7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Edited by M. Caetano and M. Painho.

Mapping trends in water table depths in a Brazilian Cerrado area

Rodrigo Lilla Manzione[1], Martin Knotters[2] and Gerard B. M. Heuvelink[2]

1 INPE (National Institute for Spatial Research) – DPI (Image Processing Division)
Avenida dos Astronautas, 1758. Caixa Postal 515, 12227-010, São José dos Campos, SP, Brazil
Tel.: + 055 12 39456481
manzione@dpi.inpe.br

2 Alterra – Soil Science Centre
Droevendaalsesteeg 3, 6708PB Wageningen, The Netherlands
Tel.: + 031 317 474240
Martin.Knotters@wur.nl, Gerard.Heuvelink@wur.nl

Abstract

The Cerrado region is the most extensive woodland-savannah in South America, situated at the central Brazilian Plateau and characterized by wet and dry periods well defined during the year. During the past 30 years, the original vegetation has been replaced by extensive cattle fields and agricultural crops, which are less adapted to drought than the Cerrado vegetation. Therefore, irrigation is increasingly applied, resulting in changes of the hydrological system. The aim of this study is to map systematic changes of the water table depths, in order to indicate areas with potential risks of future water shortage. To this purpose the PIRFICT-model is applied, a transfer function-noise (TFN) model with a Predefined Impulse Response Function In Continuous Time. Being the most important driving forces of water table fluctuation, precipitation and evapotranspiration are incorporated as exogenous variables into the model. Besides, a linear trend component is incorporated, reflecting systematic changes of water table depths over time. The linear trend parameter of the time series model is interpolated spatially using universal kriging, by utilizing actual land use derived from Landsat images as ancillary information. Series of 30 months length of semi-monthly observed water table depths are available from 40 wells in the Jardim river watershed. In this area almost all natural Cerrado vegetation has been replaced by agricultural crops, some of which are intensively irrigated. The time series models are calibrated to the 40 series, and next the trend parameter reflecting systematic changes of water table depths is mapped. The resulting map indicates potential risks of water shortage. The kriged trend parameter is evaluated by cross-validation. The significance of the interpolated trend parameter is also mapped. The uncertainties associated with the mapped trend parameter are large, suggesting that longer time series of water table depths are needed to obtain more accurate results.

Keywords: groundwater levels, phreatic levels, time series modeling, spatio-temporal modeling, trend analysis

1 Introduction

The Cerrado region is the most extensive woodland-savannah in South America. Situated at the central Brazilian Plateau, it covers 22% of the country or approximately 1.783 million km2 (Jepson, 2005). The Cerrado receives abundant rainfall (between 1100 and 1600 mm/yr), concentrated in a period of six to seven months between October and April. The rest of the year is characterized by a pronounced dry season. Hence, natural vegetation is well adapted to drought, by deep root systems that reach the water storage in deep soils present in the area.

During the past 30 years, the original Cerrado vegetation has been replaced by extensive cattle fields and agricultural crops (mainly corn, soybeans and cotton), becoming Brazil's most important grain belt and facing deforestation rates much higher than in the Amazon rainforest (Oliveira et al., 2005). A complex wood/grass ecosystem was substituted by shallow-rooted monocultures, which are less well adapted to drought. Their need of water supply by irrigation techniques is likely to change the hydrological system (Klink and Moreira, 2000). With irrigation increasing, lowering of the water table occurs, and hence risks of water shortage appear. In order to optimize and balance the interest of economical and ecological land use purposes, knowledge about the spatio-temporal dynamics of the water table is very important (Von Asmuth and Knotters, 2004).

In hydrology, water table dynamics are modeled in several ways. To describe the dynamic relationship between precipitation and the water table depth, transfer function-noise (TFN) models have been applied (Box and Jenkins, 1976; Hipel and McLeod, 1994; Tankersley and Graham, 1994; Van Geer and Zuur, 1997). Basically, these models can be seen as multiple regression methods, 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 this case the 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, since it can be assumed that the 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.

To link the response characteristics of the water table system to the dynamic behavior of the input, Von Asmuth et al. (2002) presented a method based on the use of a transfer function-noise model in continuous time, the so-called PIRFICT-model. An important advantage of the PIRFICT model as compared to discrete-time TFN-models is that it can deal with input and output series which have different observation frequencies and irregular time intervals.

