

ANÁLISE DA QUALIDADE DA ÁGUA NO RESERVATÓRIO DA UHE NOVA PONTE UTILIZANDO FERRAMENTAS ESTATÍSTICAS MULTIVARIADAS

Déborah Tavares Viana^{1}, Cristiano Cristofaro², Mônica M. D. Leão³, Silvia M.A.C. Oliveira³*

Resumo – O monitoramento dos recursos hídricos deve contemplar as mais diversas situações, exigindo a coleta em muitas estações distintas e considerando diversos parâmetros de qualidade. Nesse estudo serão avaliados dados físico-químicos de qualidade de água da sub-bacia do Araguari por meio de ferramentas estatísticas multivariadas para identificar aquelas regiões mais comprometidas pelas ações antrópicas e as possíveis fontes de alteração dos parâmetros de qualidade. Para este trabalho foram utilizados os dados de cinco estações no período de 1998 a 2010. A partir dos grupos gerados na Análise de Cluster (AC) o número de variáveis de cada grupo foi reduzido por meio da técnica da Análise Fatorial (AF), buscando ressaltar os parâmetros de qualidade mais importantes em cada grupo de estações de monitoramento. As cinco estações foram agrupadas em três grupos. A AF identificou 5 fatores significativos para os Grupos I, II, III. Os grupos encontrados na AF demonstram que as estações de monitoramento estudadas sofrem alterações advindas de fontes diversas. As técnicas multivariadas se mostraram eficientes para caracterização e identificação das principais fontes de variação na qualidade da água da sub-bacia PN2 do rio Paranaíba. Os resultados gerados podem vir a contribuir para a definição de possíveis ações de recuperação e conservação.

Palavras-Chave – Monitoramento, estatística multivariada, Araguari.

WATER QUALITY ANALYSIS IN THE HPP NOVA PONTE RESERVOIR USING MULTIVARIATE STATISTICAL TOOLS

Abstract – Water resources monitoring should include the most diverse situations, requiring the sampling in many distinct stations and considering several quality parameters. In this study physicochemical of the sub-basin Araguari will be evaluated through multivariate statistical tools to identify regions that are most affected by human actions and the possible sources of the quality parameters alteration. In this work we used five stations data in the period of 1998 to 2010. From the groups created in Cluster Analysis (CA) the number of variables in each group was reduced by the technique of Factor Analysis (FA), seeking to highlight the most important quality parameters in each group. The five stations were grouped into three groups. The AF has identified five significant factors for Groups I, II, III. Groups found in AF show that the monitoring stations undergo changes that come from various sources. The multivariate techniques were effective for the characterization and identification of the main sources of variation in the water quality of the river Paranaíba sub basin - PN2. The generated results may contribute to the definition of recovery and conservation actions.

Keywords – Monitoring, multivariate statistics, Araguari.

1 - Mestranda do Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos (PG-SMARH)

2 - Professor Adjunto, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM)

3 - Professor Adjunto/Associado, Universidade Federal de Minas Gerais (UFMG), PG-SMARH.

* autor correspondente: deborahtviana@gmail.com

INTRODUÇÃO

Com o desenvolvimento de novas tecnologias e o aumento da urbanização os recursos naturais tornam-se cada vez mais escassos. Dessa forma, é necessário investir em programas e ações que visem a conservação da fauna, recursos hídricos, dentre outros. O monitoramento da qualidade da água entra nesse contexto como uma ferramenta para auxiliar na gestão dos recursos hídricos, permitindo a caracterização e análises de tendências em bacias hidrográficas (ANA, 2013).

O comportamento do corpo d'água pode variar de acordo com diversos fatores como a carga de poluentes, usos do solo e características da região em que se encontra. Assim, o monitoramento de recursos hídricos deve contemplar as mais diversas situações, exigindo a coleta de um grande número de dados ao longo do tempo em muitas estações de coleta distintas e considerando diversos parâmetros de qualidade. Desse modo, as ferramentas estatísticas multivariadas apresentam grande relevância, pois permitem a análise simultânea e a redução de conjuntos complexos de dados. A aplicação dessas ferramentas como a Análise de Cluster ou Agrupamento (AC), Análise de Componentes Principais (ACP) e Análise de Fatores (AF) podem auxiliar na interpretação de matrizes de dados complexos para melhor entendimento da qualidade da água. Atualmente, diversos trabalhos utilizam técnicas multivariadas para a identificação das causas e fontes de poluição pontual e difusa, embasando a adoção de ações cientificamente fundamentadas e eficazes (REGHUNATH et al, 2002; LI et al, 2011; ZHOU et al, 2007; ZHANG et al, 2011).

