

## GIS applications for mapping and spatial modeling of urban-use water quality: a case study in District of Cuiabá, Mato Grosso, Brazil

Aplicações de SIG para mapeamento e modelagem espacial da qualidade de água para uso urbano: estudo de caso num distrito de Cuiabá, Mato Grosso, Brasil

Peter Zeilhofer <sup>1</sup>

Liliana Victorino Alves Corrêa Zeilhofer <sup>2</sup>

Edna Lopes Hardoim <sup>3</sup>

Zoraidy Marques de Lima <sup>2</sup>

Catarina Silva Oliveira <sup>4</sup>

### Abstract

*A cross-sectional study utilizing spatial analysis techniques was conducted to study water quality problems and risk of waterborne enteric diseases in a lower-middle-class urban district of Cuiabá, the capital of Mato Grosso State, Brazil. Field surveys indicate high rates of supply water contamination in domiciles and, conspicuously, in public and private schools. Logistic regression models developed for the variables turbidity, Escherichia coli, total coliforms, and intestinal parasite infection did not identify singular explanatory factors for the supply water conditions and elevated incidences of enteric diseases among children. The contamination problems were found to be the result of precarious conditions involving both public infrastructure and in-building sanitary installations and their maintenance. GIS methods were successfully applied to create spatial datasets for logistic regression model building and to construct risk maps using regression coefficients.*

*Water Quality; Waterborne Diseases; Spatial Analysis; Geographic Information Systems*

### Introduction

Contaminated drinking water is associated with about 80% of all diseases and one third of all premature deaths in Brazil <sup>1</sup>, making it the most serious environmental health problem in the country. Incidents of enteric diseases are said to be highly underestimated in developing countries <sup>2</sup>, particularly among children in poorer populations <sup>3</sup>. In urban areas, the potential risk of waterborne disease transmission can be increased by deficient sewage disposal, inadequate housing, precarious water supply and sanitation infrastructure, as well as poor water-related and hygienic practices <sup>4,5,6</sup>. Although the relationship between water quality and health is evident, there is still a great lack of studies that quantify the relationships between the multiple intervening factors and the socioeconomic, cultural, and environmental particularities in Brazil.

Effective monitoring of environmental health and planning of sanitation infrastructure must consider geographical factors in order to enable analysis of the spatial distribution of diseases and at-risk populations, stratification of risk factors, and targeting of interventions. Geographic information systems (GIS) provide ideal platforms for converging disease-specific information and their analyses in relation to population settlements, and natural and constructed environments. Spatial analysis, the ability to manipulate spatial data into different forms and thereby ex-

<sup>1</sup> Departamento de Geografia, Universidade Federal do Mato Grosso, Cuiabá, Brasil.

<sup>2</sup> Departamento de Engenharia Sanitária e Ambiental, Universidade Federal do Mato Grosso, Cuiabá, Brasil.

<sup>3</sup> Departamento de Biologia, Universidade Federal do Mato Grosso, Cuiabá, Brasil.

<sup>4</sup> Departamento de Física, Universidade Federal do Mato Grosso, Cuiabá, Brasil.

#### Correspondence

P. Zeilhofer  
Departamento de Geografia,  
Universidade Federal do  
Mato Grosso.  
Av. F. Correa da Costa s/n,  
Cuiabá, MT 78060-9000,  
Brasil.  
pzeilho@terra.com.br

tract additional meaning from them, is the most important capability of GIS, which encompasses techniques for visualization and exploratory data analysis, as well as model building.

Various case studies demonstrate the application of GIS and spatial analysis in health risk assessment. Njemanze et al.<sup>7</sup> conducted a spatial risk analysis of diseases associated with diarrhea in Nigeria, applying a probabilistic layer analysis. They identified an increase in health risk where populations have easy access to untreated surface water and live near human and industrial activities causing environmental hazards. In Brazil, Barcellos et al.<sup>8</sup> evaluated geographical aspects of the relationship between health risk and water supply in a case study in the city of Rio de Janeiro.

Different methodological approaches have been proposed for spatial analysis and modeling of health risks, the most important of them being reviewed in a recent, comprehensive overview by Carvalho & Souza-Santos<sup>9</sup>. Point data from studies with high sampling density can be evaluated by methods such as cluster detection, spatial autocorrelation, trend surface interpolation, and Krigging<sup>10,11,12,13,14</sup>. Alternatively, a series of modeling techniques is available to establish and quantify relations between response variables (disease occurrence, etc.) and independent covariates (socio-demographic, environmental factors, etc.)<sup>9,15</sup>. In epidemiologic studies, coefficients determined by methods such as Generalized Additive Poisson Regression<sup>16</sup> and Logistic Regression<sup>17</sup> have been successfully applied to spatially extrapolate disease risks.