The aim of this study was to map the systematic changes of water table depths in a watershed located at the Brazilian Cerrados, in order to indicate areas with potential risks of future water shortage.

2 Materials and methods

2.1 Study area

The Jardim River watershed is a representative Cerrado area in the eastern part of the Brazilian D.C., latitudes $15^0 40^{'}$S and $16^0 02^{'}$S and longitudes $47^0 20^{'}$W and $47^0 40^{'}$W. The dry and the wet season are very well defined, with the rain concentrated between October and April. During the past years, almost all natural vegetation present in this area was replaced by agricultural crops, and the use of irrigation systems has substantially increased in this region during the past years. The main cultivations present in this area are soybeans, cotton and corn crops, as well as pasture and horticultural crops.

To monitor the water table depths, 40 wells were drilled in the area. The locations were selected purposively, in order to cover the range of soil types in the area (Lousada, 2005). The water table was observed semimonthly from October 2003 until April 2006, resulting in series of 60 more or less regularly spaced semi-monthly observations. Series of precipitation and potential evapotranspiration were available from a climate station close to the basin, from 1974 until 1996 with a monthly frequency, and from 1996 until April 2006 with a daily frequency. Figure 1 shows a map of the study area and the well locations.

Ancillary information related to the sources of systematic changes can be derived from Landsat images. An image from July 23, 2005 was 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 was created based on expert knowledge and manual delineation of the classes. The class Agricultural Crops includes all kinds of agricultural products cultivated in the area: small areas cultivated with horticultural products, such as carrots, lettuce, tomatoes, and big areas cultivated with products such as corn, soybeans, cotton, coffee, sugar cane. All these crops demand more water than the original vegetation. The agricultural activities 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. Figure 1 gives the classified land use map.

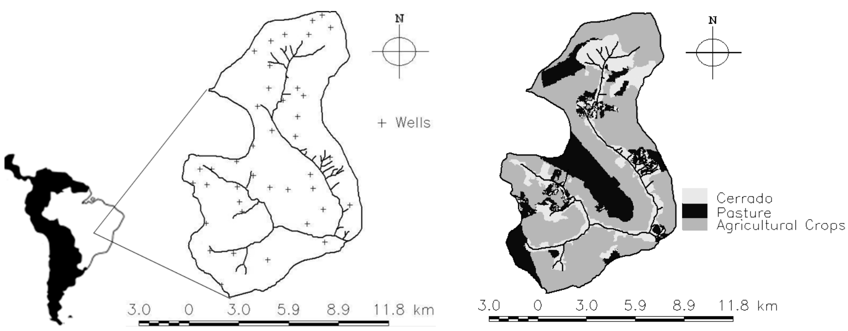

Figure 1 Jardim River watershed and location of the observation wells (+) (right) and the Image classification for actual land use (left).

The land use map shown in Figure 1 has sharp boundaries. For hydrological studies, this does not make sense because water levels do not have abrupt variations 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.

2.2 The PIRFICT-model

The behavior of linear input-output systems can be completely characterized by their impulse response (IR) function (Ziemer et al., 1998; Von Asmuth et al., 2002). The response of water table depth to impulses of precipitation series can be modeled by a transfer function-noise (TFN) model (Box and Jenkins, 1976; Hipel and McLeod, 1994; Von Asmuth and Knotters, 2004). For water table depths, the dynamic relationship between precipitation and water table depth can also be described using physical mechanistic groundwater flow models. However, by using 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 of the noise model is a series of independent and identically distributed disturbances with zero mean, and finite and constant variance, i.e., white noise. Figure 2 shows a scheme of the TFN model:

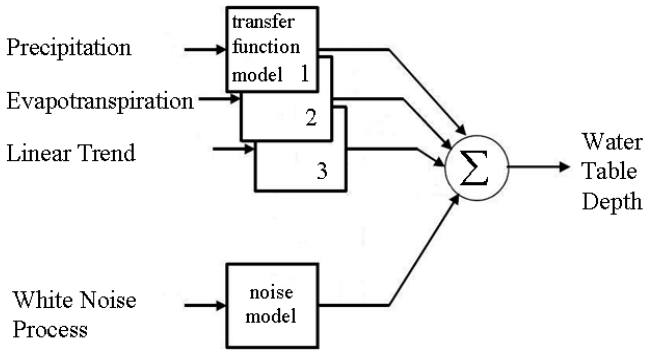

Figure 2 Schematic representation of the transfer function model with added noise for water table depths.