Nesse estudo serão avaliados dados físico-químicos de qualidade de água da sub-bacia do Araguari por meio de ferramentas estatísticas multivariadas para identificar regiões que estão mais comprometidas pelas ações antrópicas e as possíveis fontes dos parâmetros de qualidade.

METODOLOGIA

ÁREA DE ESTUDO

Para este trabalho foram utilizados os dados físico-químicos para cinco estações sub-bacia PN2 da bacia do Paranaíba no período de 1998 a 2010 (Fig.1). A bacia do Paranaíba contém alguns remanescentes do Cerrado, Cerradão e Floresta Estacional Semidecidual. Os solos apresentam baixa fertilidade natural, portando há grande uso de calagem e adubação. Predominam Latossolos e Argissolos na bacia, sendo muito comum a associação com argila. Nas áreas com altitudes menores encontra-se Latossolo Vermelho e Vermelho-Amarelo, que possuem elevada acidez, presença de alumínio e elevados teores de Fe₂O₃. Sobre o uso do solo, predominam usos diretamente relacionados à agropecuária e mineração. A atividade agrícola mais comum por toda a região é a plantação de cana de açúcar.

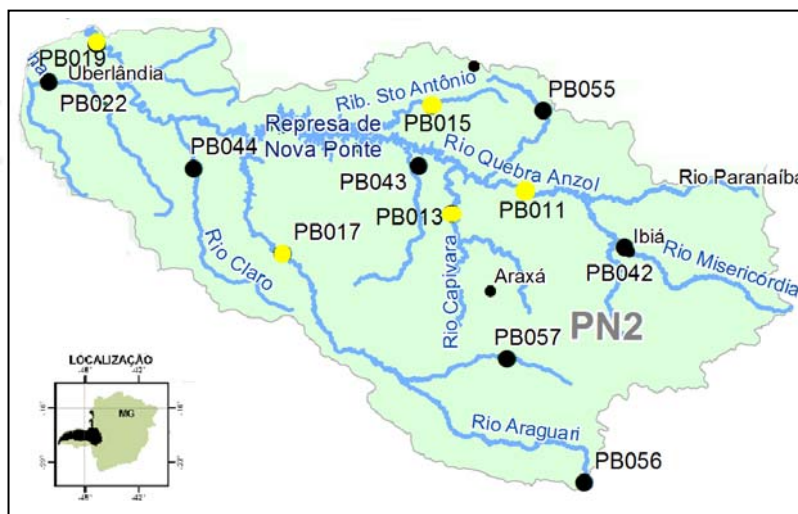


Figura 1: Mapa da Região da PN2 com a localização das estações

ANÁLISE

Foram selecionados trinta parâmetros de qualidade de água monitorados semestralmente durante doze anos. Aplicou-se uma técnica de agrupamento (Análise de Cluster – AC), com ligação completa, nos dados físico-químicos selecionados a fim de buscar grupos homogêneos de estações de monitoramento.

A partir dos grupos gerados o número de variáveis de cada grupo foi reduzido por meio da técnica da Análise Fatorial (AF), buscando ressaltar os parâmetros de qualidade mais importantes em cada grupo de estações de monitoramento. A AF foi realizada para cada um dos grupos encontrados a AC, utilizando-se do método de extração Análise de Componentes e a rotação *varimax*. A seleção do número de fatores considerou a quebra de ângulo do gráfico *Scree* (HAIR et al. 2009), de modo a selecionar apenas os fatores mais significativos de cada grupo de estações de monitoramento. Para a seleção das cargas significativas foi adotada a definição de Liu et al (2003) que considera cargas maiores que 0,75 fortes, entre 0,75-0,50 moderadas e entre 0,50-0,30 fracas. Iremos considerar portanto cargas moderadas e fortes.

Para a AC e AF os dados foram estandardizados de acordo com a escala Z (média 0 e desvio padrão 1) para eliminar as interferências das diferentes escalas dos parâmetros analisados e não resultar agrupamentos errôneos ou enviesados. As análises foram conduzidas no software STATISTICA - data analysis software system - StatSoft Inc. (2007) versão 8.0 e XLSTAT - Addinsoft SARL (2013).

