

# ANÁLISE E PREDIÇÃO DA QUALIDADE DAS ÁGUAS SUPERFICIAIS NA MICROBACIA DO CÓRREGO SEGREDO - MS.

Lucio Henrique Pereira<sup>1</sup>; Flávio Palmiro<sup>2</sup>, Ruben Barros Godoy<sup>3</sup>, João Onofre Pereira Pinto<sup>4</sup>

**RESUMO** – A ferramenta de Análise de Componentes Principais (PCA) foi empregada com o objetivo de identificar os fatores determinantes (naturais e antrópicos) da qualidade das águas superficiais na microbacia do córrego Segredo, Campo Grande - MS. Em seguida, a regressão por componente principal (PCR) foi utilizada para verificar a viabilidade de predição dos parâmetros variando nove entradas com uma saída desejada. Realizaram-se coletas trimestrais, no período de Fevereiro de 2010 a Março de 2011, em 10 pontos localizados nas áreas de maior influência urbana. Foram analisados os parâmetros pH, temperatura, turbidez, nitrogênio total, oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO), fósforo total, coliformes termotolerantes, sólidos totais e precipitação semanal sendo totalizadas 80 amostras. Através das técnicas de estatística multivariada, observou-se que as 6 primeiras Componentes Principais explicaram 90% da variância total. A correlação entre as variáveis mostrou-se coerente para bacias localizadas em centros urbanos. Utilizando a PCR todas as regressões se mostraram satisfatórias e obteve-se o menor erro quadrático médio de 0.285 para o fósforo total.

**Palavras chave:** Qualidade da água, Estatística multivariada, Recursos Hídricos .

## ANALYSIS AND PREDICTION OF SURFACE WATER QUALITY IN THE SMALL WATERSHED OF STREAM SEGREDO - MS.

**ABSTRACT:** The tool Principal Component Analysis (PCA) was used in order to identify the factors (natural and anthropogenic) of surface water quality in the Segredo stream small watershed, in Campo Grande - MS. After this, Principal Component Regression (PCR) was used to verify the feasibility for parameters prediction from the variation of nine input parameters and one desired output. The database was formed with collections from February 2010 up to March 2011, at 10 points located in areas of urban influence. The parameters were pH, Temperature, Turbidity, Total Nitrogen, Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), Total Phosphorus, Thermotolerant Coliforms, Total Solids and Weekly Precipitation, which totaled 80 samples. Through multivariate statistical techniques, it was observed that the first 6 principal components explained 90% of the total variance. The correlation among the variables was consistent for basins located in urban center. Using PCR all regressions present satisfactory and the lowest mean square error was 0.285 for Total Phosphorus.

**Keywords:** Water quality, multivariate statistics, Water Resources.

1) Engenheiro Ambiental, Mestrando em Engenharia Universidade Federal de Mato Grosso do Sul (UFMS). Rua brilhante, 1499 - Campo Grande – MS. E-mail: lucio@batlab.ufms.br

2) Engenheiro Eletricista, Mestrando em Engenharia Universidade Federal de Mato Grosso do Sul (UFMS). Rua barão do rio branco, 2350 - Campo Grande – MS. E-mail: flaviopalmiro@gmail.com

3) Prof. Dr. do Departamento de Engenharia Elétrica da Universidade Federal de Mato Grosso do Sul (UFMS), Cidade Universitária - CEP: 79070-900 - Campo Grande – MS. E-mail: ruben@batlab.ufms.br

4) Prof. Dr. Prof. Dr. do Departamento de Engenharia Elétrica da Universidade Federal de Mato Grosso do Sul (UFMS), Cidade Universitária - CEP: 79070-900 - Campo Grande – MS. E-mail: joao.pinto@ufms.br

## 1 INTRODUÇÃO

A qualidade da água é um termo usado para descrever as características químicas, físicas e biológicas da água em relação à sua aptidão para uma determinada utilização. A qualidade da água é afetada por uma ampla gama de influências naturais e antropogênicas. As águas de superfície são mais vulneráveis à poluição, devido à sua fácil acessibilidade para a eliminação de águas residuais (Singha *et al.* 2004). Processos naturais, tais como, precipitação, erosão, resistência dos materiais da crosta terrestre, bem como, as influências antropogênicas, como, atividades industriais, agrícolas e resíduos urbanos, em conjunto com o aumento da exploração dos recursos hídricos, determinam a qualidade das águas superficiais na região.

