

Full Length Research Paper

Spacial and temporal characterization of water quality in the Cuiabá River Basin of Central Brazil

Sérgio Batista de Figueiredo^{1,2*}, Ricardo Santos Silva Amorim^{1,3} and Carolina Lourencetti¹

¹Programa de Pós-Graduação em Recursos Hídricos - Universidade Federal de Mato Grosso, Campus de Cuiabá, 78060-900 Cuiabá-MT, Brazil.

²Secretaria de Estado do Meio Ambiente – Mato Grosso, Brazil.

³Departamento de Solos e Engenharia Rural, Universidade Federal de Mato Grosso, Campus de Cuiabá, 78060-900, Cuiabá-MT, Brazil.

⁴Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Campus Matão, 15991-502, Matão-SP, Brazil.

Received 1 December, 2014; Accepted 11 March, 2015

Multivariate statistical techniques were applied to evaluate the temporal and spatial variation of water quality in the Cuiabá River basin that could be affected by urban effluent discharge and diffuse pollution. After data treatment, 75.74% of the total variance was explained in the 28 parameters determined in water samples, collected at 13 stations from July 2010 to May 2011. The first component (29.08%) was represented by dissolved ions from natural processes and samples collected during rainy season, and the second component (15.42%) was related mainly to agricultural and urban effluent discharges and with samples collected during dry season (positive loadings). Hierarchical cluster analysis revealed the existence of five groups, of which three were characterized as regions of low, moderate, and high impact in the basin respectively. The other two sites, Marzagão and Porto Cercado, were considered unique groups. Marzagão is located in the headwaters of the Cuiabá River and is not subject to flow regulation caused by a reservoir, and Porto Cercado is within the flood plain of Pantanal. These data will contribute to the optimization of monitoring programs and to the understanding of other similar watersheds.

Key words: Principal component analysis, environmental monitoring, pollution, watershed, hierarchical cluster analysis.

INTRODUCTION

Decreasing quality and quantity of water resources have become a global concern, even in countries with high water potential, such as Brazil (Colletti et al., 2010). The high concentration of the urban population in Brazil has caused numerous environmental problems for water resources, including contamination of rivers by domestic and industrial sewage, flooding caused by improper

occupation of flood plain, inadequate management of urban drainage, and lack of collection and adequate disposal of urban waste (Tucci et al., 2000).

The Cuiabá River, located in Central-western portion of Brazil, is a waterway of major importance for multiple uses, such as water supply and wastewater dilution. This river is one of the main tributaries of Northern Pantanal of

*Corresponding author. E-mail: sergiobatista@outlook.com

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](http://creativecommons.org/licenses/by/4.0/)

Mato Grosso State, the largest floodplain in the world (Zeilhofer et al., 2006). Water use conflicts, changes in aquatic biota and degradation of the water quality of the aquatic environment of this river have previously been described (Figueiredo and Salomão, 2009; Laabs et al., 2002; Zeiholffer et al., 2006).

Water quality is a concern because water from this river supplies approximately one third of the population of the state and has a significant impact on the complex, rich Pantanal ecosystem. To monitor water quality, a systematic monitoring program is required to obtain accurate estimates of variations in surface water quality due to variations in physical, chemical, and biological parameters caused by seasonal effects and land use in the basin (Simeonov et al., 2003; Toledo and Nicolella, 2002). However, large amounts of data obtained in systematic studies can be difficult to interpret. Matrices with multiple dimensions are usually generated, but the most relevant environmental information is included in only a small number of variables, while the remaining variables add little to the interpretation of the results in terms of quality (Vidal et al., 2000; Simeonov et al., 2003; Toledo and Nicolella, 2002; Andrade et al., 2007b; Felipe-Sotelo, 2007).

Multivariable statistical analysis provides an alternative approach to understand the water quality of study region and identify the pollution source apportionments (Wu et al., 2009). Among these tools, factor analysis and hierarchical cluster analysis have become increasingly widespread due to the diffusion of chemometric (science of extracting information from chemical systems by data-driven means) knowledge and the availability of software that significantly reduces the labor of calculation (Mas et al., 2010).

Several studies have addressed the evaluation of water quality data using multivariate analysis in water bodies around the world (Brondnjak-Voncina et al., 2002; Felipe-Sotelo et al., 2007; Mendiguchía et al., 2004; Kazi et al., 2009; Pejman et al., 2009) including Brazil (Silva and Sacomani, 2001; Zeilhofer et al., 2006; Palácio et al., 2009, Alexandre et al., 2010; Pereira-Filho et al., 2010; Palácio et al., 2011). This study aimed to evaluate the water quality variables responsible for the spatial and temporal variation of water quality as well as the similarities and dissimilarities among stations on the River Cuiabá in the hydrological year 2010/2011 that could be affected by urban effluent discharge and diffuse pollution.

MATERIALS AND METHODS

Study area

This study was performed in the Cuiabá River basin, a sub-basin of the upper Paraguay River basin. In the section that was investigated, which has a size of approximately 680 km, the river passes through several cities, including the main municipalities in the basin, Cuiabá, the capital city of Mato Grosso, and Várzea Grande, which has a population of more than 800 million

inhabitants. Before passing the urban areas of Cuiabá and Várzea Grande, the river runs through areas of livestock, agriculture, protected areas, and small towns of low population density. In the urban area of Cuiabá, approximately 70% of the domestic effluent is discharged into the river without any treatment. After Cuiabá, there are areas of livestock and agriculture and point sources of domestic sewage discharge from the urban areas of St. Antônio do Leverger (approximately 18,500 inhabitants) and Barão de Melgaço (approximately 7,600 inhabitants). At the end of its route, the Cuiabá River enters the Pantanal region.

Water samples were collected from 13 sampling stations (Table 1) at which water quality has been monitored by the Environmental Department of Mato Grosso State (SEMA/MT) since 1995. These monitoring stations are located on stretches of the high, medium, and low Cuiabá River (Figure 1).

Sample collection and treatment

We performed six samplings campaigns, with samples taken in triplicate bimonthly between July 2010 and May 2011. This period corresponded to an annual hydrological cycle with sampling occurring in: rainy (October to April) and dry seasons (May to September). Total precipitation in the period ranged from 17.2 mm in July - September 2013, 1484 mm in October 2010 - March 2011, and 188 mm April - June 2011. Water samples were collected approximately 20 cm below the water surface using polyethylene bottles with capacities of one litre (this sample was preserved with addition of solution of 50% sulfuric acid until pH < 2) and two litre (this sample was not chemically preserved). Samples for bacteriological analyses were collected using sterilized 100 ml plastic bags. After sampling, the samples were transported under refrigerated conditions to the laboratory of SEMA-MT, where the analyses were performed within 24 h. The sampling methodology was based on the Standard Methods of Examination of Water and Wastewater (American Public Health Association, 2005) and on the Water Sample Collection Guide (CETESB, 1988).

Analyses

A total of 31 physical, chemical, and microbiological environmental parameters (air temperature, water temperature, pH, oxidation-reduction potential (ORP), dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), total solids, total dissolved solids (TDS), total suspended solids (TSS), true color, turbidity, hardness, chlorine, sulfate, fluoride, total alkalinity, phosphate, total phosphorus, sodium, potassium, magnesium, calcium, nitrite, nitrate, ammonia nitrogen, total nitrogen, salinity, electrical conductivity (EC), total coliforms, and *Escherichia coli*) were determined by employing methods recommended in the Standard Methods for the Examination of Water and Wastewater (American Public Health Association, 2005).

The field measurements for the parameters pH, water temperature, EC, DO, salinity, TDS, and ORP were conducted with an HI 9828 HANNA Multi-parameter Meter equipped with a pH/ORP/DO/EC/Temperature HI 769828 sensor with a 20 m cable. Analyses were performed using ultrapure water produced with a Milli-Q Advantage Ultrapure Water Purification System. All reagents used were analytical grade (J.T. Baker®, Chemis® and Synth® brands). The spectrophotometric readings for the parameters COD, total nitrogen, total phosphorus, orthophosphate, nitrate nitrogen, nitrite nitrogen, and sulfate were obtained with a DR 5000TM UV-Vis Hach spectrophotometer.