RESULTADOS E DISCUSSÃO

O banco de dados analisado contém 128 coletas e 3835 observações, distribuídas em 5 estações de monitoramento e 30 parâmetros. A estatística descritiva de todos os parâmetros pode ser visualizada na tabela 1. Verifica-se que os parâmetros apresentam distribuição assimétrica, portanto foram considerados não paramétricos. A curtose nos mostra que os dados possuem grande variação, já que a maioria dos parâmetros apresentou a distribuição platicúrtica, ou achatada.

A análise de cluster, utilizando o método de ligação completa das distâncias euclidianas, foi utilizada para detectar as similaridades entre as estações. As cinco estações foram agrupadas em três grupos, conforme pode ser visto no dendrograma (Fig.2). O grupo I, mais afastado dos demais, inclui o ponto PB013, o grupo II inclui os pontos PB019 e PB015 e

o grupo III inclui os pontos PB017 e PB011. Esse agrupamento aponta principalmente a diferença do PB013 com as demais estações, conforme pode ser visto no dendrograma, que mostra a que distância de ligação do PB013 é a maior dentre todas as outras distâncias.

Tabela 1: Estatística descritiva para os 30 parâmetros de qualidade da água

Parâmetro	Abreviação	Mediana	Percentil 10%	Percentil 90%	Assimetria	Curtose
Alcalinidade total	Alc	8.2	4.4	15.2	2.71	14.65
AmôniaNI	Amôn	0.0002	0.0001	0.0010	4.16	26.25
Bário total	Ba	0.04	0.01	0.20	0.66	-1.03
Cálcio total	Ca	2.18	1.20	3.48	1.51	3.33
Cloreto total	Clor	0.57	0.30	4.06	2.48	5.92
Coliformes termotolerantes	Colif	300	30	5000	9.53	99.39
Condutividade elétrica in loco	CE	20.2	11.8	52.8	2.66	7.30
Cor verdadeira	Cor	30	5	160	7.81	73.84
Demanda Bioquímica de Oxigênio	DBO	2	2	3	3.94	18.20
Demanda Química de Oxigênio	DQO	8	5	24	2.96	10.09
Dureza total	Dur	9.3	5.1	14.8	1.47	3.09
Ferro dissolvido	Fe	0.12	0.03	0.27	1.36	2.71
Fósforo total	P	0.05	0.01	0.23	3.23	15.48
Magnésio total	Mg	0.79	0.40	1.67	1.94	5.11
Manganês total	Mn	0.05	0.03	0.15	2.87	11.36
Nitrato	NO3	0.08	0.02	0.25	1.90	4.82
Nitrito	NO2	0.003	0.002	0.006	3.55	16.79
Nitrogênio amoniacal total	N-NH3	0.10	0.10	0.30	7.11	60.12
Nitrogênio orgânico	N	0.20	0.10	0.50	2.62	8.79
Oxigênio dissolvido	OD	7.2	6.3	8.2	-0.08	0.23
pH in loco	pH	6.5	5.8	7.1	-0.05	-0.89
Potássio dissolvido	K	0.77	0.54	2.32	2.53	6.59
Sódio dissolvido	Na	1.11	0.75	3.76	3.49	12.91
Sólidos dissolvidos totais	SDT	25	19	60	1.68	2.44
Sólidos em suspensão totais	SST	24	3	258	3.44	13.80
Sólidos totais	ST	53	24	295	3.34	13.10
Temperatura da água	Tag	24	21	26	0.46	1.65
Temperatura do ar	Tar	26	20	31	-0.20	-0.84
Turbidez	NTU	21.00	3.16	270.00	4.20	21.85
Zinco total	Zn	0.02	0.02	0.07	5.89	38.39

A análise de fatores (AF) foi realizada para cada um dos agrupamentos encontrados na AC com extração de análise de componentes e rotação varimax. A AF identificou 5 fatores significativos para os Grupos I, II, III. Os resultados das cargas para cada grupo pode ser visto de forma resumida na tabela 2. A partir desses resultados podemos rotular os fatores de acordo com o que os parâmetros podem indicar sobre o uso do solo e características naturais.