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/deficit. For the simple case of a linear, undisturbed phreatic system that is influenced by precipitation surplus/deficit only (Von Asmuth et al., 2002):

$$h(t) = h^*(t) + d + r(t) \tag{1}$$

$$h^*(t) = \int_{-\infty}^{t} p(\tau)\theta(t-\tau)\partial\tau \tag{2}$$

$$r(t) = \int_{-\infty}^{t} \phi(t-\tau)\partial W(\tau) \tag{3}$$

where:

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

$h^*(t)$ is the predicted water table depth at time *t* attributed to the precipitation surplus/deficit, 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/deficit 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:

$$d = \frac{\sum_{i=0}^{N} h(t_i)}{N} - \frac{\sum_{i=0}^{N} h^*(t_i)}{N} - \frac{\sum_{i=0}^{N} r(t_i)}{N} \tag{4}$$

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:

- iteratively - using correlation structures in the available data and model diagnostics.
- physically - based on insight into the behavior of the analyzed system.

Here, the second approach is followed. $\theta(t)$ is a Pearson type III distribution function (PIII df, Abramowitz and Stegun, 1964). Because of its flexible nature, this function adequately models the response of a broad range of groundwater systems. Under the assumption of 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:

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

$$\phi(t) = \sqrt{2\alpha\sigma_r^2}\, e^{-\alpha t} \tag{5}$$

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

Equation 5 and its parameters have a physical meaning that is described in Von Asmuth and 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). The parameter n denotes their number and a equals the inverse of the reservoir coefficient normally used. As Knotters and 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 the case of equation 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 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 with 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), while *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 and Knotters, 2004).

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 model systematic changes in the water table system.

2.3 Regionalizing the linear trend parameter of the time series model

PIRFICT-models were calibrated to the 40 series of water table depths, using the program Menyanthes. 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. This works as follows. Let the 'observed' trend parameters be denoted 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=40). 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):

$$\hat{z}(x_0) = \hat{m}(x_0) + \hat{e}(x_0) \tag{6}$$

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

$$\begin{aligned} \hat{z}(x_0) &= \sum_{k=0}^{p} \hat{\beta}_k \cdot q_k(x_0) + \sum_{i=1}^{n} w_i(x_0) \cdot e(x_i); \\ q_0(x_0) &= 1 \end{aligned} \tag{7}$$

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.

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

$$LTP(x_0) = \beta_0 + \beta_1 \cdot LU1(x_0) + \beta_2 \cdot LU2(x_0) + \beta_3 \cdot LU3(x_0) + e(x_0) \tag{8}$$

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 semivariogram.

3 Results and discussion

3.1 Time series modeling

Due to spatially varying hydrological conditions, a wide range of calibration results was found for the 40 observed wells. Table 1 summarizes the results of time series modeling.

Table 1 Summary of the results of PIRFICT-model calibration for 40 observation wells.

	Minimum	Maximum	Mean	Standard Deviation
R^2_{adj}	53.88	97.51	82.56	
RMSE	0.07	1.35	0.563	
A	51.26	25790	2326.35	1704.05
a	0	0.15	0.01	0.01
n	0.5	2.75	1.38	0.26
E	-4.62	2.75	0.21	1.49
LTP	-3.92	2.36	-0.15	1.40
α	6.08	12.36	9.84	13.88

R^2_{adj}=Explained Variance Percentage; *RMSE*=Root Mean Squared Error (meters); *A*=drainage resistance (days); *a*=decay rate (1/days); *n*=convection time (days); *E*=reduction factor (-); *LTP*=linear Trend Parameter (*L/T*); α=decay or memory of the white noise process (-)

Some attention should be paid on the parameter *E*, the reduction factor of evapotranspiration. For several wells we found estimates of *E* which are not very realistic, like negative values. One reason could be that the climate station is located approximately 10 km outside the study area. Another reason might be in the large temporal variation of land use, which makes these parameters difficult to estimate. It should also be noted that the relatively large memory of the system might not completely be captured by the relatively short series.

