INPE-00000-TDI/0000

REGIONALIZED SPATIO-TEMPORAL MODELLING OF WATER TABLE DEPTHS IN THE BRAZILIAN CERRADO

Rodrigo Lilla Manzione

Doctorate Thesis at Post Graduation Course in Remote Sensing, advised by Dr. Gilberto Câmara, approved at September 27, 2007.

INPE
São José dos Campos
2007

INPE-00000-TDI/0000

REGIONALIZED SPATIO-TEMPORAL MODELLING OF WATER TABLE DEPTHS IN THE BRAZILIAN CERRADO

Rodrigo Lilla Manzione

Doctorate Thesis at Post Graduation Course in Remote Sensing, advised by Dr. Gilberto Câmara.

INPE
São José dos Campos
2007

APPROVAL PAGE

"In nature nothing is created and nothing is lost, everything is transformed".

Lavoisier

To my family

ACKNOWLEDGMENTS

ABSTRACT

Water regimes governed by seasonality are sensitive to climatological disturbances and human interventions. The Brazilian Cerrados are characterized by a well pronounced dry season, from May until September. During this period, natural vegetation and agricultural crops are dependable of groundwater. The Cerrado natural vegetation is adapted to the local climate, but the cash-crops cultivated in the region not. Irrigation techniques are responsible for the maintenance of high productivities during the whole year and the availability of water resources made it possible. Nowadays, with almost all Cerrado vegetation replaced by agricultural crops, information about the spatio-temporal dynamics of the water table is important to optimize and balance the interest of economical and ecological purposes in the Brazilian agricultural frontier. The aim of this thesis is to characterize water resources in a Brazilian Cerrado area. We model the spatio-temporal variation of water table depths in the Jardim River watershed. This is a representative Cerrado area in the Brazilian Federal District, where almost all natural vegetation was replaced by agricultural crops and irrigation has substantially increased during the past years. This information is needed to monitoring the behaviour of the water levels due to land use, human interventions and climatic changes. Also, to evaluate risks associated with water levels, and to present strategies in water management. For this purpose we applied geostatistical methods and time series modelling to observed water table depths series, describing water table dynamics and accounting for uncertainty. The results presented here show how to estimate the water volume lost during a specific season and delimitation of favourable and unfavourable areas to water use, using a linear coregionalization model. We showed how to apply the PIRFICT time series model to the Cerrado situation, accounting for systematic changes in water table depths and presenting how to predict risks of extreme water levels for agriculture in the region. These results aim at contributing to improve water management, when using Cerrado areas for crop production.

MODELAGEM ESPAÇO-TEMPORAL REGIONALIZADA DE ALTURAS DE LENÇOL FREÁTICO NO CERRADO BRASILEIRO

RESUMO

Regimes hídricos governados pela sazonalidade são sensíveis a distúrbios climáticos e intervenções humanas Os Cerrados brasileiros são caracterizados por uma estação seca bem pronunciada, de maio a setembro. Durante esse período a vegetação natural e os campos agrícolas são dependentes da água subterrânea. A vegetação dos Cerrados é adaptada ao clima local, mas os campos cultivados na região não. Técnicas de irrigação são responsáveis pela manutenção das altas produtividades durante todo ano, e a disponibilidade de recursos hídricos torna isso possível. Atualmente, com quase toda vegetação de Cerrado substituída por campos agrícolas, informações sobre a dinâmica espaço-temporal dos níveis freáticos é importante para otimizar e balancear os interesses econômicos e ecológicos na fronteira agrícola brasileira. O objetivo dessa tese foi caracterizar os recursos hídricos em uma área de cerrado, modelando a variabilidade espaço-temporal das alturas de lençol freático na bacia do Rio Jardim. Essa é uma área representativa de Cerrado localizada no Distrito Federal, onde quase toda vegetação natural foi substituída por agricultura e onde a irrigação aumentou substancialmente nos últimos anos. Essa informação é necessária para monitorar o comportamento dos níveis de água em função do uso do solo, intervenções humanas e mudanças climáticas. Também para avaliar o risco associado aos níveis de água, e para apresentar estratégias para o manejo da água. Com esse propósito, aplicam-se métodos geoestatísticos e modelos de séries temporais a dados observados de alturas de lençol freático, descrevendo a dinâmica do lençol e considerando a incerteza contida nos modelos. Os resultados apresentados aqui mostraram como estimar o volume de água perdido pelo sistema aqüífero durante uma estação especifica e delimitar áreas favoráveis e desfavoráveis para o uso da água, através de um modelo linear de corregionalização. Também se apresentou a aplicação do modelo PIRFICT de séries temporais nos Cerrados, verificando mudanças sistemáticas nos níveis freáticos e apresentando como estimar riscos de níveis extremos de alturas do lençol freático para agricultura

na região. Esses resultados visam contribuir na melhora do manejo de água ao se utilizar de áreas de Cerrado para produção agrícola.

SUMMARY

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AR   Auto Regressive
BR   Block Response
CAPES   Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior
CI   Calculate first/Interpolate later
Cx   Haplic Cambisol
DEM   Digital Elevation Model
DF   Distrito Federal
DPI   Divisão de Processamentode Imagens
DSR   Divisão de Sensoriamento Remoto
EMBRAPA   Empresa Brasileira de Pesquisa Agropecuária
EV   Elevation
Fx   Haplic Plinthosol
Gx   Haplic Gleysol
INPE   Instituto Nacional de Pesquisas Espaciais
IC   Interpolate first/Calculate later
IR   Impulse Response
LCM   Linear Coregionalization Model
LDB   Local Drainage Base
LTP   Linear Trend Parameter
LU   Land Use
LV   Red Latosol
LVA   Red Yellow Latosol
NDV   Natural Discharge Volume
OUB   Ornstein-Uhlenbeck–based
PDF   Probability Distribution Function
PIRFICT   Predefined Impulse Response Function In Continuous Time
PRODETAB   Projeto de Apoio ao Desenvolvimento de Tecnologias
Agropecuárias para o Brasil
RF   Random Function
RMSE   Root Mean Square Error
RMSI   Root Mean Square Inovations
RV   Random Variable
SD   Standard Deviation
SPG   Serviço de Pós-Graduação
ST   Spatio-Temporal
TFN   Transfer Function Noise
UK   Universal Kriging
WTD   Water Table Depths

1 INTRODUCTION

1.1 Background

Managing natural resources for sustainable development is a worldwide challenge. Precious metals, gemstones, oil and gas are not the only major reasons of current conflicts. Access to land and water are just as important. The natural availability of water has decreased around the world. Many regions are experiencing water scarcity, several of them for the first time. One reason for water scarcity is the constant water demand by agriculture since modern agriculture consumes soil, seeds, fertilizers, water, energy, chemical products, people, and infrastructure. Farmers are dependent on a combination of these factors to reach high productions during all year and keep the activity profitable, even facing difficulties from climate change and competing claims on natural resources.

To support water management, we need to monitor water resources, model hydrological processes, and simulate the effects of public policies. To develop useful hydrological models, we need to consider the uncertainty of the predictions to increase the value of the models in decision-making.

In this thesis, we will consider the problem of modelling water availability for agriculture. A key idea we use is water table depth. Observations on water wells are one dimension observations on a continuous surface of phreatic water. "This surface is referred to as water table. The depth of water table below the surface is called water table depth" (Knotters, 2001). Water table depth has significant importance for agricultural and ecological potentials. Many conditions for agriculture are related to water table depths, as soil trafficability, water availability and mainly soil conditions. They influence several soil characteristics, like temperature, degree of aeration, concentration of nutrients, degree of acidity, and thickness of the root zone. In contrast with water table depths, these features are hard to be observed (Knotters, 2001).

To balance the needs of economic land use and ecological conservation, knowledge about the spatio-temporal dynamics of the water table is important (Von Asmuth and Knotters, 2004). Information on water table depths helps assessments in areas, such as on desiccation of natural areas, water shortage in agriculture, fertilizers leaching, and drinking water or urban water supply. Reliable estimates of water table depth contribute to a more sustainable and equitable use of natural resources, especially if they account for uncertainty and assess risks.

In some countries, the water table depth is intensively monitored because of its significance. In the Netherlands, for instance, the seasonal variations of water table depths influence agricultural production because of its shallow depth. Tanco and Kruse (2001) predict the seasonal water table variation in Argentina, presenting a statistical method that aims to mitigate impacts of extreme events (like floods and droughts). Leduc et al. (1997) present results about aquifer recover in southern Niger. Many times authorities just take actions to understand water table dynamics after some problem has occurred. The network of water wells and boreholes used in Leduc et al. (1997) for example, was set up after the severe drought of 1970s and early 1980s in the country. Sun et al. (2000) studied water table rises after human interventions in the wetlands of northern central Florida. In Brazil, several studies have been conducted in the different ground water systems of the country. Here, we gonna refer just to works in the study region, the Brazilian Cerrados. Souza and Campos (2001) studied the importance of the regolith in the porous domain of an aquifer and how the differences in hydraulic conductivity in this material influence the Paranoá aquifer system recharge. Lousada and Campos (2005) presented four conceptual models to explain the groundwater flow conditions in the Brazilian Federal District based in hydrogeological information. Cadamuro and Campos (2005) proposed solutions for artificial aquifer recharge in areas with problems availability after human intervention in the hydrological system.

1.2 The Brazilian Cerrado

The Cerrado is a fragile environment and the current Brazilian biggest grain belt. Human activities have changed the region during the past 40 years, and exploited its natural resources. In the following discussion, we discuss the need for information about water resources in the Cerrado region, to provide tools for sustainable water management.

1.2.1 Characterization of the Brazilian Cerrado: a survey

The Cerrado region is the most extensive woodland-savannah in South America. The Brazilian Cerrados extend from the northern margins of the Amazonian evergreen forests to outliers on the southern borders of the country with extensions into Paraguay and Bolivia. Placed at the central plateau, it covers 22% of the Brazilian territory (Figure 1.1) or around 2 million km^2 (Jepson, 2005). This ecosystem is a biodiversity hotspot (Myers, 1988, 1990).
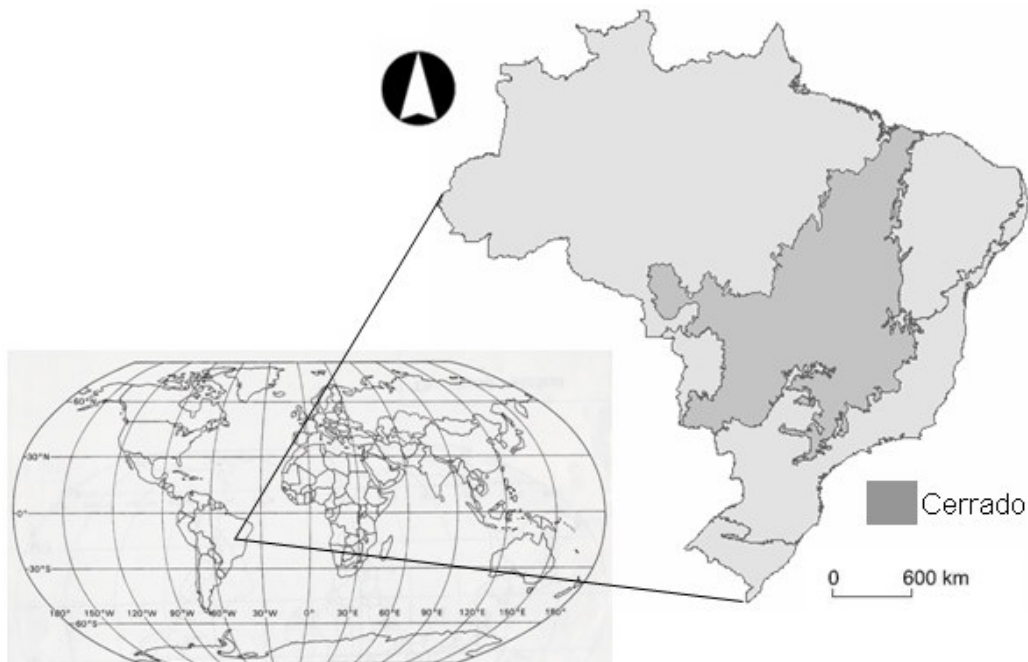
Figure 1.1 - Brazilian territory and the Cerrado area.

The Cerrado is one of 25 biologically rich areas around the world that have lost at least 70% of their original habitat. The remaining natural habitat in these biodiversity hotspots amounts to just 1.4 percent of the land surface of the planet, yet supporting nearly 60 percent of the world's plants, birds, mammal, reptile, and amphibian species. The Cerrado supports a unique array of drought and fire adapted plant species and many endemic bird species and large mammals that are competing with the rapid expansion of Brazil's agricultural frontier, which focuses mainly on soybeans and corn.

The climate in the Cerrados presents seasonal patterns. The mean annual temperature is 22 ^{0}C in the south of the region and 27 ^{0}C in the north (Goedert, 1989). The region receives annual rainfall between 1300 and 1600 mm, concentrated in six to seven months, starting in October and finishing around April. A pronounced dry season characterizes the rest of the year. As in the whole intertropical zone, dry spells (locally called "veranicos") can occur during the wet season, getting crops in early stages of development, reducing productivity and affecting the local economy (Assad et al., 1993). Due to its random nature, rain is one of the factors that most influence agricultural production (Assad, 1994).

Most of the soils under the Cerrado are highly weathered Oxisols (46%), Ultisols (15%) and Entisols (15%), posing serious limits for crop production because of low natural soil fertility (Lopes, 1996). In general the soils are deep and present high aluminium saturation and high phosphorus fixation capacities. Many crops have shallow roots due to aluminium toxicity and calcium shortage in subsurface soil layers.

The natural vegetation has developed adaptations to the seasonally wet rainfall, acidic soils, and aluminium toxicity. Plants metabolize throughout the year, drawing on soil water reserves, and can withstand short-lived fires. Fire has an important role on the dynamics, maintenance and evolution of savannas, being

in the Cerrado a major determinant of soil characteristics, nutrient cycling and plants dynamics (Mistry, 1998). The Cerrado has many types of physiognomic forms, from open formations, dominated by dense and continuous herbaceous layer nearly treeless grasslands, to closed formations, dominated by woodland of semideciduous trees (Furley, 1999). The natural vegetation is well adapted to drought presenting deep roots.

1.2.2   The Cerrado as Brazil's major grain belt

To understand the current landscape of the Cerrado it is important to consider recent agricultural expansion in Brazil. Until the mid1950s, the lack of infrastructure constrained any significant commercial development. This scenario changed drastically after the construction of Brasília, in the heart of the Cerrado region, when highways and railways linked the new capital to the main Brazilian cities. Besides, government policies gave a large boost to agricultural modernization and land transformation, stimulated by subsidized credit, tax breaks and development of new technologies (Klink and Moreira, 2002).

Expansion of modern agriculture and large-scale ranches changed the characteristic landscape and ecosystem of Cerrado. Extensive cattle fields and agricultural crops (mainly corn, soybeans and cotton) established in the Cerrado created positive socio-economic impacts. Agricultural production for Brazil's domestic and export markets has increased, with enlarged planted areas and increasing productivity. Local economies were diversified, municipal revenues increased and welfare services improved in some localities. These areas have become Brazil's most important grain belt, although facing cover change rates much higher than in the Amazonian rainforest (Oliveira et al., 2005). Current deforestation ranges from 22000 to 30000 km^2/year (Machado et al., 2004). Brazil's Forest Code requires for a rural property that only 20% of Cerrado vegetation should be preserved in its natural state as "legal reserve". In the Amazon rainforest this portion is 80% (Klink and Machado, 2005). Machado et

al. (2004) presented in a recent survey using MODIS imagery data from 2002 that 55% of Cerrado was already cleared or transformed for human uses.

Ecosystem experiments and modelling show that this change in land cover is altering the hydrology and affecting carbon stocks and fluxes (Klink and Machado, 2005). Shallow-rooted monocultures, which are less well adapted to drought, have replaced a complex wood and grass ecosystem. Cerrado agriculture is profitable, and agricultural expansion will continue to need improvements in techniques and infrastructure. Using fertilizer and lime to correct soil for loss of crop productivity caused by the high aluminium levels, farmers have converted vast areas of land to agriculture fields. Their need of water supply by irrigation techniques is likely to change the hydrological system (Klink and Moreira, 2002). With irrigation increasing, the water table falls, and so risks of water shortage appear. Preserving these resources is important also because the Cerrado connects the three main river basins of the country (Amazon, Prata and São Francisco). Rivers and brooks are particularly affected by human occupation. Rivers near cities are used as depository of wasting and sewage while rivers near agricultural areas have their banks eroded, causing a progressive decline in the flowing of water. In regions like the Federal District, groundwater has big importance on public water supply. The water exploitation increased substantially after the second half of the 80's. There are estimations about the number of deep tubular wells in the area, which increased around three times since this period, most of them to supply new condominiums, industries and communities (Campos, 2004).

Valuable knowledge gained through research can be disseminated in the region to best agricultural practices. Introduction of minimum tillage systems is a good effort to control erosion, keep soil moisture and weed control. This method prevails in the better developed agricultural Cerrado zones. Policy formulation must make use this knowledge of ecosystem functioning since landscape changes has serious and long-term implication for water, carbon cycle and

possible even for climate changes (Klink and Machado, 2005). Sound policies are needed both for conservation of natural resources and to keep the productivity of cash crops.

The above discussions point to the need for research on water management on the Cerrado as a means of supporting sustainable development. This is the motivation of the work which resulted in this thesis.

1.3 Problem definition

Risk assessment aims at finding best solutions for water management in areas of conflicting interests on the water use. Therefore, there is a demand for methods which enable to describe the water table dynamics. Deterministic methods to describe water table dynamics are less appropriate for risk assessment, because they underestimate the fluctuation since they do not account for random errors. We observe water table depth in wells and boreholes, at various locations and time steps. A time series of water table depths reflects the water table dynamics given the meteorological and hydrological conditions during a typical monitoring period of a few years. The data obtained in this period depend on the prevailing climate conditions of the monitoring period. To support strategic decisions in water policy, we need to model water table dynamics over a large time period, such as 30 years. Thus, we need to extrapolate the observed time series of water table depths to a longer series of 30 years length, from which we can calculate the characteristics of water table depths. These simulated long-term series of water table depths can be applied in risk assessments, using the uncertainty estimates.

An observed time series represents the water table conditions at a certain point in space. To support water management, we need information for any location over an area. Thus, we need to predict the spatio-temporal surface of water table depth, based on the time series of the wells and boreholes. These

predictions need to incorporate extra information related to water table depth, and to account for the uncertainty.

1.4 Hypotheses, objectives and contributions

There is a need for water management technologies in the Cerrado. These technologies should contribute to increasing the probability of success when using Cerrado areas for crop production. These challenges lead to some specific scientific questions:

a) Can the change in water volume over a certain time lag be estimated, considering the spatio-temporal correlations between time steps and observation points?

b) Can water table depths be monitored in the Cerrado to assess systematic changes of water resources due to land use?

c) Can risks of extreme water levels for agriculture be predicted, and can the uncertainty be incorporated into these predictions?

To address these questions, the following hypotheses are considered:

a) The spatio-temporal correlation from water table depths can be modelled in the scope of the linear model of coregionalization

b) The seasonal variation of water table depth in the Cerrado can be modelled using time series models, assuming a linear relationship between precipitation surplus/deficit and water table depths.

The aim is of this thesis is to characterize water table depths in the Cerrados in the watershed scale. A space-time geostatistical approach was applied using space random vector fields, and time series modelling to characterize water

table depths in the Jardim River watershed. The water table depth was monitored for more than 3 years, from October 16, 2003 to March 06, 2007. This thesis will show that geostatistical tools provided efficient estimates of water content and that time series modelling allowed assessment of change in water table depths and provided information for policy making and water management in the Brazilian Cerrados. We consider and account for model's uncertainty.

1.5 Thesis layout

This thesis is structured as follows:

a) Chapter 2 presents the study area and data set.

b) Chapter 3 presents the characterization of the aquifer volume during a season.

c) Chapter 4 presents time series modelling accounting for systematic changes in water table depths over the monitored period.

d) Chapter 5 presents a simulation study case to predict behaviour of water table depths and account for risk in water management.

e) Chapter 6 presents the conclusions of this thesis, recommendations and suggestions for future work.

2 STUDY AREA AND DATA SET

2.1 Jardim River watershed

The Jardim River micro basin is a representative Cerrado area in the eastern part of the Brazilian Federal District, between latitudes 15^040'S and 160^02'S and longitudes 47^020'W and 47^040'W. This basin is part of one of the most important basins of Brazil: the São Francisco Basin. The Jardim River waters reach São Francisco River after flowing into the Preto River and Piracatu River, subsequently. The total drainage area of this basin is 53,796 ha. The study area is the high portion of the Jardim River watershed, an area of 10,121 ha which represents 18.8% of the total basin. The main cultivations in the area are grains crops (soybeans, corn, wheat and beans), coffee, cotton, fruits, and horticulture products, as well cattle, milk and poultry (Gomes-Loebmann et al., 2005).

Following Köppen's classification, the climate is Aw (Codeplan, 1984). The dry and wet seasons are well defined, with the rainfall concentrated between October and April. The annual mean precipitation is 1386 mm, for the last 33 years. The daily average temperatures vary from 18 to 30^0 C. Steep slopes are not present in the Jardim River watershed. The landforms (Embrapa, 1999) are mostly flat (slopes varying from 0 to 3%) or gently sloping (slopes between 3 and 8%), representing 53.33 and 43.05% of the basin area, respectively. The maximum elevation observed was 1176.85 m and the minimum 889.15 m, above sea level.

Agricultural crops replaced almost all natural vegetation of the Jardim River watershed and the use of irrigation systems substantially increased in this region during the past years (Gomes-Loebmann et al., 2005). The vegetation varies from gallery forests close to the river course, some spots of woody savannah (Cerrado) and open woody savannah (campo Cerrado), scrubby savannah (campo sujo) and open grassland (campo limpo).

2.2 Geology

The geology of the Federal District was recently revised by Campos & Freitas-Silva (1998). Three lithological environments compose Jardim River watershed: Canastra, Paranoá and Bambuí groups. Canastra and Paranoá groups have meso/neoproterozoic ages, and Bambuí group has neoproterozoic age. Two mount systems acted in the geology of the region. Paranã system is responsible by the positioning of Paranoá group over Bambuí group, and São Bartolomeu/Maranhão system put Canastra group over Paranoá and consequently Bambuí groups. The environment in the area consists of low-grade metamorphic rocks. Slates of varying colours, metasiltstone, and quartzite beds are present.

2.3 Hydrogeological domains

In Federal District, this geology characterized by metamorphic rocks covered with expressive pedological layers configures two main aquifer reservoirs systems: Porous and Fractured domains (Campos, 2004). The porous domain is characterized by the geological layer where the water is storage in the empty spaces of the rocky bodies (saprolite). This layer is placed over the metasedimentary rocks of the Canastra, Bambuí and Paranoá groups, which corresponds to the fractured aquifers in the area. In the fractured domain, the water takes place in the physical discontinuities in the rocks.