The color measurements were acquired with an Aquacolor Cor PoliControl portable colorimeter. Cation analyses (sodium, potassium, magnesium, calcium, and ammonia nitrogen) were performed using an ICS-1000 Dionex ion chromatograph with an

Table 1. Locations of the sampling sites in the Cuiabá River basin.

Station	Station Name	Municipality	Altitude (m)	Latitude	Longitude
1	Marzagão	Nobres	238	14°32'31.33" S	55°50'50.5" W
2	Downstream of Nobres	Nobres	189	14°45'11.10" S	56°19'38.8" W
3	Bridge in Rosário Oeste	Rosário Oeste	186	14°49'58.19" S	56°24'51.00" W
4	Acorizal	Acorizal	173	15°12'16.22" S	56°22'0.60" W
5	Passagem da Conceição	Cuiabá	153	15°33'53.52" S	56°8'29.83" W
6	Downstream of Córrego Mané Pinto	Cuiabá	156	15°36'58.1" S	56°6'22.53" W
7	Downstream of Córrego Barbado	Cuiabá	147	15°38'25.66" S	56°4'35.18" W
8	Downstream of Córrego São Gonçalo	Cuiabá	147	15°39'0.21" S	56°4'11.61" W
9	Downstream of Córrego Ribeirão dos Cocais	Cuiabá	146	15°46'51.03" S	56°8'34.59" W
10	Santo Antônio do Leverger	Santo Antônio do Leverger	144	15°52'13.40" S	56°04'36.32" W
11	Poço Beach	Santo Antônio do Leverger	142	15°54'48.22" S	56°1'47.27" W
12	Downstream of Barão de Melgaço	Barão de Melgaço	138	16°11'43.19" S	55°58'7.27" W
13	Downstream of Porto Cercado	Poconé	122	16°31'13.17" S	56°22'31.91" W

analytical column (Dionex CS16, 3 mm, coupled to pre-column CG16, 3 mm), a 25 ml loop, 20 mM sulfuric acid eluent solution (H₂SO₄, J.T. Baker brand), an eluent flow rate of 0.36 ml min⁻¹, and a 2 mm CSRS suppressor 300 (Cation Self-Regenerating Suppressor). Unpreserved samples were filtered (0.45 µm pore size) before injection (1 ml) into the ion chromatograph. Microbiological analyses were performed according to the enzyme substrate method, using Colilert[®] culture medium and IDEXX[®] cards.

Multivariate analysis

Data treatment

The censored data, below the method's limits of detection (LOD) were replaced by values corresponding to half the LOD (LOD/2) according to procedures proposed elsewhere (Farham et al., 2002) and used by others (Navarro et al., 2006; Terrado et al., 2007; Felipe-Sotelo et al., 2008). After replacing the censored values, an analysis of the missing data was performed to evaluate the implementation of corrective actions to enable factor analysis, such as deletions of variables or cases.

To detect atypical observations, the data were converted

into standard scores (Z scores) from the univariate perspective, identifying the observations with scores higher than 3. For multivariate detection, the Mahalanobis D^2 was calculated, identifying those cases with values (D^2/df) above 3.0, where df is the number of present variables (Hair et al., 2009).

A Pearson's correlation matrix was elaborated to verify whether the data matrix had a sufficient number of correlations to justify the application of principal component analysis. Bartlett's test of sphericity and the adequacy test applied to the Kaiser-Meyer-Olkin (KMO) model were also performed. The sphericity test examines the matrix as a whole and demonstrates that the correlation matrix presents significant correlations among the variables. In the sphericity test, values that have significance ($p < 0.05$) indicate the existence of a number of correlations among variables, thus permitting the continuation of the analysis (Hair et al., 2009). The KMO test shows normalized values (between 0 and 1) and the proportion of variance that the variables have in common or the proportion due to common factors. Values less than 0.5 indicate that the method of factor analysis is inappropriate for the treatment of the data (Dziuban and Shirkey, 1974). The descriptive statistical as well as multivariate analyses were performed by employing (Statistical Package for Social Sciences (SPSS) version 19.0.

Factor analysis – principal component analysis (PCA)

After processing the data, PCA was performed. For the PCA, in this study, there were no necessity to standardize the data. The analysis was performed by transforming the correlation matrix, by estimation, into a factor matrix containing factor loadings for each variable in each factor obtained. The loads of each variable on the factors were then interpreted to identify the latent structure of the variables (Hair et al., 2009).

The latent root criterion was chosen to define the number of factors to be extracted, and all factors with eigenvalues greater than 1 were considered significant. This criterion is more reliable when the number of factors to be extracted is between 20 and 50, as in the present study (Hair et al., 2009). In this study, the VARIMAX method (orthogonal rotation) was applied to normalized data using the evaluation of the spatial and temporal variability of water quality (Andrade et al., 2007b; Palácio, 2011). Statistical software was used to obtain the rotated component matrices as well as the scores, loadings, and eigenvalues (eigenvalues).

Hierarchical cluster analysis (HCA)

In this study, the HCA aimed to observe the similarities and

Table 2. Statistical description of the chemical, physical, and biological parameters determined in the 13 stations along the Cuiabá River basin.

Variable	N	Mean	Minimum	Maximum	Standard deviation
Air temperature	201	28.6	19	39	4.7
Water temperature	201	27.3	23.7	30.6	1.6
pH	201	7.2	5.6	8.1	0.4
Salinity	201	0.04	0.02	0.013	0.02
ORP	201	72.5	-8.9	240	54.6
Conductivity	201	79	51	276	34
DO	201	6.6	1.1	9.3	1.33
BOD	201	1.2	0.1	7	0.8
Total Solids	201	136	25	3121	219
TDS	201	39	26	138	17
Color	201	33.5	4.2	134.1	23.8
Turbidity	201	71.6	1.3	337	68.4
Hardness	201	32	5.7	168.3	20.9
Chloride	201	0.72	0.1	2.54	0.49
Sulfate	201	2.84	0.26	18.97	2.36
Alkalinity	201	36	21	154	21
Phosphate	201	0.13	0	0.3	0.14
Total Phosphorus	201	0.22	0.01	2.36	0.45
Sodium	201	1.56	0.68	4.19	0.5
Potassium	201	6.42	0.64	26.29	3.69
Magnesium	201	0.55	0.13	1.52	0.16
Calcium	201	7.6	4.07	30.71	3.95
Nitrite	201	0.17	0	0.6	0.22
Nitrate	201	0.34	0	3.4	0.45
Ammonia Nitrogen	201	0.11	0.02	0.58	0.06
Total Nitrogen	201	0.72	0	5	0.75
Total Coliforms	201	11910	161	24192	9195
<i>Escherichia coli</i>	201	1994	5	24192	4283

dissimilarities among sampling stations, with the goal of clustering in each group the collection stations with similar characteristics based on the water quality data to determine which characteristics are consequences of natural environmental conditions and of land use and occupation in the Cuiabá River basin (Figure 4). To perform cluster analysis, data were standardized into standard scores (Z scores). The squared Euclidean distance was used in the HCA. Average linkage clustering and Ward's method were used as hierarchical cluster algorithms (Hair et al., 2009; Palácio et al., 2011).

To define the optimal number of clusters, the percentage variation for heterogeneity (Palácio et al., 2009; Alexandre et al., 2010; Palácio et al., 2011), with the determination of the clustering coefficient in SPSS, was adopted as a stopping rule because large percentage variations in the coefficient are used to identify stages of cluster combinations that are significantly different (Hair, 2009).

RESULTS AND DISCUSSION

Descriptive statistics and data correlation

After data treatment, the original matrix (234 cases x 31

variables) was reduced (201 cases x 28 variables). Chemical oxygen demand, fluoride and suspended solids were removed to obtain a complete matrix that could be subjected to principal component analysis. Variables exhibiting high values of standard deviation, such as turbidity, total residue and total coliforms (Table 2), displayed variability during the sampling period, which was mainly attributed to the variation in the precipitation during the study period. These variables are influenced by the amount of solid material that is carried to the river in the runoff process. Other variables such as temperature, pH, salinity, DO, BOD, and ammonia nitrogen presented very similar mean and median values, indicating an apparent symmetry of the data distribution, which can also be observed from the low values of asymmetry.