Os cursos d'água desempenham um papel importante na assimilação e transporte das águas que recebem efluentes industriais e esgotos. Variações na precipitação, escoamento superficial, fluxo de águas subterrâneas e os bombeamentos têm um forte efeito sobre a vazão do rio e, posteriormente, sobre a concentração de poluentes na água do rio (Vega *et al.*, 1998). Logo, os rios constituem uma das principais fontes de recursos hídricos para uso doméstico, industrial e irrigação. É imprescindível prevenir e controlar a poluição dos rios e dispor de informações confiáveis sobre a qualidade da água para uma gestão eficaz.

As descargas de águas residuais e industriais constituem uma fonte poluidora constante. Para entender a dinâmica do processo de uma bacia hidrográfica é necessário uma rede de monitoramento da qualidade da água bem concebida. De fato, para se avaliar o estado da qualidade das águas superficiais, é necessário que programas de monitoramento das características físicas, químicas e biológicas sejam implementados, quer em uma visão de variabilidade espacial ou temporal (Brodnjak-*et al.*, 2002; Singh *et al.*, 2005).

O resultado de programas de análise da qualidade ambiental é composto por um extenso número de informações, quase sempre de difícil interpretação. Por outro lado, há necessidade do controle das fontes de poluição das águas e a definição do manejo a ser aplicado requerem a identificação das fontes poluentes e suas contribuições qualitativas e quantitativas (Jonnalagadda e Mhere 2001). Na tentativa de minimizar ou resolver esse problema, técnicas de estatística multivariada como Análise da Componente Principal/Regressão Componente Principal vêm sendo aplicadas em diferentes partes do globo (Andrade *et al.* 2007; Oualia *et al.* 2009; Tiviroli *et al.* 2011; ).

O objetivo desse estudo é analisar e prever a qualidade das águas superficiais na microbacia do córrego segredo – MS onde os dados foram submetidos à análise estatística multivariada visando identificar os fatores determinantes, naturais e antrópicos, e suas relações.

## 2 METODOLOGIA DO ESTUDO

### 2.1 Caracterização da área de estudo

A região de estudo localiza-se na microbacia do córrego segredo no município de Campo Grande – MS (Figura 1), localizado predominantemente na Bacia Hidrográfica do Rio Paraná, com exceção de uma porção Noroeste de seu território que se situa na Bacia Hidrográfica do Rio Paraguai (PLANURB,

2006). A rede hidrográfica da área urbana é constituída por onze microbacias: Anhandui, Coqueiro, Bandeira, Bálsamo, Gameleira, Imbiruçu, Lagoa, Lageado, Prosa, Ribeirão Botas e Segredo (Consórcio RES, 2008). A atividade predominante na região são os aglomerados populacionais.