Our cross sectional community-based study addressed the following research questions for an urban district with known water supply problems in Cuiabá, Mato Grosso State, Brazil: Is contamination of the water supply a health risk? What are the risk factors for water contamination and enteric diseases? Can problems of supply water and health risks, if they exist, be spatially modeled? Can explanatory variables be identified, which can then be used to predict spatial patterns of water quality contamination and enteric diseases? To address these topics, we conducted fieldwork involving interviews to obtain social and economic data, water quality sampling, and coprological analysis, and then applied different GIS techniques such as spatial aggregation, digital elevation model, and network analysis, for the elaboration of spatially distributed explanatory data sets. Logistic regression analysis was applied to relate these independent variables to drinking water quality and incidence of symptoms of enteric, waterborne diseases in children above the age of fifteen, and to develop risk maps.

## Study area

Cuiabá, the capital of the state of Mato Grosso, has a population of about 483,000 and is located in the central west of Brazil at about 15°35' southern latitude and 56°6' eastern longitude. The tropical semi-humid climate of the region is characterized by high temperatures throughout the year (annual mean 27°C) and two clearly distinct seasons: about 80% of the mean yearly precipitation of 1,350mm is concentrated from November through April.

In 2000, the Parque Cuiabá district, which was established in the southeastern part of the city during the early 1980s and occupies an area of 256ha, had a population of 9,362 inhabitants living in 2,401 domiciles<sup>18</sup>. Mean monthly family income is between four and ten minimum salaries (about US\$ 400-1000). A single treatment station receiving water from the Cuiabá River provides the district water supply. Some deep wells, which originally complemented public water supply, were inactivated at the beginning of 2000. Two of five primary/secondary schools continue to obtain water from deep wells.

The district was selected due to its historical problems of water supply, characterized by periodic interruption due to capacity problems and the precarious state of the sanitation infrastructure.

## Methods

### Fieldwork and laboratory analysis

The cross-sectional study included socio-economic characterization of the district's households; application of questionnaires on hygiene practices, organoleptic characteristics of supply water, and occurrence of symptoms typical of enteric diseases; regular sampling of supply water quality at households and schools; as well as coprologic survey of children above 15 years of age.

To select households for socioeconomic characterization, a two-stage sampling approach was applied. 144 square city blocks ("*quadras*") were taken as clusters, and then one or two households were randomly selected from each in order to obtain an approximately 10% sample of the total 2,401 households. In each household a questionnaire was administered, identifying the number of family members, their ages, income, and education, as well as the condition of the habitation, sanitary infrastructure, hygiene practices, observed indices of symptoms of enteric diseases, and organoleptic characteristics of water supply.

From this subset, 60 households were selected for stratified water quality sampling, using as criteria homogeneous spatial distribution, socio-economic characteristics (construction characteristics), and the presence of children over the age of 15.

Water sampling was conducted on ten dates during 2002 and 2003, five from the dry period (May through October 2002) and five from the wet season (November 2002 through April 2003). A total of 409 samples were taken both at domicile connections to the distribution network and at in-house kitchen taps. For logistic regression analysis we considered only 48 households from which we could obtain at least six samples. Sample quantity from distribution network connections was low, as water supply in the district is periodically interrupted.

In addition, we monitored water quality at the five district schools, at the building connections and inside the schools at the purification equipment used to supply drinking water for children. Water quality samples were analyzed according to techniques formulated in *Standard Methods for the Examination of Water and Wastewater*<sup>19</sup>. Coprologic samples were taken from 54 children under the age of fifteen. Eggs and oocysts of main intestinal parasites were extracted by sedimentation and identified through optical microscopy.

### **Spatial data preprocessing**

Spatial data such as location of sampled domiciles, district quarters, planimetric and altimetric cadastre, water supply, and sewer network were digitalized, processed, and analyzed using the GIS software ArcInfo 7.2.1 and ArcView 3.2, both from ESRI (Redlands, United States). Geocoding of domiciles was done according to address matching and urban cadastre acquired during the socioeconomic survey.