A large number of significant correlations ($p = 0.05$) among the variables were identified (Table 3), but due to the large number of observations, the critical value of r was low (0.196 for $\alpha = 0.05$). Because the significance test for r proves only that the significant correlations differ

Table 3. Binary correlations between the studied variables determined in the Cuiabá River basin.

Variables	Air temp	Water temp	pH	Salinity	ORP	Conductivity	DO	BOD	TS	TDS	Color	Turbidity	Hardness	Cl	Sulfate	Alkalinity	Phosphate	Total phosphorus	Na	K	Mg	Ca	Nitrite	Nitrate	Ammonia nitrogen	Total nitrogen	Total coliforms	<i>Escherichia coli</i>	
Air temperature	1																												
Water temperature	0.2	1																											
pH	0.09	-0.4	1																										
Salinity	0.07	-0.2	0.31	1																									
ORP	0.34	0.36	-0.19	-0.08	1																								
Conductivity	0.05	-0.2	0.31	0.98*	-0.11	1																							
DO	-0.03	-0.64	0.41	0.08	-0.2	0.09	1																						
BOD	-0.27	0.07	-0.27	-0.09	-0.24	-0.1	-0.29	1																					
TS	-0.18	0.02	-0.08	0.07	-0.02	0.06	-0.03	0.15	1																				
TDS	0.05	-0.19	0.31	0.99*	-0.1	0.99*	0.1	-0.11	0.06	1																			
Color	-0.13	0.44	-0.41	-0.22	0.11	-0.24	-0.46	0.09	0.11	-0.23	1																		
Turbidity	-0.25	0.37	-0.46	-0.21	0.33	-0.22	-0.28	0.04	0.23	-0.21	0.64*	1																	
Hardness	0.12	-0.34	0.23	0.67*	-0.21	0.69*	0.27	0.01	0.16	0.69*	-0.33	-0.31	1																
Chloride	-0.32	-0.02	-0.3	-0.06	-0.47	-0.05	-0.11	0.18	0.1	-0.05	0.38	0.24	-0.41	1															
Sulfate	-0.25	0.59*	-0.32	-0.2	0.24	-0.2	-0.5	0.07	0.12	-0.19	0.59*	0.64*	-0.05	0.15	1														
Alkalinity	0.09	-0.28	0.38	0.93*	-0.07	0.95*	0.24	-0.15	0.01	0.95*	-0.35	-0.3	0.68*	-0.14	-0.3	1													
Phosphate	0.39	-0.62	0.51*	0.2	-0.13	0.2	0.59*	-0.24	-0.22	0.2	-0.71	-0.74	0.36	-0.31	-0.76	0.35	1												
Total Phosphorus	-0.34	-0.11	-0.07	0.05	-0.16	0.06	-0.16	0.27	0.3	0.06	0.01	0.06	0.23	0.16	0.04	-0.01	-0.2	1											
Sodium	-0.28	0.23	-0.26	-0.18	-0.1	-0.19	-0.35	0.39	0.15	-0.18	0.44	0.42	-0.18	0.47	0.45	-0.28	-0.51	0.36	1										
Potassium	0.09	-0.25	0.18	0.91*	-0.13	0.92*	0.24	-0.13	0.02	0.93*	-0.32	-0.25	0.67*	-0.04	-0.3	0.91*	0.28	-0.03	-0.27	1									
Magnesium	0.08	0.28	-0.3	-0.08	-0.07	-0.09	-0.56	0.16	0.06	-0.09	0.46	0.21	-0.1	0.43	0.31	-0.17	-0.26	0.02	0.17	-0.14	1								
Calcium	0.12	-0.25	0.36	0.97*	-0.08	0.98*	0.18	-0.15	0.03	0.99*	-0.32	-0.28	0.71*	-0.13	-0.27	0.96*	0.31	0.03	-0.27	0.93*	-0.14	1							
Nitrite	0.56*	-0.21	0.31	0.12	0.05	0.13	0.4	-0.21	-0.19	0.14	-0.63	-0.63	0.37	-0.29	-0.64	0.26	0.82*	-0.12	-0.39	0.21	-0.23	0.23	1						
Nitrate	-0.17	-0.51	0.32	-0.04	-0.11	-0.04	0.3	-0.07	-0.08	-0.06	-0.18	-0.16	-0.03	-0.07	-0.14	0.02	0.33	-0.14	-0.13	-0.06	-0.1	-0.03	-0.05	1					
Ammonia Nitrogen	-0.14	0.25	-0.06	0.03	0.03	0.03	-0.22	0.32	-0.01	0.05	0.01	0.06	-0.11	0.06	0.15	-0.03	-0.2	0	0.16	-0.02	0.1	0	-0.1	-0.14	1				
Total Nitrogen	-0.11	-0.03	0.11	-0.07	-0.05	-0.08	0.05	0.09	0	-0.08	-0.02	0.11	-0.08	0.05	0.1	-0.09	-0.07	-0.02	0.06	-0.1	0.05	-0.09	-0.17	0.17	0.18	1			
Total Coliforms	-0.36	0.02	-0.03	-0.16	-0.04	-0.14	0.02	0.12	0.06	-0.14	0.08	0.27	-0.17	0.19	0.21	-0.14	-0.31	-0.02	0.21	-0.15	-0.11	-0.19	-0.37	0.04	0.15	0.17	1		
<i>Escherichia coli</i>	-0.27	0.14	0.03	0	-0.11	0	-0.08	0.25	0	0.01	0.07	0.14	-0.13	0.19	0.17	-0.07	-0.22	-0.07	0.16	-0.09	0.09	-0.06	-0.18	-0.07	0.52*	0.1	0.47	1	

* Correlations > 0.5 (absolute value). Significant values are indicated in bold characters (significance level = 95%, p < 0.05). ORP: oxidation-reduction potential, DO: dissolved oxygen, BOD: biochemical oxygen demand, TDS: total dissolved solids. TS: total solids. Cl: chloride. Na: sodium. K: potassium. Mg: magnesium. Ca: calcium.

from zero, it is important to focus on correlations that are greater than 0.5 ($p = 4.4 \times 10^{-5}$), as indicated by Helena et al. (2000) and Andrade et al. (2007b).

According to the Brazilian legislation (CONAMA Resolution nº 357/2005), in raining season there were values in parameters like turbidity, total phosphorous and *Escherichia coli* above the maximum limits, mainly in the sampling sites near the urban area. Sewage effluents from industrial and domestic sources and surface runoff of the waste could explain the increase of these parameters values. Dissolved oxygen presented lower values during rainy season, only in the sampling sites Barão de Melgaço and Porto

Cercado due to the increasing the organic matter near the Pantanal Floodplain.

Among the observed correlations, there were associations between salinity and conductivity ($r = 0.98$), salinity and total dissolved solids ($r = 0.99$), and total dissolved solids and conductivity ($r = 0.99$); these variables are strongly correlated because they represent the presence of salts in the water. These salts may be natural due to the geological and soil conditions in the region, particularly in the upper course of the river. The salts may also have anthropogenic origins, such as the silting process in the water bodies or the diffuse pollution that occurs in the basin. Similar results were found in a study developed in the basin of

the upper Acarajú, Ceará (Andrade et al., 2007b). From this correlation matrix, KMO (value of 0.773) and Bartlett's sphericity (approximate chi-square of 6526.300, significance of 0.000) tests revealed that there were sufficient correlations to apply the factor analysis.

Factor analysis – PCA

Considering the latent root criterion, 7 components were retained, which were composed of 19 of 28 variables from the original data matrix. The 7 components combined explained approximately 75.74% of the total variance.

In this matrix (Table 4), factor loadings represent

Table 4. Stopping rule for the HCA – Ward's Algorithm Method.

