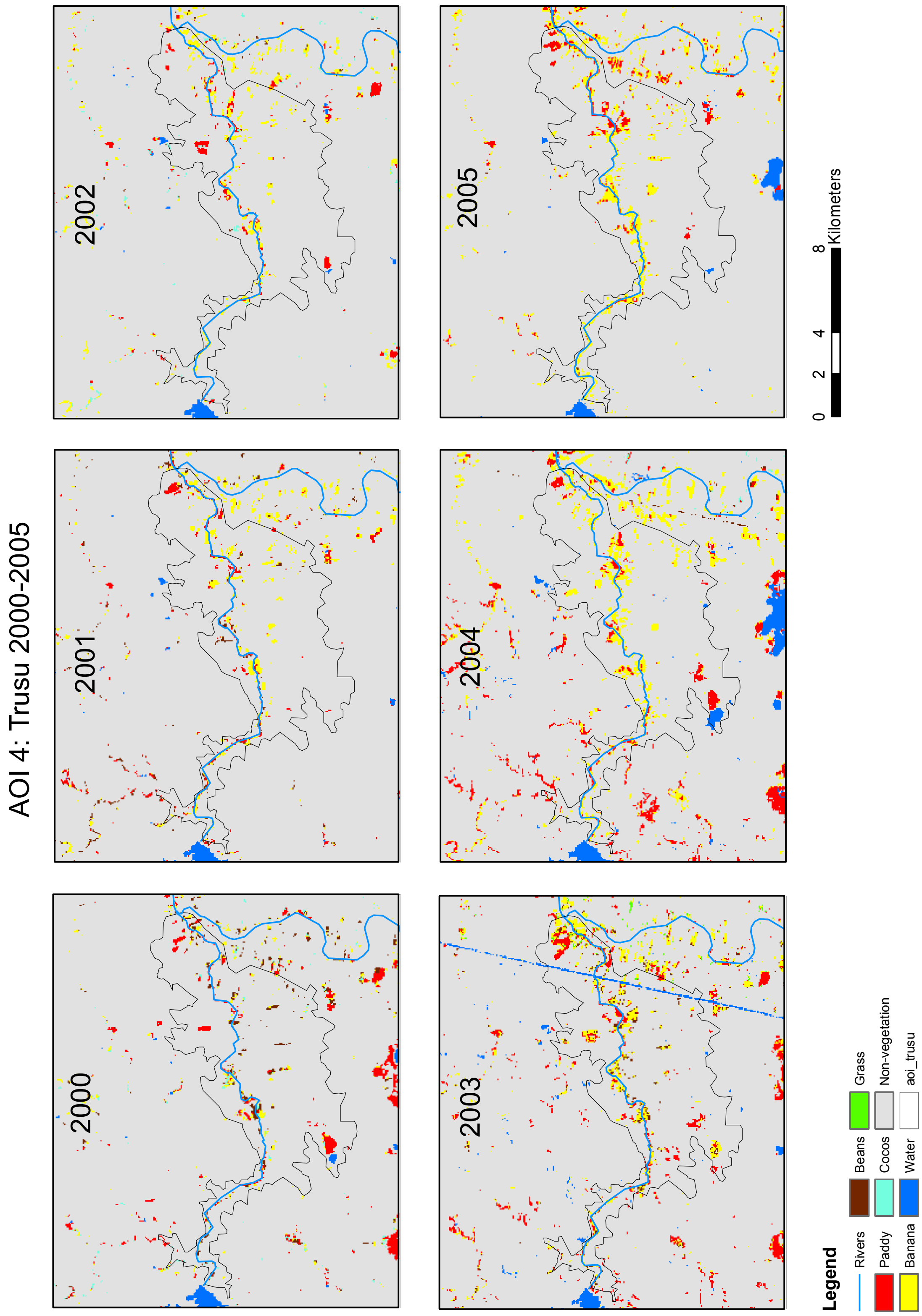


figure 4.6: Results classification AOI 4

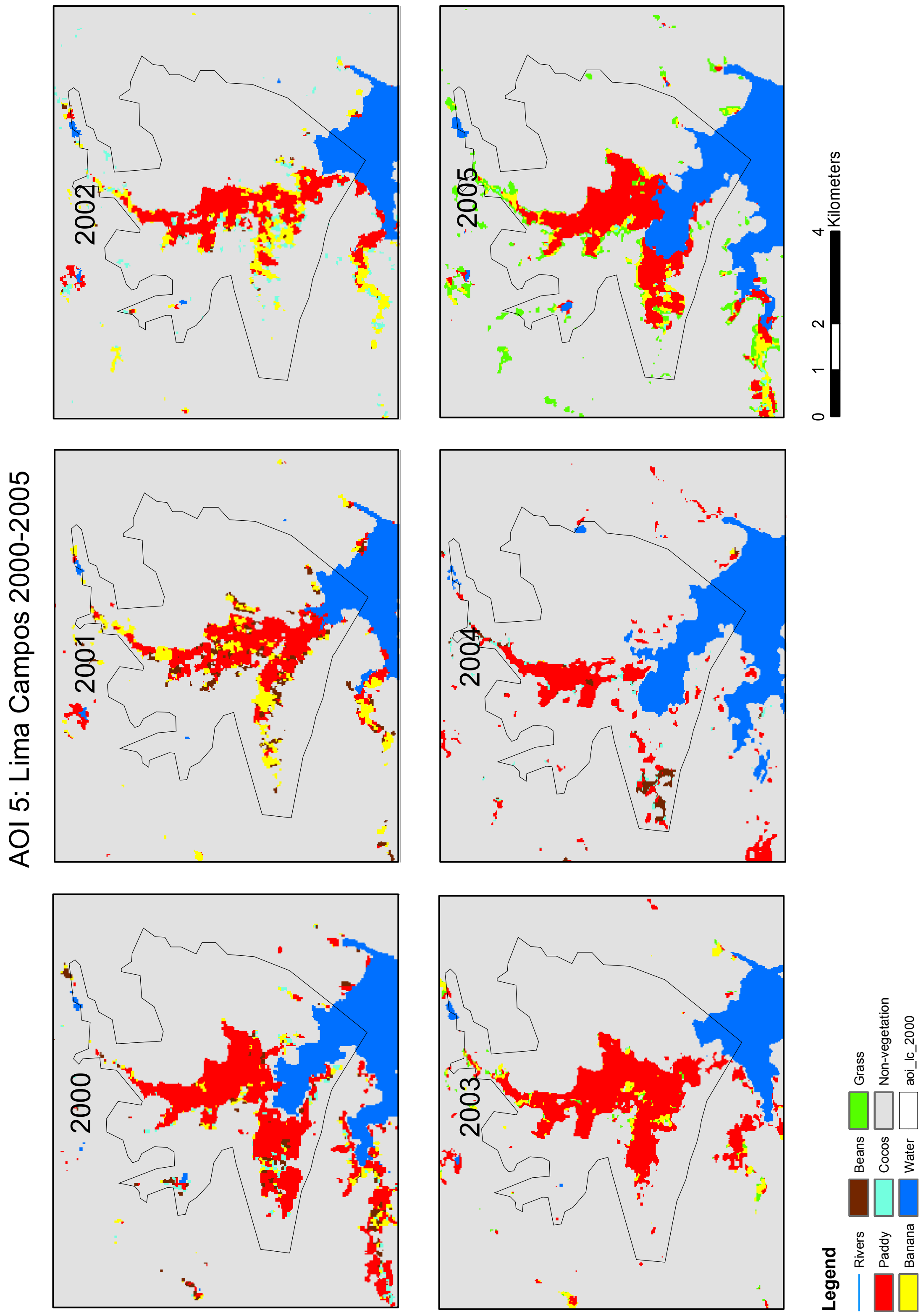


figure 4.7: Result classification AOI 5