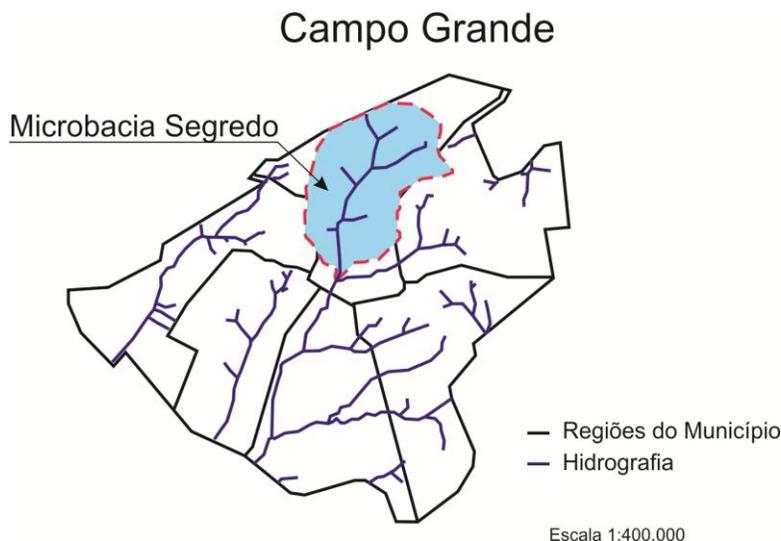


Figura 1: Localização geográfica da Microbacia do córrego segredo.

## 2.2 Monitoramento e análise da qualidade da água

As coletas ocorreram durante o período de jan/2010 a dez/2011, sendo realizadas trimestralmente em 10 pontos distintos, totalizando 8 campanhas e 80 amostras, em pontos de densidade urbana (Figura 2). De montante para jusante os pontos são: SEG 01, SEG 08, SEG 06, SEG 07, SEG 09, CAS 01, CAS 02, SEG 02, SEG 03 E SEG 04 . Os parâmetros estudados foram: pH, temperatura, turbidez, nitrogênio total, dissolvido (OD), demanda bioquímica de oxigênio (DBO<sub>5,20</sub>), fósforo total, coliformes totais e sólidos totais. Outro fator estudado foi a precipitação, pois a mesma tem uma influência direta na qualidade dos recursos hídricos.

Os dados foram retirados do projeto córrego limpo realizado pela prefeitura municipal de Campo Grande. As análises das amostras de água seguiram a metodologia apresentada por (APHA 1998). Dados pluviométricos foram extraídos da estação DNOS - 8.DRS, gerenciada pela Agência Nacional de Águas (ANA) , os dados utilizados foram o somatório da precipitação semanal antecedente a dada da coleta.

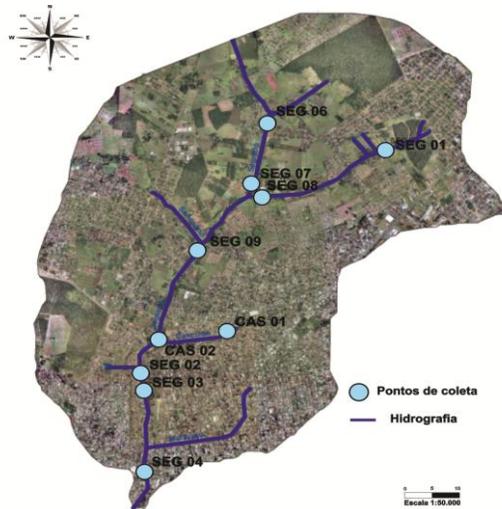


Figura 2: Pontos de coleta na microbacia do córrego segredo.

### 2.3 Componente Principal (PCA)

A associação entre as variáveis que influenciam a qualidade das águas superficiais foi identificada mediante as técnicas de estatística multivariada, Análise Fatorial/Análise do Componente Principal (AF/ACP) com o *software* Matlab<sup>R</sup>.

Os principais objetivos das componentes principais são reduzir o número de variáveis e analisar quais as variáveis ou quais conjuntos de variáveis explicam a maior parte da variabilidade total, revelando que tipo de relacionamento existe entre eles (Bouroche e Saporta 1982)

As novas variáveis, as CP's, são combinações lineares das variáveis originais e derivadas em ordem decrescente de importância tal que, por exemplo, a primeira componente principal é a combinação linear normalizada com variância máxima (Jonhson e Wichern 1988).

### 2.4 Regressão componente principal (PCR)

Em estatística, PCR é uma análise de regressão que usa análise de componentes principais para estimar coeficientes de regressão. É um procedimento utilizado para superar os problemas que surgem quando as variáveis experimentais estão perto de serem colineares.