Stage	Stopping rule			
	Hierarchical Process		Clustering coefficient	
	Number of Groups		Value	Percentage increase for the next stage
Before clustering	After clustering			
5	9	8	213.499	29.41
6	8	7	276.29	28.16
7	7	6	354.093	30.57
8	6	5	462.337	25.72
9	5	4	581.266	29.88
10	4	3	754.94	27.91
11	3	2	965.667	34.21
12	2	1	1296	-

represent the degree of association (correlation) of each variable with each factor. The 1st component explains the greatest amount of variance, and this factor is composed of 7 variables containing high loads (greater than 0.50). Factor loadings of ± 0.50 or greater are considered practically significant (significance level $\alpha = 0.05$ and power level of 80% for sample size equal to or greater than 120), and loads greater than 0.70 are considered indicative of defined structure (in this case, the factor explains 50% of the variance of the variable).

The 1st and 2nd components together explain approximately 44.49% of the variance of the data. The 1st component (PC1) was assigned to the variables TDS, conductivity, salinity, calcium, alkalinity, hardness, and potassium. This 1st component is basically a component of dissolved salts, particularly potassium and calcium salts. The occurrence of similar variables in the 1st component and the correlation of these variables with the natural process of weathering of the geological components of the soil have also been reported in other studies (Singh et al., 2004; Andrade et al., 2007b).

The 2nd component (PC2) of this study was attributed to the parameters turbidity, sulfate, and color, which were positively correlated with the factor, and to the parameters phosphate and nitrite, which were negatively correlated. Samples collected during the rainy season are related to the positive values of component 1 (Figure 1) and to variables with negative loadings in component 2 (PC2), while samples collected during the dry season are related to the variables with positive loadings in component 2 (PC2). Observing the sample distribution as a function of the collection station in Figure 2, the highest scores recorded for the 1st component are related to Station 1 (Marzagão). This station is located in the municipality of Nobres, a region with geological records (the Araras, Raizama, and Puga Formations) that are composed mainly of sediments and rocks rich in calcium and magnesium carbonates, such as sandstones (Figueiredo and Salomão, 2009). When incorporated in

the water, these compounds increase the concentration of calcium ions and carbonates, increasing the hardness and alkalinity and, therefore, the conductivity and salinity.

During the rainy season, there is an increase in turbidity and color due to surface runoff. This runoff is aggravated by the removal of riparian vegetation in many parts of the basin and by localized processes of siltation. Large amounts of sediments and organic matter reach the tributaries and flow into the Cuiabá River, which results in increased color and turbidity. The transport of sediment in the Cuiabá River basin has already been reported, and the increase in sediment transport was attributed to urbanization, with sediment deposition in the surroundings of St. Antônio do Leverger, at the beginning of the Pantanal Plain (Figueiredo and Salomão, 2009).

Sulfate can be derived from anthropogenic activities, mainly from fertilizers and water and wastewater treatment, which occur closer to the urban areas of Cuiabá and Várzea Grande. The 3rd component, which accounted for 9.60% of the explained variance, can be attributed to the discharge of domestic and industrial effluents and the leaching of organic matter from livestock areas in the sub-basin because this component was composed of the variables *E. coli* and ammonia nitrogen. These variables presented high values, mainly at the Marzagão Station (Station 1), the stations in the urban areas of Cuiabá and Várzea Grande (Stations 7, 8, and 9), and the St. Antônio do Leverger station (Station 10).

The other components each explain approximately 3 to 7% of the total variance. The 4th component is positively related to the variable chloride and negatively related to the variable ORP. The influence of chloride is related to domestic effluent discharges in urban areas (Helena et al., 2000; Andrade et al., 2007a). The increase in the concentration of chlorides (highly reactive chemical species) decreases the ORP (which indicates which half of the chemical species is reduced). The 5th component can be explained by the presence of limestone rich in

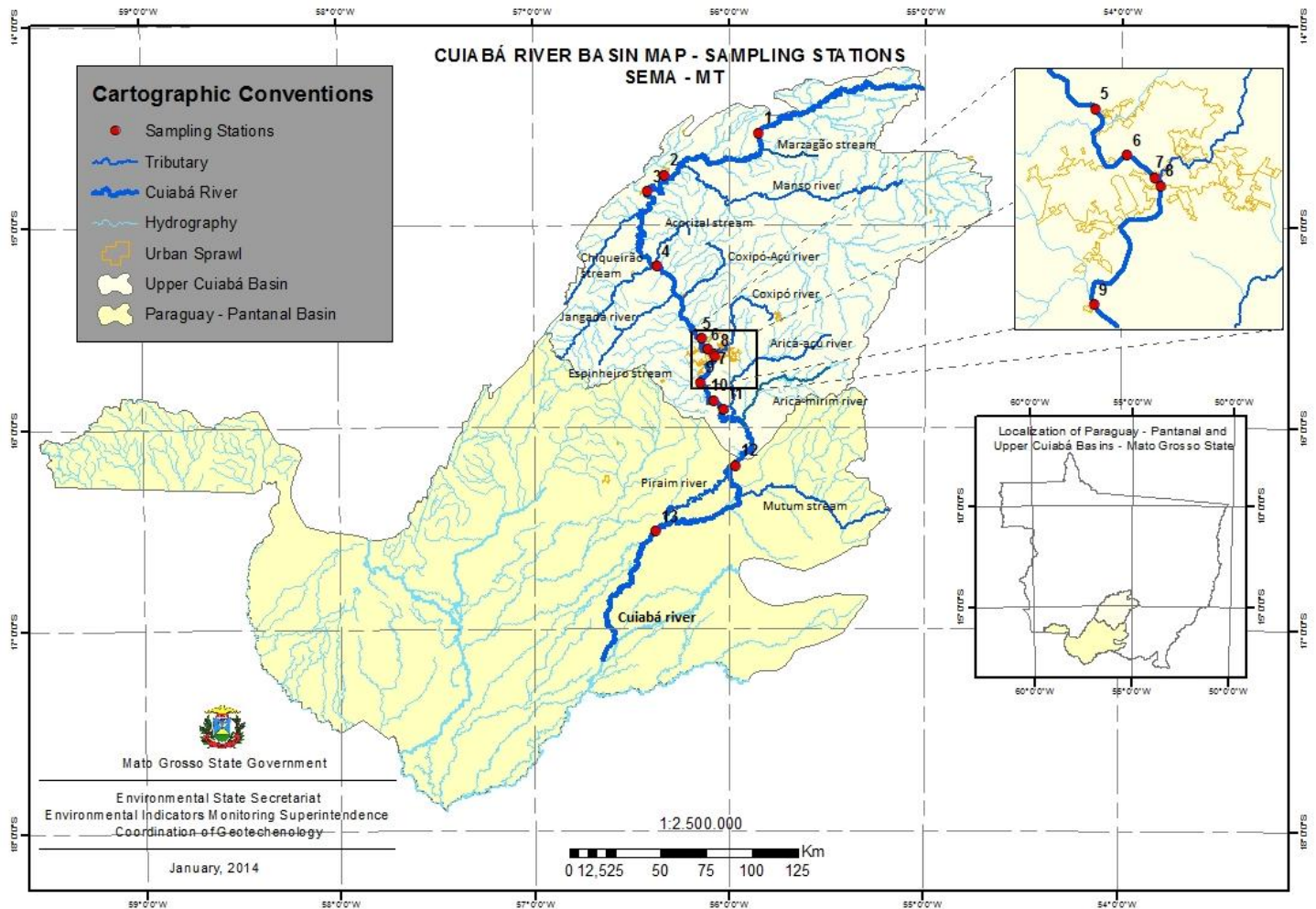


Figure 1. Location of the sampling stations in the Cuiabá River basin.

magnesium, particularly in Província Serrana, where the upper course of the Cuiabá River is located (Figueiredo and Salomão, 2009). The 6th and 7th components are represented by the variables total phosphorus and nitrate. Both variables are indicative of agricultural and urban pollution processes (Andrade et al., 2007a; Campanha et al., 2010; Pereira-Filho et al., 2010). However, because they are different components, they may reveal different aspects of this pollution, as nitrates more closely represents point sources, while total phosphorus represents diffuse sources or a different combination of both sources.