3.2 Spatial Interpolation

Including the land use variables into the geostatistical model caused a decrease in the semivariance, as can be seen in Figure 3. 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 semivariogram reflects the precision of the *LTP* and the short-distance spatial variation in *LTP*.

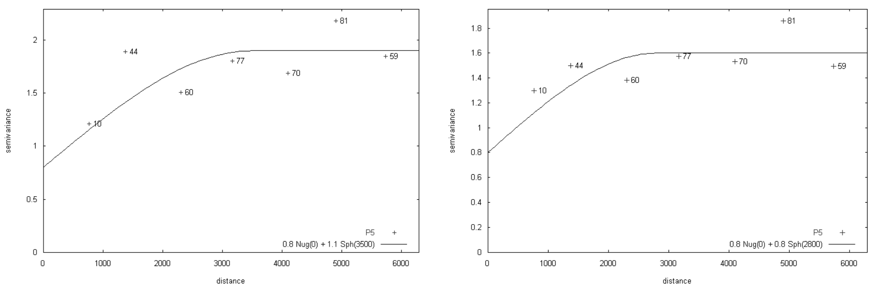

Figure 3 Semivariograms fitted for the linear trend parameter without including a trend that depends on land use (right) and with including a trend.

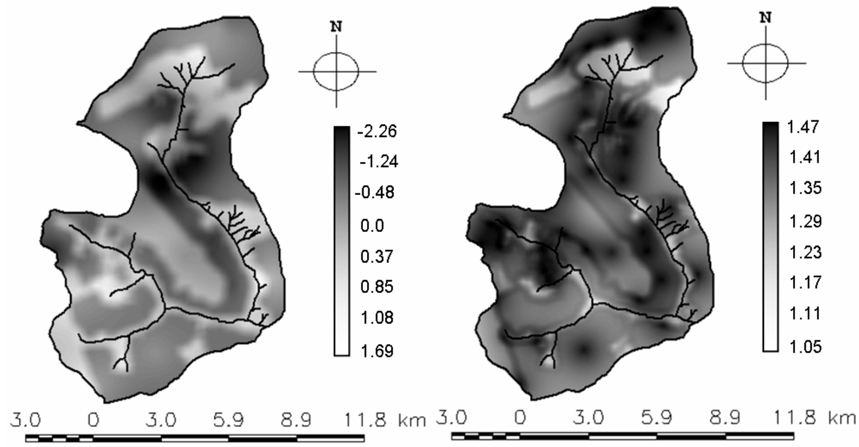

Figure 4 Systematic changes of the water table depths during the period from October 2003 to April 2006 (Left), and the corresponding kriging standard deviations (Right).

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). Also, this region has a barrier to stop the river flow and to create a water reservoir for the irrigation systems. For some areas systematic rising of the water table depths was estimated. These risings can be explained as follows. 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 31 years, respectively. Therefore, the groundwater system could recharge during the latter period in some areas, resulting in rising water tables.

In the northern part of the basin some areas are found where the Cerrado vegetation still remains, and where systematic risings are indicated. 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 same is true for the area with risings in the eastern part of the basin, which belongs to the same geological system. The degradation of the Cerrado vegetation in these areas could also be a reason of systematic rising of the water table depths, because the degradated vegetation does not use all the water volume that could be explored by the original biomass.

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 which were estimated from relatively short time series. The large uncertainty implies that observed lowerings and risings of the water table depth may not be statistically significant.

The results of the universal kriging were evaluated by cross-validation. Table 2 gives the results.

Table 2 Cross-validation for the spatial interpolation of *LTP*.

	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 quantil; 3rd Q=Third quantil; Max=Maximum; SD=Standard deviation; *Z*-score=(Pred-Obs) / Kriging variance.

The cross-validation results indicate large interpolation errors, because the standard deviation of the prediction errors is only a bit smaller than that of the observations themselves. These errors can be explained from the uncertainty about the *LTP* parameters at the 40 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.

The significance of the systematic changes mapped was checked at a 90% confidence level. Figure 5 shows the areas were the water table has changed significantly.

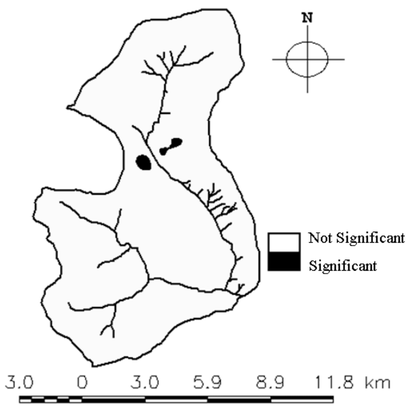

Figure 5 Areas with significant changes in the water table depths.