Multidimensional data sets on socioeconomic characteristics, hygiene practices, water quality, etc., were stored in an Access database (Microsoft Inc., United States) and associated with domiciles or, in some cases, to district quarters after averaging. Hydrographic network, streets, quarter divisions, and topographic information were digitalized from the urban cadastre available in a 1:10.000 scale. As the urban cadastre was made in the early 1990s, information on transportation infrastructure, land use, occupation, and green areas was updated through aerial photographs from the year 1999, available in a 1:8000 scale. Water supply and sewer network layout, including information about when they were constructed, were digitalized from cadas-

tre of the municipal sanitation agency SANECAP, available in a 1:10.000 scale.

Local pressure in the water supply network is influenced by topography. Low pressure or interruption of fluxes favors infiltration of contaminated water, particularly because the sewer network in the district is partially installed in the same valleys as supply infrastructure. Therefore, a digital elevation model was interpolated and resampled to a 10m horizontal resolution, using the ArcInfo implementation of the Topogrid algorithm<sup>20</sup>, based on planialtimetric data from the urban cadastre 1:10.000 with isolines of 5m vertical resolution.

We also evaluated the hypothesis that domicile distance from the treatment station and age of the water supply infrastructure influences supply water quality. Domicile distance within the supply network was therefore modeled using the ArcView Network Analyst extension and water supply network was classified as original or recently renewed (last 4 years).

### **Spatial modeling**

Moran's index was calculated to test point-related attributes for spatial autocorrelation and evaluate the possibility of interpolation.

Logistic regressions were applied to test for associations between response variables (dependent) obtained by questionnaire, water quality, and coprologic analysis, as well as occurrence of symptoms of waterborne diseases and predictive environmental, social, economic, and infrastructural factors. Regression coefficients were then applied for spatial modeling.

In logistic regression, response variables are dichotomous, assuming 1 for the event and 0 for the nonevent. Being a nonlinear estimation method, the predictive variables can have a metric or non-metric scale. Logistic regression is a nonlinear transformation of a linear regression and its results can be interpreted as an estimate of the probability of the occurrence of an event.

The model of logistic regression is given by:

$$\ln[p/(1-p)] = a + BX + e \quad [1]$$

where:

p: probability of an event,  $p(Y=1)$ ;

$p/(1-p)$ : odds ratio (probability of the event divided by the probability of a nonevent);

$\ln[p/(1-p)]$ : natural logarithm of the "odds ratio", also called "logit";

a, B: estimated logit coefficient;

X: explanatory variable;

e: model constant.

In logistic regression, the estimated slope coefficient is interpreted as the rate of change in the "log odds" as X changes, instead of representing

the rate of change in Y (the dependent variables) as X changes, as it is in linear regression models, for example.

The null hypothesis, that the coefficient of an explanatory variable is significantly different from 0, can be tested by the Wald statistic, which is given by:

$$\text{Wald} = [B/SE.B]^2 \quad [2]$$

where:

B: estimated logit coefficient;

SE.: standard error of the coefficient.

Classification tables plot the observed versus the predicted responses of logistic regression models. The ratio between correctly classified events and total number of events is a measure of the overall accuracy of the model.

Because response variables must be coded as binaries, a threshold must be selected to convert measured values or counts into predictions of critical versus non-critical situations (what kind of water quality alteration or frequency of enteric disease symptoms are to be considered abnormal or critical, etc.). In our study, we synthesized water quality data for each household, coding a household as positive if one of at least six samples had turbidity values above 5uT or bacteria counts above 0 (total coliforms, *Escherichia coli*), respectively. A coprologic sample was coded positive if at least one intestinal parasite was detected.

Uni- and multivariate statistics were done with the ADE-4 software package, version 2001 (CNRS, Lyon, France). Spatial models were calculated and visualized with ArcView, version 3.2.