Na PCR em vez de regredir a variável dependente das variáveis independentes diretamente, são utilizados os CP's das variáveis independentes. Tipicamente utiliza apenas um subconjunto dos CP's na regressão, fazendo uma espécie de estimativa regularizada. Na maioria das vezes os CP's com a maior variância são seleccionados.

A PCR é um método de regressão que pode ser dividido em duas etapas. O primeiro passo é executar uma PCA em cima das mesmas das variáveis explicativas, o segundo passo é executar uma regressão dos mínimos quadrados ordinários sobre as componentes seleccionados: os fatores que estão mais correlacionados com a variável dependente serão seleccionados.

Os dados coletados foram separados em duas categorias: dados de treinamento, que são utilizados para o treinamento dos dados e os de teste, que são utilizados para verificar sua performance sob condições reais de utilização. A porcentagem escolhida foi de 70% para treinamento e 30% para o teste, sendo os mesmos escolhidos aleatoriamente. A regressão foi realizada para todos os parâmetro em busca do melhor resultado.

### 3 RESULTADOS E DISCUSSÕES

Os dados, por apresentarem variação nas unidades de medida e no desvio padrão, foram normalizados. A normalização utilizada foi a de variância 1 e média 0.

A matriz de correlação gerada possui todos os 10 parâmetros analisados, explicando assim a variância total (Tabelas 2). Pela referida matriz, observa-se que apenas 16% das variáveis estudadas apresentaram correlação superior a 0,5 com pelo menos uma variável. Um coeficiente de correlação superior a 0,5 expressa uma forte relação. (Helena *et al.* 2000)

O coliforme total foi o parâmetro de maior correlação com os demais, sendo este fortemente correlacionado com a DBO, e os nutrientes (fósforo e nitrogênio). A maior correlação foi entre a turbidez e sólidos suspensos, de quase 0,7, pois a turbidez das águas é devida à presença de partículas em estado coloidal, de natureza orgânica ou inorgânica e outros organismos microscópicos. O pH obteve os menores índices de correlação, correlacionando-se somente com a temperatura.

Tabela 1: Matriz de correlação das variáveis indicadoras da qualidade.

	Coliformes	pH	DBO	Fósforo Total	Nitrogênio Total	Temperatura	Turbidez	Sólidos Totais	OD	Precipitação
Coliformes	1,00									
pH	-0,02	1,00								
DBO	<b>0,65</b>	0,01	1,00							
Fosforo Total	<b>0,50</b>	-0,01	<b>0,64</b>	1,00						
Nitrogenio Total	<b>0,53</b>	0,05	0,45	0,43	1,00					
Temperatura	0,17	<b>-0,54</b>	0,16	0,08	0,43	1,00				
Turbidez	0,20	-0,35	-0,01	0,04	0,34	0,36	1,00			
Sólidos Totais	0,10	-0,21	0,16	0,17	0,42	0,35	<b>0,69</b>	1,00		
OD	-0,13	0,20	-0,17	-0,02	-0,26	-0,38	0,07	0,01	1,00	
Precipitação	0,09	-0,14	0,19	0,04	0,14	<b>0,52</b>	0,22	0,31	-0,17	1,00

A precipitação traz outra constatação importante para o estudo da região pois a mesma está predominantemente relacionada com a temperatura, sólidos totais e turbidez; mostrando que as chuvas carregam consigo a sujeira das cidades porém essa não possui grande quantidade de matéria orgânica e nutrientes.

A evolução da porcentagem explicada por cada componente principal foi de: 33 % PC 1, 20 % PC 2, 13 % PC 3, 9 % PC 4, 8 % PC 5, 7 % PC 6, 3 % PC 7, 3 % PC 8, 2 % PC 9, 2 % PC 10 respectivamente. Logo, as 6 primeiras componentes explicam 90% dos dados. O peso fatorial das variáveis nas 6 primeiras CP's selecionadas (Figura 3) traz a influência dos parâmetro em cada nova componente. Os valores elevados dos pesos fatoriais sugerem quais são as variáveis mais significativas em cada fator.