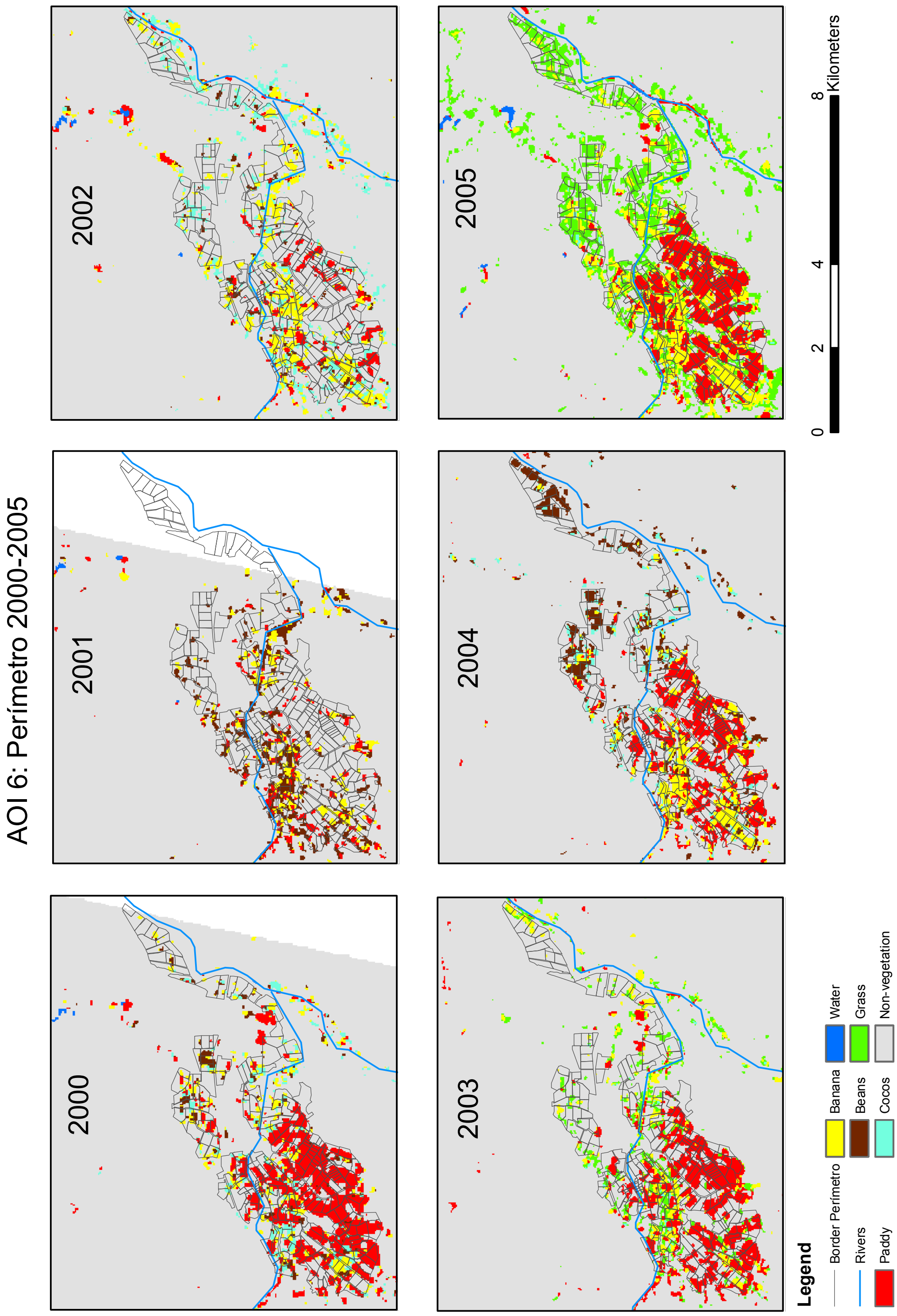


figure 4.8: Land use classification of AOI 6: Perímetro

4.5 Comparison numerical data with classification results