## Results

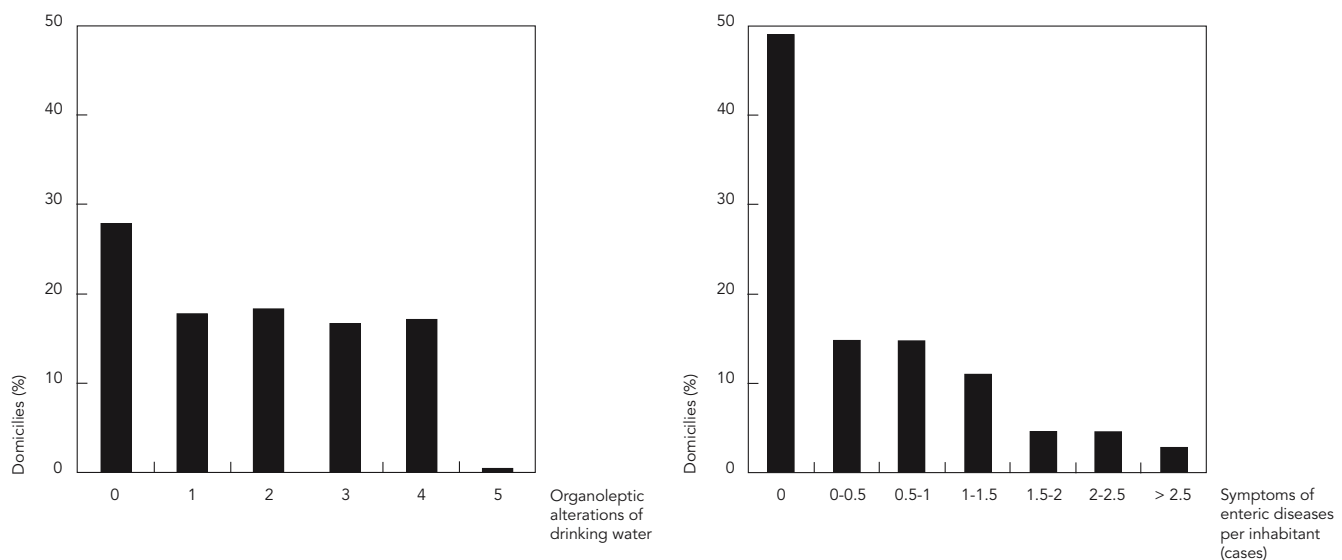
### Questionnaire

Figure 1 shows the questionnaire results, including rates of complaints about organoleptic characteristics of supply water such as odor, taste, color, turbidity, and others, as well as observations of symptoms of enteric diseases such as diarrhea, stomach pain, abdominal colic, vomiting, nausea, fever, or skin conditions.

Only about 28% of the interviewed residents had not perceived alterations in water quality, while more than 50% related having noticed two or more alterations, indicating, at least, a lack of confidence in water quality in the district. In more than 50% of the households one or more symptoms of enteric diseases were observed during the previous six months. No spatial trends in socioeconomic variables and complaints about water quality or enteric disease symptoms were observed. For household characteristics such as

Figure 1

Organoleptic alterations of drinking water and symptoms of enteric diseases per inhabitant during the last 6 months, according to questionnaire (n = 208).



number of dependents, education and income levels, and construction type, Moron Indexes (MI) of less than 0.05 were obtained. Complaints of unpleasant attributes of supply water as well as of prevalence of disease symptoms had MI values of less than 0.15.

### Drinking water quality

In this analysis of drinking water quality, we focus on the variables turbidity, *Escherichia* sp., and total coliform counts. According to von Sperling<sup>21</sup>, turbidity in supply water can be caused by clay, sludge, or domestic or industrial sewage, and may be associated with the occurrence of toxic compounds and pathogens. The Brazilian Ministry of Health, through governmental decree 1,469/2000, has defined as tolerable limits for supply water a turbidity of 5uT. Coliform bacteria must be absent.

From a total of 94 samples taken from the distribution network at-home connections (hc), turbidity exceeded legal limits in 41.5% of cases (Figure 2), while 27.7% of the samples presented contamination with total coliforms, and 2.1% with *E. coli*. In some cases, maximum counts of total coliforms and *E. coli* reached more than 2.4 E4 MPN/100ml. Turbidity values were frequently higher than 10uT, and in one case reached 71uT