2.3.1 Fractured domain

The Fractured domain in the Jardim River watershed comprises four systems: Paranoá (ubsystems R3/Q3 and R4), Canastra (subsystem F) and Bambuí (Figure 2.1).

a) Paranoá subsystem R3/Q2

36

System Paranoá is situated in the northern part of the basin, and can be divided into two subsystems. Subsystem R3/Q2 consists of sandy metarhythmites and medium quartizites. These materials favour to keep open cracks in the system. Thus, the high number of water entrances causes high water fluxes in the section, resulting in a high local hydrogeological importance.

b) Paranoá subsystem R4

Subsystem R4 consists of pelitic rocks and clay. These materials provide low the water flux. The local hydrogeological importance is moderate.

c) Canastra subsystem F

The Canastra system is represented by subsystem F. It covers an extensive area in the western part of the Jardim River basin. This system is lithologically associated to phyllits with high porosity and presents bedrock fractures and foliation. It facilitates water percolation, resulting in areas with big aquifer recharge potential. However, the local hydrogeological importance is moderate.

d) Bambuí system

This system is characterized metasiltites and clay metasiltites, and is present in the eastern part of the basin. The aquifer has moderate local hydrogeological importance.
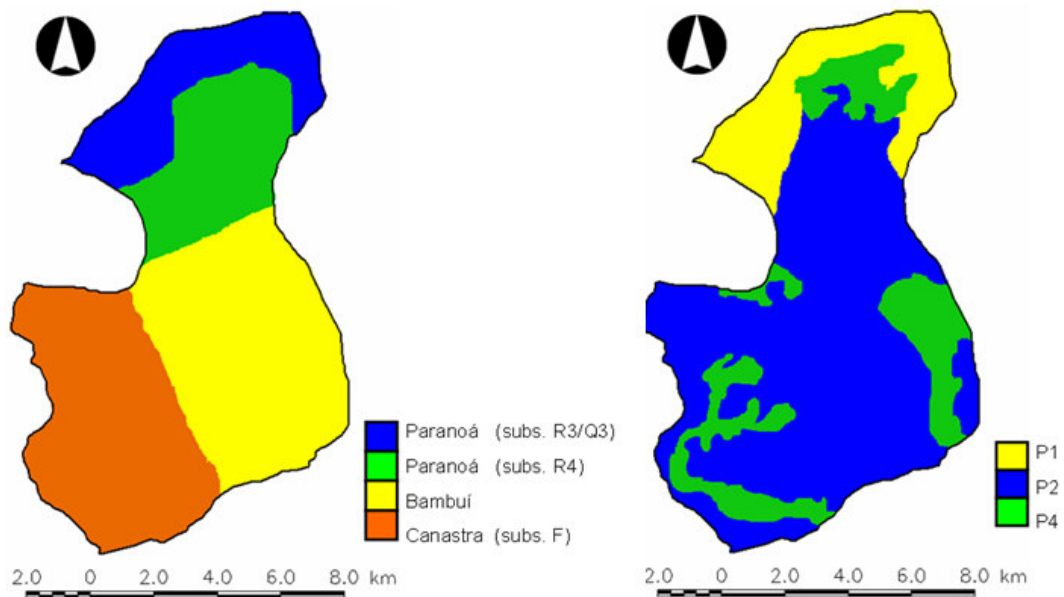
Figure 2.1 - Fractured (left) and Porous (right) hydrological domains on Jardim River watershed (Lousada, 2005).

2.3.2 Porous domain

The Porous domain comprises three systems: P1, P2 and P4 (Figure 2.1).

a) System P1

Located in the extreme north part of the basin, this system is associated with sandy metarhythmites and medium quartizites from Paranoá group. The resulting soils are Red Yellow Latosols and sandy soils in small portions. The aquifer has high local hydrogeological importance. The hydrological importance is measured at aquifer recharge and water exploration potentials.

b) System P2

This system comprises the major part of the basin, and is associated with pelitic rocks. This material is highly susceptible to chemical weathering and it favours

to decompose the underlying rock, originating a thick pedologic covering. The resulting soils are Red Latosols, sandy to loamy, with moderate porosity. The aquifer has moderate local hydrogeological importance.

c) System P4

This system is associated to pelitic rocks (schists and phyllits) which produced shallow and rock soils (mostly Cambisols) close to drainages and waterheads. The local hydrogeological importance is very low.

2.4 Soils

Reatto et al. (2000) presented a large-scale soil survey of Jardim River watershed in the scale of 1:50,000. The soil classes identified in the high portion of the basin are Red Latosol, Red Yellow Latosol, Haplic Cambisol, Haplic Plinthosol, Haplic Gleysol and Quartzarenic Neosol. These soil classes are presented in Figure 2.2.
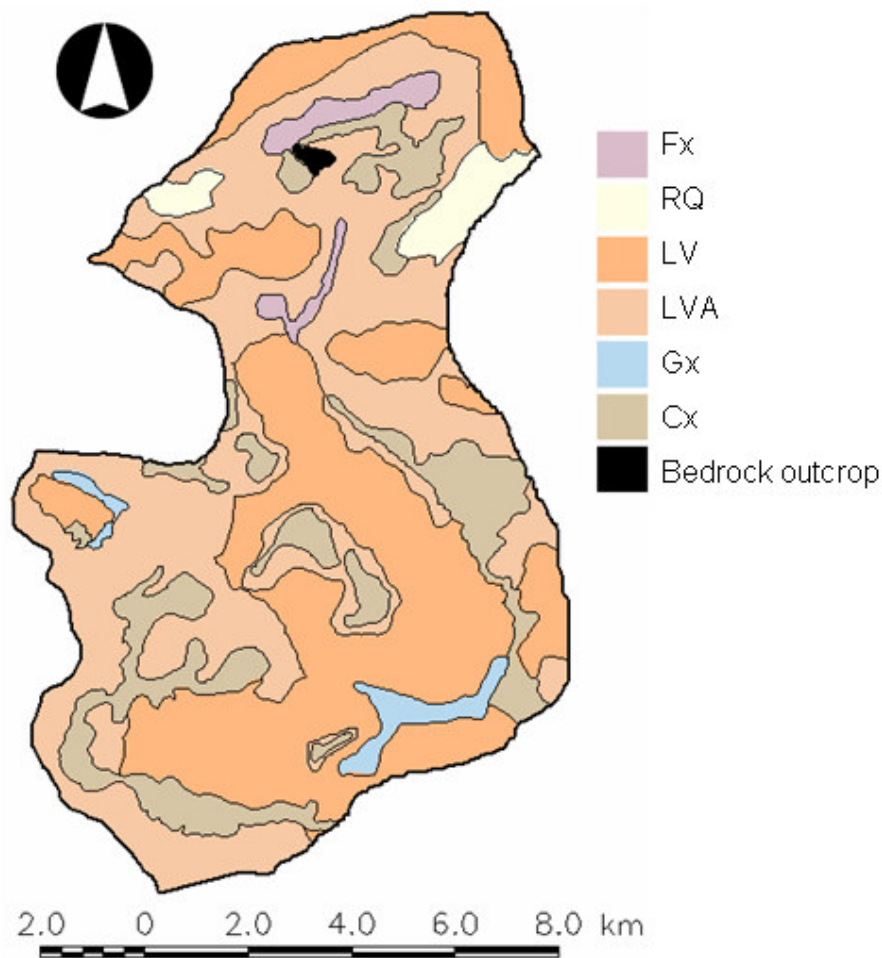
Figure 2.2 - Soil types over the Jardim River watershed (Reatto et al., 2000).

Red Latosols (LV) are mineral soils, acid (pH around 5), deep (>1,5 m), with thick B horizon (>50 cm), high water permeability and well drained. Clay content varies from 67 to 75%, base cation concentration between 4 and 49% and aluminium concentration between 2 and 58%.

Red Yellow Latosols (LVA) have almost the same characteristics as LV, but are less clayey. The clay content is between 38 and 71%. The organic matter in Latosols is low.

Haplic Cambisols (Cx) are shallow soils, with incipient B horizon, low weathering degree, moderately drained, and texture and chemical characteristics variable. The clay content is between 46 and 63%. Mostly, the cambisols of Jardim River watershed are poor soils, degradated, with low water storage capacity and so low agricultural potential.

Haplic Plinthosols (Fx) in the area are soils with slow water movement, flooding susceptible, and imperfect drainage. Its plintic horizon, with red spots from iron reduction, can be a limit for water permeability and root development occasionally. Clay content is around 25%, base cations concentration around 6% and aluminium concentration around 32%.

Haplic Gleysols (Gx) are hydromorphic soils, close to bad drained or flood susceptible areas. These soils are not well developed. The water table is close to the surface almost all year. It presents a gray layer from redox reactions in the soil and low organic matter content. The clay content is between 48 and 53%, concentration of base cations between 3.47 and 11.37% and aluminium concentration between 81.68 and 95.33%.

Quartzarenic Neosols (RQ) are deep soils (more than 2 m) with sandy texture (maximum 15% of clay), originated by quartzo minerals, absente of B horizon and well drained. The base cations concentration is around 4% and aluminium concentration around 79%.

2.5 Data set

2.5.1 Water table depths

To monitor water table depths (WTD), 40 wells were originally drilled in the area (Figure 2.3). The locations were selected purposively, to cover the range of soil types in the area and in a try to characterize the different responses of water table depths (Lousada, 2005). Some wells closed and others presented dryness

after the installation. It can be result of a hydraulic barrier between porous and fractured domains in areas where there is clay vertical migration causing narrowing of the water flux in the fractures and closing the connection of the systems (Lousada and Campos, 2005). These wells were excluded from further analyses.
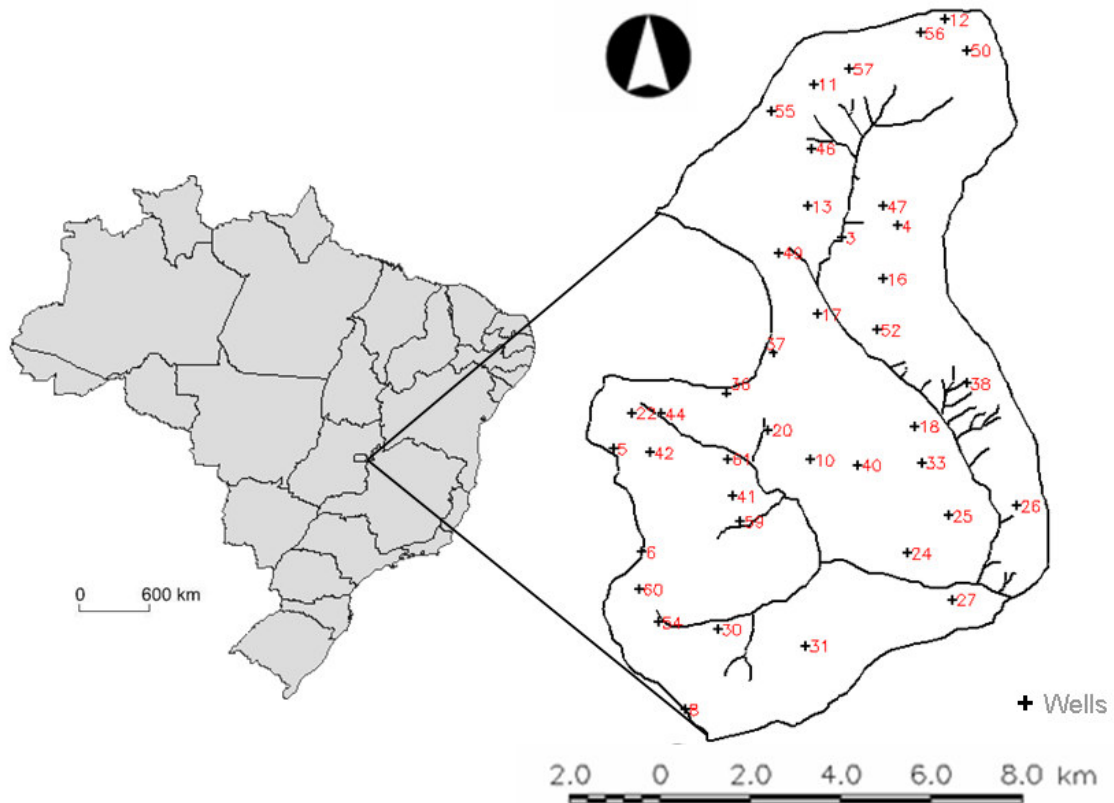
Figure 2.3 - Observation wells distributed over the Jardim River watershed.

The water table depths were observed semi-monthly from October 11, 2003 until March 06, 2007, resulting in series of 55 more or less regularly spaced semi-monthly observations, during a monitoring period of 1240 days. The wells filters levels were variable with the soil depth.

2.5.2 Meteorological time series

Series of 33 years length of precipitation and potential evapotranspiration were available from the EMBRAPA Cerrados climate station. These data were available at www.agritempo.gov.br, from 1974 until 1996 in a monthly frequency and from 01/11/1996 until the present day in a daily frequency and from This climate station is located at 10 km distance from the basin. For the period from 1974 until 1996 a monthly data are available, and for the period from 1996 until March 06, 2007 daily data is available. Climate reports for agriculture about the region were frequently consult at www.cpac.embrapa.br/tempoeagri/ tempoeagri.html. Every week new bulletins report the climatological water balance and soils conditions for agricultural practices.

2.5.3 Digital maps

The digital soil map from Reatto et al. (2000) is the pedological base for this study (Figure 2.2). The hydrogeological features (Figure 2.1) the digital elevation model (DEM) with resolution of 15m (Figure 2.4) are available from Lousada (2005).

Ancillary information related to the actual land sources was derived from Landsat 5 images. Three images from July 23, 2005, July 26, 2006 and January 18, 2007, orbit/point 221/71, were inspected and used to classify the actual land use in the region. Using maximum likelihood classification and expert knowledge, the supervised image classification results in a land use surface, divided in three classes: Agricultural Crops, Pasture and Cerrado Vegetation. The class Agricultural Crops includes all kinds of agricultural products that are cultivated in the area: small areas cultivated with horticultural products, such as carrots, lettuce and tomatoes, and big areas cultivated with products such as corn, soybeans, cotton, coffee and sugarcane. All these crops demand more water than the original vegetation. The agricultural practices are intensive, resulting in three production cycles during one year when irrigation is applied. Also, the land use in the class Agricultural Crops is very dynamic because of

agronomical recommendations, rotation schemes or simply prices. The class Pasture is considered to be less water demanding than Agricultural Crops, but more demanding than the natural Cerrado vegetation. These areas are not as dynamic in land use changes as the Agricultural Crops. Figure 5 presents the classified land use map. This surface is considered as the land use for the region during the monitoring period, ignoring possible changes in the established classes of land use.
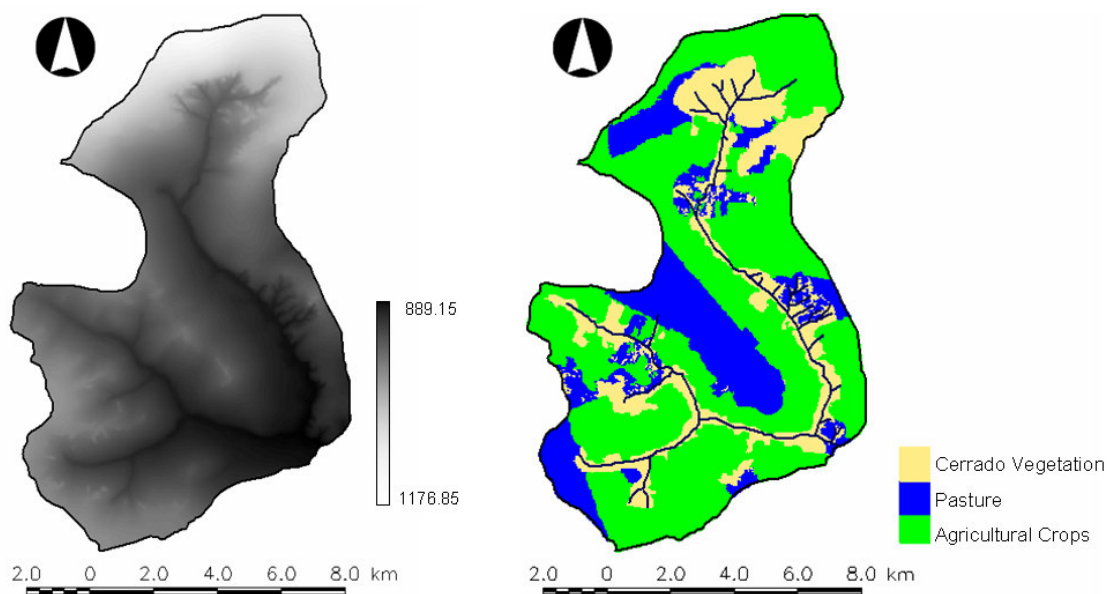
Figure 2.4 - Digital elevation model (left) and image classification for actual land use (right) on Jardim River watershed.

2.6 Monitoring scheme

This section presents the monitoring scheme followed in this research. We refer to scheme as the entire plan, all the decisions and all information relevant to data acquisition, data recording and data processing. De Gruijter et al. (2006) presented issues to develop a complete sampling and monitoring scheme. The authors present seven principals considered essential for a good design scheme for survey and monitoring. These principles are i) develop a complete

scheme; ii) structure the design process; iii) pay attention to practical issues; iv) employ prior information on variation; v) balance the various sources of error; vi) anticipate changing conditions during monitoring; and vii) calculate the sample size correctly.

2.6.1   Detailed analysis and specification of the objective

a) Target universe: Jardim River watershed.

b) Domains of interest: high portion of the watershed (local scale).

c) Target variable: water table depths.

d) Target parameters: water table depths at locations and given time steps, temporal trends, probabilities of exceedance of critical levels and critical days.

e) Target quantity: water volume ($m^{3.}ha^{-1}$) and water table depths (meters bellow surface) that are exceeded at specific time steps at predefined probability levels.

f) Type of result: Quantitative (spatial prediction of water table depths and estimation of water table statistical characteristics).

2.6.2   Quality measure

Error variance in spatial prediction and half-width of a 95% confidence interval in estimation of statistical characteristics.

2.6.3   Constrains

Limited budget to drill wells.

2.6.4   Prior information

a) Sampling frame: soil map scale 1: 50.000 (paper and digital).

b) Miscellaneous information: Digital Elevation Model (DEM) with 15 m resolution, geological maps and a climate station (precipitation, evapotranspiration, temperature).

c) Model of variation of the target variable: linear model of coregionalization and regionalized time series models for water table depths, developed from available sample data.

2.6.5 Sample support

Constant and continuous (predefined wells to measure the spatial variation and semi-monthly for time sampling water table depths).

2.6.6 Assessment method

Field measures directly in the wells, from the surface using in situ groundwater probe.

2.6.7 Composite sampling

No composite sampling, measures taken in situ.

2.6.8 Design-based or Model-based inference

Model based. The choice between Design-Based and Model-Based inference lies in the fact we need the values distribution over the entire universe, or in other words, mapping.

2.6.9 Sampling pattern type

Forty fixed wells at selected locations and covering specific strata (soil classes) derived from a soil map (scale 1: 50.000), where measures are taken with semi-monthly frequency.

2.6.10 Identification of the actually selected sample

Reference to a map with the wells locations and a table of sampling times.

2.6.11 Protocols on data record and fieldwork

Data record directly on a spreadsheet in paper or electronic in a palm top.

2.6.12 Method to be used for statistical inference

Linear model of coregionalization, time series modelling, regionalization of the calibrated models and calculation of the percentiles prediction intervals for water table depths kriging predictions and associated kriging variances.

2.6.13 Prediction of operational costs and quality of results

Ex-ante evaluation.

3 UNCERTAINTY MODELING IN SPATIO-TEMPORAL ANALYSIS OF WATER TABLE DEPTHS IN A WATERSHED[*]

3.1 Introduction

Changes in water resources availability have direct impacts on environment and agricultural land use. The aquifer exploitable volume is a decision-making variable to be measured as part of local water management planning. Many times water resources present influences not just inside a watershed, but also regional, municipal and even state importance (Hoffmann and Jackson, 2000). The fragility in water resources exploration lies in groundwater recharge, which sometimes does not present the same speed of extraction. It can result in exploration of aquifer permanent reserves, with risks even of exhaustion when the caption is not properly monitored. Water loss and aquifer recharge present variability in space due to the absorption and water retention differences in the soil (given by soil classes, texture, infiltration rates, for example). Also, these variables present variability in time, from climatic and seasonal effects (precipitation, evapotranspiration) and from the water use.

Groundwater has significant importance in the Brazilian Cerrado agriculture, since the region is characterized by wet and dry periods well defined over the year. Aquifer regulatory reserves are responsible for supplying water for irrigation in the dry period. The reserves behaviour depends on seasonal effects of rainfall. How much water is lost by the system at the dry season and how much it recovers in the rainy period are fundamental subjects for water management. We are considering a water volume decrease in the dry season

and recharge during the wet season (Oliveira et al., 2005). These reserves are important not only for irrigation but also for natural vegetation maintenance and for the regional aquifer system (Klink and Machado, 2005). The water volume recharge quality during the wet period is an important aspect and should be analyzed. The risk can be accessed by modelling the uncertainty associated with water estimations, allowing analysts to verify possible extreme situations (Goovaerts, 2002).

The aim of this study is to analyse how water table depths vary in space and time, and modelling their dynamics over a certain period in time. The target variable is the increases or decrease of water levels (in meters) from one month to another at each observed well. The work used space-time geostatistical procedures (Goovaerts and Sonnet, 1993; Papritz and Fluhler, 1994; Kyriakidis and Journel, 1999) which consider the coregionalization structure between monthly water table depths differences during the dry season of 2004 (from May until September) in the Jardim River watershed. In the Jardim River watershed, the water resources have been used intensively by irrigation systems and without an equal division of the available water. Irrigation systems are installed even in places with low superficial water availability or close to river springs, what increase risks of aquifer recharge (Dolabella, 1996). These techniques allow to estimate the explored aquifer volume during the period. Also, they allow to account for uncertainty to enable risk modelling (Goovaerts, 1997) and to create scenarios for water management and reserves evaluation.

3.2 Materials and methods

One of the biggest challenges in groundwater modelling is water table characterization in space and time. Geostatistics provide three conceptual points of views to solve problems with space-time (ST) indexation (Kyriakidis and Journel, 1999): i) models of simple random functions integrating the space and time components (Christakos and Raghu, 1996); ii) vectors of space

random functions (Goovaerts and Sonnet, 1993); iii) vectors of time series (Rouhani and Wackernagel, 1990).

Multivariate geostatistics present ST procedures which priory the domain in which information is more abundant (Goovaerts and Sonnet, 1993). In our case study, the field of largely available information is the spatial. The adopted hypotheses were: i) at a time instant, the water table depths behave as a regionalized variable realization of a random field, with a constant mean and a variance described by a stochastic component. Also, it is supposed space correlation among data. ii) at subsequent time instants, the realizations of the random field are correlated and modelled as a finite collection of space random functions temporally correlated.