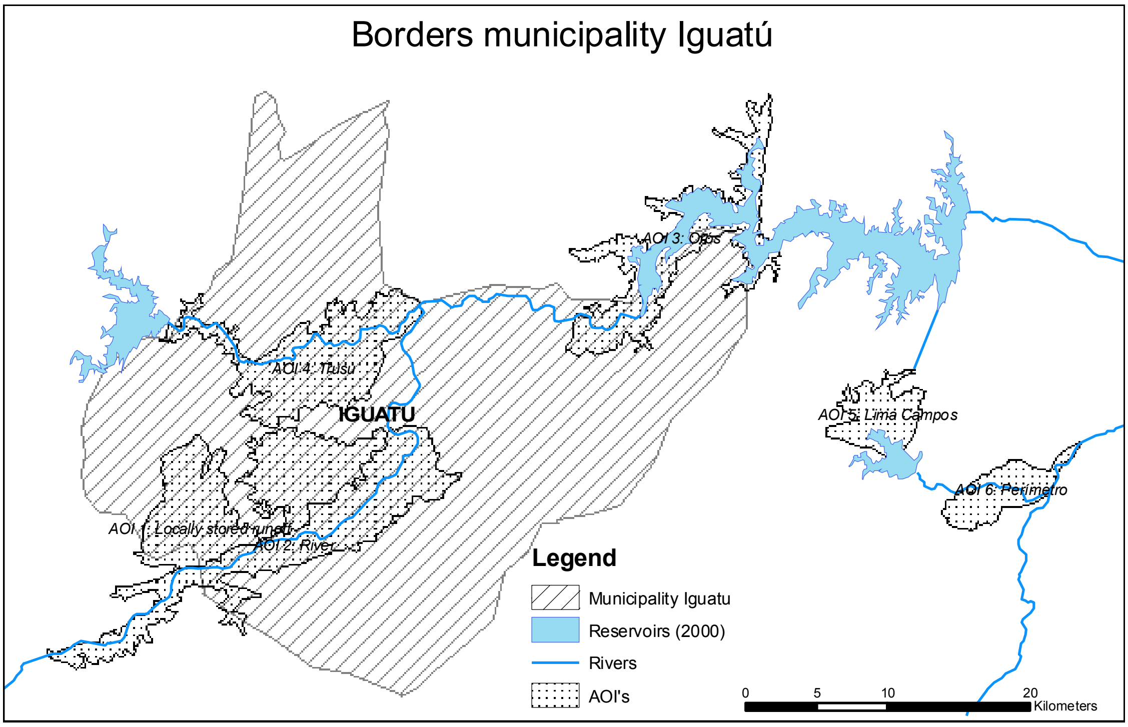


figure 4.9: Borders Iguatú according to WAVES data

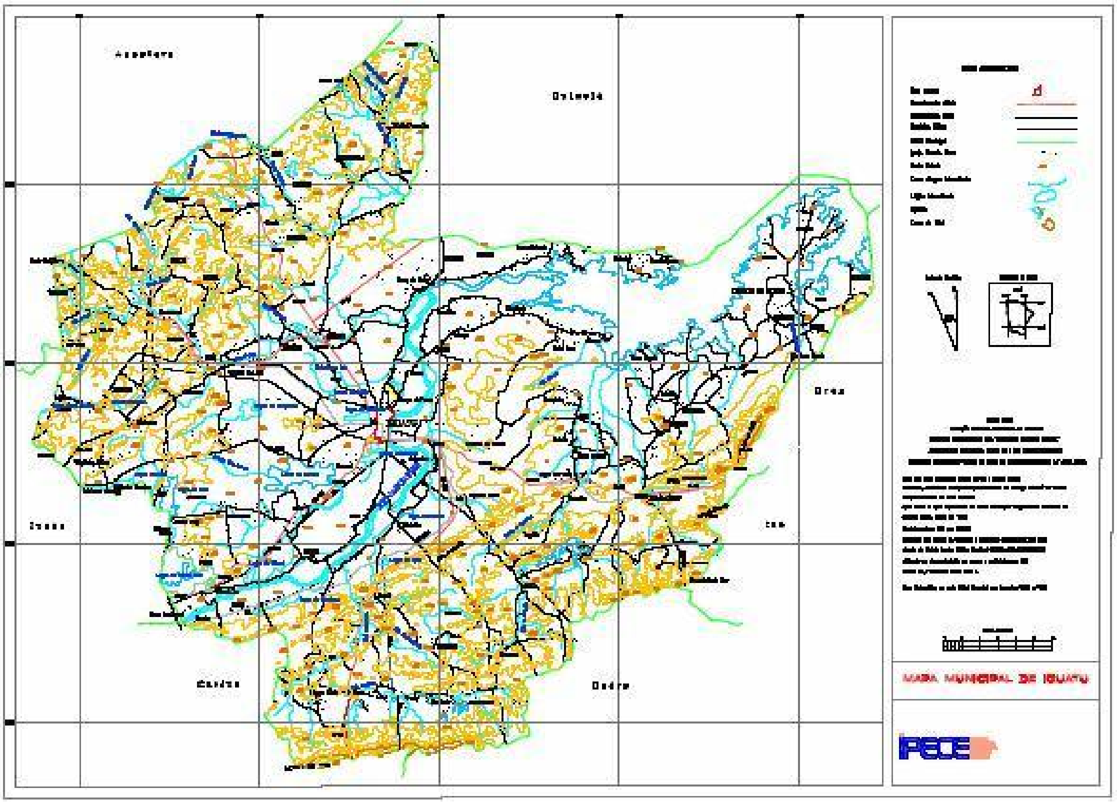


figure 4.10: Borders Iguatú according to IPECE (1998)