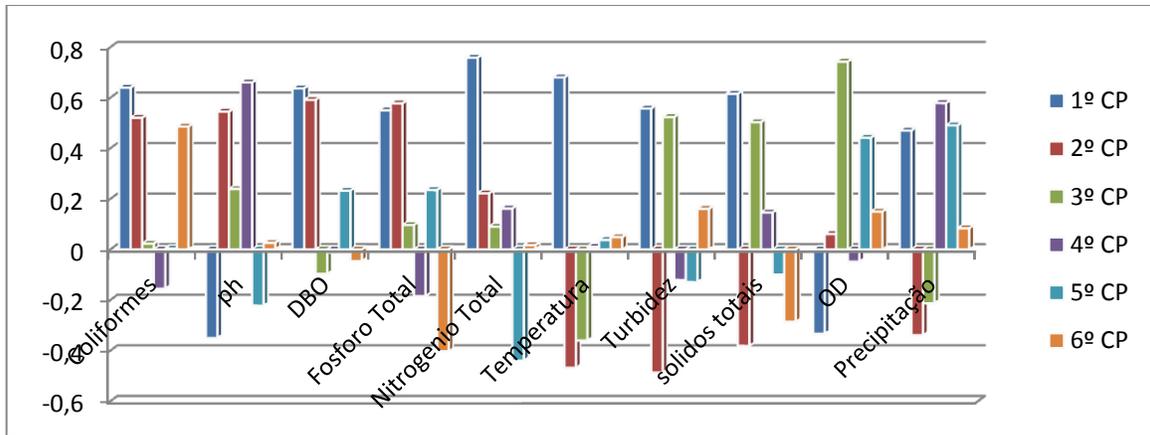


Figura 3: Peso das variáveis em cada componente.

Observa-se que as 9 entradas estão bem representadas na primeira componente, devido a variabilidade das informações, conseqüentemente nenhuma entrada é irrelevante para compor a base de dados. O nitrogênio, entrada 5, foi o mais representativo na 1º CP, fato esse que pode ser explicado devido a microbacia estar localizada em um centro urbano e exposta a lançamentos de poluentes, seguido da DBO e dos coliformes, porém o mesmo é o segundo menos significativo na 2º CP e diminui consecutivamente, fato natural para decomposição em CP's. Já na 2º CP foi observado que a variavel turbidez, entrada 7, foi a mais representativa, e a menor foi o oxigenio dissolvido, porém para a 3º CP o mesmo foi o de maior importância para sua composição.

O gráfico biplot (Figura 4) apresenta a formação de 3 grupos distintos de correlação positiva, sendo o fósforo, nitrogênio, DBO e coliformes no 1º quadrante; precipitação, turbidez, sólidos totais e temperatura no 4º quadrante e o pH e o OD no 2º quadrante. Fica clara também a relação de positividade do OD com a DBO, pois, na maioria das vezes onde há uma elevação da DBO ocorre um decréscimo da disponibilidade de oxigênio (von Sperling, 2005).

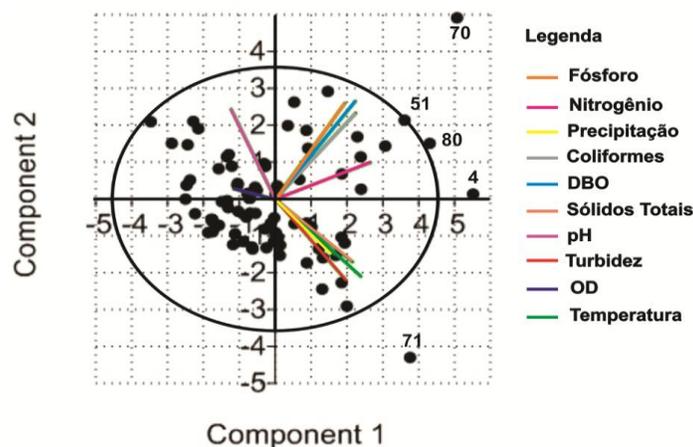


Figura 4: Gráfico biplot da primeira pela segunda

As amostras 4, 51, 70, 71 e 80 ultrapassaram a distribuição normal com intervalo de confiança de 95%. Estas estão divididas entre os pontos de coleta SEG 01, SEG 04 e CAS 01. onde devarão ser