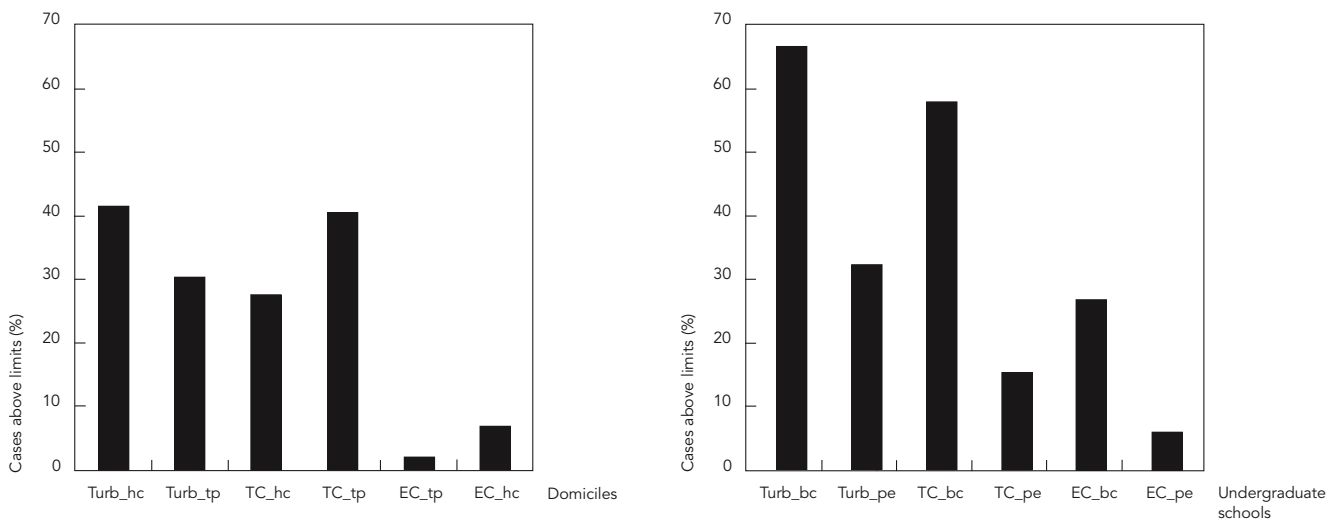
– more than fourteen times higher than permitted by law. Analysis of in-house supply water (kitchen sink) showed a significant increase in total coliforms and *E. coli* contamination when compared to sampling from home connections (t-test of paired samples,  $p < 0.05$ ), indicating that pollution is frequently caused by problems in the reservoirs and hydraulic installations. In 40.6% of the samples, total coliform bacteria were found, and positive cases of *E. coli* more than tripled (7.1%) compared with measures from home connections. Only the cases of elevated turbidity decreased slightly, to 30.4%.

Drinking water quality in the district's public and private schools is extremely precarious. Rates of total coliform contamination reached 100% of analyzed samples at one institution. On average, 58% of cases showed total coliform contamination (Figure 2), while 26.9% of samples tested positive for *E. coli*. Samples from school building connections were found to exceed turbidity limits in 66.7% of cases.

Results of samples taken from water purification equipment in the schools were alarming. Water from taps used for human consumption was contaminated with total coliforms in 15.5% of cases, and with *E. coli* in 5.9%, respectively. Turbidity samples exceeded the 5uT limit in 32.4% cases.

Figure 2

Water samples of Turb, TC, and EC above legal limits at domiciles.



hc: house connection, n = 94; tp: kitchen tap, n = 313 and schools; bc: building connection, n = 48; pe: purification equipment, n = 59.

### Spatial modeling

Following logistic model building, the dependent variables water quality and coprological analysis were tested for temporal differences between dry (2002) and wet season samples (2002-2003). In a paired sampled t-test, no significant differences were observed for these variables between the rainy and dry seasons ( $p < 0.05$ ), neither for water quality variables, nor for results of the coprologic study. In addition, district data from municipal health authorities for the years 2002 and 2003 do not indicate significant differences in enteric disease prevalence during the dry and wet seasons. Therefore, in the consecutive analysis, data sets from both periods were joined.

As sampling frequency of supply quality at house connections ( $n = 96$ ) was less than 35% of sample frequency at kitchen taps ( $n = 313$ ), only regression models for in-house samples are presented.

Table 1 shows summarized logistic regression results for the prediction of in-house water turbidity and *E. coli* above the limits set by the Brazilian normative decree 1,469/2000, as well as for the prediction of intestinal parasites in children

over fifteen. For the variable total coliforms, no significant model could be developed. Figure 3 shows the spatial models developed from logistic regression for the three evaluated dependent variables that showed significance.

Turbidity showed a significant positive correlation with the explanatory variable "elevation" ( $p < 0.05$ ), and a slightly significant negative correlation with the variable "number of rooms" ( $p < 0.10$ ). Domiciles at higher elevations as well as in poorer areas of the districts seem to be more affected. The classification table indicates good model sensitivity (accurate classification of positive cases), but unsatisfactory sensibility (underestimates of negative cases). Spatial simulation in accordance with the cut-off value set by Brazilian law (5uT) predicts high probabilities ( $> 90\%$ ) of contamination in the higher areas located in the eastern parts of the district. It should be mentioned that simulated high probabilities are affected by the tendency of the model to overestimate positive cases.

The logistic regression model for *E. coli* confirms the results of the exploratory data analysis, suggesting that this kind of bacterial contamination is primarily due to in-house reservoirs lacking protective features. Infrastructural or envi-

Table 1

Logistic regression model results for the relationship between turbidity, *Escherichia coli*, and intestinal parasites and significant exploratory variables.