4.6 Land use input data for calculation water use

AOI	Crop	2000 A (ha)	2000 A (%)	2001 A (ha)	2001 A (%)	2002 A (ha)	2002 A (%)	2003 A (ha)	2003 A (%)	2004 A (ha)	2004 A (%)	2005 A (ha)	2005 A (%)
1	Banana	111	10	123	28	111	38	297	22	426	38	165	72
	Beans	108	10	1	0	11	4	71	5	3	0	13	5
	Paddy	861	80	319	72	167	58	984	73	682	61	51	22
	Total	*1079*	*100*	*443*	*100*	*289*	*100*	*1352*	*100*	*1111*	*100*	*229*	*100*
2	Banana	660	58	935	87	616	72	850	52	1123	84	1011	88
	Beans	253	22	2	0	96	11	292	18	16	1	33	3
	Paddy	218	19	139	13	141	17	499	30	192	14	110	10
	Total	*1130*	*100*	*1077*	*100*	*854*	*100*	*1642*	*100*	*1330*	*100*	*1153*	*100*
3	Banana	262	20	217	33	174	23	85	13	229	13	338	20
	Beans	67	5	0	0	67	9	61	9	26	1	233	14
	Paddy	987	75	448	67	517	68	535	79	1468	85	1154	67
	Total	*1317*	*100*	*665*	*100*	*758*	*100*	*681*	*100*	*1723*	*100*	*1725*	*100*
4	Banana	111	28	271	66	250	55	379	49	547	71	528	72
	Beans	125	31	0	0	47	10	171	22	2	0	6	1
	Paddy	163	41	141	34	154	34	228	29	225	29	196	27
	Total	*400*	*100*	*412*	*100*	*451*	*100*	*778*	*100*	*775*	*100*	*730*	*100*
5	Banana	33	7	99	28	65	18	29	6	1	0	62	13
	Beans	23	5	0	0	53	15	11	2	29	13	65	13
	Paddy	436	88	260	72	243	67	466	92	188	86	367	74
	Total	*493*	*100*	*360*	*100*	*362*	*100*	*506*	*100*	*217*	*100*	*494*	*100*
6	Banana	150	14	128	32	176	54	140	20	219	25	395	32
	Beans	350	33	169	43	18	6	82	12	301	34	328	27
	Coco	6	1	8	2	8	2	8	1	8	1	12	1
	Paddy	553	52	92	23	125	38	462	67	354	40	481	40
	Total	*1060*	*100*	*397*	*100*	*326*	*100*	*692*	*100*	*882*	*100*	*1216*	*100*

Notes:

- The area of 'beans' in AOI 1 – 5 is composed of the classification results of beans and grass
- The areas of 'beans' and 'coco' in AOI 6 are taken from DNOCS data
- The remaining values are derived from the land cover classification

5 AOI's

Procedure to determine the relative elevation:

Firstly, the elevation of each point in the river is determined by using the DEM. Secondly, the line of the river is transformed into a point map, with each point a value for the elevation. With the interpolation method 'moving average' the surrounding area could be assigned to a certain elevation. Moving average is a point interpolation which performs a weighted averaging on point values. Because all the points are in a line, the surrounding pixels of that line are the most influenced by the riverpoint that is perpendicular on the flow direction of the river. In this way, a contour around the river emerges with the same elevation distribution as the river itself. When this map is subtracted from the DEM, the relative elevation compared to the elevation of the river is the result (see figure 5.1)