obeservados com maior atenção auxiliando na busca de medidas preventivas para lançamentos clandestinos, causas naturais ou problemas de amostragem e medição.

Na análise da regressão para pevisão da qualidade da água, foram utilizadas somente as 6 primeiras PC's. Sempre usando 9 entradas e uma saída, variando-as com o objetivo e observar a que melhor se adaptaria ao processo. O erro quadratico médio (EQM) variou de 0.28 à 0.81(Tabela 2). O EQM foi calculado com os dados ainda normalizados. Os parâmetros que apresentaram melhor resultado foram: fósforo total, DBO<sub>5,20</sub>, nitrogênio e turbidez respectivamente (Figura 5).

Tabela 2: Erro Quadrático Médio

	Coliformes	pH	DBO	Fosforo Total	Nitrogênio	Temperatura	Turbidez	Solidos Totais	OD	Precipitação
EQM	0.704	0.784	0.347	0.285	0.355	0.464	0.396	0.666	0.607	0.807

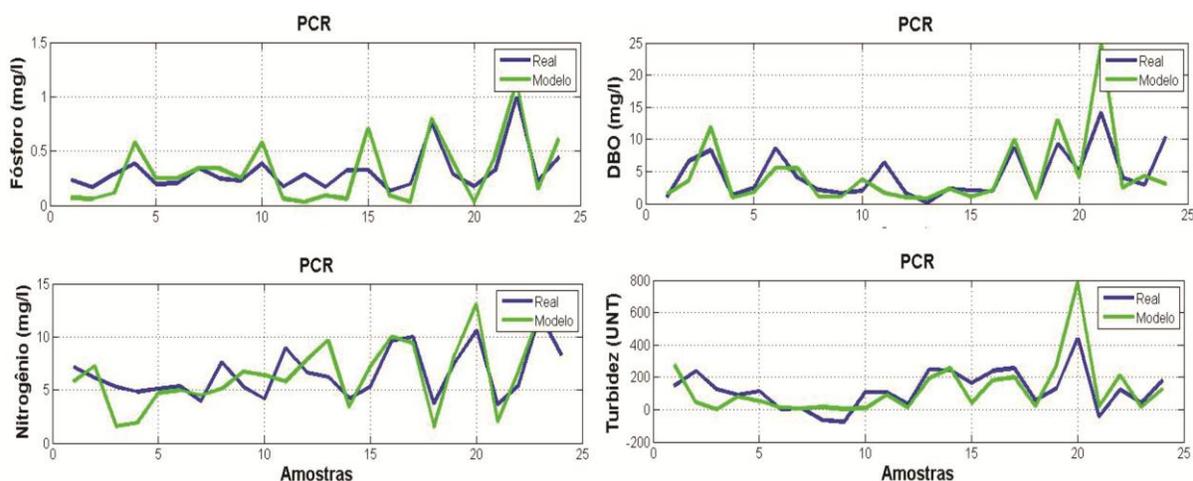


Figura 5: Regressões PCR para os 4 melhores resultados.

Os resultados acima estão desnormalizados, expressos em escala real. Apesar da previsão não ser exata a ferramenta mostrou-se eficaz para prever as tendências dos parâmetros. Mediante a impossibilidade ou erros de amostragem e medição a modelagem PCR pode ser uma saída eficaz para o problema. Se faz necessário uma maior quantidade de dados para aprimorar uma função que melhor defina o possível comportamento da qualidade da água do córrego segredo.