Turbidity * (> 5uT)	B	S.E.	Wald	Sig.	Exp (B)	Classific. Table	Estimated	Correctly classified	Accuracy
Elevation	0.199	0.072	7.565	0.006	1.220	Obs.	0	1	
Number of rooms	-1.026	0.523	2.555	0.094	0.359	0	7	9	43.8
Constant	-29.984	12.196	6.044	0.014	0.000	1	3	29	90.6
						Overall			75.0
EC (> 1NMP/100ml)									
Reservoir protection	-1.438	0.825	3.043	0.081	0.237	Obs.	0	1	
Reservoir cleaning	-2.303	0.963	5.721	0.017	0.100	0	34	3	91.9
Constant	1.555	1.084	2.056	0.094	4.734	1	6	5	45.5
						Overall			81.3
Intestinal parasites **									
Distance	0.002	0.001	2.868	0.090	1.002	Obs.	0	1	
New supply network	-2.542	1.103	5.314	0.021	0.079	0	37	2	94.9
Constant	-2.584	1.330	3.775	0.052	0.075	1	8	7	46.7
						Overall			81.5

\* One case above limits (at least six samples);

\*\* At least one intestinal parasite detected (children under the age of 15).

Elevation: absolute altitude of domicile (GIS DEM analysis); Number of rooms: number of dependencies in the domicile (field survey); Reservoir protection: measures for protection of the in-house water reservoir(s) (questionnaire); Reservoir cleaning: regular cleaning of the in-house water reservoir(s) (questionnaire); Distance: distance between household and treatment station (GIS network analysis); New supply network: domicile connected to recently renewed supply network (GIS analysis).

Figure 3

LR-based spatial models for prediction of positive cases of water supply turbidity above 5uT, EC contamination and cases of intestinal parasites among children under the age of fifteen. Parque Cuiabá district, Cuiabá, Brazil.



ronmental variables were not found to have significant correlations with *E. coli*. Overall accuracy for the *E. coli* model was found to be higher than 80%, but the classification table indicates limited performance in predicting positive cases (45.5%). As both exploratory variables have no significant spatial autocorrelation (Moran Index < 0.1), simulation results showed a heterogeneous spatial pattern.

Distance ( $p < 0.10$ ) and condition of supply network ( $p < 0.05$ ) were found to predict significantly the probability of cases of intestinal parasites in children. Overall model performance can be considered satisfactory (> 80%). Generally low probabilities in the spatial simulation are believed to be influenced by model coefficients, which tend to underestimate positive cases: eight observed positive cases were predicted to be negative (Table 1).

## Discussion

This study examines relationships between socioeconomic, environmental, and infrastructural variables and supply water quality in an urban district of Cuiabá and discusses the applicability of spatial analysis of the epidemiological and socioeconomic, environmental, and infrastructural data obtained for risk assessment.

Results of the questionnaire, water quality monitoring, and coprologic analysis suggest that supply water quality problems and related health problems are not restricted to lower class urban populations: mean family income in 2,000 Brazilian cities between 100,000 and 500,000 habitants is R\$ 810 per month, R\$ 755 in Mato Grosso state, and about R\$ 924 in the city of Cuiabá<sup>18</sup>. Accordingly, the Parque Cuiabá district can be considered typical of the Brazilian working class ("*classe média baixa*").

Complaints about organoleptic alterations of supply water as well as about symptoms of enteric diseases were poignant. It must be assumed, however, that responses to such items in the questionnaire had some bias, considering that residents were keenly aware of the district's sanitation problem since it had caused repercussions in the local media. A direct comparison between questionnaire responses and water quality monitoring using logistic regression reinforces this assumption concerning the subjectivity of interview responses: complaints of organoleptic alterations predicted none of the three monitored water quality parameters significantly. Questionnaire results in combination with informal interviews, however, suggest that water quality and related diseases may be critical in the district and

that these problems are recognized by the population.

Questionnaire results pertaining to complaints about water quality were confirmed by supply water monitoring. For the evaluated variables turbidity, *E. coli*, and total coliforms, between 20% and 45% of collected samples were above tolerable limits. Comparison of samples from household connections and kitchen taps indicates that pollution problems of drinking water in the study area are not only caused by public infrastructure, as frequently postulated by the interviewed population. Many cases of contamination by total coliforms and *E. coli* occurred inside the households. These findings are comparable to those by Moraes et al.<sup>22</sup> for a suburban area of the city of Salvador, capital of the northeastern state of Bahia, which showed total coliform contaminated samples at 24% of at-house connections and 35% at in-house taps.