The results obtained here to explain the influence of organic pollution near the urban stations of Cuiabá River and the improvement of the water quality downstream from the urban area of Cuiabá and Várzea Grande are similar to those of a previous study (Zeilhofer et al., 2006). Considering the 17 parameters that were employed in our PCA matrix, the weight of organic pollution (discharge of domestic and industrial effluents)

was minimized due to the influence of dissolved salts and surface runoff. The absence of the parameter COD also explains some of the differences among the results of our study and those of previous studies (Zeilhofer et al., 2006).

HCA

After analyzing the diagrams obtained, the dendrogram obtained with Ward's method presented a greater segregation of clusters in the middle stages and was therefore selected for the HCA. Evaluating the percentage difference in the clustering coefficient to estimate the number of clusters formed (stopping rule), the highest percentage difference was observed to occur between stages 11 and 12 (Table 5). However, solutions of a cluster resulting from the union of 2 very heterogeneous groups presented high values for the stopping rule that are unacceptable (Hair et al., 2009). Thus, the increase

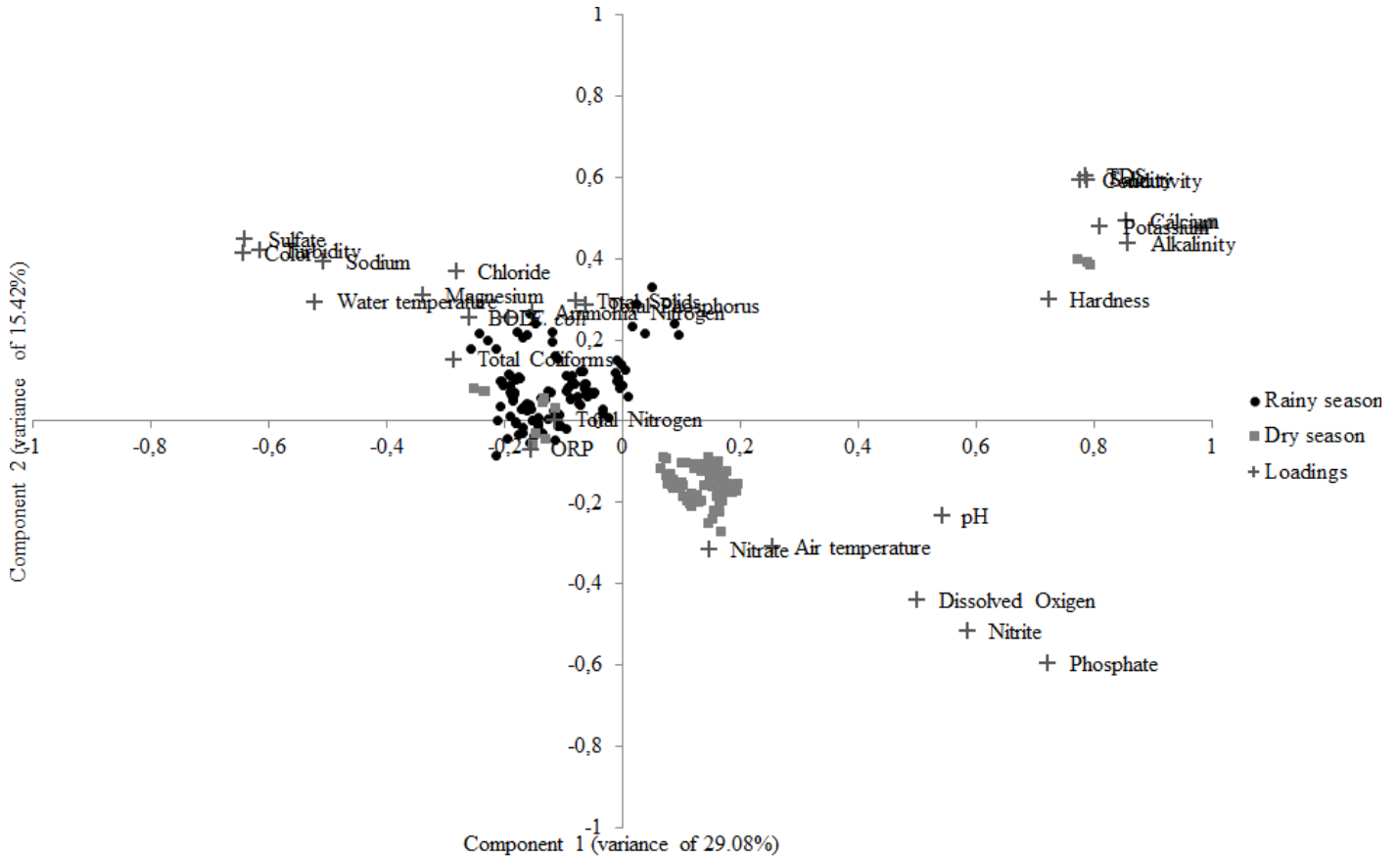


Figure 2. PCA biplot (scores and variables), classified according to seasonality.

registered between stages 7 and 8 was considered to be the highest, with a variation of 30.57%. The cutoff mark was the distance at approximately 225 in the dendrogram obtained by Ward’s method (Figure 3), resulting in 5 groups.

Station 1 (Marzagão) differs from the other stations in the dendrogram (Figure 4). There were significant variations in the values of turbidity, color, and total phosphorus, as well as in the concentrations of dissolved salts, as indicated by PCA. These variations, which are most likely attributable to seasonality, contribute to the characterization of a unique station (Figure 5).

The beginning of the operation of the APM-Manso water reservoir in 2001 regulated the water flow in the Cuiabá River. Station 1, which is located upstream from this reservoir, is not affected by this regulation, which has a strong influence on the water quality parameters at sites located downstream from the reservoir, mainly Stations 2-6. At these locations (Stations 2-6), the values of the analyzed parameters were low in comparison to those of other stations. These results are a reflection of the land use and occupation in this stretch, which still has significant conservation of riparian forests and a low degree of anthropization along the banks.

Proximity to point sources of effluent discharge

influenced the water quality parameters at Stations 7 and 8, which are located downstream from 3 polluted tributaries that serve as a flow channel for the untreated effluents from Urban Agglomeration. There are also considerable industrial effluent discharges as well as solid waste disposal along this stretch, which suggests that among the uses of this water body, the dilution of effluent is prominent (Zeilhofer et al., 2006). It can be explained by the influence observed in PCA related to component 3 due to high values of *E. coli* and ammonia nitrogen measured in these stations.

Stations 9 to 12 are located along a stretch of the Cuiabá River located downstream from the Urban Agglomeration. These stations presented high concentrations of organic matter but lower concentrations than those recorded in the urban area, which may indicate autodepuration of the organic load coming from the Urban Agglomeration of Cuiabá and Várzea Grande. This phenomenon may occur due to a reduction in the quantity of point sources of effluent discharge in this stretch as well as the small contribution from diffuse sources of pollution, as there are few human activities with significant impact potential in this region.

Station 13 (Porto Cercado) differs from the other stations in this cluster solution because it is located within

Table 5. Explained variance and component matrix of PCA (Varimax rotation) applied for 28 chemical, physical, and biological parameters determined in the 13 stations along the Cuiabá River basin.