The model starts with observing $\{z(\mathbf{u}_\alpha, t), \alpha = 1, ..., n; t = 1, ..., T\}$ at n points $\mathbf{u}_\alpha$ in space and T points t in time. Each observed value is understood as a realization of a random variable (RV). Each set of realizations of RV's in certain time step is a particular realization of a ST random function (RF) $\{Z(\mathbf{u}, t), \text{ where } (\mathbf{u}, t) \in D \times T\}$. D refers to spatial domain and T to temporal domain where this function takes values. Based on the previous hypothesis, this function was modelled as a finite realization for each time step of T spatial random functions, temporally correlated (Goovaerts and Sonnet, 1993; Papritz and Fluhler, 1994). In other words, for each t there is a RF representing the target variable and these functions are correlated amongst themselves. The following representation is adopted for the ST processes $Z(\mathbf{u}, t)$:

$$Z(\mathbf{u}, t) = [Z_t(\mathbf{u}), t = 1, ..., T] = [Z_1(\mathbf{u}), ..., Z_T(\mathbf{u})]'$$

(3.1)

Considering a set of T RF, each one represented by a time step t, where $t = 1, ..., T$. Supposing intrinsic stationarity, the process is first order stationary and the variogram can be calculated for each t RF as:

51

$$E[Z_t(\mathbf{u}) - Z_t(\mathbf{u} + \mathbf{h})] = 0$$

and (3.2)

$$\gamma_{t_i t_j}(\mathbf{h}) = \frac{1}{2} E\left[\{Z_{t_i}(\mathbf{u}) - Z_{t_i}(\mathbf{u} + \mathbf{h})\}\{Z_{t_j}(\mathbf{u}) - Z_{t_j}(\mathbf{u} + \mathbf{h})\}\right] \quad t_i, t_j = 1, \ldots, T$$

It means that correlation between any data, in two different time steps and separated by the same spatial lag $\mathbf{h}$ is written as a variogram matrix $\Gamma(\mathbf{h})$:

$$\Gamma(\mathbf{h}) = \begin{bmatrix} \gamma_{t_1 t_1}(\mathbf{h}) & \cdots & \gamma_{t_1 t_j}(\mathbf{h}) & \cdots & \gamma_{t_1 t_T}(\mathbf{h}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma_{t_i t_1}(\mathbf{h}) & \cdots & \gamma_{t_i t_i}(\mathbf{h}) & \cdots & \gamma_{t_i t_T}(\mathbf{h}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \gamma_{t_T t_1}(\mathbf{h}) & \cdots & \gamma_{t_T t_i}(\mathbf{h}) & \cdots & \gamma_{t_T t_T}(\mathbf{h}) \end{bmatrix}$$ (3.3)

The values of variogram matrix are estimated by:

$$\hat{\gamma}_{t_i t_j}(\mathbf{h}) = \frac{1}{n} \sum_{\alpha=1}^{n} \left[z_{t_i}(\mathbf{u}_\alpha) - z_{t_i}(\mathbf{u}_\alpha + \mathbf{h})\right] \left[z_{t_j}(\mathbf{u}_\alpha) - z_{t_j}(\mathbf{u}_\alpha + \mathbf{h})\right]$$ (3.4)

The variogram matrix $\Gamma(\mathbf{h})$ is the covariance matrix from temporal increments. It describes the correlation structure between subsequent time steps $(t_i, t_j) = 1, \ldots, T$ (Goovaerts, 1992). Supposing no correlation between RV at long space distances (the range of spatial continuity is lower than the biggest distances of the domain), the variogram matrix converges to a unique covariance matrix $\mathbf{V}$:

$$\Gamma(\mathbf{h}) \rightarrow \mathbf{V} \quad to \quad |\mathbf{h}| \rightarrow \infty$$ (3.5)

Here, spatial variograms from the difference between the present water table depth and the water table depth at a past time instant were calculated from observed water table depths. The time lag considered calculate the water table depth to increase/decrease between time steps was one month, covering five time instants in a period from May to September, 2004.

3.2.1 Coregionalization analysis

Analyzing the spatial distribution at different scales is one aspect that contributes to a good understanding of the phenomenon. In this work, the term scale is used in an intuitive way, without any direct relation with cartographic scale. Microscale refers to random errors not captured by the sample grid (nugget effect). Mesoscale refers to regional variations, with short range; and macroscale refers to global variation, in long range. The spatial phenomenon can occur in different ways at different scales. These different scales refer only to spatial scale (geographic), and can be modelled by the variance/covariance matrix $\Gamma(\mathbf{h})$ which describes the correlation structure from each distance (lag) $\mathbf{h}$ modelled.

In the scope of the linear coregionalization model (LCM), described at Goovaerts (1997), it is supposed that the spatial distribution of the variable is a result of different process interactions, independently acting in different spatial scales. The variogram matrix can model these scales using nested models with S variogram functions. All $T(T+1)/2$ direct and cross variograms are modelled by a linear combination of standard variograms for the same spatial range. When $\Gamma(\mathbf{h})$ has spatial dependence in different spatial scales, the matrix $\Gamma(\mathbf{h})$ can be decomposed in basic variograms functions $g^s(\mathbf{h})$, where $s = 1,...,S$ represent the number of spatial scales modelled (Castrignanò et al., 2000):

$$\Gamma(\mathbf{h}) = \sum_{s=1}^{S} B^s g^s(\mathbf{h}) \qquad (3.6)$$

where, $B^s = \left[b_{t_i t_j}^s \right]$ is the positive semi-defined coregionalization matrix $T \times T$. This matrix presents more precise descriptions of the correlation structure $\Gamma(\mathbf{h})$ at different spatial scales. The coefficients $b_{t_i t_j}^s$ express the relative importance of the basic variogram function $g^s(\mathbf{h})$ in $\Gamma(\mathbf{h})$. Under second order stationarity hypothesis, $g^s(\mathbf{h}) \rightarrow 1$ when $|\mathbf{h}| \rightarrow \infty$, and the variance/covariance matrix $\mathbf{V}$ equals to the sum of the coregionalization matrix:

$$\mathbf{V} = \sum_{s=1}^{S} B^s \qquad (3.7)$$

When correlations change in function of the spatial scale, coregionalization matrixes B^s describe the correlations better than variograms matrixes $\Gamma(\mathbf{h})$ because measure just the spatial contribution of each variable to the variogram function (Goovaerts, 1992).

The use of a LCM lies in considering that each RF $Z_t(\mathbf{u})$ can be decomposed in a set of uncorrelated RF $\left\{ Y_v^s(\mathbf{u}), v = 1,...,T; s = 1,....,S \right\}$ with transformation coefficients $a_{t_v}^s$:

$$Z_t(\mathbf{u}) = \sum_{s=1}^{S} \sum_{v=1}^{T} a_{t_v}^s Y_v^s(\mathbf{u}) \qquad (3.8)$$

where $Y_v^s(\mathbf{u})$ are regionalized factors from v components at the spatial scale s. To an established s, the T RF $Y_v^s(\mathbf{h})$ has the same variogram function $g^s(\mathbf{h})$. Clustering the regionalized factors with the same variogram functions $g^s(\mathbf{h})$, each RF $Z_t(\mathbf{u})$ can be written as the sum of orthogonal RF $Z_t^s(\mathbf{u})$:

$$Z_t(\mathbf{u}) = \sum_{s=1}^{S} Z_t^s(\mathbf{u}) \tag{3.9}$$

Each spatial component $Z_t^s(\mathbf{u})$ represents the behaviour of the RF $Z_t(\mathbf{u})$, which means the behaviour of Z at time t in scale s. Decomposing each $Z_t(\mathbf{u})$ in different components $Z_t^s(\mathbf{u})$, it is possible to study the temporal behaviour of Z at different scales, and for a known scale s, the correlation between different time steps is measured by the structural correlation coefficient:

$$r_{t_i t_j}^s = \frac{b_{t_i t_j}^s}{\sqrt{b_{t_i t_i}^s \, b_{t_j t_j}^s}} \tag{3.10}$$

These coefficients measure correlations among variables for each spatial model basic structured, showing phenomenon's behaviours in scales previous defined. Many times these behaviours are not detected by methods which do not consider the phenomenon as variant in different scales (Goovaerts, 1992).

In this work, we applied principal component analyses (PCA) to decompose the matrixes B^s in matrixes A^s (Castrignanò et al., 2000). These procedures explore ST relations of the study variable for the time steps analyzed. The matrixes A^s are set by the coefficients $a_{t_v}^s$, described by equation (3.8).

Linear combinations of the original functions generate orthogonal functions with coefficients as the eigenvectors of matrix **V**. Each principal component explains a percentage of the total variance, which corresponds to the ratio between principal component eigenvalue and the total variance. The presence of a LCM lies in the hypothesis that the process can be explained by independent factors $Y_v^s(\mathbf{u})$. Eigenvalues calculated from the coregionalization matrix correspond to the variable weights at the principal components. The correlation

graphic shows in the Y axis the first principal component eigenvalues and in the X axis the second principal component eigenvalues, for each variable. These variables projections in factorial space are defined for the two firsts principal components in the graph, showing the behaviour of the temporal structure of the variables in study.

3.2.2 Estimation and uncertainty measure

The value $Z_t(\mathbf{u_o})$ is estimated to calculate the aquifer natural discharge volume (NDV) and assess the risk of water loss involved by water use during the dry season. This volume evaluates the aquifer regulatory water reserves, which represents the amount of free water storage in the unsaturated zone with natural recharge under effect of seasonal precipitation (Costa, 2000). Two scenarios are presented to evaluate risks in water management: the first with a conservative water use and the second with an optimistic strategy (Goovaerts, 1997).

These RV values $Z_t(\mathbf{u_o})$ are estimated by cokriging. The aim here is to infer values of these variables at unvisited locations $\mathbf{u}_o$, using the available information at T time steps of the considered phenomenon. At a given time step $t', n \times T$, the RF $Z(\mathbf{u}, t)$ is estimated as:

$$\hat{Z}_t(\mathbf{u_o}) = \sum_{t=1}^{T}\sum_{\alpha=1}^{n} \lambda_{\alpha t} z_t(\mathbf{u}_{\alpha}) \qquad (3.11)$$

It is a ST weighted mean, where different weights $\lambda_{\alpha t}$ are attributed in function of the ST dependence estimated in the variograms. NDV maps $V_t(\mathbf{u}_{\alpha})$ and standard deviations for May, June, July, August and September were calculated from $\hat{z}_t(\mathbf{u}_{\alpha}), \alpha = 1,...,N$ estimations. In this case, the estimator is adapted from Costa (2000):

56

$$\hat{V}_t(\mathbf{u}_\alpha) = A.p_{ef}.\hat{Z}_t(\mathbf{u}_\alpha), \alpha = 1,..,N \tag{3.12}$$

$$\hat{\sigma}_t(\mathbf{u}_\alpha) = A.p_{ef}.\hat{\sigma}_{Z_t}(\mathbf{u}_\alpha), \alpha = 1,..,N \tag{3.13}$$

where A is the pixel area of the inferential regular grid and p_{ef} is the effective porosity for each considered soil class, described at Table 3.1 (Lousada, 2005).

Table 3.1 - Estimated effective porosity for the soil types presented over the Jardim River watershed.

Soil Type	Minimum	Maximum
Red Latosol (LV)	7%	9%
Red Yellow Latosol (LVA)	7%	12%
Haplic Cambisol (Cx)	3%	7%
Quartzarenic Neosol (RQ)	12%	18%
Haplic Gleysol (Gx)	7%	10%
Haplic Plinthosol (Px)	4%	8%

Source: adapted from Lousada (2005)

For each grid cell we calculate a probability density function (PDF). Percentiles of the PDF are selected assuming a normal distribution of water table depths over the area. We generated scenarios characterizing risks associated with these estimations from the percentile values. Percentiles below the mean are estimations with high probability to incur in errors of overestimation. Percentiles above the mean are estimations with high probability to incur in errors of subestimation (Goovaerts, 1997). The critical limits established were the percentiles 0.10 and 0.90:

$$\text{Prob}\left[\hat{V}_t(\mathbf{u}_\alpha) > v_{q(0.1)}\right] = 0,90 \tag{3.14}$$

$$\text{Prob}\left[\hat{V}_t(\mathbf{u}_\alpha) > v_{q(0.9)}\right] = 0{,}10 \qquad (3.15)$$

These limits were considered to be convenient for the main sources of uncertainty associated to this modelling, and reasonable for water management. The main sources of uncertainty considered were the model of spatio-temporal variation of WTD, erros from observations, soil properties and soil maps, and interpolation erros. We generated quantitative measures of aquifer natural discharge volume from water gain/loss scenarios of the dry season.

3.3 Results and discussion

The descriptive statistitics of the water table depths increase/decrease during the dry season are presented in Table 3.2.

Table 3.2 - Descriptive statistics of the Jardim River watershed water table depths increase/decrease data from May to September, 2004, in meters.

Variable	Mean	SD	Variance	Skewness	Kurtosis	CV	P-value	Frequency Distribution
May	-0,29	0,70	0,48	-0,77	4,46	-2,39	0,032	NN
June	-0,54	0,67	0,45	-1,43	5,26	-1,24	0,000	NN
July	-0,37	0,50	0,25	-2,10	7,49	-1,35	0,000	NN
Aug.	-0,55	0,64	0,41	-1,98	7,00	-1,18	0,000	NN
Sept.	-0,47	0,44	0,19	-1,10	3,33	-0,93	0,000	NN

Unites: May, June, July, August and September=meters, SD=Standard deviation CV=coefficient of variation, P-value=Anderson-Darling test for normality at 5% probability; N-normal distribution, NN- not normal distribution.

The frequency distribution of the water table depths increase/decrease has high negative skew and kurtosis coefficients. It is a result of extreme values found

from the different responses of water table depths over the area which influence the frequency distributions. At some places, water table depths had centimetric variations. At other sites, the water table depths varied some meters during the monitored period. From a visual inspection of the frequency histograms, the frequency distribution presented deviates when compared with a normal distribution. The Anderson-Darling test (Parkin and Robinson, 1992) with 0.05% probability accounts for normality. The results presented non-normal distributions for the months under investigation. The distributions started to present shapes close to a normal distribution after removing some outliers (possibly originated from error measurements). This procedure reduced drastically skewness and kurtosis.

The direct and cross covariance structure (equation 3.4) were estimated by the LCM (Goulard and Voltz, 1992) presented at equation 3.6. Two structures were modelled: first, a nugget effect (B^0) or random component of microscale errors, and second a Gaussian model which explains the spatial dependence in regional scale with a 2,130.95 m common range:

$$B^o = \begin{bmatrix} 0.0562 & & & & \\ 0.0065 & 0.049 & & & \\ -0.0138 & -0.0142 & 0.0076 & & \\ -0.0341 & -0.014 & 0.0001 & 0.1482 & \\ -0.0363 & -0.0285 & 0.0064 & 0.1281 & 0.1169 \end{bmatrix}$$

$$B^1 = \begin{bmatrix} 0.3196 & & & & \\ 0.3911 & 0.4985 & & & \\ 0.1618 & 0.2289 & 0.2606 & & \\ 0.17251 & 0.2312 & 0.2007 & 0.3354 & \\ 0.1643 & 0.1974 & 0.1517 & 0.1560 & 0.1308 \end{bmatrix}$$

where B^s, s=0,1 are the positive semi-defined coregionalization matrixes for the nugget effect and Gaussian structure of the model $\gamma_{ij}(\mathbf{h})=B^0 g^0(\mathbf{h})+ B^1 g^1(\mathbf{h})$. The coregionalization matrixes give the contribution of each month under study as variances and covariances for the modelled coregionalization structured. The 15 direct and cross adjusted variograms (Figure 3.1) present the modelling. The solid lines represent the theoretical model adjusted to the sample variance. Direct variograms contributions for each month are in the diagonal of the matrix, and out of the diagonal are the contributions of the cross variograms (covariances between months). The highest contributions for the model were given by May, June and August, which had the highest values for the Gaussian structure (mesoscale). August and September had the highest nugget effects, being the months which more influence the model random component. The temporal correlation is highest at small temporal lags. The correlation is lost in time because of the changes in water table patterns during the dry season. The structural correlation coefficients are in Table 3.3.

Table 3.3 - Structural correlation coefficient from the dry season at microscale (nugget effect) and mesoscale (Gaussian model with 2,130.95m range).

Microscale	May	June	July	August	September
May	1.00	-	-	-	-
June	0.13	1.00	-	-	-
July	-0.66	-0.75	1.00	-	-
August	-0,37	-0.16	0.03	1.00	-
September	-0,45	-0,41	0,22	0.98	1.00
Mesoscale	May	June	July	August	September
May	1.00	-	-	-	-
June	0.82	1.00	-	-	-
July	0,56	0.64	1.00	-	-
August	0.53	0.57	0.70	1.00	-
September	0.80	0.76	0.79	0.8	1.00
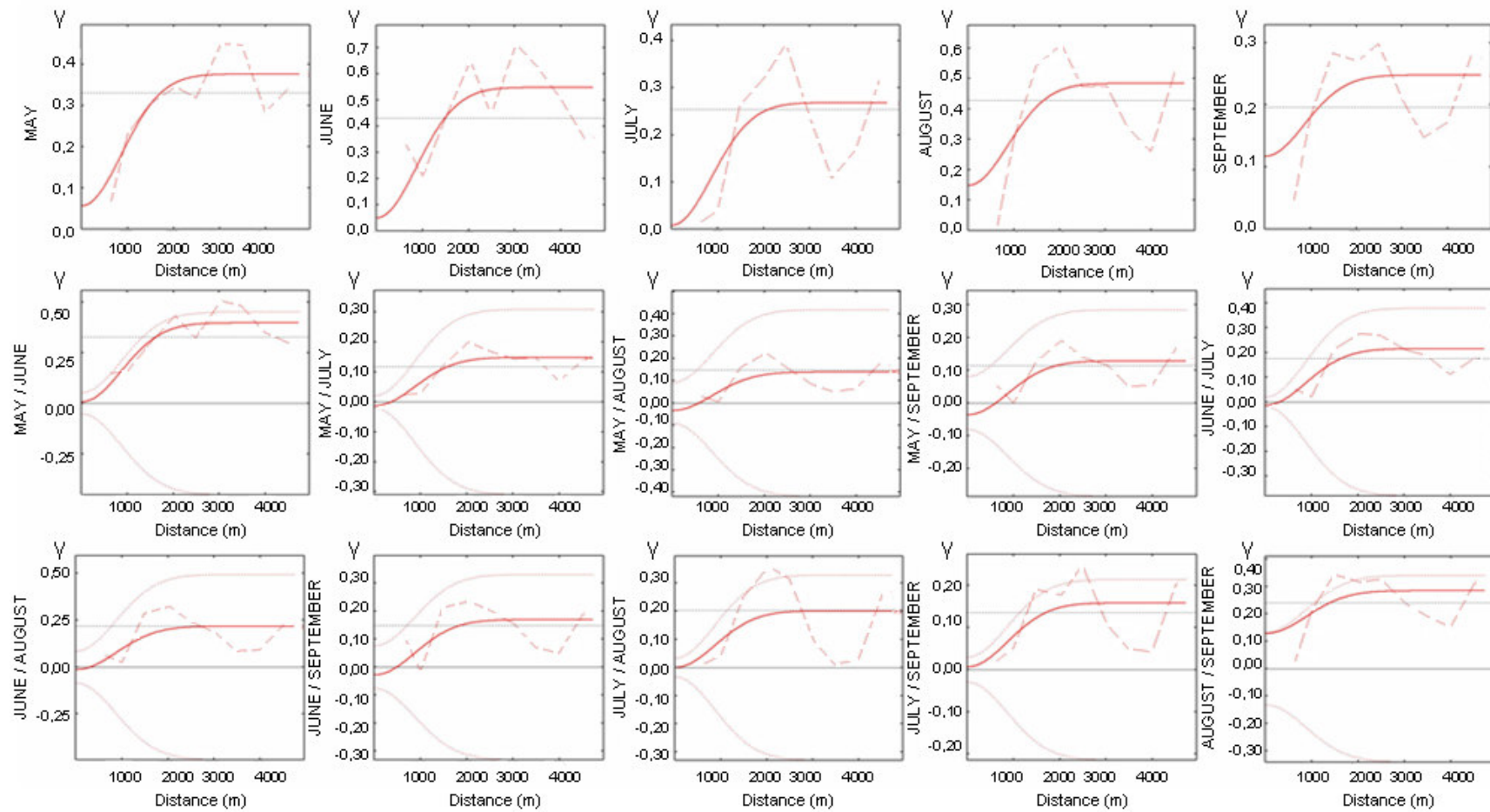
Figure 3.1 - Direct and cross experimental (dash line) and theoretical (solid line) variograms adjusted for the dry season of 2004 in the Jardim River watershed.

61

The scenario for September (end of dry season) is very different from May (beginning of the season), not only in water table depths but also in water availability. It can be checked at Figure 3.2, which presents the increase/decrease of water table depths estimated by cokriging. Some areas present systematic decreases in water table depths, verified over the months.
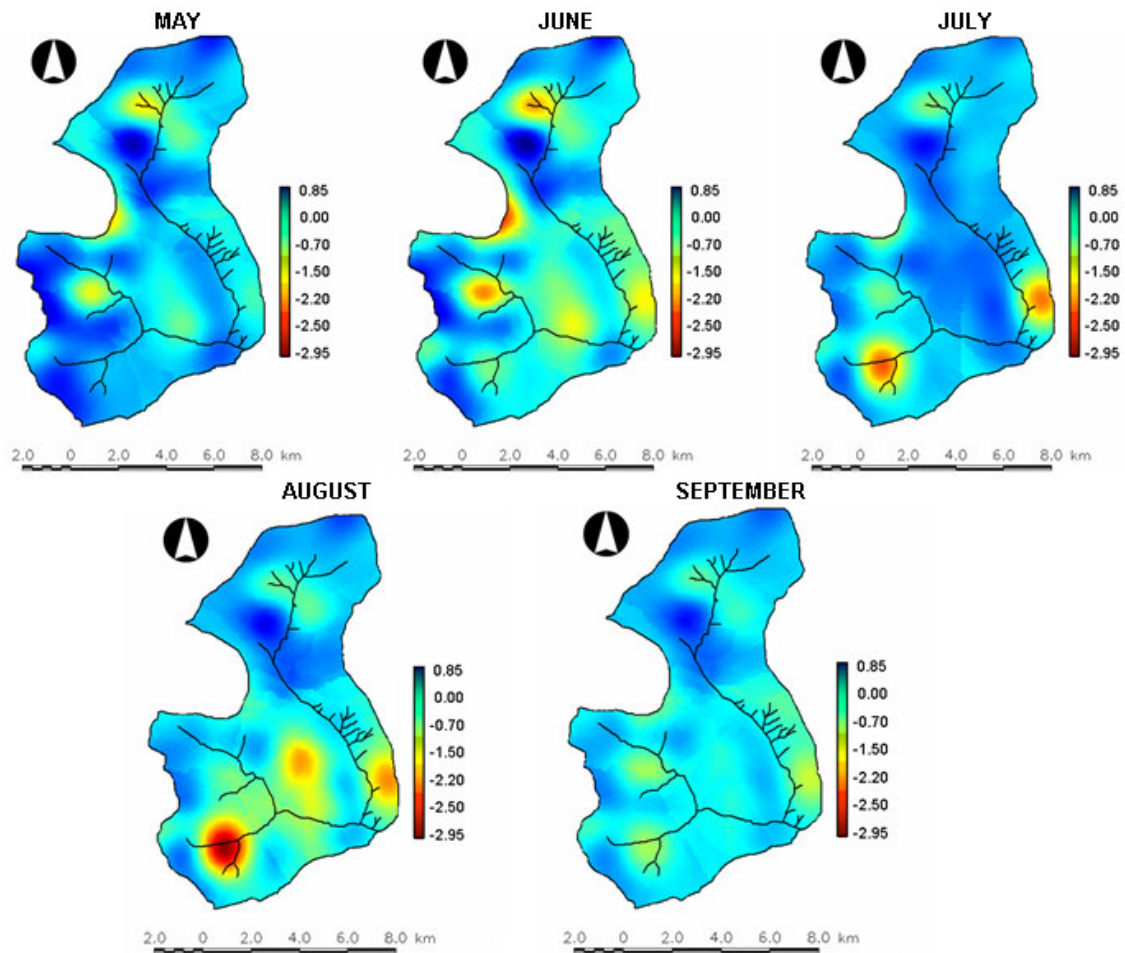
Figure 3.2 - Increase/decrease of water table depths (meters) at Jardim River watershed estimated by cokriging for May, June, July, August and September, 2004.