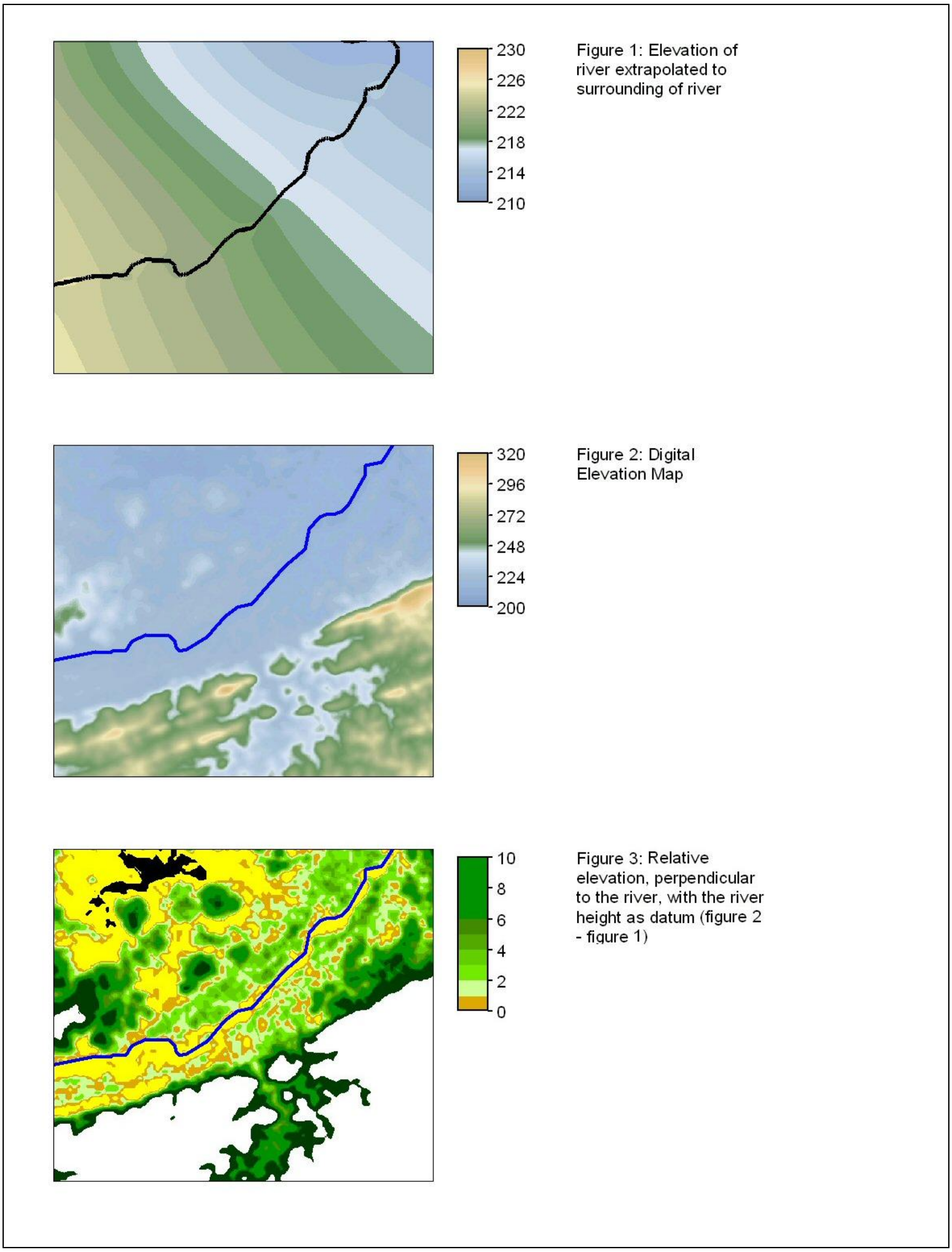


figure 5.1: Procedure to determine relative elevation from river

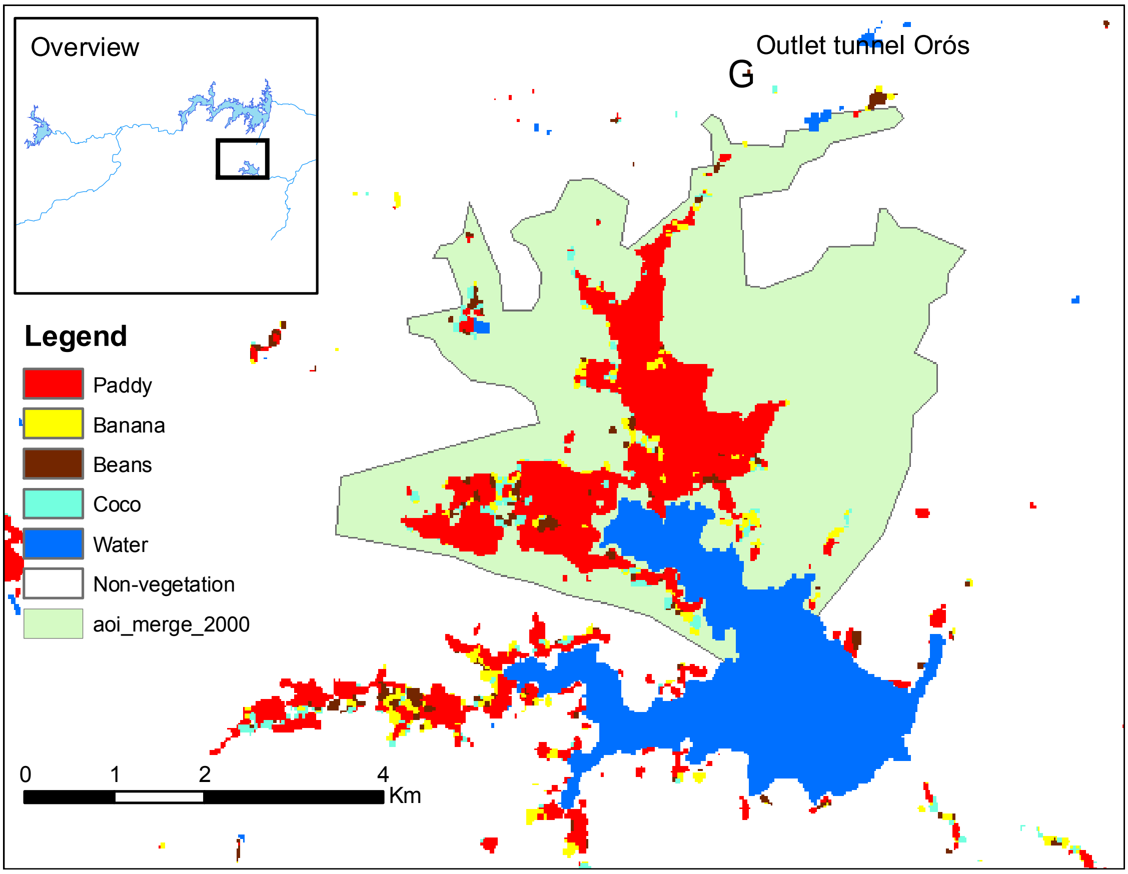


figure 5.2: Boundaries AOI 5

6 Water availability