Variable	Components							Communalities
	1	2	3	4	5	6	7	
TDS**	0.994*	-0.058	0.03	-0.002	0.001	0.012	0	0.992
Conductivity	0.990*	-0.064	0.022	0.004	0.006	0.013	0.016	0.986
Salinity	0.982*	-0.053	0.016	-0.02	0.025	0.015	0.029	0.97
Calcium	0.980*	-0.161	-0.028	-0.051	-0.019	-0.019	0.018	0.99
Alkalinity	0.948*	-0.185	-0.03	-0.045	-0.073	-0.048	0.039	0.945
Potassium	0.933*	-0.144	-0.049	0.045	-0.064	-0.056	-0.074	0.908
Total hardness	0.701*	-0.341	-0.128	0.15	-0.012	0.236	-0.089	0.711
Turbidity	-0.155	0.839*	-0.017	-0.066	-0.087	0.067	-0.091	0.753
Color	-0.19	0.769*	-0.091	0.177	0.273	-0.042	-0.115	0.756
Sulfate	-0.153	0.816*	0.128	-0.139	0.153	0.035	0.025	0.756
Ammonia nitrogen	0.026	0.035	0.817*	-0.078	0.149	0.052	0.039	0.701
<i>Escherichia coli</i>	-0.006	0.15	0.797*	0.184	-0.192	-0.137	-0.011	0.747
Chloride	-0.029	0.293	0.055	0.834*	0.119	0.052	-0.127	0.819
Magnesium	-0.058	0.298	0.029	0.349	0.735*	-0.098	0.017	0.765
Total phosphorus	0.047	0.022	-0.049	0.107	0.017	0.834*	-0.054	0.715
Nitrate	-0.067	-0.126	-0.201	0.127	-0.169	-0.116	0.753*	0.687
Phosphate	0.154	-0.880*	-0.198	-0.007	-0.082	-0.221	0.164	0.92
Nitrite	0.091	-0.846*	-0.087	-0.169	0.067	-0.153	-0.226	0.839
ORP**	-0.08	0.24	-0.079	-0.779*	0.051	-0.148	-0.138	0.721
BOD**	-0.127	-0.011	0.471	0.22	0.225	0.538	0.011	0.627
Total solids	0.09	0.206	-0.101	-0.046	-0.096	0.575	-0.026	0.404
Total nitrogen	-0.072	0.101	0.236	-0.06	0.102	0.039	0.659	0.521
Dissolved oxygen	0.093	-0.461	-0.224	0.112	-0.65	-0.157	0.162	0.758
Total coliforms	-0.124	0.333	0.439	0.156	-0.541	-0.054	0.043	0.641
pH	0.277	-0.412	0.011	-0.103	-0.22	-0.125	0.471	0.543
Air temperature	0.043	-0.36	-0.194	-0.382	0.425	-0.382	-0.254	0.705
Water temperature	-0.171	0.471	0.284	-0.359	0.413	-0.093	-0.403	0.803
Sodium	-0.184	0.436	0.182	0.277	0.107	0.425	-0.084	0.533
Sum of squares (eigenvalues)	8.142	4.316	2.687	1.876	1.709	1.383	1.094	
Percentage of trace***	29.078	15.416	9.596	6.702	6.105	4.938	3.909	75.742
Percentage of accumulated trace	29.078	44.494	54.09	60.792	66.897	71.835	75.744	

*Factor loads that indicate a definite structure (value > 70); **ORP: oxidation-reduction potential, BOD: biochemical oxygen demand, TDS: total dissolved solids, ***Trace = 28.0 (sum of eigenvalues).

the floodplain in the Pantanal of Mato Grosso. Among the results presented for this station, low concentrations of DO and low levels of coliforms are highlighted. The increase in organic matter is significant, particularly during the floods in this plain, due to sedimentation of solid particles caused by the decrease in slope, which occurs downstream of the town of Santo Antonio do Leverger (Figueiredo and Salomão, 2009).

The stretches indicated by the HCA are consistent with the spatial gradient indicated in another study (Figueiredo and Salomão, 2009), in which 4 distinct stretches were

indicated: I) from Cuiabazinho Springs to Manso Junction (where Station 1 is located); II) after Manso until the beginning of the urban areas of Cuiabá and Várzea Grande (Stations 2 and 6); III) urban areas (Stations 7 and 8); and IV) Pantanal (Station 13). However, the present study suggests a better segmentation of the stretch that begins after the urban area along the Cuiabá River. This study considers a stretch that begins downstream from the urban areas of Cuiabá and Várzea Grande and extends to the urban area of Barão de Melgaço (Stations 9-12) and a stretch that begins within

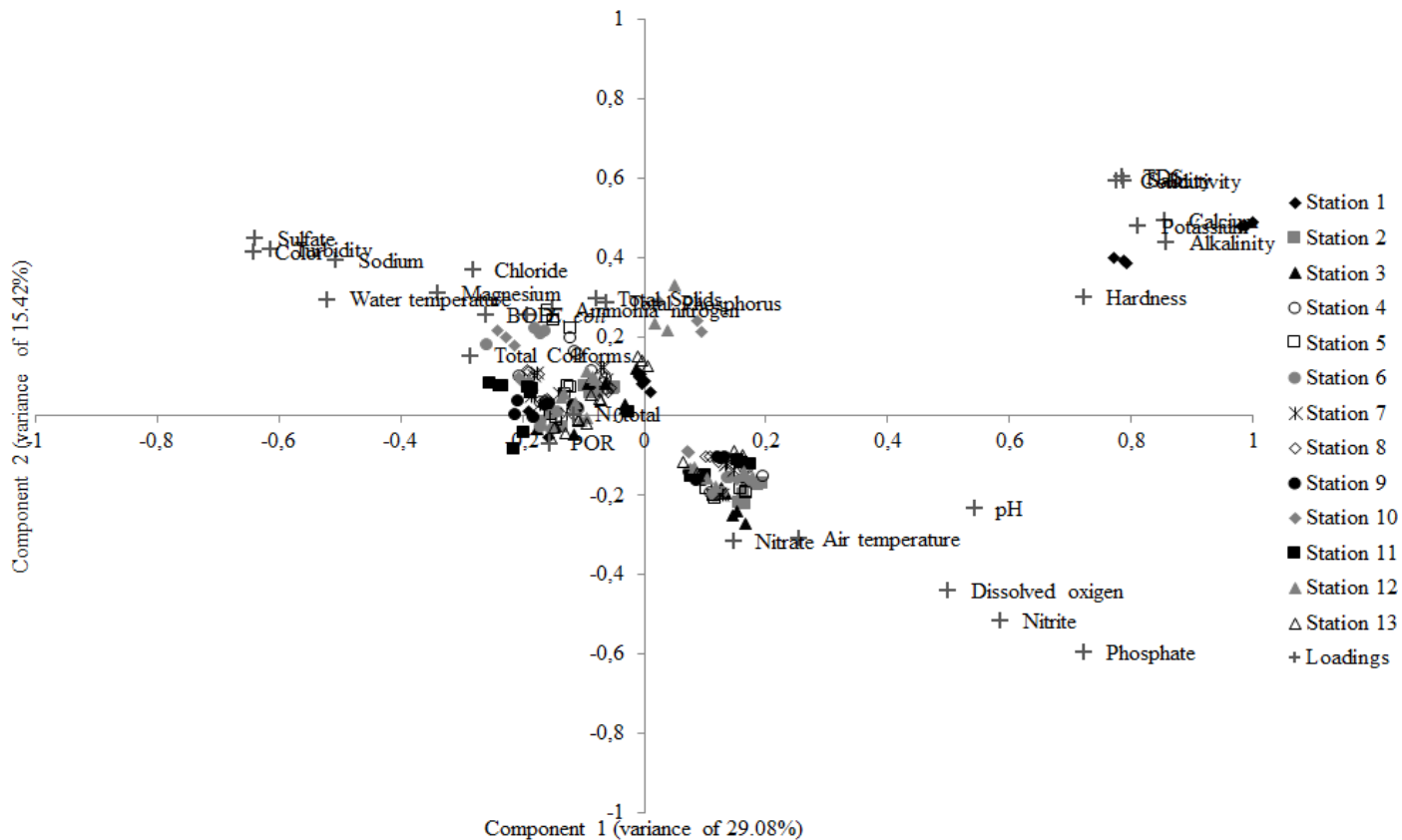


Figure 3. PCA biplot (scores and variables), classified according to the collecting station.