Principal components analyzes applied on B^s, $s=0,1$, converged the matrix V to $\sum_{s=0}^{S} B^s$. The first two principal components resume 87.96% of nugget effect variation at microscale. The first two principal components resume 92.33% of

the Gaussian model variation at mesoscale. Analyzing both scales together, the nugget effect corresponds to 19.66% and the Gaussian model to 80.34% of the LCM variation.

The correlation graph (Figure 3.3) shows the temporal relation among the variables. For nugget effect, August and September form a cluster because have high contribution in the model, as well the May, June and July that had lower contribution to this part of the model. At regional scale May and June form a cluster because of its high contribution to the model, as well July, August and September (low model contribution).
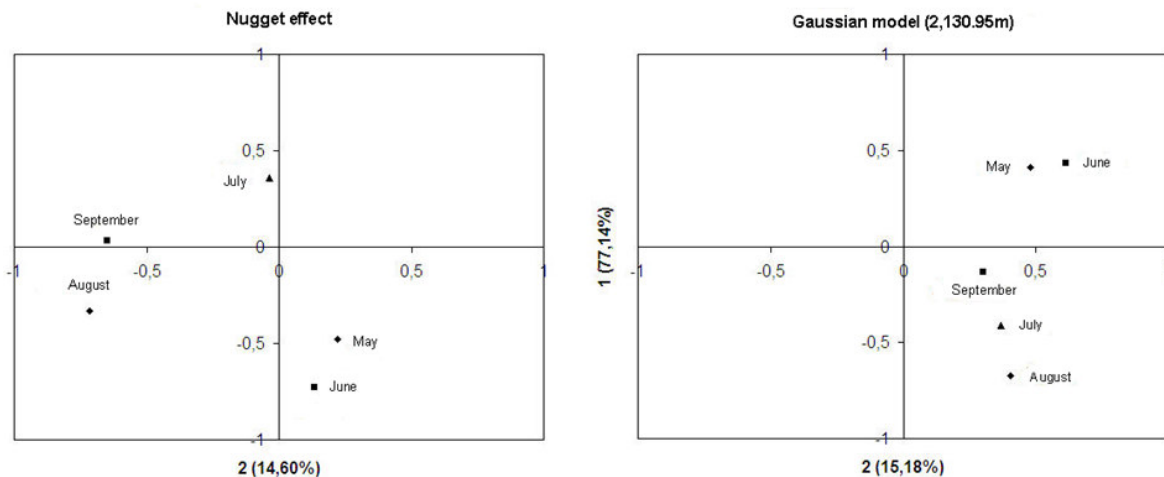
Figure 3.3 - Projection in the space defined by the two first principal components showing the modelled structures (nugget effect and Gaussian model) for the dry season in the Jardim River watershed.

At August and September water table depths start to evidence the dry season effects. The water table depths are still under the effect of the rains of March and April at May, June and July because of the long memories of the hydrological system. These results ratified the temporal correlation described by the structural correlation coefficients of each modelled scale before the decomposition of the coregionalization matrix in independent factors (Table

3.3), and by the spatial patterns at the cokriging maps (Figure 3.2). The correlation is strong at small time lags and is lost in the elapsed dry season.

The cokriging maps were sumed in order to obtain the total increase/decrease of water table depths at Jardim River watershed during 2004 dry season. Table 3.4 presents the values of NDV for the whole area during the monitored period.

Table 3.4 - Natural discharge volume of the Jardim River watershed aquifer during the dry season of 2004.

	Total (m3)	Hectare (m3 ha-1)	Day (m3 ha-1 dia-1)	Dia (L m-2 dia-1)
Mean volume (X)	-22963100	-2069,07	-13,52	-1,35
Percentile 0.1	-51541600	-4643,97	-30,35	-3,03
Percentile 0.9	5603300	504,87	3,29	0,33

Neosols class had high decreases of water table depths. The sandy soil texture and consequently large amount of macropores in these areas result in a fast water movement and infiltration, less water storage by capillarity and fast dryness during long dry periods (Resende et al., 2002).

Red Yellow Latosols presented larger decreases than Red Latosols because of the texture not so clay. The water is storage by soil capillary and pores retention inside porous and fractured rocks. Part of this water infiltrates by saturation, staying in the soil and inside the rocks, and the other part is lost by evaporation and evapotranspiration (Lousada, 2005). However, Red Latosol areas presented both decreases and increases of water volume. This variability came from structural differences from Latosol mineralogy (Reatto et al., 2000) and land uses that this areas have being submitted over the past years (Dolabella, 1996). Water flow and water retention on soil are dependable on soil depth, texture, structure, porosity and pedoform (Resende et al., 2002). All these

variables are essential to predict soil water retention in different soil classes and can interact in several ways. Identification and evaluation of these attributes subsidize the prediction of possible impact in the soil-plant-atmosphere system.

Haplic Cambisols presented low water table depths decrease. In general, Haplic Plinthosols presented increase, and Haplic Gleysols presented decrease in water table depths. These areas in which the water table depths continue increasing during the dry season are concentrated close to the river springs.

The scenarios created from the selected percentile values calculated from the PDF's are interesting for water management because from these estimations account for subestimation and overestimation errors (Goovaerts, 1997). These errors produce different consequences in water management. The variation among the percentiles 0.1 and 0.9 were 124% when compared with the mean estimated scenario (Figure 3.4).
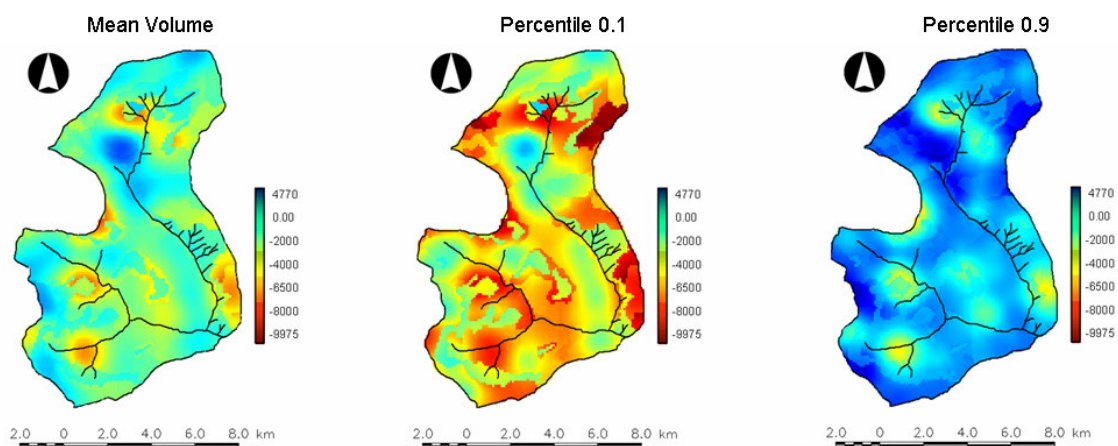
Figure 3.4 - Estimated scenarios accounting for risk of Natural Discharge Volume (NDV) at the dry season of 2004 in the Jardim River watershed (m^3 ha^{-1}).

Percentile 0.1 is a subestimated map when compared with the mean NDV. These values are sensitive to overestimations, creating a pessimistic scenario where the water loss in the dry season is considered maximum. These map values have 90% probability to be higher than the presented and just 10% probability to be lower. This map is indicative for areas where the water table depths decreases are not so big and where the water loss is low.

In the same way, the map of percentile 0.9 is overestimated. These values are sensitive to subestimation, creating an optimistic scenario where the water loss in the dry season is considered minimum. The values of this map have just 10% probability to be higher than the presented and 90% probability to be lower. Areas with deceases in water table depths have high probability to loose water volume. Areas with higher water loss (areas with low values in the map) are areas with higher water table depths variation. This map indicates areas in which the water table depths are bigger and important regions for local drainage and system recharge. These areas are more sensitive to long dryness periods.

Areas with decreases in water table depths have high probability to loose water volume. Optimistic evaluations at areas with big variations of water table depths can incur in agricultural production crashes and errors choosing areas to install irrigation systems. Considering uncertainty contributes to preserve important areas for aquifer recharge. It allows for realistic impact evaluations of new agricultural system implementation which needs high water demand.

3.4 Conclusions

The methodology allows to create uncertainty maps about aquifer NDV considering the ST dependence between water table depths observations.

Percentiles maps 0.10, 0.50 and 0.90 generated from the same ST correlation structure presenting big differences in the NDV.

Percentile map 0.10 indicated favourable areas for water use, where water table depths varied less; and percentiles map 0.90 indicated areas with big water table depths oscillations, which contribute more to the local drainage and aquifer recharge.

4 MONITORING SYSTEMATIC CHANGES IN WATER TABLE DEPTHS IN A BRAZILIAN CERRADO AREA[†]

4.1 Introduction

The rapid expansion of modern agriculture and the establishment of large-scale ranches have affected the characteristic landscape and ecosystem of Cerrado during the past 40 years. A complex wood/grass ecosystem was substituted by shallow-rooted monocultures, which are less well adapted to drought. Cerrado agriculture is profitable, and agricultural expansion is expected to continue needing improvements in techniques and infrastructure. Their need of water supply by irrigation techniques is likely to change the hydrological system (Klink & Moreira, 2002). With irrigation increasing, lowering of water table can occur, and so risks of water supply appear.

The Jardim River watershed is one of the most important regions which supply the Brazilian Federal District with agricultural products. Dolabella (1996) already predicted that the actual water need from irrigated areas in the Jardim River watershed is not guarantied in the future. An over exploration of the water resources in this area is reported and a decrease in the areas irrigated with central pivots suggested, even been considered not too much probable by economic interests. The author also made an alert about the small viability of building barrages as water reservoirs along the river course. Gomes-Loebman et al. (2005) analyzed the increasing of irrigated areas in the Jardim River watershed during a period from 1984 until 2002. They used a time series of Landsat images, classified by their spectral mixture. The agricultural crops

represented 61% of the basin in 1984 and irrigation techniques were not applied. In 2002, their represented 73% of the area and 18% use irrigation techniques in their cultivations (most central pivots).

To optimize and balance the interest of economical and ecological land use purposes, knowledge about the spatio-temporal dynamics of the water table is important (Von Asmuth and Knotters, 2004). In hydrology, water table dynamics are modelled in several ways. In the field of time series analysis, transfer function-noise (TFN) models have been applied to describe the dynamic relationship between precipitation and the water table depths (Box and Jenkins, 1976; Hipel and McLeod, 1994; Tankersley and Graham, 1994; Van Geer and Zuur, 1997). Basically, these methods are multiple regression models where the system is seen as a black box that transforms series of observations on the input (the explanatory variable) into a series of the output variable (the response variable). In our case, water table depth.

The parameters of time series models can be regionalized using ancillary information related to the physical basis of these models (Knotters and Bierkens, 2000, 2001). This approach can be used to describe the spatio-temporal variation in the water table depths. We assume that spatial differences in water table dynamics are determined by the spatial variation in the system properties, while its temporal variation is driven by the dynamics of the input into the system.

An important application of time series analysis is estimating the effects of hydrological interventions. As examples of interventions we can name the clearing or transformation of forests, rise retained surface water levels, operation of pumping wells, construction of barrages or ditches, and so on. When such interventions occur over a certain period of time, we can analyze it by trend analysis. Studying the possible effects of a natural event or a human interventions on the object is the aim of trend monitoring (De Gruijter et al.,

2006), not only to find out whether there has been a change, but also to find out whether the change was caused by a specific event or measure, supporting decision-making to reduce societal and agricultural vulnerability to periods of water shortages. Linear trends can reveal if the water levels are descending or increasing, without precisely knowing if the groundwater level is changing and why. To link the response characteristics of the water table system to the dynamic behaviour of the input, Von Asmuth et al. (2002) presented a method based on the use of a specific TFN model, the Predefined Impulse Response Function In Continuous Time (PIRFICT) model.

The aim of this study was to map and monitor systematic changes in water table depths in a watershed located at the Brazilian Cerrados. We verified linear trends in water table depths and suggest areas with potential risks of future water availability.

4.2 Materials and methods

4.2.1 Time series modelling

The behaviour of linear input-output systems can be completely characterized by their impulse response (IR) function (Ziemer et al., 1998; Von Asmuth et al., 2002). For water table depths, the dynamic relationship between precipitation and water table depth can also be described using physical mechanistic groundwater flow models. However, much less complex TFN models predictions of the water table depth can be obtained which are often as accurate as those obtained by physical mechanistic modeling (Knotters, 2001). In TFN models one or more deterministic transfer components and a noise component are distinguished. These components are additive. A transfer component describes the part of the water table depth that can be explained from an input by a linear transformation of a time series of this input. The noise model describes the autoregressive structure of the differences between the observed water table depths and the sum of the transfer components. The input

71

of the noise model is a series of independent and identically distributed disturbances with zero mean, and finite and constant variance, that is the white noise. Figure 4.1 shows a scheme of the TFN model for water table depths.
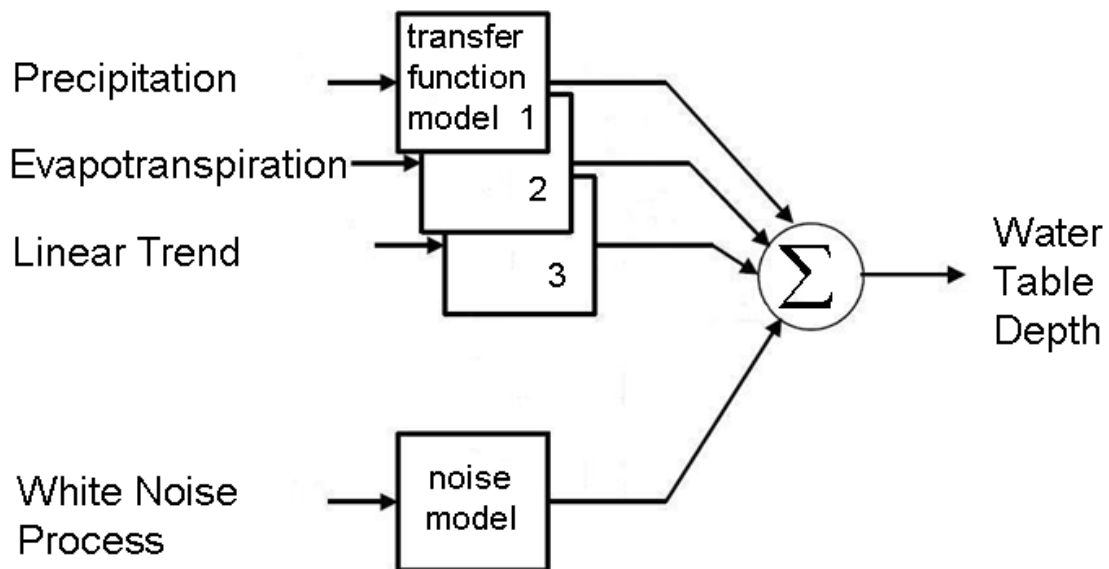
Figure 4.1 - Schematic representation of the transfer function model with added noise for water table depths.

4.2.2 The PIRFICT-model

The PIRFICT-model, introduced by Von Asmuth et al. (2002), is an alternative to discrete-time TFN models. In the PIRFICT-model a block pulse of the input is transformed into an output series by a continuous-time transfer function. The coefficients of this function do not depend on the observation frequency. The following single input continuous TFN model can be used to model the relationship between water table dynamics and precipitation surplus.

4.2.2.1 Estimation of response characteristics of groundwater systems

Under the assumption of linearity, a time series of water table depths is a transformation of a time series of precipitation surplus. The transformation of precipitation surplus series into a series of water table depths is completely governed by the IR function. For the simple case of a linear, undisturbed phreatic system that is influenced by precipitation surplus only, the following single input continuous TFN model, written as a convolution integral, can be used to model the relationship between water table dynamics and precipitation surplus (Von Asmuth et al., 2002):

$$h(t) = h^*(t) + d + r(t) \tag{4.1}$$

$$h^*(t) = \int_{-\infty}^{t} p(\tau)\theta(t-\tau)\partial\tau \tag{4.2}$$

$$r(t) = \int_{-\infty}^{t} \phi(t-\tau)\partial W(\tau) \tag{4.3}$$

where:

$h(t)$ is the observed water table depth at time t [T];

$h^*(t)$ is the predicted water table depth at time t credited to the precipitation surplus, relative to d [L];

d is the level of $h^*(t)$ without precipitation, or in other words the local drainage level, relative to ground surface [L];

$r(t)$ is the residuals series [L];

$p(t)$ is the precipitation surplus intensity at time t [L/T];

$\theta(t)$ is the transfer Impulse Response (IR) function [-];

$\phi(t)$ is the noise IR function [-];

$W(t)$ is a continuous white noise (Wiener) process [L], with properties $E\{dW(t)\}=0$, $E[\{dW(t)\}^2]=dt$, $E[dW(t_1)dW(t_2)]=0$, $t_1 \neq t_2$.

The local drainage level *d* is obtained from the data as follows:

$$d = \frac{\sum_{i=0}^{N} h(t_i)}{N} - \frac{\sum_{i=0}^{N} h^*(t_i)}{N} - \frac{\sum_{i=0}^{N} r(t_i)}{N} \tag{4.4}$$

with *N* the number of water table depth observations.

TFN models are identified by choosing mathematical functions which describe the impulse response and the autoregressive structure of the noise. This identification can be done in two ways:

a) Iteratively: using correlation structures in the available data and model diagnostics.

b) Physically: based on insight into the behaviour of the analyzed system.

Here, the second approach is followed. The IR function describes the way in each the water table responds to an impulse of precipitation. In that respect it is similar to the instantaneous unit hydrograph used in surface water hydrology (Von Asmuth & Maas, 2001). A typical IR function looks like a very skew probability distribution function. The form and area of the impulse response function depends strongly on the hydrological circumstances in situ. Where, for instance, the flow resistance to the nearest drainages is low, the water table will drop quickly after a shower and consequently the area of the IR function will be small. $\theta(t)$ is a Pearson type III distribution function (PIII df, Abramowitz & Stegun, 1964). The option for this function is because of its flexible nature that can adequately model the response of a broad range of groundwater systems. Assuming linearity, the deterministic part of the water table dynamics is

completely determined by the IR function moments. In this case, based on Von Asmuth et al. (2002), the parameters can be defined as:

$$\theta(t) = A \frac{a^n t^{n-1} e^{-at}}{\Gamma(n)}$$

$$\phi(t) = \sqrt{2\alpha\sigma_r^2}\, e^{-\alpha t}$$

(4.5)

where A, a, n, are the parameters of the adjusted curve, $\Gamma(n)$ is the Gamma function and α controls the decay rate of $\phi(t)$ and σ_r^2 is the variance of the residuals.

The PIII df can take shapes gradually ranging from steeper than exponential, via exponential to Gaussian (Figure 4.2). Equation 4.5 and its parameters have a physical meaning that is described in Von Asmuth & Knotters (2004). The physical basis of the PIII df lies in the fact that it describes the transfer function of a series of linear reservoirs (Nash, 1958). For the purpose of general catchment responses modelling, idealization of the catchment as a linear storage reservoir is the most elementary of the various levels of conceptualization that are involved (Bodo & Unny, 1987).

The parameter n shows the number of linear reservoirs and a equals the inverse of the reservoir coefficient normally used. As Knotters & Bierkens (2000) explain, a single linear reservoir (a PIII df with $n=1$) equals a simple physical model of a one-dimensional soil column, discarding lateral flow and the functioning of the unsaturated zone. The extra parameter A is necessary because in equation 4.5, where a precipitation and evapotranspiration series are transformed into a water table depths series, the law of conservation of mass does not apply. The extra parameter adjusts the area of the PIII df to describe the response of the water table to precipitation surplus.
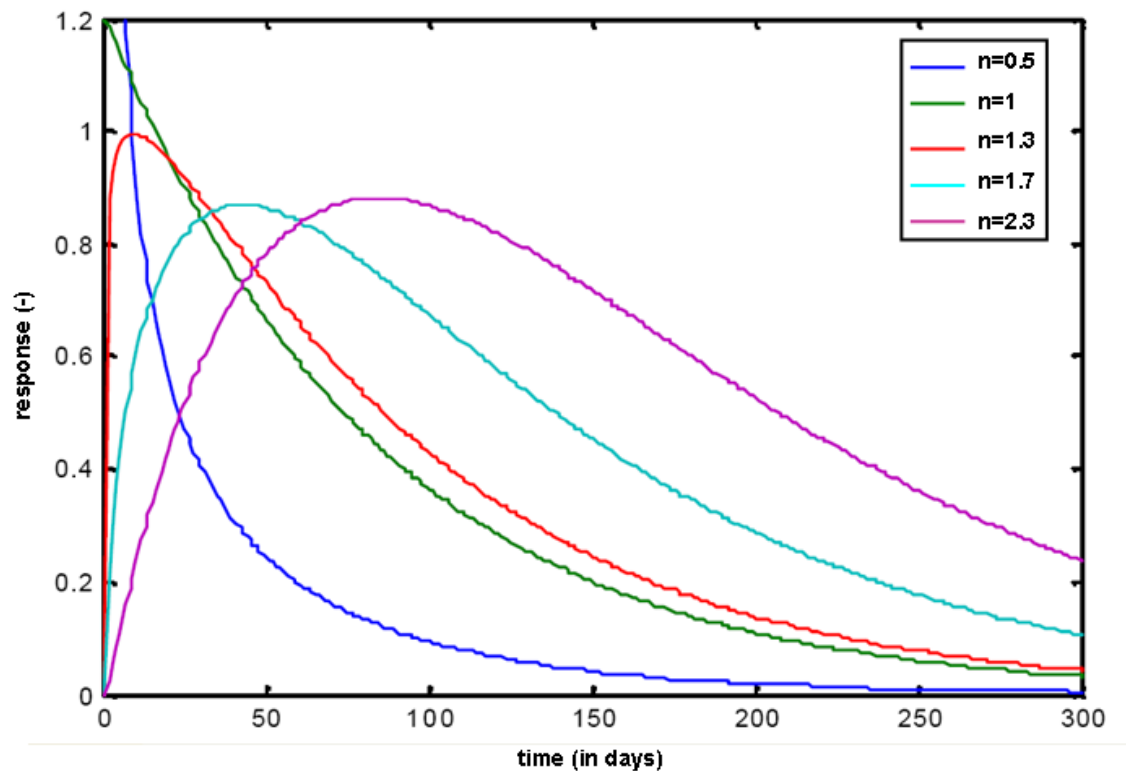
Figure 4.2 - Example of the range of shapes that Pearson type III df can take with (n=
[0.5, 1, 1.3, 1.7, 2.3], *A=n*×100, *a*=0.01).

The PIII df has shown to be able to model fluctuations of water table closely and comparably to Box-Jenkins TFN models with many more parameters (Von Asmuth et al., 2002). The parameter *A* is related to the local drainage resistance (the area of the IR function equals the ratio of the mean height of the water table to the mean water table recharge). *Aa* is determined by the storage coefficient of the soil and *n* as the convection and dispersion time of the precipitation through the unsaturated zone. However, care should be taken when interpreting the parameters of the PIII df, or any other time series model for that matter, in physical sense, because of their lumped and empirical nature (Von Asmuth & Knotters, 2004).