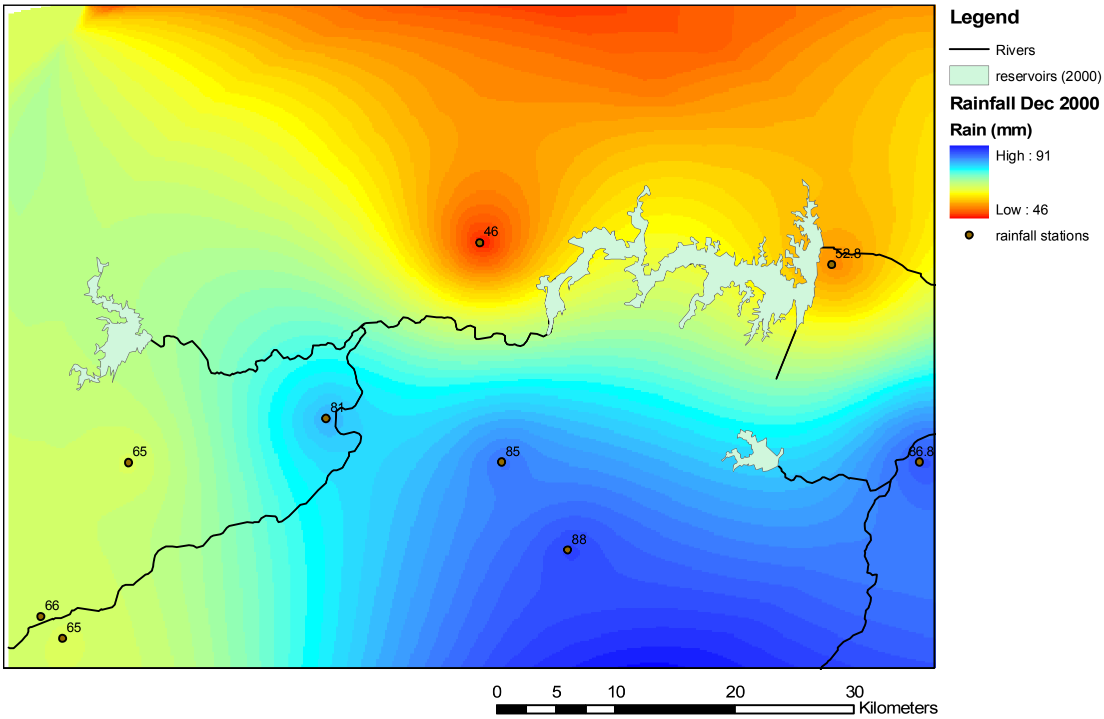


figure 6.1: Example of interpolation rainfall stations (Dec 2000)

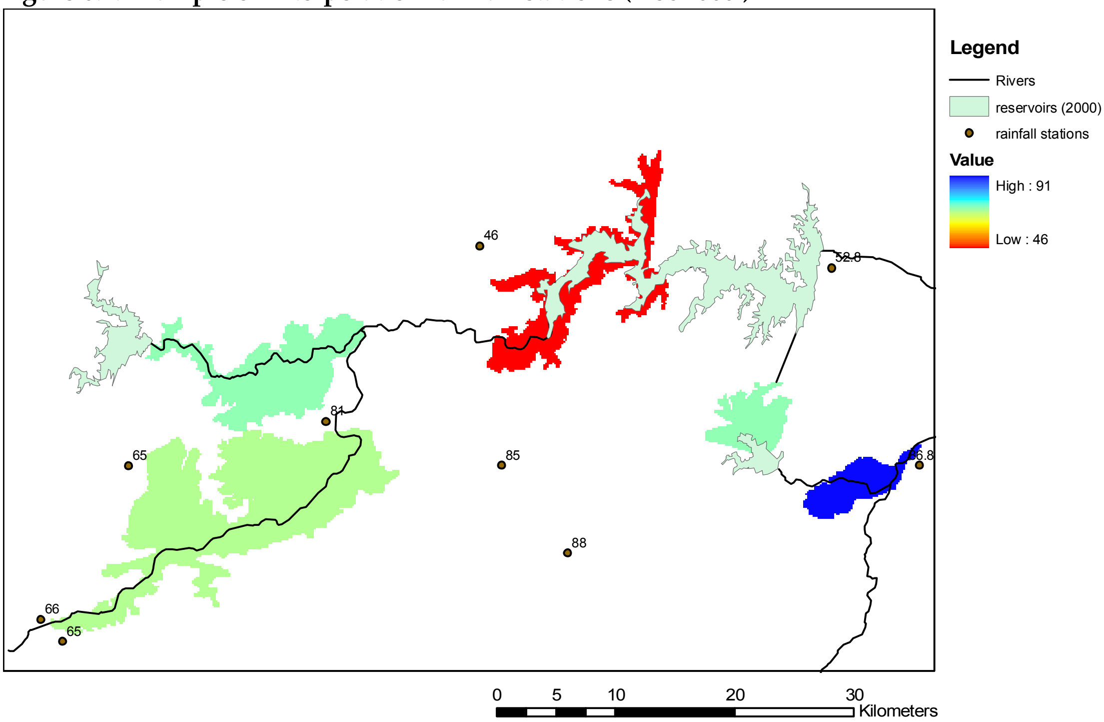


figure 6.2: Example of mean monthly rainfall per AOI (Dec 2000)