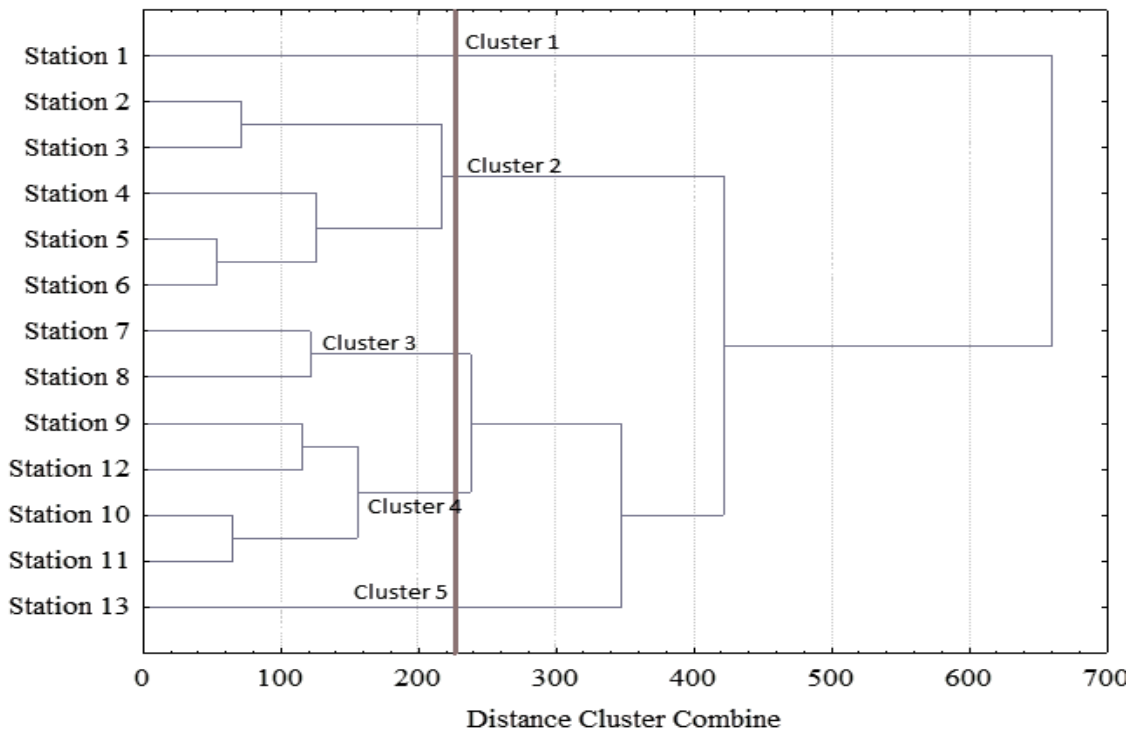


Figure 4. Dendrogram of the 13 sampling stations studied in the Cuiabá River basin.

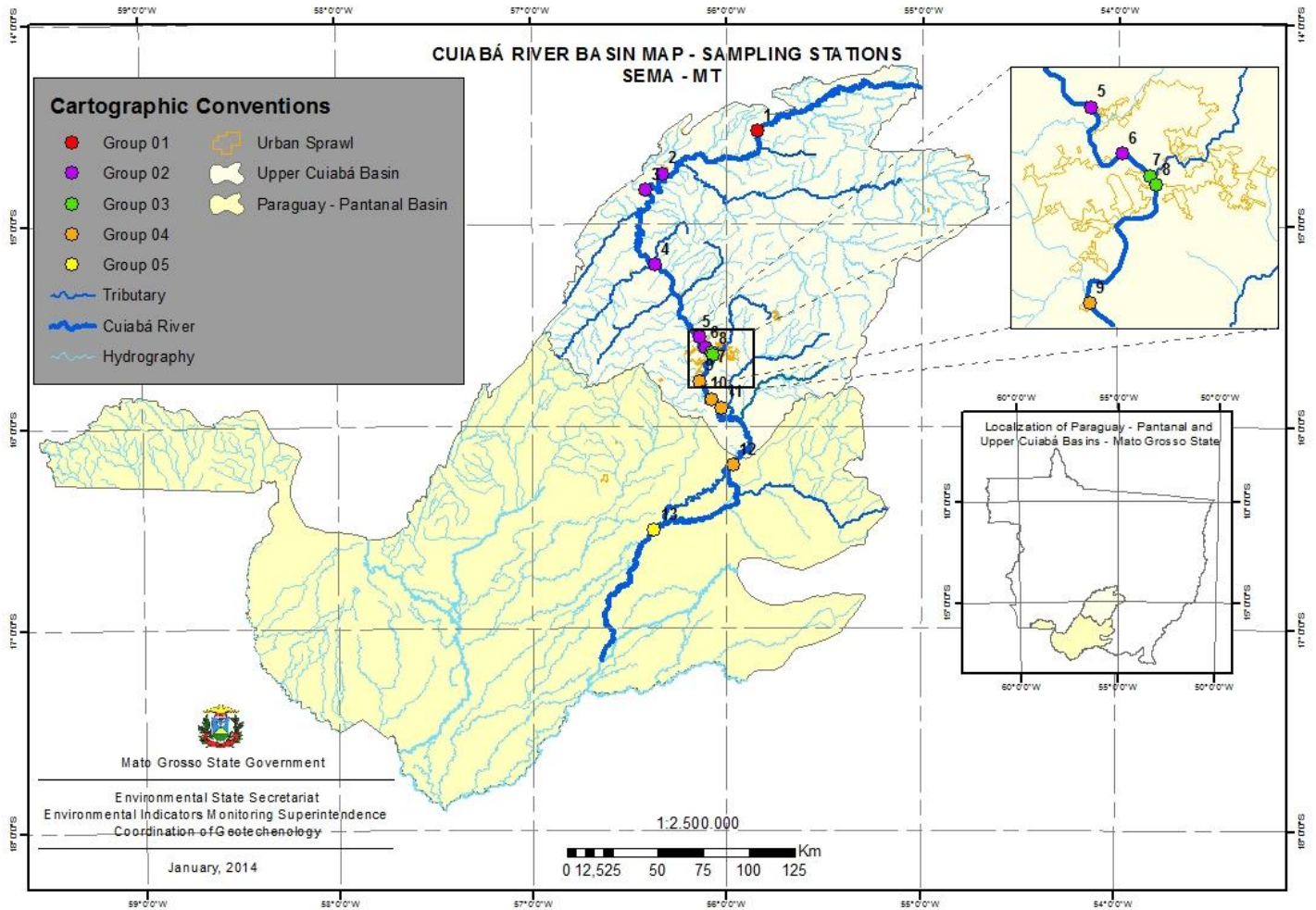


Figure 5. Sampling stations classified according dendrogram in the Cuiabá River basin.

the floodplain due to the peculiar behavior of the analytical parameters presented at Station 13. These results will contribute to improved management of the environmental monitoring of the water quality of the studied basin by the Department of Environment of Mato Grosso State (SEMA-MT) and may be useful for a better understanding of other watersheds.

Conclusions

PCA identified 7 components composed of 19 variables that are responsible for the variation in the data from the Cuiabá River basin and that together accounted for 75.74% of the data variation. The 1st axis (29.08%), composed of 7 variables, represents the component dissolved ions, thus reflecting the natural process of weathering of geological components of the soil, particularly in the stretch of the upper Cuiabá River that is characterized by regions containing limestone rocks. The 2nd component (15.42%), composed of 5 variables, was

defined by a component associated with human activities that produce diffuse pollution (agricultural and urban) and point sources of effluent discharge. Samples of this sub-basin presented a wide variation in color and turbidity parameters, influenced mainly by the carrying of sediments, which occurs principally during the rainy season, aggravated by the processes of human occupation in the basin.

HCA revealed the existence of 5 clusters of stations that present homogeneous characteristics. Stations 1 and 13 (Marzagão and Porto Cercado) present a high degree of dissimilarity when compared to the other stations and are considered unique clusters. Three other homogeneous groups were identified: II) after the mouth of the Manso River until the beginning of the urban areas of Cuiabá and Várzea Grande (Stations 2-6); III) the urban area (Stations 7 and 8); and IV) after the urban area until the vicinity of Pantanal (Stations 9-12); these groups characterize regions of low, high, and moderate pollution in the basin, respectively. The water quality in Cuiabá River is influenced mainly by natural phenomena,

but processes guided by urban agglomerations and diffusion pollution in basin are increasing and contributing for quality degradation. This kind of information is important to improve water management by environmental state agency and regulate the occupation and wastewater discharge in basin.