4.2.2.2 Model evaluation, parameter estimation and diagnostic checking of PIRFICT- model

After selecting an IR function that represents the underlying physical process, the available time series have to be transformed to continuous series. First, in order to characterise the variability of precipitation and evaporation, we rely on a simple but effective method to estimate the average precipitation surplus intensity and its annual amplitude. When precipitation surplus data is only available at discrete intervals, the continuous series $p(\tau)$ cannot be reconstructed exactly, but it can be approximated by assuming that the distribution of $p(\tau)$ is uniform during the period t_{pb} to t_{pe} (Ziemer et al., 1998). The average level $\overline{p}$ of the precipitation surplus is obtained as:

$$\overline{p} = \frac{\int_{t_{pb}}^{t_{pe}} p(\tau)\partial\tau}{t_{pe} - t_{pb}}$$

(4.6)

with t_{pb} and t_{pe} denoting the start and end of the period over which the meteorological characteristics are calculated. Next, time is split into year Y and the Julian day D, and the precipitation surplus is averaged over Y, which effectively filters out its yearly course $\widetilde{p}$:

$$\widetilde{p}(D) = \frac{\sum_{Y_{pb}}^{Y_{pe}} p(Y,D)}{Y_{pe} - Y_{pb}}, \quad 1 \leq D \leq 365$$

(4.7)

Because the temperature largely determines the annual evaporation cycle and is more or less harmonic, so is the precipitation surplus and the annual amplitude can be obtained by matching a sine to the yearly course (Von Asmuth & Knotters, 2004).

Using equation 4.6, the transfer model (equation 4.2) can be evaluated using a block response (BR) function $\Theta(t)$. The BR function can be obtained convoluting the IR function with a block pulse of precipitation surplus with unit intensity over a period Δt, as follow:

$$\Theta(t) = \int_{t-\Delta t}^{t} \theta(\tau)\partial\tau$$

(4.8)

Because $\Theta(t)$ is a continuous function, $h^*(t)$ itself is also continuous, and for every observation of $h(t)$ a sample of the residual series $r(t)$ can be obtained. Next, the noise model (equation 4.3) is evaluated in order to obtain a series of innovations $v(t)$. The noise model weights the individual elements of the noise series according to their variance, which are the innovations. The variance of the innovations is a function of the time step (Von Asmuth & Bierkens, 2005). To evaluate the noise model we will derive a direct relation between the residuals $r(t)$ and the innovations $v(t)$. Consider the series $v(t)$ as the nonequidistantly sampled change in the solution to the stochastic integral describing the residual series:

$$v(t) = \int_{t-\Delta t}^{t} \phi(t-\tau)\partial W(\tau)$$

(4.9)

with $\phi(t)$ from equation 4.5 as the noise IR function, we can rewrite equation 4.3 as:

$$r(t) = e^{-\alpha\Delta t} r(t-\Delta t) + \int_{-\infty}^{t} \sqrt{2\alpha\sigma_r^2}\, e^{-\alpha(t-\tau)} \partial W(\tau)$$

(4.10)

which is know as an Ornstein-Uhlenbeck process (Uhlenbeck and Ornstein, 1930; Gardiner, 2004).

78

The use of a simple Ornstein-Uhlenbeck–based (OUB) noise model provides an elegant solution for modelling irregularly spaced observations and data with mixed frequencies (Von Asmuth & Bierkens, 2005). The OUB model is equivalent to an autoregressive model with order 1 (AR(1) model), which is often used to model the residuals in hydrological applications. In a comparison of the AR(1) model, conventional or embedded in a Kalman filter, and the OUB noise model, Von Asmuth & Bierkens (2005) showed that their equations are mathematically equivalent, yielding almost the same estimates. The continuous equations of OUB noise model, however, are more general, give an exact solution and are computationally more efficient as they are not evaluated recursively. Also, Von Asmuth et al. (2002) argue that the Kalman filter approach does not offer a satisfactory solution for time series of slow systems with a nonexponential response because when it is applied to a large extent this problem is alleviated for simple exponential systems. A restriction of the OUB model in its present form is that it is limited to processes that show exponential decay (Von Asmuth & Bierkens, 2005). Combining equations 4.9 and 4.10, we obtain the innovation series calculated from the available data:

$$v(t) = r(t) - e^{-\alpha \Delta t} r(t - \Delta t) \tag{4.11}$$

Subsequently, an estimative of model parameters set $\beta=(A, a, n, \alpha)$ is made with the aid of a Levenberg-Marquardt algorithm, which numerically minimizes a weighted least squares criterion based on the likelihood function of the noise model. Finally the accuracy and validity of the model are checked using the auto and cross-correlation functions of the innovations, the covariance matrix of the model parameters and the variance of the IR functions. For a complete overview of the PIRFICT-model formulation, applications and study cases we refer to Von Asmuth & Maas (2001), Von Asmuth et al. (2002), Von Asmuth & Knotters (2004) and Von Asmuth & Bierkens (2005).

4.2.2.3 Advantages of the PIRFICT-model

Time series of the groundwater level are often collected manually, tend to be nonequidistant and containing missing data (Von Asmuth et al., 2002). As verified above, the PIRFICT-model can be calibrated on data at any frequency available because it operates in a continuous time domain and the time steps of the output variable are not coupled to the time steps of the input variables. Compared to the combined AR model and Kalman filter, the PIRFICT-model offers a further extension of the possibilities of calibrating TFN models on irregularly spaced time series, because the shape of the transfer function is not restricted to an exponential (Von Asmuth & Bierkens, 2005).

In the continuous case, the model order is defined by choosing continuous mathematical functions to represent the IR functions. The mathematical functions are selected on physical grounds, by an iterative procedure of model identification, estimation and diagnostic checking, or with the use of automatic model selection criteria. However, Von Asmuth et al. (2002) presented several important differences from the discrete model identification procedure. First of all, when chosen carefully, a continuous IR function can have a flexible shape and be equivalent to a series of autoregressive/moving average (ARMA) transfer functions. Secondly, the model identification procedure is simplified, because the model frequency does not interfere with the model order and parameter values, and the flexibility of a single continuous IR function can be such that it comprises a range of ARMA transfer functions. Thirdly, the model can be readily identified using physical insight. A continuous IR function can be objectively chosen as the function that best represents the physics of the analyzed system. A physically based IR function on the one hand reduces the sensitivity of the model to coincidental correlations in the data, but on the other hand it can reduce the fit if for some reason the physical assumptions prove to be incorrect (Von Asmuth et al., 2002).

4.2.2.4 Summary of the method

In summary, the method described consists of the following steps:

a) For every input series an IR function is chosen;

b) The input series are assumed to be uniformly distributed in between the time steps and transformed into continuous series using equation 4.6;

c) The transfer convolution integral (equation 4.2) is evaluated using impulse response functions for every block pulse, to obtain a continuous prediction of the output series;

d) A sample of the innovation series is obtained (equation 4.11) for every observation of the output series;

e) Model parameters are estimated with the aid of a Levenberg-Marquardt algorithm;

f) Accuracy and validity of the model are checked using the auto and cross-correlation functions of the innovations, the covariance matrix of the model parameters and the variance of the IR functions.

The PIRFICT-model was applied in this study because the model can describe a wide range of response times with differences in sampling frequency between input series and output series. Being the most important driving forces of water table fluctuation, precipitation and evapotranspiration are incorporated as exogenous variables into the model. Besides precipitation and evapotranspiration, a linear trend component is incorporated to verify systematic changes in the water system. The response of this new impulse in the PIRFICT-model is itself an impulse. That is, the input series is not distorted

by the model, but is, apart from a scaling factor, directly added to the simulated groundwater level itself (Von Asmuth et al., 2002).

4.2.3 Regionalizing the linear trend parameter of the time series model

From the original 40 wells, the PIRFICT-models were calibrated to 37 series of water table depths, using the program Menyanthes. Three wells were considered outliers and removed from the further analyses. Next, the trend parameters reflecting systematic changes of water table depths were mapped. The trend parameter of the PIRFICT-model was interpolated spatially using Universal Kriging (Matheron, 1969; Pebesma, 2004). This works as follows. Let the 'observed' trend parameters be meant as $z(x_1)$, $z(x_2)$, …, $z(x_n)$, where x_i is a (two-dimensional) well location and n is the number of observations (i.e., n=37). At a new, unvisited location x_0 in the area, $z(x_0)$ is predicted by summing the predicted drift and the interpolated residual (Odeh et al., 1994; Hengl et al., 2004):

$$\hat{z}(x_0) = \hat{m}(x_0) + \hat{e}(x_0) \tag{4.12}$$

where the drift m is fitted by linear regression analysis, and the residuals e are interpolated using kriging:

$$\hat{z}(x_0) = \sum_{k=0}^{p} \hat{\beta}_k \cdot q_k(x_0) + \sum_{i=1}^{n} w_i(x_0) \cdot e(x_i); \tag{4.13}$$
$$q_0(x_0) = 1$$

Here, the β_k are estimated drift model coefficients, $q^k(x_0)$ is the kth external explanatory variable (predictor) at location x_0, p is the number of predictors, $w_i(x_0)$ are the kriging weights and $e(x_i)$ are the zero-mean regression residuals.

Ancillary information related to the actual land sources was derived from Landsat 5 images. Three images from July 23, 2005, July 26, 2006 and January 18, 2007, orbit/point 221/71, were inspected and used to classify the actual land use in the region. The image classification results in a land use surface, divided in three classes: Agricultural Crops, Pasture and Cerrado. This classification is based on expert knowledge and a manual outline of the classes. The class Agricultural Crops includes all kinds of agricultural products that are cultivated in the area: small areas cultivated with horticultural products, such as carrots, lettuce and tomatoes, and big areas cultivated with products such as corn, soybeans, cotton, coffee and sugarcane. All these crops demand more water than the original vegetation. The agricultural practices are intensive, resulting in three production cycles during one year when irrigation is applied. Also, the land use in the class Agricultural Crops is very dynamic as a result of agronomical recommendations, rotation schemes or simply prices. The class Pasture is considered to be less water demanding than Agricultural Crops, but more demanding than the natural Cerrado vegetation. These areas are not as dynamic in land use changes as the Agricultural Crops. The classification of both images result in the same grid of ancillary information because the land use did not change in the region during the monitoring period.

The general universal kriging technique was used to interpolate the linear trend parameter (*LTP*) of the PIRFICT-model. The classified Land Use (LU) map was used as predictor. The model was formulated as follows:

$$LTP(x_0) = \beta_0 + \beta_1 \cdot LU1(x_0) + \beta_2 \cdot LU2(x_0) + \beta_3 \cdot LU3(x_0) + e(x_0) \quad (5.14)$$

where *LU*1 is land use class 1 (Agricultural Crops), *LU*2 is land use class 2 (Pasture), *LU*3 is land use class 3 (Cerrado Vegetation) and *e* is a zero-mean spatially correlated residual. Its spatial correlation structure is characterized by a variogram. The land use map has sharp boundaries. For hydrological studies, this does not make sense because water levels do not have abrupt variations

83

related with land use. Therefore, the land use map was smoothed by computing the average presence of land use within a window with 500m radius. The choice of radius was based on expert knowledge and chosen after several tests.

4.3 Results

4.3.1 Time series modelling

Due to spatially varying hydrological conditions, a wide range of calibration results was found for the 37 observed wells. The accuracy and validity of the model were checked using the auto;correlation and cross-correlation functions of the innovations, the covariance matrix of the model parameters and the variance of the IR functions. Table 4.1 summarizes the results of time series modelling.

Table 4.1 - Summary of the statistics of PIRFICT-model calibrations.

	Min	1st Q	Med	3rd Q	Max	Mean	SD
R^2_{adj}	69.27	75.45	79.94	85.11	94.06	80.03	6.56
RMSE	0.13	0.32	0.66	0.90	1.37	0.65	0.35
RMSI	0.13	0.26	0.50	0.73	1.27	0.52	0.29

R^2_{adj}= Percentage of explained variance; RMSE=Root Mean Squared Error (meters); RMSI=Root Mean Squared Innovation (meters); Min=minimum; 1st Q=first quartile; Med=median; 3rd Q=third quartile; Max=maximum; SD= standard deviation

The percentage of variance accounted for each model indicated a good fit of the PIRFICT-model to the data. Low percentages might be caused by errors in the data or lack of data, or possibly other inputs that affect the groundwater dynamics are not incorporated into the model (Von Asmuth et al., 2002).

The RMSE is the average error of the transfer model (Root Mean Squared Error). An individual check at each well denoted small errors, considered reasonable for the target variable (water table depths).

The RMSI is the average innovation or error of the combined transfer and noise model (Root Mean Squared Innovation). Observations that are taken soon after each other are not independent, but contain almost the same information about the state of the system. Increasing sampling time frequency, the variance between observations increases and the noise model gives higher weights to the individual elements of the innovations series. The autocorrelation function of the innovations did not reveal seasonal patterns in the autocorrelation. The autocorrelation functions were rather smooth and the accompanying confidence interval narrow. It indicates that the white noise assumption holds. The noise model was effective to remove the autocorrelation in the time series (Von Asmuth et al., 2002).

As an example of the varying hydrological conditions, we selected four wells distributed over different soil types, hydrogeological systems and land uses. As can be seeing in Figure 4.3, the water table depths respond different from the same inputs.
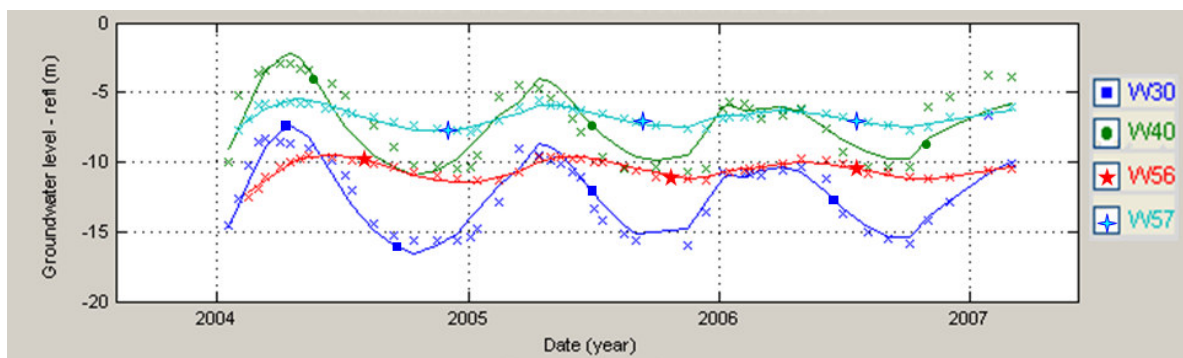
Figure 4.3 - Examples of PIRFICT-model calibrations on WTD time series at four well over the area.

Well 30 (W30) locates in the western part of the basin, under fractured hydrogeological subsystem F, porous hydrogeological system P2, Red Latosol and pasture (Figures 2.1, 2.2, 2.3 and 2.4). Well 40 (W40) locates in the central part of the basin, under fractured hydrogeological subsystem Bambuí, porous hydrogeological system P2, Cambisol and pasture. Well 56 (W56) locates in the extreme northern part of the basin, under fractured hydrogeological subsystem R3/Q3, porous hydrogeological system P1, Red Latosol and agriculture. Finally, Well 57 (W57) locates in the northern part of the basin, under fractured hydrogeological subsystem R3/Q3, porous hydrogeological system P4, Red Yellow Latosol and Cerrado vegetation.

The water table depths variations have a relatively large amplitude for W30 and W40. This can be explained from the hydrogeological subsystem P2 which consists of deep soils. W30 varies deeper because is located under more developed soil types. The soils where W56 and W57 are located present larger water retention capacity due to its mineralogy. The water flux is intense in the north part of the basin, but the subsystem R4 located below this region presents slow water movement. It makes that the water table depths vary less in these wells, even with the well developed and sandy Latosols that are present there. At W57, the levels are more superficial due to the rocky nature of the local geology. These wells are examples of how different the water table depths react in a small catchment like the Jardim river watershed. The flexibility of Pearson III df makes it possible to adjust the PIRFICT-model closely to these different responses.

A continuing monitoring strategy is essential for a good time series model calibration. The calibrations here, for a period of 1240 days finishing at March 03, 2007, are more accurate when comparing with the calibration results found at Manzione et al. (2006) and Manzione et al. (2007b). The monitoring period in Manzione et al. (2006) covered 908 days, finishing at April 05, 2005. In Manzione et al. (2007b), the monitoring period covered 1092 days, finishing at

October 06, 2006. In all situations, the monitoring period starts at October 11, 2003. When the relationship between inputs and output in a TFN model is not totally representing the underlying processes, adjust must be made in the sampling frequency or in the length of the time series. For the long memories of the aquifer system verified in the Cerrados, the semi-monthly sampling frequency here is enough to cover the response time of the system. In the case of length of the monitoring period, when it is not fully covering the correlation length, the monitoring should be continued at least until the correlation length or response time is completely covered (De Gruijter et al., 2006). Manzione et al. (2006) verified that the lengths of time series of the water table depth are too short to obtain reliable results for all sites when it covers only 908 days. Therefore it is recommended to continue monitoring in order to obtain more accurate estimates of the trends in future. Here it is verified in a monitoring period almost one year longer. Our baseline for such assumption is the results of the R^2_{adj}, which are getting better when long time series are analyzed. The values of RMSE and RMSI are subjective and not used for compare model calibrations because these values refers to specific fluctuations of water table depths during the monitoring periods.

4.3.2 Physical interpretation of the PIRFICT-model

The physical plausibility of the results of a TFN model can be judged, for instance, by checking the IR functions, which are equivalent to the cross-correlation function. We check if the memory of the hydrological system, indicated by the time lag where the IR function approximates to zero, is covered by the monitoring period (De Gruijter et al., 2006). Manzione et al. (2006) and Manzione et al. (2007b) found problems with the calibration of the PIRFICT-model for periods of 908 and 1092 days respectively, checking the impulse response function for each well. For both cases, the IR functions did not cover the monitoring period at some locations. The relative short time series were not long enough to characterize the long memories of the hydrological system.

These problems were solved after monitoring WTD for 1240 days. The parameters of PIRFICT-model are summarized on Table 4.2.

Table 4.2 - Summary of calibrated parameters of the PIRFICT-model.

	Min	1st Q	Med	3rd Q	Max	Mean	SD
IR	71	348	442	719	1362	557.24	319.84
A	146.8	552.5	687.8	1101.0	1782.0	860.61	433.72
a	0.0035	0.0102	0.0138	0.0170	0.0449	0.0145	0.0075
n	1.06	1.35	1.75	1.93	3.85	1.78	0.57
E	-2.09	-0.30	0.43	1.42	2.12	0.38	1.18
LTP	-1.83	0.53	1.14	2.19	5.83	1.29	1.52
α	14.48	28.13	37.22	65.19	121.01	49.48	29.37
LDB	-24.24	-13.21	-8.66	-5.31	-0.39	-9.99	5.94

A=drainage resistance (days); *a*=decay rate (1/days); *n*=convection time (days); *E*=evaporation reduction factor (-); *LTP*=Linear trend parameter (meters); *α*=decay or memory of the white noise process (-); IR=Impulse Response (days); LDB=Local Drainage Base (meters); Min=minimum; 1st Q=first quartile; Med=median; 3rd Q=third quartile; Max=maximum; SD= standard deviation

Parameters *A*, *a* and *n* are calculated and regard the shape of the IR function from the pulse of precipitation as input. We found large values of *A* at sites where water table depths varies over a larger range. Following Von Asmuth & Knotters (2004), if the *A* parameter is related with the drainage resistance, "*a*" parameter related with storage coefficient (or porosity) and *n* parameters related with the number of linear reservoirs, they are measures of the memory of the system. In other words, the water table will respond slowly in sites where the values of *A* and *n* are high and *a* has low values, and the other way round. Comparing with Manzione et al. (2006) and Manzione et al. (2007b), parameters set (*A*, *a*, *n*) are much better calibrated, as we update the time series with new data. Continuing monitoring collaborated for more reliable

results about the relationship between WTD and precipitation surplus in the Cerrados. Figure 4.4 gives the IR functions for precipitation impulse for the wells 30, 40, 56 and 57. The functions have higher response factors for wells with large water table depths amplitudes (W30 and W40).

Parameter *E* is part of the IR function of evapotranspiration. The response to evaporation can be considered the same as the response to precipitation, apart from a variable evaporation reduction factor. This value should be between 0 and 1. For some wells we found estimates of *E* which are not realistic. Also, the SD values are high. One reason could be that the climate station, located around 10 km outside the study area, does not represent the meteorological circumstances at all wells locations. Another reason might be in the large temporal variation of land use, which makes these parameters difficult to estimate.
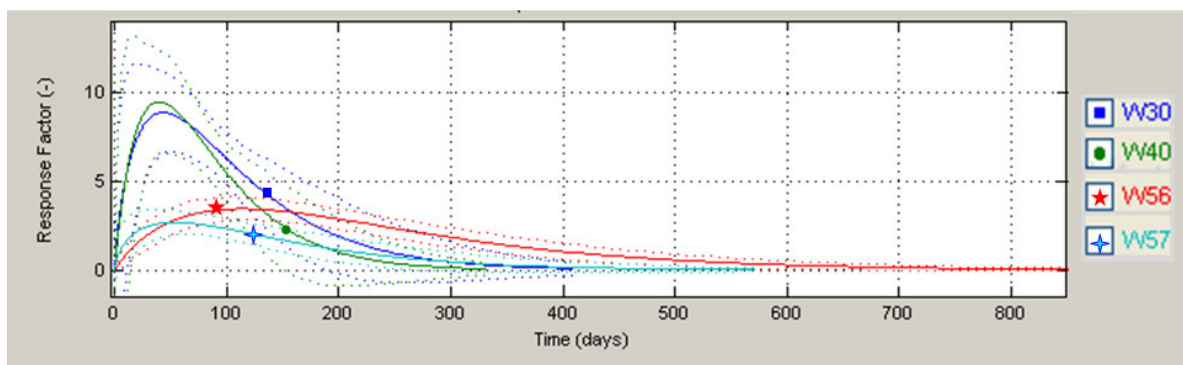
Figure 4.4 - Adjust of the IR functions for the input series of precipitation.

Parameter *α* from the noise model was well calibrated and statistically significant for all wells, confirming the assumption of exponential decay of the noise model. Wells which had longer IR functions also had bigger values of *α*. LDB is the local drainage base, or more correctly the level of the groundwater level without precipitation or other influence. The values were well calibrated, being under the fluctuation range of WTD variation for all wells.

4.3.3 Mapping trends in water table depths

4.3.3.1 Spatial modelling

Including the land use variables into the geostatistical model caused a decrease in the variance, as can be seen in Figure 4.5. The spatial dependence at small distances is poorly estimated because of the small number of observation wells that are fairly uniformly spread across the area. The nugget parameter of the variogram reflects the precision of the *LTP* and the short-distance spatial variation in *LTP*.
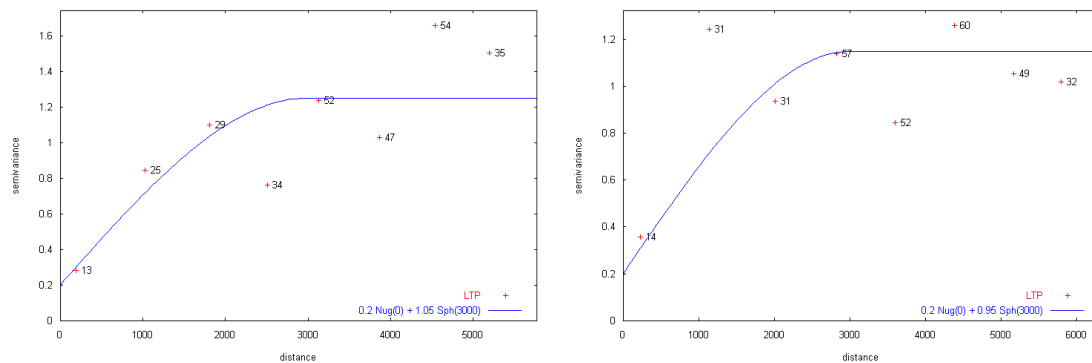
Figure 4.5 - Variograms fitted for the linear trend parameter without including a trend that depends on land use (left) and with including a trend (right).