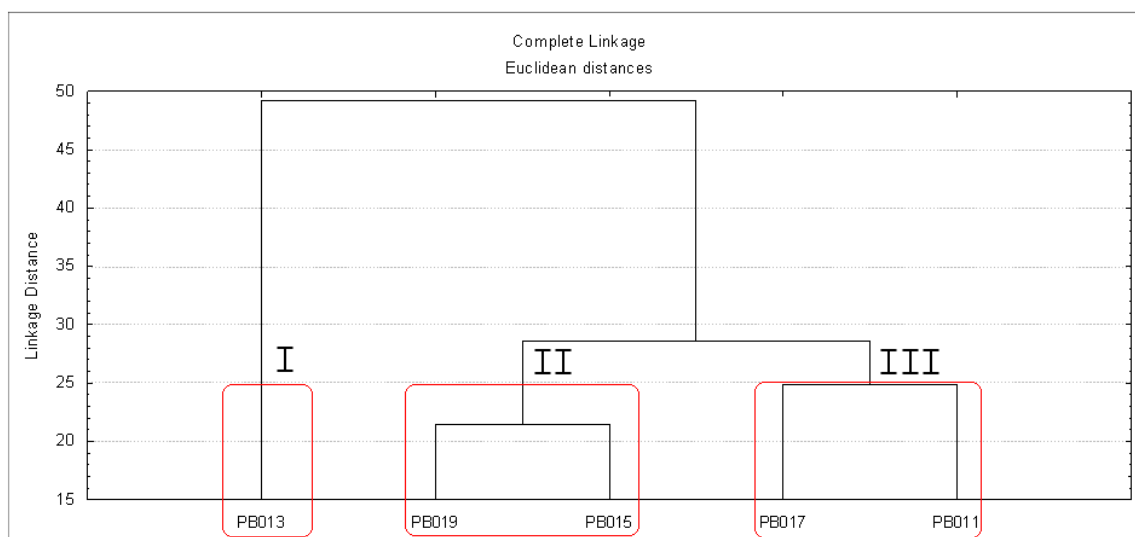


Figura 2: Dendrograma das distâncias de ligação da análise de cluster

No grupo I o fator 1 explica 29,8% da variância total, podendo estar relacionado a fontes difusas que causariam o arraste superficial de materiais particulados e matéria orgânica do solo. O fator 2 responde por 19,9% da variância e pode ser caracterizado como a lixiviação de materiais inorgânicos dissolvidos, que podem ser provenientes de fontes naturais. O fator 3 (11,3% da variância) pode ser caracterizado por fonte antrópica mista com influência de esgotos e fertilizantes. O fator 4 (7,8% da variância) pode indicar características naturais da região. Por último o fator 5 (5,3% da variância) pode ser caracterizado como componentes orgânicos advindos de fontes pontuais de esgotos.

Para o grupo II o fator 1 explica 23,1% da variância total e possivelmente originados de fontes naturais. O fator 2 (12,3% da variância total) é composto por partículas sólidas e contaminação fecal, indicando a participação de fonte antrópica contendo efluentes sanitários. O fator 3 (11,3% da variância total) pode ser considerado como uma fonte mista com parâmetros de origem natural na região, representados pelo Ba e Zn e fontes antrópicas de DBO. O fator 4 (7,8% da variância total) está relacionado com processos naturais de dissolução de O₂ na água, já que a temperatura altera de forma inversa a dissolução de O₂ na água. Por último o fator 5 (6,3% da variância total) está relacionado à presença de compostos nitrogenados, possivelmente relacionados a fontes agropecuárias.

Para o grupo III o fator 1 explica 30,1% da variância total e pode ser interpretado como um grupo composto por fontes difusas antrópicas, relacionadas com materiais em suspensão e dissolvidos na água, e fontes pontuais, relacionadas a compostos orgânicos. O fator 2 (10,1% da variância total) pode ser interpretado como poluição advinda de esgoto. O fator 3 indica uma baixa influência de fontes de amônia, cuja concentração é diretamente afetada pelo pH. O fator 4 (6,8% da variância total) pode ser denominado características naturais das rochas locais, indicando baixa dureza e alcalinidade. O fator 5 (6,0% da variância total) demonstra a influência da temperatura local e da água. e o último fator como a temperatura.

Tabela 2: Cargas dos parâmetros por fator (os valores em negrito indicam cargas negativas)

Cargas	Parâmetros				
	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5
Grupo 1					
Forte	Cor	CE	Clor	Tar	Colif
	SST	Dur	NO3		DBO
	ST	P	K		
	NTU	Na			
		Zn			
Moderada	DQO	Ca	Fe	Ba	Amôn
	Mn	SDT	N-NH3		DQO
	N				Mn
					N
					pH
% Variância	29.8	19.9	11.3	7.8	5.3
% Var. Acumulada	29.8	49.7	61.0	68.8	74.1
Grupo 2					
forte	Alcal	DQO		Tág	Amôn
	Clor	SST			N -NH3
	CE	ST			
	K	NTU			
moderada	Ca	Colif	Ba	Tar	
	Na	Cor	DBO	OD	
			pH		
			Zn		
% Variância	23.1	12.3	9.5	7.8	6.2
% Var. Acumulada	23.1	35.4	44.9	52.7	58.9
Grupo 3					
forte	DQO	Clor	pH	Dur	Tág
	P	K		Mg	Tar
	Mn				
	SDT				
	SST				
	ST				
	NTU				
moderada	Alcal		Amôn	Alcal	
	Ba			Ca	
	Ca				
	CE				
	N				
% Variância	30.1	10.1	7.8	6.8	6.0
% Var. Acumulada	30.1	40.2	48.0	54.8	60.8