## 4 CONCLUSÕES

A Análise de Componente Principal se mostrou uma importante ferramenta auxiliando na melhor compreensão da interação dos parâmetros de qualidade da água do córrego lagoa em Campo Grande - MS. A técnica indicou os parâmetros que mais influenciam na qualidade da água e a formação de três grupos bem correlacionados. Pôde-se perceber também a influência direta das chuvas na temperatura, e no transporte de nutrientes e matéria orgânica.

O emprego da PCA também promoveu uma redução para 6 CP's das 10 variáveis iniciais, que explicam 90% da variância total dos dados. A predição através do método PCR mostrou-se satisfatória, principalmente para o fósforo total, com um EQM de 0.285, demonstrando precisão em prever a tendência do mesmo, utilizando-se das outras 9 características como entradas para o sistema. Sugestões: Utilizadas os dados de outros cursos d'água e outras regressões para outros parâmetros.

## REFERÊNCIAS

- ANDRADE, E. M.;1; PEREIRA, L. A.; ROSA, M. F.;GOMES, B.; LOBATO, F. A. (2007) Fatores determinantes da qualidade das águas superficiais na bacia do Alto Acaraú, Ceará, Brasil. *Ciência Rural*, Santa Maria, v. 37, n. 6.
- APHA, A. WPCF, 2005. Standard methods for the examination of water and wastewater, 21.
- BOUROCHE, J- M; SAPORTA, G. (1982) *Análise de Dados*. Rio de Janeiro, Zahar, PP. 117
- BRODNJAK-VONCINA, D.; DOBČNIKA, D. NOVIČB, M.; ZUPANBET, J. (2002) Chemometrics characterisation of the quality of river water. *Analytica Chimica Acta*, v.462, p.87-100.
- HELENA B.; PARDO, F; VEJA, M.; BARRADO, H.; FERNANDE J. M.; FERNANDEZ L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal component analysis. *Water Research*, v.34, n.3, p.807-816, 2000.
- JOHNSON, R. A., WICHERN, D. W. (1988). *Applied multivariate statistical analysis*. 2. ed. Englewood Cliffs, Prentice Hall, pp. 608.
- JONNALAGADDA, S.B.; MHERE, G (2001) Water quality of the Odzi river in the eastern highlands os Zimbabwe. *Water research* , v.35, n.10, pp. 2371 – 2376.
- OUALIA, A., AZRIB, C.; MEDHIOUBB, K.; GHRABIA, A. (2009) Descriptive and multivariable analysis of the physico-chemical and biological parameters of Sfax wastewater treatment plant. *Desalination*, v.246, pp. 496-505.
- PLANURB – Instituto Municipal de Planejamento Urbano e Meio Ambiente. *Perfil socioeconômico de Campo Grande*, Edição 2006 - 2007. 13 ed. ver. Campo Grande: PMCG, 193p. 2006.
- SINGH, K.P.; MALIKA, A.;MOHANA, D.; SINHAB S. (2004) Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. *Water Research*, v.38, pp. 3980–3992.
- SINGH, K.P; MALIKA,A.; SINHABET, S. (2005). Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques-a case study. *Analytica Chimica Acta*,v.515, p.143-149.
- TIVIROLI, V. A.;ROCHE, K. F.; FREIRE, H. M. F.; MARTINS R. M. Padrões interanuais, espaciais e sazonais da qualidade da água na bacia do rio taquari – MS . In *Anais do XIX Simpósio de Recursos Hídricos*, Maceió, Nov. 2011.
- VEGA, M.;PARDO, R.; BARRADO, B.; DEBÁN L. (1998) Assesment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research*, v.32, n.12, p.3581-3592.
- VON SPERLING, M. (2005). *Introdução à qualidade da água e ao tratamento de esgotos*. v. 1. Belo Horizonte: DESA-UFMG.