4.3.3.2 Spatial interpolation

The interpolation results of the *LTP* of the PIRFICT-model using universal kriging and Land Use classification as a drift are in Figure 4.6. Positive values in the interpolated map of systematic changes in water table depth indicate a rise of the water table during the last three years, and negative values lowering. The map shows a large area near the river where systematic lowering occurs. These areas are covered with traditional agricultural crops, using irrigation systems that catch water directly from the river (surface water).
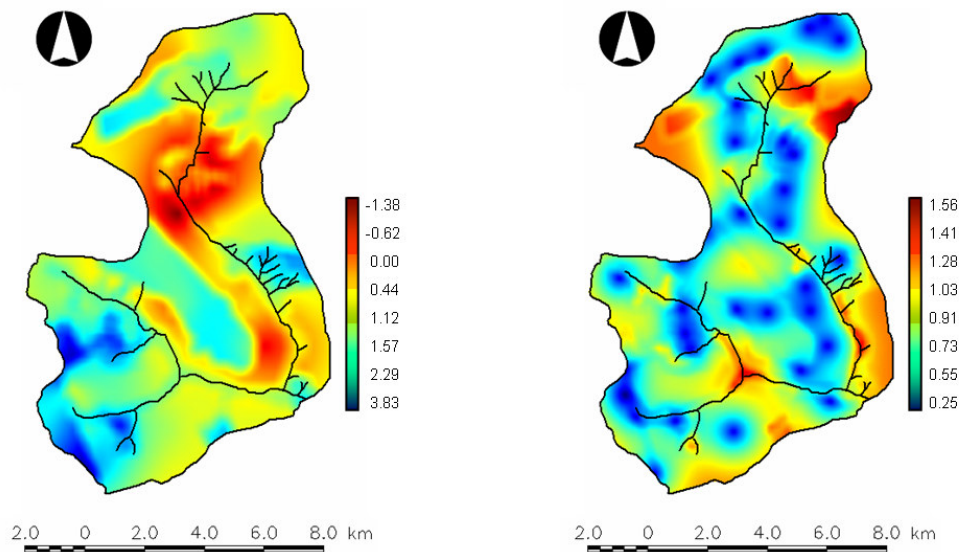
Figure 4.6 - Estimated *LTP* of the PIRFICT-model (meters) during the period from October 2003 to March 2007 (Left) and the corresponding kriging standard deviations (Right).

For some areas systematic risings of the water table depths were estimated. These risings can be explained as follows. After three consecutive dry years (2001, 2002 and 2003) with less rainfall than the annual average over the last 33 years, the Jardim river area had two consecutive wet years (2004 and 2005), with more rainfall than the annual average. The years 2001, 2002 and 2003 were very dry with 24.4, 41.02 and 33.2% less rainfall than the annual average over the last 31 years, respectively. During 2004 and 2005, rainfall was 8.54 and 4.6% larger than the annual average of the last 33 years, respectively.

In the northern part of the basin, the areas with significant risings of water table are under the influence of irrigated crops there located. The areas with risings in the eastern part of the basin belong to the porous hydrogeological system P4. These locations have shallow soils, with slightly fluctuating water tables close to the ground surface. The contribution of this subsystem to the groundwater system of the Jardim river watershed is restricted (Lousada, 2005). The degradation of the Cerrado vegetation in these areas could be a reason of

91

systematic rising of the water table depths. This vegetation does not explore the water reserves during all year long as agricultural crops, given time to aquifer recharge. In the western part of the basin, the risings are associated with the fractured hydrogeological system Canastra. The porosity of the bedrock material acts as a sponge, holding the water that is not use by the degradated pastures areas presented there. Also, this vegetation does not have root systems long enough to reach and use this water. The map of the kriging standard deviations reflects the accuracy of the predicted systematic changes of water table depth. The large standard deviations reflect the large uncertainty in the *LTP* parameters. Large uncertainty implies that observed lowerings and risings of the water table depth may not be statistically significant.

4.3.3.3 Cross-validation

The results of the universal kriging were evaluated by cross-validation. Table 4.3 gives the results.

Table 4.3 - Cross-validation for the spatial interpolation of *LTP* (meters).

	Observed	Predicted	Pred. – Obs.	Pred. SD	Z-score
Min	-3.9200	-1.6020	-2.9240	1.1580	-2.2710
1st Q	-0.8952	-0.5655	-0.8108	1.2120	-0.6629
Median	-0.1595	-0.1963	-0.2134	1.2440	-0.1727
3rd Q	0.9463	0.2274	1.0150	1.2920	0.7775
Max	2.3600	1.1220	3.4180	1.3460	2.7380
Mean	-0.1456	-0.1456	0.0007	1.2490	0.0005
SD	1.4040	0.6120	1.3890	0.0511	1.1130

Pred.=Predicted; Obs.=Observed; Min=Minimum; 1st Q=First quantile; 3rd Q=Third quantile; Max=Maximum; SD=Standard deviation; Z-score=(Pred-Obs) / Kriging variance.

The cross-validation results indicate large interpolation errors. These errors can be explained from the uncertainty about the *LTP* parameters at the 37 well locations, the poor relationship between land use and *LTP* and the poor spatial correlation structure in the stochastic residual of the universal kriging model. The *Z*-score mean and standard deviation of the *Z*-score indicate a good performance of the kriging system, with values close to zero and one, respectively. Compared with Manzione et al. (2006) and Manzione et al. (2007b) the interpolation performed better here with the parameters of the model better adjusted.

4.3.3.4 Statistical significance of *LTP*

The systematic changes were mapped and the significance of these estimations checked at 5 and 10% probability levels. In Manzione et al. (2006) were not found significant risings of the water table. Two spots with systematic lowering were indicated. These two spots are areas with intensive irrigation (Figure 4.7, left).
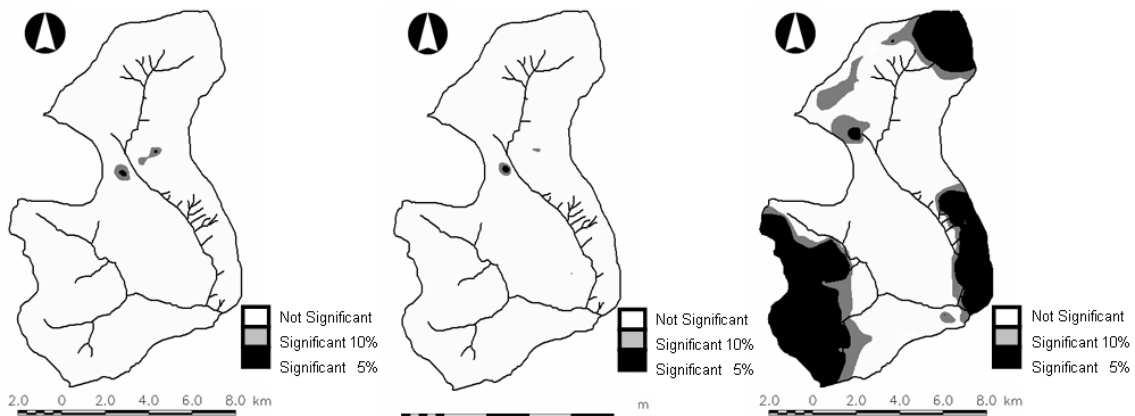
Figure 4.7 - Significant lowerings of WTD verified for 908 days monitoring (left) and 1240 days monitoring (center) and significant risings of WTD for 1240 days monitoring (right).

93

Again, following the recommendation of Manzione et al. (2006) we continued monitoring to obtain more accurate estimates of the trends in future. We investigate systematic changes in the water table depths for a period of 1240 days, from October 11, 2003 until March 03, 2007. As can be seeing in Figure 4.7 (center), extending the analyses for a longer period, it confirms spots where the water levels are decreasing. These areas in the course of the river with significant systematic lowerings deserve attention because a few years ago a barrage was constructed there as water reservoir to supply irrigation of small crops. The trend analyses for a monitored period of 1240 days also reveal areas significant systematic risings of the water table depths (Figure 4.7, right). Due to the slowness and long memories, the groundwater system could recharge during the monitored period in some areas, resulting in rising water tables.

4.4 Discussion

Human interventions and the changes in the Cerrado environment over the last decades can change the hydrological system, affect water availability and lead to water shortage. Following the evolution of trends during time is essential to water policy and decision making. Areas with systematic risings in water table depths indicate favourable areas for water use, where the levels are increasing and the aquifer recharging. Attention should be paid in order to explore these water reserves because this rising could be just an indication that the areas are under recovering water reserves. If the levels were too low, the areas could be just returning to original water table levels. Continuing monitoring these effects is important to distinguish systematic changes in water table depths and short term variations due to climatological conditions over a period.

4.5 Conclusions

Time series modelling is an effective method to characterize the seasonal patterns of water table depths in the Cerrado. The PIRFICT-model adjusts IR

functions for different responses of the hydrological system over the Jardim River watershed and the calibrated parameters interpreted physically. Calibrations were getting more accurate adjusting the length of the time series to cover the long memories of the water system.

The water table depths in the study area appear to have changed systematically between October 2003 and March 2007. Trend monitoring is important in evaluating changes in water table depths. Time series models are important tools in trend monitoring, especially in areas like the Cerrado which are affected by seasonality and under pressure of an intense use of water resources. The monitoring in the Jardim river watershed showed persistent patterns of significant systematic lowering in water table depths as well effect of rainy years in the aquifer recharge.

Ancillary information on land use reduced the variance and incorporated physical meaning in the spatial prediction model. Uncertainty about the spatial interpolation was reduced analyzing longer time series.

For more reliable results, which can be utilised in water management and policy making, we recommend to continue monitoring.

5   PREDICTIVE RISK MAPPING OF WATER TABLE DEPTHS IN A BRAZILIAN CERRADO AREA[‡]

5.1   Introduction

Time series models provide a systematic empirical method of estimating and predicting the temporal behaviour of dynamic hydrological phenomenon (Box & Jenkins, 1976; Hipel & McLeod, 1994). Time series modelling allow us to simulate and forecast the behaviour of hydrologic systems and to quantify the expected accuracy of these forecasts (Tankersley & Graham, 1993; Salas & Pielke, 2003). In areas where the water resources are under effect of seasonal influences, like the Cerrado, an especially useful application of TFN models is when we are interested in the probability of extremes. Such probabilities are underestimated when using only the deterministic model (Knotters & Van Walsum, 1997). A stochastic model considers the unexplained part or noise model component. This is interesting for estimations of extreme water table depths because uncertainties can be quantified. Using a time series model with added noise it is possible to simulate over periods that do not have observations, as long as data on explanatory series are available.

Models using precipitation surplus/deficit as an input variable, calibrated on time series of water table depths of limited length, enable us to simulate series of extensive length (Knotters, 2001). WTD values not influenced by the particular weather circumstances during the monitoring period of water table depths (like short-term variations) are estimated from these extensive series. These estimations can be made for any specific date in any future year, given the

---

[‡] A previous version of this chapter was presented in the 5th International Symposium on Spatial Data Quality and published in the proceedings, Manzione, R. L., Knotters, M., Heuvelink, G. B. M., Von Asmuth, J. R., Câmara, G. CDROM

prevailing hydrological and climatic conditions. These may include expected values of WTD at certain times, as the start of the growing season, or probabilities that critical levels are exceeded at certain times or during certain periods. These target parameters are estimated with the purpose of obtaining characteristics of a certain time.

The aims of this study are to estimate and map the expected WTD in a watershed located at the Brazilian Cerrados. We calculated and measured risks that critical levels are exceeded, simulating realizations of the PIRFICT-model. These estimations are made for specific dates in any future year, to support decision making in long-term water policy and indicate areas with potential risks of future water shortage and shallow water table depths. In addition, the uncertainty associated with the estimated WTD is quantified.

5.2   Material and methods

5.2.1   Simulating water table depths

Time series models using precipitation surplus/deficit as input variable, calibrated on time series of WTD with limited years, enable us to simulate series of extensive length (Knotters & Van Walsum, 1997). Statistics of WTD are estimated from extensive series. These statistics will represent the prevailing hydrological and climatic conditions rather than specific meteorological circumstances during the monitoring period of WTD.

To link the response characteristics of the water table system to the dynamic behaviour of the input, Von Asmuth et al. (2002) presented a transfer function-noise model in continuous time, the so-called PIRFICT-model. The simulation of WTD presented here is based on a time frequency filtering of the PIRFICT-model performed as a convolution in the time frequency domain. This procedure considers the shape of the PIII df adjusted from the parameters of each model. Interacting the precipitation input signal and the aquifer system can be regarded

98

as an operation in the time frequency domain between the time frequency expansion of the signal and the time frequency response of the system. An important advantage in the use of the PIRFICT-model here, compared with discrete-time TFN-models, is that it can deal with input and output series which have different observation frequencies and irregular time intervals.

These models contain a dynamic component, describing the dynamic relationship between the input and the output, either physically or empirically. But variation of the water table cannot be completely explained from the precipitation and evapotranspiration series. So, the models must contain a noise component, which describes the part of water table variation that cannot be explained with the used physical concepts or empirically from the input series. The noise component has to be considered in the simulation procedure, since we are interested in statistics of extremes, like the probabilities that critical levels are exceeded. Details about simulations for water resources can be found in Hipel & McLeod (1994). Here, the following steps are followed:

a) After modelling the relationship between precipitation surplus/deficit and WTD using the PIRFICT-model, series of water table depths are extrapolated to 30 years. It is assumed the average weather conditions during the last 30 years represent the prevailing climate. As result, deterministic series of predicted WTD are generated.

b) Realizations of the noise process are generated by stochastic simulation and next added to deterministic series resulting in realizations of series of WTD. Realizations of the noise process can be generated either by random sampling from a normal distribution with zero mean and residual variance, or by resampling from the fitted residuals.

c) From the previous steps, *N* realizations of the stochastic simulation are generated. Statistics representing the prevailing hydrologic conditions are calculated from the WTD PDF for each *t* instant.

In this study, we applied random sampling from a normal distribution. We extrapolate WTD series from a monitoring period of 1240 days (from October 11, 2003 until March 06, 2007) to 30 years length series using available 30 years length series of precipitation and potential evapotranspiration.

5.2.2 Risk assessment of water table depths

We calculate statistics about WTD to set up two risk situations: risk of water shortage and risk of shallow WTD. The estimations of process characteristics were made for two specific interesting dates in the Cerrado region:

a) Beginning of the planting season: This date has special interest for the Cerrados. The wet season usually starts around this period and farmers use to start cultivations just after the first rains of the season. We considered October 1st as a hypothetical date for the beginning of the planting season.

b) End of wet season: This is another important date in the Cerrados because the water table reaches the highest levels. Also it is the date when the cultivations ends and harvest operations start in the fields. We considered April 30 as a hypothetical date for this scenario.

PDF's were calculated from the simulated WTD series. We selected the 0.05 percentile for water shortage. We can say that the area have just 5% probability to have lower WTD than these values and 95% probability to have higher values. The limits established for risks of water shortage were the depth of the

wells, with dry wells characterizing a scenario of water shortage in the area. Moisture stress at every stage of plant development can reduce seed yields, but the extent of yield decrease from water stress varies with stage development. At the beginning of the planting season it can be a problem during for plants development, affecting water availability and resulting in production losses. In the end of the wet season, close to the end of the corn productive cycle, for example, it can reduce pollination successes; affect kernel development and grain fill. For soybeans, it can abort flowers, reduce pod set and pod filling.

For risk of shallow water tables we selected the 0.95 percentile. We can say that the area have just 5% probability to have higher WTD than these values and 95% probability to have lower values. The limits established for risks of shallow water table depths were 0.5 meters bellow the ground surface. Shallow water tables can be a problem for machinery. At the beginning of the planting season shallow water tables are problematic because they can make it impossible to plow and execute planting operations. Also it can influence soil conditions, decreasing the soil redox potential, increasing the pH in acid soils and decreasing it in alkaline soils, and increasing the conductivity and ion exchange reactions. These changes in the system have direct influence in plant growth, by affecting the availability and toxicity of nutrients, regulating uptake in the rhizosfere. In the end of wet season shallow WTD can impossible harvest or even delay conditions for plants maturation, for example.

5.2.3 Regionalizing simulated water table depths

The results of WTD simulations are interpolated spatially using Universal Kriging (Matheron, 1969; Pebesma, 2004). A digital elevation model (DEM) with 15 meters resolution provides ancillary information related to local geomorphology (Lousada, 2005). The use of exhaustive information on elevation is interesting because it can decrease the variance and the uncertainty in the spatial prediction model. Also, when the ancillary information

is physically related to target variable it can incorporate physical meaning to the predictions (Manzione et al., 2007a). In our case areas with relatively low elevation and close to drainages present relatively shallow water tables, whereas in areas with relatively high elevation and far from drainages the water table is relatively deep (Furley, 1999).

DEM can be incorporated as a drift (Odeh et al., 1994; Knotters et al., 1995) in the spatial prediction model. Let the estimated characteristic, e. g. probability of exceedance, be given as $z(x_1)$, $z(x_2)$, …, $z(x_n)$, where x_i is a (two-dimensional) well location and n is the number of observations (i.e., n=37). At a new, unvisited location x_0 in the area, $z(x_0)$ is predicted by summing the predicted drift and the interpolated residual (Odeh et al., 1994; Hengl et al., 2004):

$$\hat{z}(x_0) = \hat{m}(x_0) + \hat{e}(x_0) \tag{5.1}$$

where the drift m is fitted by linear regression analysis, and the residuals e are interpolated using kriging:

$$\hat{z}(x_0) = \sum_{k=0}^{p} \hat{\beta}_k \cdot q_k(x_0) + \sum_{i=1}^{n} w_i(x_0) \cdot e(x_i);$$
$$q_0(x_0) = 1 \tag{5.2}$$

Here, the β_k are estimated drift model coefficients, $q_k(x_0)$ is the kth external explanatory variable (predictor) at location x_0, p is the number of predictors, $w_i(x_0)$ are the kriging weights and $e(x_i)$ are the zero-mean regression residuals. In this case, for WTD, the model was formulated as follows:

$$WTD(x_0) = \beta_0 + \beta_1 \cdot EV(x_0) + e(x_0) \tag{5.3}$$

where *EV* is the elevation value for each location and *e* is a zero-mean spatially correlated residual. Its spatial correlation structure is characterized by a semivariogram.

5.2.4 Summary of the method

In summary, the method described consists of the following steps:

a) Calibration of the PIRFICT-model;

b) Stochastic simulation of WTD series by using the PIRFICT-model (*N*=1000) and input series of 30 years length;

c) Pick all selected date values of WTD generated by stochastic simulation;

d) Create a probability distribution function (PDF) of these values;

e) Selection of percentile values for WTD;

f) Repeat steps above for all wells;

g) Model the spatial structure of the percentile values with geostatistical techniques;

h) Finally, we use these percentiles to create risk maps of water levels that could be exceeded at the selected date with extreme values of the PDF.

5.3 Results

5.3.1   Simulation with the PIRFICT-model

Stochastic simulation with the PIRFICT-model was performed for the 37 wells. For ten wells the simulation results showed that the supposed stationarity conditions were not met. The distribution functions of the simulated WTD for these wells were bimodal.

Possibly these results can be explained from the relatively short length of the water table time series which did not completely cover the response time of the hydrological system. Another reason could be the strong influence of long-term WTD variations in the time series used to extrapolate the results. For example, the dry years of 2001, 2002 and 2003 might have a long-term effect on water table in systems with long hydrological memory. This effect acts different over the basin, due the presence of different geological systems (Lousada, 2005). Continuing monitoring WTD would enable us to clarify these questions (De Gruijter et al., 2006). Manzione et al. (2007b) found the same problem simulating WTD series from October 11, 2003 until October 06, 2006. Increasing the length of the time series in 5 months (until March 03, 2007) was not enough to clarify this question. These ten wells were excluded from calculating the PDF's and spatial interpolation.

5.3.2 Risk Mapping

5.3.2.1 Spatial modelling

The WTD PDF's for October 1st and April 30 are created from the simulated data for the remaining wells. The WTD's which are expected to be exceeded with 5% and 95% probability, have a spatial dependence which was modelled by semivariograms. Ancillary information was used in spatial predictions, because the number of observation points was relatively low. Including elevation as a spatial drift into the geostatistical model caused a decrease in the semivariance as observed by Manzione et al. (2007a). Table 5.1 summarizes the adjusted semivariograms.

The spatial dependence at short distances is poorly estimated because of the small number of observation wells that are fairly uniformly spread across the area. The nugget parameter of the semivariogram reflects the measurement precision of the WTD and the short-distance spatial variation in WTD.

Table 5.1 - Parameters of the adjusted semivariograms for the selected percentiles from the simulated WTD PDF's at October 1st and April 30.

WTD	Model	Nugget	Sill	Contribution	Range (m)
P 0.05 Oct 1st	Spherical	4	26	22	2400
P 0.95 Oct 1st	Spherical	3	18	15	2400
P 0.05 Apr 30	Spherical	5	19	14	2400
P 0.95 Apr 30	Spherical	4	11.5	7.5	2400

5.3.2.2 Spatial interpolation

Universal kriging interpolation resulted in maps with a physical meaning related to the local drainage. These maps were applied to quantify the risks to be utilised in water management to evaluate water shortage and wet conditions for agricultural purposes. Figures 5.1 and 5.2 give the results for October 1st and Figures 5.3 and 5.4 the results for April 30.
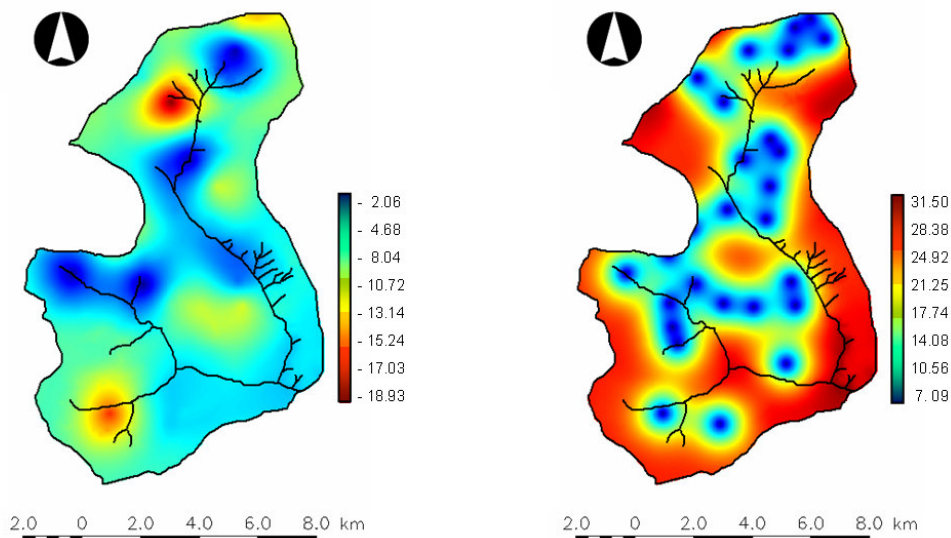
Figure 5.1 - Map of WTD levels (meters) that will be exceeded with 95% probability at
October 1st (left) and the corresponding kriging variance (right).

Figure 5.2 - Map of WTD levels (meters) that will be exceeded with 5% probability at
October 1st (left) and the corresponding kriging variance (right).