Os grupos encontrados na AF demonstram que as estações de monitoramento estudadas sofrem alterações advindas de fontes diversas. O grupo I apresentou nos seus fatores mais significativas características que podem ser relacionadas com arraste de fertilizantes e particulados de agropecuária. Esse fato condiz com a realidade da região já que essa é extensivamente ocupada por plantações e pastagens (IBGE, 2013).

Já o grupo II teve no seu fator mais representantes de íons e características naturais das águas, enquanto que o segundo fator mais importantes indicou a influência de fontes antrópicas pontuais urbanas. Esse fato pode ser devido ao ponto PB019 ser um rio de curso largo, o que causa a diluição dos poluentes e o ponto PB015 estar numa região mais preservada, permitindo certo grau de autodepuração de suas águas.

Por fim, o grupo III apresentou nos primeiros componentes os parâmetros relacionados com material particulado de arraste dos solos e material dissolvido de características naturais. O ponto PB011 está ao lado de uma extração de areia, o que causa revolvimento do solo, liberando partículas suspensas. Já o ponto PB017 está ao lado de uma queda d'água, o que pode também ser fonte de partículas suspensas do solo.

CONCLUSÃO

As técnicas multivariadas, Análise de Agrupamentos e Análise dos Fatores, se mostraram eficientes para caracterização e identificação das principais fontes de variação na qualidade da água da sub-bacia PN2 do rio Paranaíba. A AC permitiu agrupar os pontos que apresentavam pressões semelhantes em três grupos distintos facilitando assim a análise posterior. A Análise de Fatores permitiu um melhor entendimento dos motivos da diferenciação entre os grupos, fornecendo informações sobre as possíveis fontes mais influentes em cada um. Assim, o grupo I foi composto por uma estação sob maior influência de fontes difusas agrícolas. O grupo II foi formado por estações que se encontram sob influência de áreas urbanas. O terceiro grupo mostrou ter influência de atividades de mineração. Desse modo, verifica-se que as análises multivariadas apresentam Desse modo, o estudo demonstrou a importância das análises multivariadas na análise e interpretação de bancos de dados complexos, permitindo uma avaliação satisfatória da qualidade da água na bacia estudada. Os resultados gerados podem vir a contribuir para a definição de possíveis ações de recuperação e conservação.

REFERÊNCIAS

- AGÊNCIA NACIONAL DAS ÁGUAS - ANA. *Portal da Qualidade das Águas*. Disponível em < http://pnqa.ana.gov.br/rede/rede_avaliacao.aspx>. Acesso em: 30 abr. 2013.
- IBGE. Censo Agropecuário. 1995-2011, Instituto Brasileiro de Geografia e Estatística, 2013. Disponível em: < <http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/default.shtm> >
- KAZI, T. G. et al. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. **Ecotoxicology and Environmental Safety**, v. 72, n. 2, p. 301-309, 2// 2009. ISSN 0147-6513. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0147651308000705> >.

LIU, Chen-Wuing. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. **The Science of The Total Environment**, Taiwan, n. 313, p.77-89, 2003.

REGHUNATH, Rajesh et al. The utility of multivariate statistical techniques in hydrogeochemical studies:an example from Karnataka, India. **Water Research**, India, v. 36, n. , p.2437-2432, 2002.

SINGH, K. P. et al. 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, n. 18, p. 3980-3992, 11// 2004. ISSN 0043-1354. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0043135404003367> >.

ZHANG, Xuan et al. Application of multivariate statistical techniques in the assessment of water quality in the Southwest New Territories and Kowloon, Hong Kong. **Environ Monit Assess**, China, n. 173, p.17-27, 2011.

ZHOU, Feng et al. Application of Multivariate Statistical Methods to Water Quality Assessment of the Watercourses in Northwestern New Territories, Hong Kong. **Environ Monit Assess**, China, n. 132, p.1-7, 2007.