Overall logistic regression model performance for predicting supply water contamination has to be considered low for turbidity (75%) and intermediate for *E. coli* and intestinal parasites (> 80%)<sup>23</sup>. No significant model was obtained for the prediction of total coliforms cases. In evaluating this model, it should be noted that validation was internal and not done through independent observations due to limited sampling quantity. The hypothesis that several exploratory variables could explain the dependent variables must be rejected. None of the exploratory variables were found to be significant for the prediction of more than one dependent variable. Positive cases of turbidity above legal limits and cases of intestinal parasites seem to be more strongly influenced by infrastructural and environmental factors than by household reservoir protection or hygienic behavior. Domiciles at higher elevations, at greater distances from the treatment station, and with old pipe systems are believed to have lower pipe pressures that favor infiltration.

Surprisingly, only one of the socioeconomic indicators (number of rooms) was found to have a slightly significant correlation with water quality contamination or cases of parasite infection. We suggest that these results can be explained above all by the socioeconomic characteristics of the district, where inhabitants generally have access to basic education and infrastructure.

Cifuentes et al.<sup>5</sup> point out, based on their cross-sectional study in Mexico City, that supply water contamination is not necessarily appropriate for predicting health risk. In logistic regression models of the Parque Cuiabá district, water quality characteristics as independent variables did not predict well the occurrence of intestinal parasites among children. Only total coliforms



was found to have a weak significant relationship ( $p < 0.10$ ), but prediction is only reliable for nonevents. When we used water quality data from schools as a covariate, prediction results improved slightly. We suggest, therefore, that cross-sectional studies of the prevalence of enteric diseases must include public facilities that inhabitants visit frequently.

## Conclusions

According to the results presented above, we conclude that in the district of Cuiabá, drinking water quality is precarious due to both problems of public supply infrastructure and inadequate household reservoir protection or sanitary installations. Bacterial contamination is particularly

critical in public and private schools, and less so in the domiciles.

The effects of external factors on water quality were found to be complex. We surmise that sampling frequency of water quality (2%, on average sampled about six times during one year) should be increased to develop highly significant empirical models.

GIS methods can be efficiently applied to complement and extend results obtained from logistic regression analysis: spatial data analysis can provide additional relevant inputs for statistical analysis (in our case, altitude and distance between domiciles and treatment station). Spatial extrapolation can significantly improve visualization of health risks, supporting environmental and infrastructure planning and health care measures.

## Resumo

*Foi realizado um estudo transversal incluindo técnicas de análise espacial para avaliar problemas de qualidade de água de consumo e riscos de doenças entéricas em um bairro da classe média baixa na cidade de Cuiabá, capital de Mato Grosso, Brasil. Os trabalhos de campo indicam altos índices de contaminação nos domicílios e, particularmente, nas escolas públicas e particulares. Modelos de regressão logística, desenvolvidos para as variáveis turbidez, Escherichia coli, coliformes totais e parasitas intestinais não puderam ser relacionados a fatores singulares que explicassem riscos de comprometimento da água de consumo e de infecção com parasitas em crianças. Os problemas detectados, entretanto, podem ser ligados às condições precárias da infra-estrutura pública de abastecimento e das instalações sanitárias e sua manutenção nos domicílios. Técnicas de geoprocessamento foram aplicadas com sucesso para a elaboração de planos de informação espaciais, utilizados na geração dos modelos de regressão logística e no mapeamento de risco com base nos coeficientes da regressão logística.*

*Qualidade de Água; Doenças de Origem Hídrica; Análise Espacial; Sistemas de Informação Geográfica*

## Contributors

P. Zeilhofer participated in study design, statistical analysis, spatial data analysis, article drafting, and final approval. L. V. A. C. Zeilhofer, E. L. Hardoim, Z. M. Lima, and C. S. Oliveira participated in study design, field sampling, laboratory analysis, data interpretation, revision, and final approval.

## Acknowledgments

This study was financed by Fundação Nacional de Saúde (Brazilian National Health Foundation).