7 Land availability

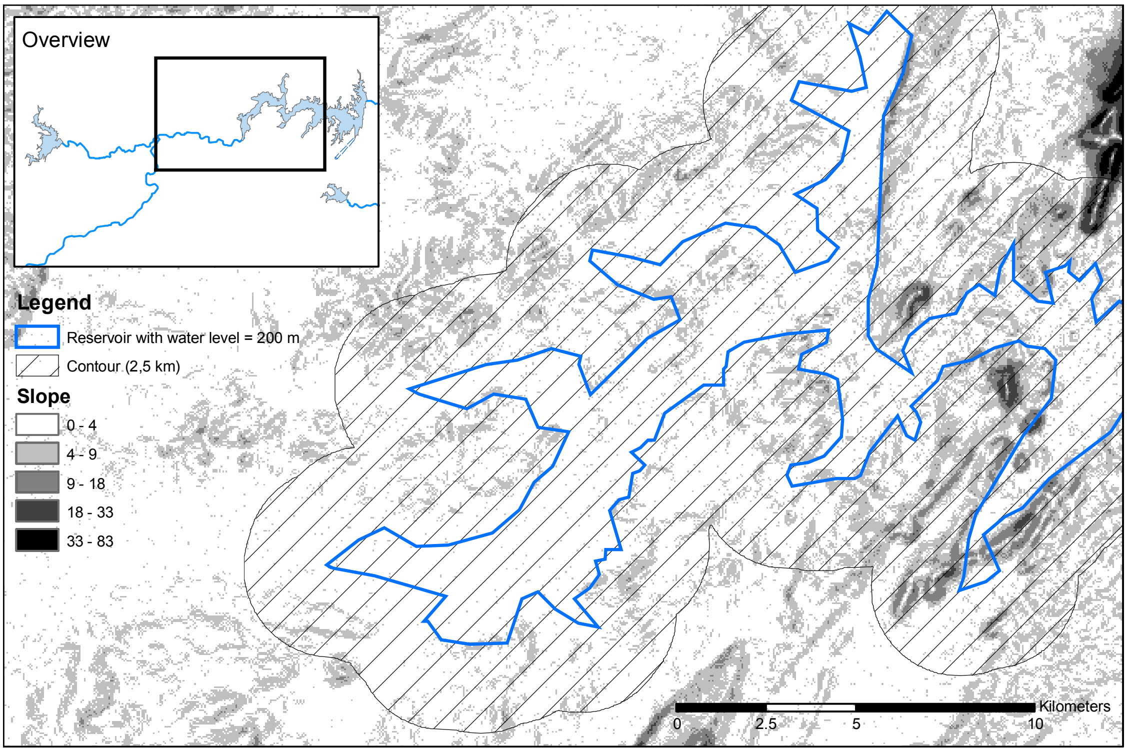


figure 7.1: Slope map

8 Cropwat

8.1 Input

Climate data

Monthly climate data of Iguatú, the main municipality in the study area, was available in Climwat (1993). It includes the climatic data for calculating the reference evapotranspiration (ET_0) according to the Penman-Monteith formulas.

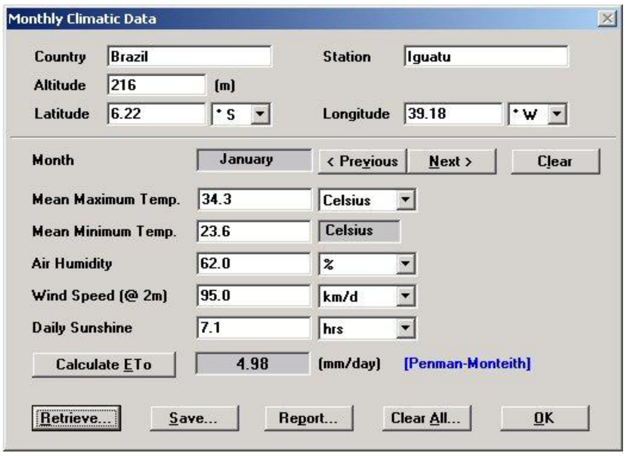


figure 8.1: Climate data

Rainfall data

Monthly rainfall data of each AOI was available, as described in paragraph 3.4.1 of the thesis.

Crop data

In figure 8.2 the input parameters of the crop data are shown. The K_c values are used to calculate the crop water requirement (ET_{crop}) by multiplying it with the reference evapotranspiration (calculated with the climate parameters) with the formula: $ET_{crop} = K_c \times ET_0$. In table 8.1 the K_c-values and the according grow periods are listed. A preference is given to values used in projects in the same study area, since these values are more specified on the local situation. When no 'local' values could be found, the Climwat database is used.

References of table 8.1:

- K_c-values and periods of paddy: Colares (2004). This is study about water consumption of paddy in a region close to the study area.
- K_c-values and periods of Beans: Döll et al (2003). This is a study to future water use in Piauí and Ceará.
- The other values are retrieved from the 'Climwat for Cropwat' database (Climwat, 1993)

For the parameters in the root zone (root depth, depletion and K_y-values), standard parameters of Climwat are used.

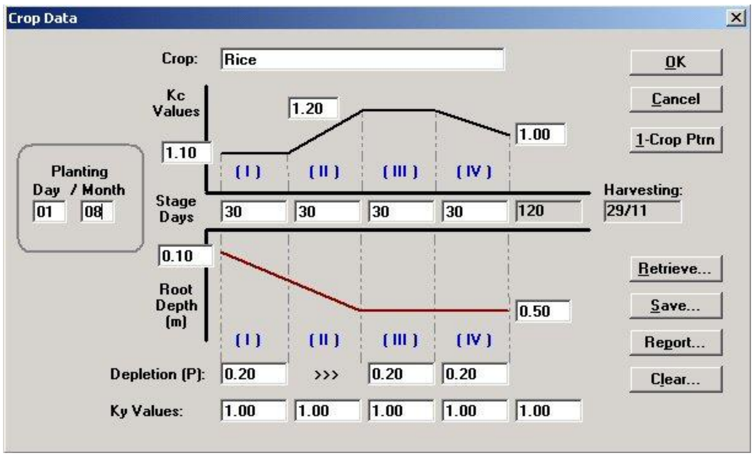


figure 8.2: Crop data

Crop	Period 1 (days)	Period 2 (days)	Period 3 (days)	Period 4 (days)	Period total (days)	Kc 1	Kc 2	Kc 3	Kc 4
Paddy	30	30	30	30	120	1.10	1.15	1.20	1.00
Beans	20	30	30	10	90	0.35	0.75	1.15	0.70
Banana	135	60	140	30	365	1.00	1.10	1.20	1.00
Coco (date palm)	140	30	150	45	365	0.90	0.93	0.95	0.90

table 8.1: Grow periods and K_c-values per crop (Colares, 2004), (FAO, 1992), (Döll et al, 2003)

Cropping pattern

Cropwat allows the user to specify a crop pattern of an area over the whole year. A cropping pattern describes which crops are grown in which period. In figure 8.3 a random example of a cropping pattern is shown.