Significant rising of the water table was not found. Two spots with systematic lowering were indicated. These two spots are areas with intensive irrigation.

4 Conclusions

The water table depths in the study area appear to have changed systematically between 2001 and 2004, although the length of the water table depth time series is too short to obtain significant results. Significant lowerings were indicated only for two spots. These areas deserve attention, because of a potential risk of water shortage. Large uncertainties were found, due to the restricted length of the monitoring period (30 months). Therefore, we recommend to continue monitoring in order to obtain more accurate estimates of the trends in future.

Acknowledgements

CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) enabled this study by a scholarship and financial support. Also, we are grateful to Kiwa Water Research for the use of the program Menyanthes on time series analysis, and Jos Von Asmuth (Kiwa Water Research) for his advices.

References

Abramowitz, M. and Stegun, I. A., 1964, *Handbook of mathematical functions*. New York: Dover Publications Inc.

Box, G. E. P. and Jenkins, G. M., 1976, *Time series analysis: forecasting and control*. San Francisco: Holden-Day.

Hengl, T., Heuvelink, G. B. M. and Stein, A., 2004, A generic framework for spatial prediction of soil properties based on regression-kriging. *Geoderma*,120, pp. 75-93.

Hipel, K. W. and McLeod, A. I., 1994, *Time series modeling of water resources and environmental systems.* Amsterdam: Elsevier.

Jepson, W., 2005, A disappearing bioma? Reconsidering land-cover change in the Brazilian savanna. *The Geographical Journal*, 2, p. 99-111.

Klink, C. A. and Moreira, A. G., 2002, Past and current human occupation, and land-use. In: *The Cerrado of Brazil: Ecology and Natural History of a Neotropical Savanna*, P. S. Oliveira and R. J. Marquis (Org.), pp. 69-88, New York: Columbia University Press.

Knotters, M., 2001, *Regionalised time series models for water table depths*. PhD thesis, Wageningen: Wageningen University.

Knotters, M. and Bierkens, M. F. P., 2000, Physical basis of time series models for water table depths. *Water Resour. Res.*, 36, pp.181-188.

Knotters, M. and Bierkens, M. F. P., 2001, Predicting water table depths in space and time using a regioinalised time series model. *Geoderma*, 103, pp. 51-77.

Lousada, E. O., 2005 *Estudos hidrológicos e isotópicos no Distrito Federal: Modelos conceituais de fluxo.* PhD Thesis, Brasília: University of Brasília.

Nash, J. E., 1958, Determining runoff from rainfall. *Proc. Inst. Civ. Eng*., 10, pp.163-184.

Odeh, I., McBratney, A., Chittleborough, D., 1994, Spatial prediction of soil properties from landform attributes derived from a digital elevation model. *Geoderma*, 63, pp. 197-214.

Oliveira, R. S., Bezerra, L., Davidson, E. A., Pinto, F., Klink, C. A., Nepsttad, D. C. and Moreira A., 2005, Deep root function in soil water dynamics in cerrado savannas of central Brazil. *Functional Ecology*, 19, pp. 574-581.

Tankersley, C. D. and Graham, W. D., 1994, Development of an optimal control system for maintaining minimum groundwater levels. *Water Resour. Res.*, 30, pp. 3171-3181.

Van Geer, F. C.and Zuur, A. F., 1997, An extension of Box-Jenkins transfer/noise models for spatial interpolation of groundwater head series. *J. Hydrol.* 192, pp. 65-80.

Von Asmuth, J. R., Bierkens, M. F. P. and Maas, C., 2002, Transfer function noise modeling in continuous time using predefined impulse response functions. *Water Resour. Res.*, 38 (12), pp. 23.1-23.12.

Von Asmuth, J. R. and Knotters, M., 2004, Characterising groundwater dynamics based on a system identification approach. *J. Hydrol.* 296, pp. 118-34.

Ziemer, R. E., Tranter, W. H. and Fannin, D. R., 1998, *Signals and systems: continuous and discrete.* Upper Saddle River: Prentice-Hall.