## References

1. Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento. Agenda 21. Brasília: Senado Federal; 1996.
2. Neumann NE, Smith DW, Belosevic M. Waterborne disease: an old foe re-emerging? *Journal of Environmental Engineering and Science* 2005; 4:155-71.
3. Gundry S, Wright J, Conroy R. A systematic review of the health outcomes related to household water quality in developing countries. *J Water Health* 2004; 2:1-13.
4. Brick T, Primrose B, Chandrasekhar R, Roy S, Muliyl J, Kang G. Water contamination in urban south India: household storage practices and their implications for water safety and enteric infections. *Int J Hyg Environ Health* 2004; 207:473-80.
5. Cifuentes E, Mazari-Hiriart M, Ferreira F, Bianchi F, Gonzalez D. The risk of enteric diseases in young children and environmental indicators in sentinel areas of Mexico City. *Int J Environ Health Res* 2002; 12:53-62.
6. Martins G, Boranga JA, Latorre MRDO, Pereira HASL. Impacto do saneamento básico na saúde da população de Itapetininga/SP, de 1980 a 1997. *Eng Sanit Ambient* 2002; 7:161-88.
7. Njemanze PC, Anozie J, Ihenacho JO, Russell MJ, Uwaeziozi AB. Application of risk analysis and geographic information system technologies to the prevention of diarrheal diseases in Nigeria. *Am J Trop Med Hyg* 1999; 61:356-60.
8. Barcellos C, Coutinho K, Pina ME, Magalhães MMAE, Paola JCMD, Santos SM. Inter-relacionamento de dados ambientais e de saúde: análise de riscos à saúde aplicada ao abastecimento de água no Rio de Janeiro utilizando sistemas de informações geográficas. *Cad Saúde Pública* 1998; 14:597-605.
9. Carvalho MS, Souza-Santos R. Análise de dados espaciais em saúde pública: métodos, problemas, perspectivas. *Cad Saúde Pública* 2005; 21:361-78.
10. Odoi A, Martin W, Michel P, Middleton D, Holt J, Wilson J. Investigation of clusters of giardiasis using GIS and a spatial scan statistic. *Int J Health Geogr* 2004; 3:11. <http://www.ij-healthgeographics.com/content/3/1/11>.
11. Kleinschmidt I, Bagayoko M, Clarke GPY, Craig M, Le Sueur D. A spatial statistical approach to malaria mapping. *Int J Epidemiol* 2000; 29:355-61.
12. Tango T. Comparison of general tests for spatial clustering. In: Lawson AB, Biggeri A, Böhning D, Lesaffre E, Viel J-F, Bertollini R, editors. *Disease mapping and risk assessment for public health*. New York: John Wiley & Sons; 1999. p. 111-7.
13. Timander LM, McLafferty S. Breast cancer in West Islip, NY: a spatial clustering analysis with covariates. *Soc Sci Med* 1998; 46:1623-35.
14. Gatrell AC, Bailey TC, Diggle PJ, Rowlingson BS. Spatial point pattern analysis and its application in geographical epidemiology. *Trans Inst Br Geogr* 1996; 21:256-74.
15. Vine ME, Degnan D, Hanchette C. Geographic information systems: their use in environmental epidemiologic research. *Environ Health Perspect* 1997; 105:598-605.
16. Gouveia N, Hajat S, Armstrong B. Socioeconomic differentials in the temperature-mortality relationship in São Paulo, Brazil. *Int J Epidemiol* 2003; 32:390-7.
17. Bakker MI, Hatta M, Kwenang A, Faber WR, van Beers SM, Klatser PR, et al. Population survey to determine risk factors for *Mycobacterium leprae* transmission and infection. *Int J Epidemiol* 2004; 33:1329-36.
18. Instituto Brasileiro de Geografia e Estatística. Perfil dos municípios brasileiros. <http://www.ibge.gov.br> (accessed on 03/Nov/2000).
19. American Public Health Association. Standards methods for the examination of water and wastewater. 20<sup>th</sup> Ed. Washington DC: American Public Health Association; 1998.
20. Hutchinson ME. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology* 1989; 106:211-32.
21. von Sperling MV. Introdução à qualidade da água e ao tratamento de esgotos. 2<sup>a</sup> Ed. Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental, Universidade Federal de Minas Gerais; 1996.
22. Moraes SL, Borja PC, Tosta CS. Qualidade de água da rede de distribuição e de beber em assentamento periurbano: estudo de caso. In: Anais do 20<sup>o</sup> Congresso Brasileiro de Engenharia Sanitária e Ambiental. Rio de Janeiro: Associação Brasileira de Engenharia Sanitária e Ambiental; 1999. p. 1462-73.
23. Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989.

---

Submitted on 01/Dec/2005

Final version resubmitted on 12/Jul/2006

Approved on 01/Aug/2006