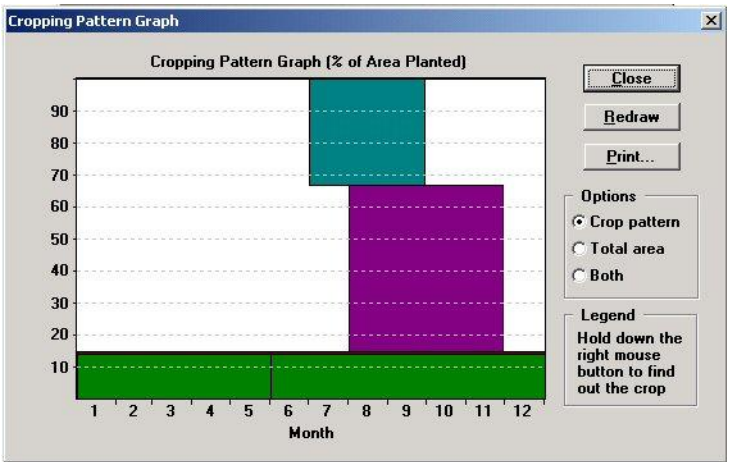


figure 8.3: Random example of a Cropping Pattern (upper: beans, middle: paddy, under: bananas)

Soil data

The standard options in the Cropwat program to describe the soil is: 'light soil', 'medium soil' and 'heavy soil'. The differences between these soils is determined by the initial available soil moisture. A heavy soil can contain more soil moisture than a light soil. Since all AOI's contain mainly alluvial soils and vertisoils (see figure 3.3 of the thesis), known as heavy soils, the parameters belonging to heavy soils are used (see figure 8.4).

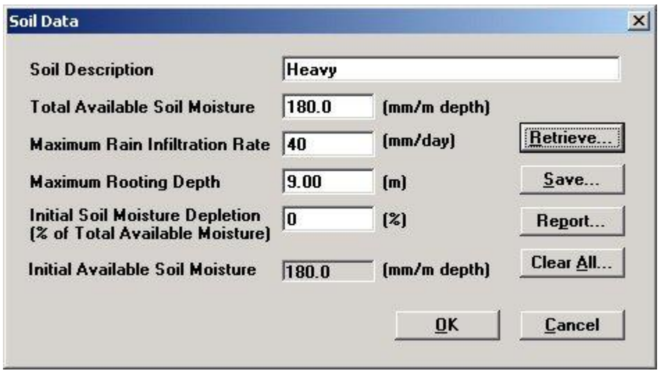


figure 8.4: Soil data

8.2 Output

The used output of Cropwat is a Crop Water Requirement table/graph (see table 8.2 for a random example). It shows per 10 days how much irrigated water have to be add to the agricultural field to keep the soil moisture between the critical SMD (soil moisture deficit) and the field capacity. To calculate the real amount of water that has to be irrigated (the field water supply, FWS) an irrigation efficiency value can be specified. This irrigation efficiency is derived from Colares (2004): 66%.

Crop Water Requirements Table

[All Crops] Time Step (Days): 10 Update Report...

[All Blocks] Irrigation Efficiency (%): 66 Close

Date	ETo (mm/period)	Crop Area (%)	Crop Kc	CWR (ETm) (mm/period)	Total Rain (mm/period)	Effect. Rain (mm/period)	Irrig. Req. (mm/period)	FWS (l/s/ha)
31/5	41.96	10.00	0.10	4.20	1.62	1.28	2.92	0.05
10/6	43.21	10.00	0.10	4.32	1.06	0.96	3.36	0.06
20/6	44.65	10.00	0.10	4.47	0.75	0.75	3.72	0.07
30/6	46.26	19.00	0.13	6.09	1.39	1.39	4.69	0.08
10/7	47.99	20.00	0.14	6.48	2.03	2.03	4.45	0.08
20/7	49.80	20.00	0.15	7.32	3.16	3.01	4.32	0.08
30/7	51.63	84.00	0.88	45.45	20.41	17.58	27.87	0.49
9/8	53.44	100.00	1.08	57.74	32.50	26.46	31.28	0.55
19/8	55.17	100.00	1.10	60.41	40.63	31.77	28.64	0.50
29/8	56.77	100.00	1.10	62.71	45.60	34.74	27.96	0.49
8/9	58.18	100.00	1.13	65.80	44.12	33.14	32.67	0.57
18/9	59.36	100.00	1.14	67.52	31.96	24.07	43.44	0.76
28/9	60.27	91.00	1.07	64.28	6.46	5.22	59.07	1.04
8/10	60.85	90.00	1.06	64.53	0.00	0.00	64.53	1.13

table 8.2: Random example of a Crop Water Requirements Table

9 Inter AOI analysis

year	Water use AOI 5+6 (*10^6 m3)	Perc. of capacity volume Orós (m3)	Volume Orós reservoir at 1 June (*10^6 m3)	level Orós reservoir at 1 June (m)	decrease water level Orós reservoir (m)	Level Orós reservoir after water use AOI 5+6 (m)	available land at 1 june (m2)	available land after water use AOI 5+6 (m2)	difference in land avail (m2)	Relative difference in land avail. (%)
2000	15.35	0.52	1013.48	193.3	0.13	193.17	4180	4050	-130	-3.1
2004	12.58	1	1949	198.8	0.07	198.73	7167	7139	-28	-0.39
2005	16.79	0.8	1559.2	197.5	0.1	197.4	6585	6540	-45	-0.68

References

[Only the references not mentioned in the report are listed].

FAO (1993), *Climwat for Cropwat - A climatic database for irrigation planning and management* by M. Smith, FAO Irrigation and Drainage Paper No. 49, Rome: FAO

Colares D.S. (2004), *Análise Técnico-económica do Cultivo do Arroz no Perímetro Irrigado Morada Nova, Ceará*, Fortaleza, Brazil: Universidade Federal do Ceará, Centro de Ciências Agrárias, Departamento de engenharia agrícola curso de mestrado em irrigação e drenagem

Döll P., Hauschild M., Mendiondo E.M., Carlos de Araújo J., *Modeling of Present and Future Water Use in Piauí and Ceará as a Basis for Water Resources Planning*, Berlin: Springer