Conflict of Interest

The authors have not declared any conflict of interest.

ACKNOWLEDGMENTS

This study was conducted through a partnership between the Environmental Secretariat of Mato Grosso (SEMA-MT) and the Federal University of Mato Grosso (UFMT). The study was funded by SEMA-MT, CNPq, and CAPES-PNPD-PRODOC. The authors would like to thank everyone who assisted with the sampling campaigns and sample analyses.

REFERENCES

- Alexandre DMB, Andrade EM, Lopes FB, Araújo H, Palácio Q, Ferreira ACS (2010). The water quality investigation using GIS and multivariate analysis in a semiarid region reservoir. *Rev. Ciênc. Agron.* 41(4):554-561.
- Andrade EM, Araújo LFP, Rosa MF, Gomes RB, Lobato FAO (2007). Assessment of the surface water quality in the upland of Acaraú watershed, Ceará, Brazil. *Cienc. Rural.* 37(6):1791-1797.
- Andrade EM, Araújo LF, Rosa MF, Disney W, Alves AB (2007). Surface water quality indicators in low Acaraú basin, Ceará, Brazil, using multivariable analysis. *Eng. Agríc.* 27(3):683-690.
- APHA (2005). (Standard methods for the examination of water and wastewater., AWWA, WEF: Washington DC.
- Brondnjak-Voncina D, Dobcnik D, Novic M, Zupan J (2002). Chemometrics characterization of the quality of river water. *Anal. Chim. Acta.* 462(1):87-100.
- Campanha MB, Melo CA, Moreira AB, Ferrarese RFMS, Tadini AM, Garbin EV, Bisinoti MC (2010). Spatial and seasonal variability of physico-chemical parameters in Turvo, Preto and Grande rivers on São Paulo state, Brazil. *Quim. Nova.* 33(9):1831-1836.
- CETESB (1988). Coleta e Preservação de Amostras de Água. São Paulo (SP), Brazil.
- Colletti C, Testezlaf R, Ribeiro TAP, Souza RTG, Pereira DA (2010). Water quality index using multivariate factorial analysis. *Rev. Bras. Eng. Agr. Amb.* 14(5):517-522.
- Dziuban CD, Shirkey E (1974). When is a correlation matrix appropriate for factor analysis? *Psychol. Bull.* 81(6):358-361.
- Farnham IM, Singh AK, Stetzenbach KJ, Johannesson KH (2002). Treatment of nondetects in multivariate analysis of groundwater geochemistry data. *Chemom. Intell. Lab. Syst.* 60:265-281.
- Felipe-Sotelo M, Andrade JM, Carlosena A, Tauler R (2007). Temporal characterization of river Waters in urban and semi-urban áreas using physico-chemical parameters and chemometric methods. *Anal. Chim. Acta.* 583(1):128-137.
- Felipe-Sotelo M, Tauler R, Vives I, Grimalt JO (2008). Assessment of the environmental and physiological processes determining the accumulation of organochlorine compounds in European mountain lake fish through multivariate analysis (PCA and PLS). *Sci. Total Environ.* 404(1):148-161.
- Figueiredo DM, Salomão FXT (2009). Bacia do Rio Cuiabá: uma abordagem socioambiental. Cuiabá (MT), Brazil: Entrelinhas EdUFMT. 216 pp.
- Hair Jr JF, Tatham RL, Black WC, Anderson RE, Babin BJ (2009). *Análise Multivariada de Dados*, 6th ed. Translation Sant'Anna, AS. Bookman: Porto Alegre, Brasil.
- Helena B, Pardo R, Vega M, Barrado E, Fernandez JM, Fernandez L (2000). Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Res.* 34(3):807-816.
- Kazi TG, Arain MB, Jamali MK, Jalbani N, Afridi HI, Sarfraz RA, Baig JA, Abdul QS (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicol. Environ. Saf.* 72(2):301-309.
- Laabs V, Amelung W, Pinto AA, Wantzen M, da Silva CJ, Zech W (2002). Pesticides in Surface Water, Sediment, and Rainfall of the Northeastern Pantanal Basin, Brazil. *J. Environ. Qual.* 31:1636-1648.
- Mas S, Juan A de, Tauler R, Olivieri AC, Escandar GM (2010). Application of chemometric methods to environmental analysis of organic pollutants: A review. *Talanta* 80(3):1052-1067.
- Mendiguchía C, Moreno C, Galindo-Riño MD, García-Vargas M (2004). Using chemometric tools to assess anthropogenic effects in river water. A case study: Guadalquivir River (Spain). *Anal. Chim. Acta.* 515(1):143-149.
- Navarro A, Tauler R, Lacorte S, Barceló D (2006). Chemometrical investigation of the presence and distribution of organochlorine and polyaromatic compounds in sediments of the Ebro River Basin. *Anal. Bioanal. Chem.* 385(6):1020-1030.
- Palácio HAQ, Andrade EM, Lopes FB, Alexandre DMB, Arraes FDD (2009). Similarity of surface water quality at the Curu watershed, Ceará, Brazil. *Cienc. Rural* 39(9):2494-2500.
- Palácio HAQ, Neto JRA, Meireles ACM, Andrade EM, Santos JCN, Chaves LCG (2011). Similarity and determining factors in the salinity of surface waters in Ceará by multivariate techniques. *Rev. Bras. Eng. Agr. Amb.* 15(4):395-402.
- Pejman AH, Nabi Bidhendi GR, Karbassi AR, Mehrdadi N, Bidhendi ME (2009). Evaluation of spatial and seasonal variations in surface water quality using multivariate statistical techniques (2009). *Int. J. Environ. Sci. Tech.* 6(3):467-476.
- Pereira-Filho J, Rörig LR, Schettini CAF, Soppa MA, Santana BL, dos Santos JE (2010). Spatial Changes in the water quality of Itajaí-Açu Fluvial-Estuarine System, Santa Catarina, Brazil. *An. Acad. Bras. Ciênc.* 82(4):963-982.
- Silva AMM, Sacomani LB (2001). Using chemical and physical parameters to define the quality of Pardo River water (Botucatu-SP-Brazil). *Water Resour.* 35(6):1609-1616.
- Simeonov V, Stratis JA, Samara C, Zachariadis G, Voutsas D, Anthemidis A, Sofoniou M, Kouimtzis T (2003). Assessment of the surface water quality in Northern Greece. *Water Res.* 37(17):4119-4124.
- Singh KP, Malik A, Mohan D, Sinha S (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India): A case study. *Water Res.* 38(18):3980-3992.
- Terrado M, Kuster, M, Raldúa, D, Alda, ML, Barceló D, Tauler R (2007). Use of chemometric and geostatistical methods to evaluate pesticide pollution in the irrigation and drainage channels of the Ebro River Delta during the rice-growing season. *Anal. Bioanal. Chem.* 387(4):1479-1488.
- Toledo LG, Nicoletta G (2002). Water quality index for agricultural and urban watershed use. *Sci. Agric.* 59(1):181-186.
- Tucci CEM, Hespanhol I, Cordeiro Netto OM (2000). Scenarios of Water Management in Brazil: A contribution to the World Water Vision. *Rev. Bras. Rec. Hídric.* 5:31-43.
- Vidal M, López A, Santoalla MC, Valles V (2000). Factor analysis for the study of water resources contamination due to the use of livestock slurries as fertilizer. *Agric Water Manage.* 45(1):1-15.
- Wu B, Zhao D, Zhang Y, Zhang X, Cheng S (2009). Multivariate statistical study of organic pollutants in Nanjing reach of Yangtze River. *J. Hazard. Mater.* 169:1093-1098.
- Zeilhofer P, Rondon-Lima EB, Lima GAR (2006). Spatial patterns of water quality in the Cuiabá River Basin, Central Brazil. *Environ. Monit. Assess.* 123:41-62.