Figure 5.4 - Map of WTD levels (meters) that will be exceeded with 5% probability at Apr 30 (left) and the corresponding kriging variance (right).

Kriging variance maps present different results. The variances presented in the maps result of the spatial model adjusted for each situation. Even adjusted for the same range, the semivariogram parameters influence the estimations. The semivariograms with low nugget values produced more accurate estimates of WTD. The semivariograms with high sill values produced maps with high uncertainty about the estimates.

5.3.2.3 Cross-validation

The results of spatial interpolation were evaluated by cross-validation. Tables 5.2 and 5.3 give the results for October 1st and April 30, respectively.

Table 5.2 - Cross-validation for spatial interpolation of WTD (meters) for October 1st.

	Observed	Predicted	Pred. – Obs.	Pred. SD	Z-score
P 0.05					
Min	-20.76	-14.44	-10.01	3.99	-2.18
1st Q	-11.47	-9.76	-2.01	4.54	0.40
Median	-8.49	-8.93	1.15	4.78	0.24
3rd Q	-7.36	-7.86	2.64	5.04	0.58
Max	-0.67	-5.64	12.06	5.55	2.39
Mean	-9.03	-9.08	-0.05	4.78	-0.005
SD	4.76	1.79	5.38	0.42	1.09
P 0.95					
Min	-16.66	-11.11	-7.29	3.36	-1.90
1st Q	-8.87	-6.90	-3.03	3.79	-0.83
Median	-5.81	-6.13	0.90	3.98	0.24
3rd Q	-3.08	-5.17	2.29	4.20	0.56
Max	-0.11	-2.03	10.53	4.62	2.50
Mean	-6.10	-6.07	0.03	3.99	0.004
SD	4.21	1.79	4.66	0.34	1.14

Pred.=Predicted; Obs.=Observed; Min=Minimum; 1st Q=First quantile; 3rd Q=Third quantile; Max=Maximum; SD=Standard deviation; Z-score=(Pred-Obs) / Kriging variance.

The SD-values of the observations are much higher than those of predictions, pointing to smoothed interpolation values (For example. 4.76 vs. 1.79 and 4.21 vs. 1.79 in Table 5.2 or 4.19 vs. 3.47 and 3.28 vs. 1.86 in Table 5.3). The errors can be explained from uncertainty about the calibrated models, a poor relationship between elevation and WTD and from a poor spatial correlation structure in both kriging models.

Table 5.3 - Cross-validation for spatial interpolation of WTD (meters) for April 30.

	Observed	Predicted	Pred. – Obs.	Pred. SD	Z-score
P 0.05					
Min	-18.53	-12.45	-8.95	3.66	-2.21
1st Q	-9.77	-8.29	-3.02	4.02	-0.79
Median	-6.87	-7.06	-0.03	4.18	-0.01
3rd Q	-4.49	-5.84	2.81	4.35	0.72
Max	-0.49	-3.47	11.32	4.76	2.59
Mean	-7.22	-7.26	-0.04	4.19	-0.004
SD	4.19	3.47	4.61	0.29	1.09
P 0.95					
Min	-11.35	-7.49	-5.06	2.93	-1.60
1st Q	-5.29	-4.36	-2.49	3.14	-0.75
Median	-1.48	-2.29	-0.97	3.26	-0.27
3rd Q	-0.06	-1.61	2.35	3.36	0.72
Max	0.10	-0.24	8.12	3.66	2.41
Mean	-2.90	-2.92	-0.01	3.27	-0.002
SD	3.28	1.86	3.39	0.19	1.04

Pred.=Predicted; Obs.=Observed; Min=Minimum; 1st Q=First quantile; 3rd Q=Third quantile; Max=Maximum; SD=Standard deviation; Z-score=(Pred-Obs) / Kriging variance.

Predictions on Tables 5.2 and 5.3 had the mean WTD value respected and showed small mean interpolation errors (-0.05 and 0.03 meters at Table 5.2, and -0.04 and -0.01 at Table 5.3). The mean and standard deviation of the Z-score had values close to zero and one, respectively, which suggests a good performance of the kriging systems. In both cases, the model with high nugget values produced maps with high mean Z-score, reflecting higher errors in the

spatial model and spatial interpolation. Comparing with Manzione et al. (2007b) the interpolation performed better for October 1st here but the results were smoothed.

5.3.2.4 Risk areas

Manzione et al. (2007) applied the PIRFICT-model in simulations, to predict risk of water shortage and shallow water table depths for October 1st. They found a negligible risk of shallow water tables or water shortage for this date. We found the same results simulating the PIRFICT-model with a long WTD time series.

For April 30, there is also a negligible risk of water shortage. On the other hand, we detected 3 areas with potential risk of shallow water table depths (Figure 5.5). Fortunately, these areas are close to the drainages and under legal protection. The Brazilian legislation fixes maintenance of gallery forest along rivers courses. Remove this vegetation characterize an environmental crime, so farmers avoid cultivations in such areas.
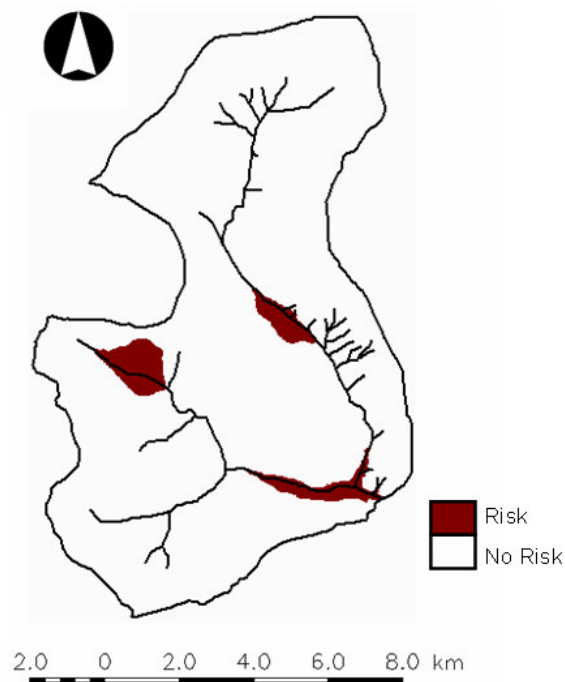
Figure 5.5 - Areas with risks of shallow water table depths at Apr 30.

The risk spot in the western part of the basin deserve some attention. Crops locate inside and close to this area present potential risk of shallow water table depths.

5.4 Discussion

Information about hydrological systems and water table variations is important for water management and is needed to assess choices for long-term water management policy. Preferably, uncertainty about the future water table depths is quantified to enable risk assessment.

Time series modelling using the PIRFICT-model was efficient to model a wide range of different responses of the hydrological system presented over the basin. Policy makers can and should use these results to optimize water use and to regulate the competing claims for water resources that often occur between small farmers, big farmers with irrigated crops and water withdrawal for human use. However, the results reflect uncertainties from different sources: uncertainties related to the data (observed WTD, climatic database, DEM), uncertainty associated with time series modelling and with the model of spatial variation.

The quality of the map was restricted by effects of relatively short time series that did not satisfactory characterize long memory systems. The quality of the time series models depends on both sampling frequency and length of series. The quality of the model of spatial structure depends on the number of wells and their sample configuration. The use of DEM as ancillary information slightly improved the quality of the final risk maps.

5.5 Conclusions

Simulating water table depths at long periods using a stochastic model enabled us to predict water levels without influences of isolated short-term climatic disturbances. The results support decision making in long-term water policy and suggest areas with potential risks of shallow or deep water table depths.

For October 1st there is a negligible risk of water shortage and shallow water table depths that could affect agriculture in someway. For April 30 there is negligible risk of water shortage, but 3 risk spots of shallow water table depths were found. The analysis should be extended to other dates and periods that are critical to water supply. The method presented in this study enables this extension to predict risks at any date in future.

Given the long memories in the hydrological system of the study area, we recommend to continue monitoring water table depths to obtain more reliable results in future.

6 CONCLUSIONS AND FUTURE WORK

6.1 Introduction

This thesis presents methods to support water management by modelling the spatio-temporal dynamics of water table depths. We show how to estimate the water volume lost during a specific season, using a linear coregionalization model. We show how to apply the PIRFICT time series model to an area in the Brazilian Cerrado. We present how to predict risks of extreme water levels for agriculture in the region. These results aim at contributing to improve water management for crop production in Cerrado areas. In this final chapter, we discuss the results presented in chapters 3, 4 and 5, report problems found during this research, and point out some possible improvements and future work.

6.2 Jardim River watershed

Monitoring water table depths for agro-ecological purposes in Brazil is uncommon. Monitoring data are scarce mainly because of the costs to set up a sampling network and the lack of human resources to perform field observations frequently. The Jardim River watershed is among the few watersheds which present several observation wells at this scale.

In this study, the observation wells are part of a subproject of Agricultural Technology Development Project for Brazil (PRODETAB) entitled "Water resources characterization, availability and demand for irrigated agriculture in the Jardim River watershed". This project is an agreement between the Brazilian Government and the World Bank, managed by EMBRAPA (Brazilian Agricultural Research Corporation) and including University of Brasília (UnB) and INPE. The project ended in 2006 after three years of financial support.

After the project finished, the monitoring continues, and the data are shared by EMBRAPA Cerrados and UnB.

At first, the plan to drill 60 wells was considered unfeasible. After several meetings, the leaders of the project agreed on the number and the locations of the wells. Some wells got dry after some weeks. Probably the bedrock under pressure or impermeable layers obstructed the water flow. This study used data from 40 wells, excluding outliers or erratic information. This thesis is a result of these investments. Lousada (2005) presented the first result that used the project's data.

6.3 Data modelling in space and time

6.3.1 Geostatistical modelling

We used geostatistical tools to describe the spatial distribution of water table depths in the watershed. The available data was used to describe the spatio-temporal variability of water table depths with a linear coregionalization model. Accounting for uncertainty in this modelling prevents for overestimation or underestimation of water resources for agriculture. Characterization of the natural discharge helps in preventing water shortage and improving land use for agricultural and ecological purposes.

If more information were available in the spatial and time domains, different geostatistical methods could be applied. Space-time variograms (Kiriakidis & Journel, 1999; Heuvelink & Webster, 2000; Bechini et al., 2000; Snepvangers et al., 2003) can account for spatial and temporal correlation among time steps but need a large amount of data and are difficult to adjust. An important disadvantage of the purely geostatistical approaches to spatio-temporal modelling is that it is not easy to incorporate physically based knowledge in the model (Heuvelink & Webster, 2001).

Geostatistical approaches make use of assumptions like stationarity, in either space or time. Essentially, these methods are based on data, not necessarily incorporating physical knowledge into the models. When geostatistical methods are applied to describe the spatiotemporal variation of water properties over long periods (longer than we present in chapter 3, for example), it is difficult to hold such strong assumptions without incorporating any physical knowledge to describe the processes.

Also, with more available information from observation wells, some stratification could be performed for spatial prediction, isolating effects from the different hydrogeological domains presented in the Jardim River watershed.

One point that could be improved in future situations is the lack of validation points. Validation experiments are important sources of information in the applicability of the models in quantification of uncertainty. Simple statistical measures to evaluate model quality have been applied in several studies (Knotters, 2001; Mardikis et al., 2005; Leterme et al., 2007). We advise validation in studies for practical purposes. In our situation, a good design for a validation set should be the installation of validation wells covering all soil types, hydrogeological domains and land uses of the basin, to contrast with modelling results of the observation wells. The number of wells would be restricted to the budget to drill new wells.

6.3.2 Time series modelling

The regionalized time series models that were applied in this study were devised for studies at regional scale where available information is scarce (Knotters, 2001). Stochastic methods were applied to account for uncertainty in the modelling of water table depths. Transfer function models are in general considered as "black box models", because they only describe the empirical relationship between precipitation surplus and water table depths without

physical knowledge about the process. Models are considered "white box models" when the models base on specific knowledge of the physical system (Knotters, 2001). The PIRFICT-model can be considered a middle term between these models, a "grey box model". The choice for the PIRFICT model is motivated by finding an optimum between purely physical-deterministic models which need a lot of inputs and with which quantification of uncertainty is complex, and purely data-based black-box models that describe statistical relationships but make no physical sense. Models like PIRFICT cover the most important physical processes, and enable us to quantify uncertainty.

The PIRFICT-model is based on the impulse response function, reflecting the physical relationship between precipitation surplus and water table depth, and calibrated on observed data. This linear relationship is only an approximation. Threshold nonlinearities in transfer functions arise from different soil layers and drainage levels. The results show the different responses of the hydrological systems in the Jardim River watershed as well the long memories of the system in some areas. The modelling reveals systematic changes in water table depth due to human interventions in the watershed and recharge of the aquifer over the monitoring period.

Time series modelling and geostatistical methodology were combined to describe the spatio-temporal dynamics of water table depth. The use of universal kriging enables us to incorporate ancillary information from digital elevation models and land use in the spatial prediction models, improving the estimations by reducing the variance on the spatial model.

6.3.3 Monitoring water table depths

For a monitoring period of 1240 days, model calibrations were better, model parameters better adjusted and more significant than for a monitoring period of 908 or 1092 days (Manzione et al., 2006; Manzione et al., 2007b). Long-term

116

monitoring confirmed the hypothesis about the long memories of the water system. The mapping of the linear trend parameter of the PIRFICT-model (presented on Chapter 4) showed aquifer recharge at 2006 after the wet years of 2001, 2002 and 2003.

The improvements in the model results achieved by using a longer time series express the dependence of the PIRFICT-model on data input. Larger improvements can even be expected, if a more reliable local climatological monitoring in the Jardim River watershed were available. Therefore, installation of a climate station inside the basin is recommended. Another suggestion is to unify farmers located inside the basin to measure and record local climatological events and report them to public institutions. To obtain more reliable information, we recommend that monitoring of water table depth continues.

6.3.4 Estimating characteristics of water table depths for risk analyses and water management

Estimating characteristics of water table depths through stochastic models representing the prevailing climatic and hydrological conditions enabled risk assessment to support water management. These characteristics are, in this study, water table depths that are exceeded with a certain probability at a crucial day in the season, probabilities of exceedance, temporal trends and expected water table depths. We express the risks by probabilities, not involving costs on the analysis. We detected areas and delimitated risk areas of water table depths for the fixed limits. The extension of these analyses to other dates and with different limits can be applied in water policy, water resources management and water planning. The results support decision making and should be used in maintaining productivity in the agricultural crops, in natural resources preservation and in poverty mitigation in areas affected by seasonal climatic effects. However, it should be emphasized that the uncertainty is

117

described by models. Parameter uncertainty or model uncertainty (arising from the choice for a specific model) are not accounted for. Therefore, validation sets are needed to evaluate and, if necessary, to improve the uncertainty quantification.

In the simulation experiment we have decided to first calculate the model for each well and later interpolate the outputs. This question was addressed at Stein et al. (1991) and Bosma et al. (1994) as interpolate first/calculate later (IC) or calculate first/interpolate later (CI). Heuvelink & Pebesma (1999) recommend IC procedure as a preferable choice because it can be a more efficient way to characterize the spatial distribution of individual inputs. The authors argue that the kriging step of CI does not fully exploit all available information. There is not an agreement about this topic. Stein et al. (1991) found that CI performs better than IC to simulate soil moisture. Heuvelink & Pebesma (1999) find the opposite in a case study using a linear pedotransfer function. Sinowski et al. (1997) present regionalized soil retention curves in which IC performed better than CI. In IC procedures, all model inputs need to be interpolated, what can increase errors in model application (Bechini et al., 2000). Bechini et al. (2003) present that even if less powerful in theory, CI performed better than IC for simulating soil physical properties for irrigation, when comparing the results with an independent data set. Leterme et al. (2007) found higher values for pesticide leaching using CI than using IC, but the spatial patterns did not change significantly.

In our case, the model parameters did not show a strong spatial structure. The IC procedure should be applied only if high quality and reliable spatial and temporal structures are identified for model inputs (Bechini et al., 2000). The PIRFICT-model is calibrated for each well to improve the local estimations. The impulse-response function in the PIRFICT-model reflects the physical relationship between the input on precipitation surplus and the output on water table depth. The impulse-response function works as a filter, reconstructing the

time series of water table depths from the inputs. The uncertainty about the true water table depths is described by the stochastic component of the PIRFICT-model. This uncertainty arises from the schematization of the model, measurement errors etc. Also, individual information about the local system is aggregated to the prediction. The physical bases are improving the predictions. Here, the choice for calculate the model first and then interpolate is a better procedure to explore the physical bases of the time series model. Alternatively, if the model parameters present strong spatial correlation, interpolate them and then calculate the models should bring some benefits.

The PIRFICT-model has five parameters, what can propagate errors arising from parameter uncertainty. In some wells, the parameters of the PIRFICT-model were still not reasonably calibrated, due to the long memories of the hydrological system and the relatively short observed time series of water table depths. Also, calculating the simulations was a computational intense task. Leterme et al. (2007) explained how IC could have much lower computational costs. IC procedures should perform better for nonlinear models (Bechini et al., 2000). For linear relations between model inputs and output, IC and CI should produce similar estimations (Heuvelink & Pebesma, 1999)). The PIRFICT-model describes the linear relationship between precipitation surplus and water table depths. This relationship is established by an IR function, which represent the memory of the system. Finding out if IC and CI perform similar to simulations of the PIRFICT-model is an interesting topic for further analysis. To explore all available information about the processes under study, we opted for universal kriging as a method to incorporate ancillary information as a drift in the spatial model. This drift was a digital elevation model with plenty information for the whole area.

6.3.5 Uncertainty measures

Various sources of uncertainty can be distinguished: data (observed water table depths, climatic database, image classification, digital elevation model), uncertainty associated with time series modelling and with the model of spatial variation. Uncertainty is inherent on data and on process modelling. Policy makers should consider uncertainty, probabilities and risks in plans for water management. Stochastic methods, such as applied in this study, enable to account for uncertainty in decision making.

Utilisation of statistical information on uncertainty in the practice of water management is a big challenge. Reducing uncertainty and optimizing monitoring strategies are part of a successful plan for water management, but the costs involved should be considered. A balance should be found between efforts to reduce uncertainty, such as collecting more data and applying more complex models, and the risks which can be avoided by these efforts.

6.4 Some topics for future work

Interesting topics for future research arise from the discussions in the previous sections. It would be interesting to continue monitoring water table depths to analyze longer time series and obtain better calibrations of the models. It can lead to improve estimates of the spatial structure of model parameters. IC methods can produce different results from those presented in Chapter 5. The results of IC and CI should be compared by validation. The behaviour of IR functions adjusted by the PIRFICT-model for each grid cell through an IC procedure is a still unknown question. Modelling procedures which enable to interpolate the PIRFICT-model parameters before calculates and simulates water table depths can integrate these questions (Carneiro, 2006).

Integration of models that describe the spatio-temporal variation of water table depths with weather forecast methods should enable the development of warning systems. These systems predict future water table depths, or statistical

characteristics describing future fluctuations of water table depths, from a simulated input on weather. For instance, such a system can inform us when some critical level will be reached. From this point of view, as described at De Gruijter et al. (2006), the water table depths monitoring would involve not only the status of the water system (status monitoring) and check for trends presented in the universe (trend monitoring) as was demonstrated in this thesis, but also decide whether the universe satisfies regulatory conditions (compliance monitoring). Since 2000 the European community is following the precautionary principle (European Commission, 2000) to establish standards on water use. When these standards about water levels, pollution, availability and other risk situations are met, actions are taken to the burden of proof falls on those who would advocate taking the interventions in the hydrological system. This regulatory information could substantially improve water management and prevent for risks related to water table depths, at local and regional scales. Warming systems could alarm when these standards are met.

REFERENCES

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical functions**. New York: Dover Publications Inc, 1965. 1046p.

ASSAD, E. D.; SANO, E. E.; MASUTOMO, R.; CASTRO, L. H. R.; SILVA, F. A. M. Veranico na região dos cerrados: Freqüência e probabilidade de ocorrência. **Pesquisa Agropecuária Brasileira**, Brasília, v. 28, n. 9, p. 993-1003, 1993. (In Portuguese).

ASSAD, E. D. **Chuva no Cerrado: Análise e espacialização.** Brasília: Embrapa Cerrados, 1994. 423p.

BECHINI, L.; DUCCO, G.; DONATELLI, M.; STEIN, A. Modelling interpolation and stochastic simulation in space and time for global solar radiation. **Agriculture, Ecosystems and Environment**, v.81, p.29-42, 2000.

BECHINI, L.; BOCCHI, S.; MAGGIORE, T. Spatial interpolation of soil physical properties for irrigation planning. A simulation study in northern Italy. **European Journal of Agronomy**, v.19, p.1-14, 2003.

BODO, B. A.; UNNY, T. E. On the outputs of the stochsticized Nash-Dooge linear reservoir cascade. In: MACNEILL, I. B.; UMPHREY, G. J. (Eds.). **Stochastic Hydrology**. Dordretch: D. Reidel Publishing Company, 1987. p. 131-147.

BOSMA, W.; MARINUSSEN, M.; VAN DER ZEE, S. Simulation and areal interpolation of reactive solute transport. **Geoderma**, v.62, p.217-231, 1994.

BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. 2.ed. San Francisco: Holden-Day, 1976. 575p.

CAMPOS, J. E. G.; FREITAS-SILVA, F. H. Hidrogeologia do Distrito Federal. In: **Inventário hidrogeologico e dos recursos hídricos superficiais do Distrito Federal.** Brasília: IEMA/SEMATEC/UnB, 1998, 85 p.

CARNEIRO, T. G. S. **Nested-CA: a foundation for multiscale modelling of land use and land cover change**. 2006. 114 p. Thesis (PhD in Computer science) – Brazilian Institute for Spatial Research (INPE), São José dos Campos, 2006.

CASTRIGNANÒ, A.; GIUGLIARINI, L.; RISALITI, R.; MARTINELLI, N. Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate geostatistics. **Geoderma**, v.97, p.39-60, 2000.

CHRISTAKOS, G.; RAGHU, R. Dynamic stochastic estimation of physical variables. **Mathematical Geology**, v.28, p.341-365, 1996.

COSTA, W. D. Uso e gestão de água subterrânea. In: FEITOSA, F. A. C.; MANOEL FILHO, J. (Ed.) **Hidrogeologia:** conceitos e aplicações. Fortaleza: CPRM, 2000. p.314-367.

DE GRUIJTER, J. J.; BRUS, D. J.; BIERKENS, M. F. P.; KNOTTERS, M. **Sampling for natural resources monitoring**. Berlin: Springer-Verlag, 2006. 332p.

DOLABELLA, R. H. C. **Caracterização agroambiental e avaliação da demanda e disponibilidade dos recursos hídricos para agricultura irrigada na bacia hidrográfica do Rio Jardim-DF.** 1996. 109 p. Dissertation (Master in Agronomy) - University of Brasília (UnB), Brasília, 1996.

FURLEY, P. A. The nature and diversity of neotropical savannah vegetation with particular reference to the Brazilian Cerrados. **Global Ecology and Biogeography**, v. 8, p. 223-241, 1999.

GARDINER, C. W. **Handbook of stochastic methods: for physics, chemistry and the natural sciences.** 3. ed. Springer series in Synergeticss, v.13. New York: Springer, 2004. 415p.

GOEDERT, W. Região do Cerrado: Potencial agrícola e politica para seu desenvolvimento. **Pesquisa Agropecuária Brasileira**, v. 24, p. 1-17, 1989. (In Portuguese)

GOMES-LOEBMANN, D.; GUIMARÃES, R. F.; BETTIOL, G. M.; FREITAS, L. F.; REDIVO, A. L.; CARVALHO JUNIOR, O. A. Mistura espectral de imagens LANDSAT para análise multitemporal de uso da terra nas diferentes unidades pedológicas da bacia do Rio Jardim, DF. In: BRAZILIAN SYMPOSIUM ON REMOTE SENSING, 12, 2005, Goiânia. **Proceedings…** São José dos Campos: INPE, 2005. p. 557-564. (In Portuguese) Available at: <marte.dpi.inpe.br/rep-/ltid.inpe.br/sbsr/2004/11.21.20.01>. Access at: 06 Feb. 2007.

GOOVAERTS, P. Factorial kriging analysis: a useful tool for exploring the structure of multivariate spatial soil information. **Journal of Soil Science**. v.43, p.597-619, 1992.

GOOVAERTS, P. Geostatistical modelling of spatial uncertainty using p-field simulation with conditional probability fields. **International Journal of Geographical Information Science**, v.16, p.167-178, 2002.

GOOVAERTS, P. **Geostatistics for natural resources evaluation.** New York: Oxford University Press, 1997. 483p.

GOOVAERTS, P., SONNET, P. Study of spatial temporal variation of hydrogeochemical variables using factorial kriging analysis. In: SOARES, A. (Ed.) **Geostatistical Tróia 92**. Dordrecht: Kluwer, 1993. v.2, p.745-756.

GOULARD, M.; VOLTZ, M. Linear coregionalization model: a toll for estimation and choice of cross-variogram matrix. **Mathematical Geology**, v.24, p.269-286, 1992.

HENGL, T.; HEUVELINK, G. B. M.; STEIN, A. A generic framework for spatial prediction of soil properties based on regression-kriging. **Geoderma**, v.120, p.75-93, 2004.

HEUVELINK, G. B. M.; WEBSTER, R. Modelling soil variation: past, present and future. **Geoderma**, v.100, p.269-301, 2001.

HEUVELINK, G. B. M.; PEBERSMA, E. Spatial aggregation and soil process modelling. **Geoderma**, v.89, p.47-65, 1999.

HIPEL, K. W.; MCLEOD, A. I. **Time series modelling of water resources and environmental systems**. Amsterdam: Elsevier, 1994. 1013p.

HOFFMANN, W. A.; JACKSON, R. B. Vegetation-climate feedbacks in the conversion of tropical savanna to grassland. **Journal of Climate**, v.13, p.1593-1602, 2000.

JEPSON, W. A disappearing bioma? Reconsidering land-cover change in the Brazilian savannah. **Geographical Journal**, v. 2, p. 99-111, 2005.

KLINK, C. A.; MACHADO, R. B. Conservation of the Brazilian Cerrado. **Conservation Biology**, v. 19, p. 707-713, 2005.

KLINK, C. A.; MOREIRA, A. G. Past and current human occupation and land-use. In: OLIVEIRA, P. S.; MARQUIS, R. J. (Eds.). **The Cerrado of Brazil: Ecology and Natural History of a Neotropical Savannah**. New York: Columbia University Press, 2002. p. 69-88.

KNOTTERS, M. **Regionalised time series models for water table depths**. 2001. 152p. Thesis (PhD in Hydrology) - Wageningen University (WUR), Wageningen, 2001.

KNOTTERS, M.; BIERKENS, M. F. P. Physical basis of time series models for water table depths. **Water Resources Research**, v. 36, p.181-188, 2000.

KNOTTERS, M.; BIERKENS, M. F. P. Predicting water table depths in space and time using a regioinalised time series model. **Geoderma**, v. 103, p.51-77, 2001.

KNOTTERS, M.; BRUS, D. J.; OUDE VOSHAAR, J. H. A comparation of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizont depth with censored observations. **Geoderma**, v.67, p.227-246, 1995.

KNOTTERS, M.; VAN WALSUM, P. E. V. Estimating fluctuation quantities from time series of water-table depths using models with a stochastic component. **Journal of Hydrology**, v.197, p.25-46, 1997.

KYRIAKIDIS, C. P.; JOURNEL, A. Geostatistical time models: a review. **Mathematical Geology**, v.31, p.651-684, 1999.

LEDUC, C.; BROMLEY, J.; SCHROETER, P. Water table fluctuation and recharge in semi-arid climate: some results of the HAPEX-Sahel hydrodynamic survey (Niger). **Journal of Hydrology**, v. 188-189, p. 123-138, 1997.

LETERME, B.; VANCLOOSTER, M.; VAN DER LINDEN, A. M. A.; TIKTAK, A.; ROUNSEVELL, M. D. A. The consequences of interpolating or calculating first on the simulation of pesticide leaching at the regional scale. **Geoderma**, v.137, p.414-425, 2007.

LOPES, A. S. Soils under Cerrado: A success story in soil management. **Better Crops**, v. 10, n. 2, p. 9-13, 1996.

LOUSADA, E. O**. Estudos hidrogeologicos e isotópicos no Distrito Federal: modelos de fluxo conceituais**. 2005. 127 p. Thesis (PhD in Geosciences) - University of Brasília (UnB), Brasília, 2005. (In Portuguese)

MACHADO, R. B.; RAMOS NETO, M. B.; PEREIRA, P.; CALDAS, E.; GONÇALVES, D.; SANTOS, N.; TABOR, K.; STEININGER, M. **Estimativas de perda de área no Cerrado brasileiro**. Brasília: Conservation International Brazil, 2004. 28p. (In Portuguese) Available at: <http://www.conservation.org.br/arquivos/RelatDesmatamCerrado.pdf>. Access at: 06 Feb. 2007.

MANZIONE, R. L.; KNOTTERS, M.; HEUVELINK, G. M. B. Mapping trends in water table depths in a Brazilian cerrado area. In: CAETANO, M.; PAINHO, M. (Eds.) **Procedings of Accuracy 2006: 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Enviromental Sciences**. Lisboa: Instituto Geográfico Português, 2006. p.449-458. Available at: <http://www.spatial-accuracy.org/2006/PDF/Manzione2006accuracy.pdf >. Access at: 15 Jul. 2007.

MANZIONE, R. L.; KNOTTERS, M.; HEUVELINK, G. M. B. Incorporation of ancillary information derived from satellite images applied on environmental variables evaluation. In: BRAZILIAN SYMPOSIUM ON REMOTE SENSING, 13, 2007a, Florianópolis. **Proceedings…** São José dos Campos: INPE, 2007a. p. 3437-3444. Available at: <http://marte.dpi.inpe.br/col/dpi.inpe.br/sbsr@80/2006/11.14.22.15/doc/3437-3444.pdf>. Accessed at: 15 Aug. 2007.

MANZIONE, R. L.; KNOTTERS, M.; HEUVELINK, G. M. B.; VON ASMUTH, J. R.; CÂMARA, G. Predictive risk mapping of water table depths in a Brazilian Cerrado area. In: INTERNATIONAL SYMPOSIUM ON SPATIAL DATA QUALITY, 5, 2007b, Enschede. **Proceedings…** Enschede: ITC, 2007b. (CDROM). Available at: <http://www.itc.nl/issdq2007/proceedings/session%202%20Spatial%20Statistics/Paper%20Manzione.pdf>. Access at: 15 Jul. 2007.

MARDIKIS, M. G.; KALIVAS, D. P.; KOLLIAS, V. J. Comparison of interpolation methods for the prediction of reference evapotranspiration – An application in Greece. **Water resources management**, v.19, p.251-278, 2005.

MATHERON, G. **Le krigeage universel. Cachiers du Centre de Morphologie Mathematique**, v. 1. Fontainebleau: Ecole des Mines de Paris, 1969. 83p.

MISTRY, J. Fire in the Cerrado (savannas) of Brazil: an ecological review. **Progress in Physical Geography**, v. 22, n. 4, p. 425-448, 1998.

MYERS, N. Threatened biotas: 'hotspots' in tropical forests. **Environmentalist**, v. 8, p. 187-208, 1988.

MYERS, N. The biodiversity challenge: expand hotspots analysis. **Environmentalist**, v. 10, p. 243-256, 1990.

NASH, J. E. Determining runoff from rainfall. **Proc. Inst. Civ. Eng.**, v.10, p.163-184, 1958.

ODEH, I.; MCBRATNEY, A.; CHITTLEBOROUGH, D. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. **Geoderma**, v. 63, p.197-214, 1994.

OLIVEIRA, R. S.; BEZERRA, L.; DAVIDSON, E. A.; PINTO, F.; KLINK, C. A.; NEPSTAD, D. C.; MOREIRA A. Deep root function in soil water dynamics in cerrado savannas of central Brazil. **Functional Ecology**, v. 19, p. 574-581, 2005.

PAPRITZ, A.; FLÜHLER, H. Temporal change of spatially autocorrelated soil properties: Optimal estimation by cokriging. **Geoderma**, v.62, p.29–43, 1994.

PARKIN, T. B.; ROBINSON, J. A. Analysis of lognormal data. **Advances in Soil Science**, v.20, p.191-235, 1992.

PEBESMA, E. J. Multivariable geostatistics in S: the Gstat package. **Computers & Geosciences**, v.30, p.683-691, 2004.

REATTO, A.; CORREIA, J. R.; CERA, S. T.; CHAGAS, C. S.; MARTINS, E. S.; ANDAHUR, J. P.; GODOY, M. J. S.; ASSAD, M. L. C. L. **Levantamento semidetalhado dos solos da Bacia do Rio Jardim_DF, escala 1:50000.** Planaltina: Embrapa Cerrados, 2000, 63 p. CD-ROM (Embrapa Cerrados Research Bulletin, 18) – (In Portuguese)

RESENDE, M.; CURI, N.; REZENDE, S. B.; CORRÊA, G. F. **Pedologia: Base para distinção de ambientes.** 4.ed. Viçosa: NEPUT, 2002. 339p. (In Portuguese)

ROUHANI, S.; WACKERNAGEL, H. Multivariate geostatistical approach to space-time data analysis. **Water resources research**, v.26, p.585-591, 1990.

SALAS, J. D.; PIELKE SR., R. A. Stochastic characteristics and modelling of hydroclimatic processes. In: POTER, T. D.; COLMAN, B. R. (Eds.) **Handbook of weather, climate and water: Atmospheric chemistry, hydrology and social impacts**. New York: John Wiley & Sons, 2003. p.587-605.

SINOWSKI, W.; SCHEINOST, A.; AUERSWALD, K. Regionalization of soil water retention curves in a highly variable landscape, II. Comparison of regionalization procedures using a pedotransfer function. **Geoderma**, v.78, p.145-159, 1997.

SNEPVANGERS, J. J. J. C.; HEUVELINK, G. B. M.; HUISMAN, J. A. Soil water content interpolation using spatio-temporal kriging with external drift. **Geoderma**, v.112, p.253-271, 2003.

STEIN, A.; STARITSKY, I.; BOUMA, J.; VAN EIJNSBERGEN, A.; BREGT, A. Simulation of moisture deficits and areal interpolation by universal cokriging. **Water resources research**, v.27, n.8, p.1963-1973, 1991.

SUN, G.; REIKERK, H.; KORNHAK, L. V. Ground-water-table rise after harvesting on Cypress-Pine flatwoods in Florida. **Wetlands**, v. 20, n. 1, p. 101-112, 2000.

TANCO, R.; KRUSE, E. Prediction of seasonal water-table fluctuations in La Pampa and Buenos Aires, Argentina. **Hydrogeology Journal**, v. 9, p. 339-347, 2001.

TANKERSLEY, C. D.; GRAHAM, W. D. Development of an optimal control system for maintaining minimum groundwater levels. **Water Resources Research**, v. 30, p. 3171-3181, 1994.

UHLENBECK, G. E.; ORNSTEIN, L. S. On the theory of Brownian motion. **Physical Review**, v. 36, p. 823-841, 1930.

VAN GEER, F. C.; ZUUR, A. F. An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series. **Journal of Hydrology**, v. 192, p. 65-80, 1997.

VON ASMUTH, J. R.; BIERKENS, M. F. P.; MAAS, C. Transfer function noise modelling in continuous time using predefined impulse response functions. **Water Resources Research**, v. 38 (12), p. 23.1-23.12, 2002.

VON ASMUTH, J. R.; MAAS, C. The method of impulse response moments: a new method integrating time series, groundwater and eco-hydrological modelling. In: GEHERLS, J. C.; PETERS, N. E.; HOEHN, E.; JENSEN, K.; LEIBUNDGUT, C.; GRIFFIOEN, J.; WEBB, B.; ZAADNOORDIJK, W. J. (Eds.) **Impact of human activity on groundwater dynamics**. Wallingford: IAHS Publication, 2001. p.51-58.

VON ASMUTH, J. R.; BIERKENS, M. F. P.; MAAS, C. Transfer function noise modelling in continuous time using predefined impulse response functions. **Water Resources Research**, v. 38 (12), p. 23.1-23.12, 2002.

VON ASMUTH, J. R.; KNOTTERS, M. Characterising groundwater dynamics based on a system identification approach. **Journal of Hydrology**, v. 296, p. 118-34, 2004.

VON ASMUTH, J. R.; BIERKENS, M. F. P. Modelling irregularly spaced residual series as a continuous stochastic process. **Water Resources Research**, 41, p. W12404, 2005.

ZIEMER, R. E.; TRANTER, W. H.; FANNIN, D. R. **Signals and systems: continuous and discrete**. Upper Saddle River: Prentice-Hall, 1998. 622p.

ANNEX A – DESCRIPTION OF OBSERVATION WELLS LOCATIONS CONSIDERING SOIL TYPE, HYDROGEOLOGICAL DOMAINS AND ACTUAL LAND USE

Wells Number	Soil Types	Hydrological Domains Porous	Fractured	Actual Land Use
W12	LV	P1	R3/Q3	AG
W56	LVA	P1	R3/Q3	AG
W50	LVA	P1	R3/Q3	AG
W57	LVA	P1	R3/Q3	AG
W11	LVA	P1	R3/Q3	AG/P
W55	LVA	P1	R3/Q3	AG
W46	Cx	P4	R4	AG/P/CV
W13	LV	P2	R4	AG
W47	LVA	P2	R4	CV
W04	Cx	P2	R4	P
W03	Fx	P2	R4	CV
W16	LV	P2	R4	AG
W52	LVA	P2	Bambuí	AG
W17	LV	P2	R4	AG
W49	LVA	P2	R4	AG
W37	Cx	P2	R4	P
W36	Cx	P4	Canastra	P
W38	Cx	P4	Bambuí	CV
W18	LV	P2	Bambuí	AG
W33	LV	P2	Bambuí	AG
W25	LV	P2	Bambuí	AG
W26	Cx	P4	Bambuí	CV/AG
W27	Gx	P2	Bambuí	CV/AG
W24	LV	P2	Bambuí	AG/P
W40	Cx	P2	Bambuí	P
W10	LVA	P2	Bambuí	P/AG
W20	LV	P2	Canastra	P/AG
W44	Gx	P2	Canastra	AG/CV
W22	LV	P2	Canastra	AG/CV
W42	Gx	P2	Canastra	AG/CV
W05	LVA	P2	Canastra	AG
W61	LVA	P2	Canastra	AG/P/CV
W41	Cx	P4	Canastra	AG
W59	LVA	P2	Canastra	AG/P/CV
W06	LVA	P2	Canastra	AG/P
W60	Cx	P4	Canastra	P/AG
W30	LV	P2	Canastra	AG/P/CV
W54	LVA	P2	Canastra	CV/P
W08	LVA	P2	Canastra	P/AG
W31	LV	P2	Canastra	AG

Legend: LV-Red Latosol, LVA-Red Yellow Latosol, Cx-Haplic Cambisol, Gx-Haplic Gleysol, Fx-Haplic Plinthosol, AG-Agriculture, P-Pasture, CV-Cerrado Vegetation.

ANNEX B – FLOWCHART OF THE PRESENTED METHODOLOGY FOR WATER TABLE MODELING IN THE CERRADOS
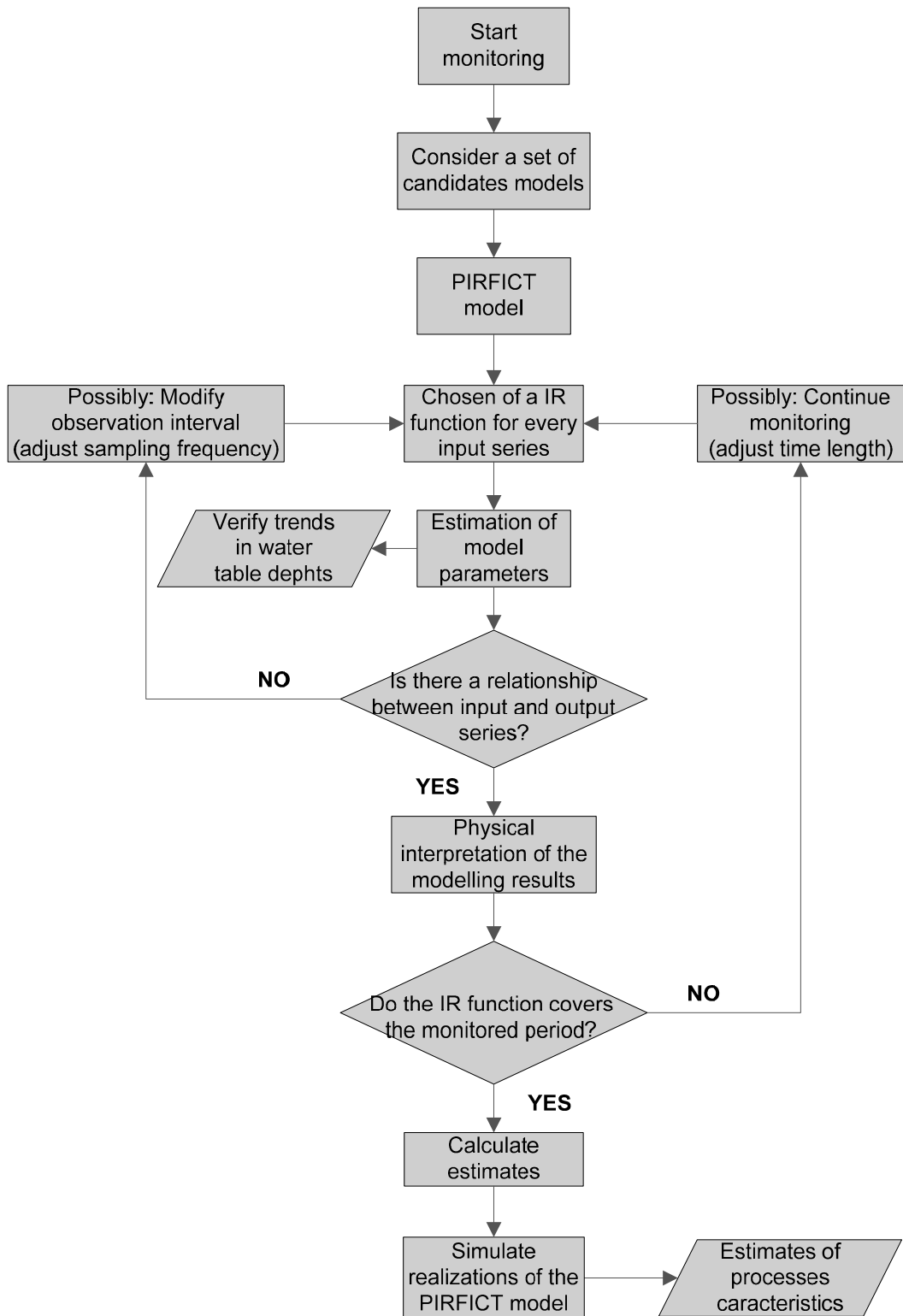
ANNEX C - IMPLEMENTED CODE FOR PIRFICT-MODEL SIMULATION AND ESTIMATION OF WATER TABLE CHARACTERISTICS

```matlab
%Stochastic Simulation of water table depths from Menyanthes export files
%
%
%load Menyanthes export file
load W01.mat%name of the file
%
%
% function [hsim,ssim,pint,d1,error] = simulate(m,simdates,hini,usethresh,stoch)
%
% output variables:
%  hsim      = simulated groundwater level [dates values]
%  ssim      = simulated noise [dates values]
%  pint      = 95% prediction interval
%  d1        = regional drainage level in non-linear model
%  error     = possible error message
% input variables:
%  m         = structure with model data
%  simdates  = column vector with dates for which the simulation should be calculated
%  hini      = initial groundwater level for non-linear model
%  usetresh  = boolean defining if the threshold is used in non-linear models
%  stoch     = boolean defining if the simulation is stochastic or not
%
%
%define simulated dates interval(31/01/1974-06/08/20006) in numbers as Menyanthes export
D=[721932:732895]'
%
%
%number of simulations
N=1000;
%
%
%loop for run the simulate function N times
for i=1:N
   hsim(:,i)=simulate(M,D,0,0,1);
end
%
%remove warming up period to let the hsim matrix with 30 years
H30=hsim(913:11876,1:1000)
%sampling the new simulation matrix H30
Day='01'
Month='Oct'
count1=0;
s=datestr(D,1);
for j=1:length(D)
   if (min(s(j,1:2)==Day>0.5))&(min(s(j,4:6)==Month)>0.5)
      count1=count1+1
      Daux(count1)=D(j);
      Haux(count1,:)=H30(j,:);
   end
end
```

```
%Calculation of Water Table characteristics after Stochastic simulation
%
%
%calculate percentiles
Hdef=reshape(Haux,size(Haux,1)*size(Haux,2),1)
X=sort(Hdef);
L=length(Hdef);
Phat=((1:L)-0.5)/L;
P=[0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95];
Qhat=interp1(Phat,X,P)
%
%Plot results
%Normal Probability Plot
normplot(Hdef)
%Frequecy Histogram
[E,Y]=hist(Hdef,15)
bar(Y,E,1)
title('Histogram of Simulated WTD for Oct 1st')
ylabel('Frequency')
xlabel('WTD (m)')
%
%
%Verify shallow condition (# of events)
count2=zeros(1,size(H30,2));
for j=1:length(D)
    cond1=H30(j,:)>-0.5;%upper limit
    count2=count2+cond1;
end
PSH=sum(count2)/(j*N);%probability of event during the simulated period
%
%
%Verify dry well condition (# of events)
count3=zeros(1,size(H30,2));
for j=1:length(D)
    cond2=H30(j,:)<-10;%lower limit
    count3=count3+cond2;
end
PDR=sum(count3)/(j*N);%probability of event during the simulated period
%
%
%Verify consecutive days for some event
%Dry wells
count4=zeros(1,size(H30,2));
MW=10; %moving window (size=#days for specified event)
for j=1:size(H30,1)-MW+1
    cond3=H30(j:j+MW-1,:)>-0.5 %verify if the indicated condition is satisfyed
    soma3(j,:)=sum(cond3)
    aux1=soma3(j,:)==MW
    aux2=aux1*0
    if j>1
        aux2=soma3(j,:)==soma3(j-1,:)&soma3(j,:)==MW
    end
    count4=count4+aux1-aux2
end
%
```

```
%Shallow water table
count5=zeros(1,size(H30,2));
MW=10; %moving window (size=#days for specified event)
for j=1:size(H30,1)-MW+1
    cond4=H30(j:j+MW-1,:)<-10 %verify if the indicated condition is satisfyed
    soma4(j,:)=sum(cond4)
    aux3=soma4(j,:)==MW
    aux4=aux3*0
    if j>1
        aux4=soma4(j,:)==soma4(j-1,:)&soma4(j,:)==MW
    end
    count5=count5+aux3-aux4
end
```