



Aplicação da análise estatística multivariada no estudo da qualidade da água do Rio Pomba, MG

**Hugo A. S. Guedes², Demetrius D. da Silva², Abrahão A. A. Elesbon²,
Celso B. M. Ribeiro³, Antonio T. de Matos² & José H. P. Soares³**

RESUMO

O objetivo deste estudo foi avaliar a qualidade da água e identificar os grupos de poluição presentes no médio Rio Pomba, por meio de técnica estatística multivariada. Duas campanhas no período de out/2008 a jan/2009 foram realizadas em nove pontos georreferenciados, localizados ao longo do médio Rio Pomba, compreendendo a análise de 18 variáveis de qualidade de água. A técnica estatística multivariada por meio da aplicação da análise fatorial/análise de componentes principais promoveu a redução no número de variáveis de qualidade de água, uma vez que o melhor comportamento das variáveis ocorreu com a inclusão de 15 das 18 variáveis analisadas. Pelo emprego da análise fatorial/análise de componentes principais identificou-se que o melhor comportamento das variáveis de qualidade das águas do médio Rio Pomba foi aquele composto por três fatores, explicando 74,30% da variância total. As variações na qualidade da água no médio Rio Pomba foram definidas por um grupo de nutrientes associado ao esgoto doméstico e à poluição difusa; por um grupo orgânico, causado pelo lançamento de esgoto doméstico no curso de água e por um grupo de sólidos em suspensão, expressando o processo de erosão hídrica na bacia.

Palavras-chave: análise multivariada, aspectos limnológicos, análise fatorial, análise de componentes principais, qualidade de água em rios

Application of multivariate statistical analysis in the study of water quality in the Pomba River (MG)

ABSTRACT

The aim of this study was to evaluate the water quality and identify groups of pollution in the Middle Pomba River through multivariate statistical technique. There were two campaigns from Oct/2008 to Jan/2009 in nine geo-referenced points along the Middle Pomba River including the analysis of 18 variables of water quality. The multivariate statistical technique, through the application of factor analysis/principal component analysis, caused a decrease in the number of variables of water quality, since the best performance of the variables occurred with the inclusion of 15 of the 18 variables. By the use of factor analysis/principal components analysis, it was found that the best behavior of the variables of the water quality of the Middle Pomba River waters was the one composed of three factors, explaining 74.30% of the total variance. Changes in water quality of the Middle Pomba River were defined by a group of nutrients associated with sewage and diffused pollution; for an organic group, caused by untreated sewage in the water stream, and by a group of suspended solids, expressing the process of water erosion in the basin.

Key words: multivariate analysis, limnological aspects, factor analysis, principal component analysis, river water quality

¹ Parte da Dissertação de Mestrado desenvolvida pelo primeiro autor. Projeto financiado pela CAPES, FAPEMIG e CNPq

² DEA/UFV, Av. PH Rolfs s/n, Campus Universitário, CEP 36570-000, Viçosa, MG. Fone: (31) 3899-3471. E-mail: hugo.guedes@ufv.br; demetrius@ufv.br; abrahaoelesbon@hotmail.com; atmatos@ufv.br

³ ESA/UFJF, Campus Universitário, Faculdade de Engenharia, CEP 36036-900, Juiz de Fora, MG. Fone: (32) 2102-3419, Ramais: 23 e 30. Fax: (32) 2102-3402. E-mail: celso.bandeira@ufjf.edu.br; homero.soares@ufjf.edu.br

INTRODUÇÃO

Os recursos naturais usados pelo homem em atividades agrícolas, áreas urbanas e na indústria, aliados aos processos naturais, como a variação da precipitação, intemperismo das rochas e erosão, alteram a qualidade das águas tornando-as inadequadas para usos mais exigentes (Mendiguchía et al., 2004).

Analisando as fontes de poluição, pode-se dizer que a agricultura e outras atividades agropecuárias, principalmente os resíduos da bovinocultura e suinocultura (Moraes & Paula Júnior, 2004), têm importante influência na contaminação dos mananciais, compreendendo atividades de alto potencial degradador (Merten & Minela, 2002). Cabe ressaltar, também, que o uso e o manejo do solo em áreas agrícolas acarretam alterações na qualidade das águas superficiais (Andrade et al., 2007a).

O monitoramento da qualidade da água é um dos principais instrumentos de sustentação de uma política de planejamento e gestão de recursos hídricos, visto que funciona como um sensor que possibilita o acompanhamento do processo de uso dos corpos hídricos, apresentando seus efeitos sobre as características qualitativas das águas, visando subsidiar as ações de controle ambiental. De acordo com Simeonov et al. (2003), um programa de monitoramento ambiental inclui, em geral, coletas frequentes nos mesmos pontos de amostragem e análise em laboratório de grande número de variáveis, resultando em matriz de grandes dimensões e complexa interpretação. Muitas vezes, um pequeno número dessas variáveis contém as informações mais relevantes enquanto a maioria das variáveis adiciona pouco ou nada à interpretação dos resultados, em termos qualitativos.

Normalmente, o monitoramento ambiental gera grandes quantidades de dados que são difíceis de analisar e interpretar, pois as relações entre as variáveis são complexas (Liao et al., 2008). Deste modo, abordagens multivariadas, como Análise Fatorial (AF) e Análise de Componentes Principais (ACP), têm sido utilizadas com sucesso para apoiar a gestão dos recursos hídricos e extrair informações significativas a partir dessas bases de dados (Lambrakis et al., 2004; Zeng & Rasmussen, 2005; Papatheodorou et al., 2006; Melo Júnior et al., 2006; Andrade et al., 2007a,b; Liao et al., 2008; Zhang et al., 2010).

Os métodos de análise de dados multivariados são amplamente utilizados quando se deseja promover a redução do número de variáveis com o mínimo de perda de informação (Andrade et al., 2007a; Helena et al., 2000). O método dos componentes principais é um dos mais usados para resolver problemas clássicos de análise fatorial. Esta análise permite a redução do número de variáveis facilitando a extração de informações que serão de grande relevância na avaliação da qualidade das águas e no manejo da bacia (Andrade et al., 2007a).

Deste modo, aplicou-se a análise estatística multivariada - Análise Fatorial/Análise de Componentes Principais, com o objetivo de avaliar a qualidade da água e identificar os grupos de poluição presentes no médio Rio Pomba.

MATERIAL E MÉTODOS

Região de estudo

Esta pesquisa foi desenvolvida na Bacia hidrográfica do Rio Pomba (Figura 1), afluente da margem esquerda do Rio Paraíba do Sul. Esta bacia se localiza na região Sudeste, nos estados de Minas Gerais, MG e Rio de Janeiro, RJ, e detém uma área de drenagem de 8.519,94 km² entre as coordenadas de 20° 52' e 21° 43' S e 41° 59' e 43° 38' W, abrangendo 38 municípios mineiros e 3 municípios fluminenses, nos quais vive uma população de aproximadamente 600 mil habitantes.

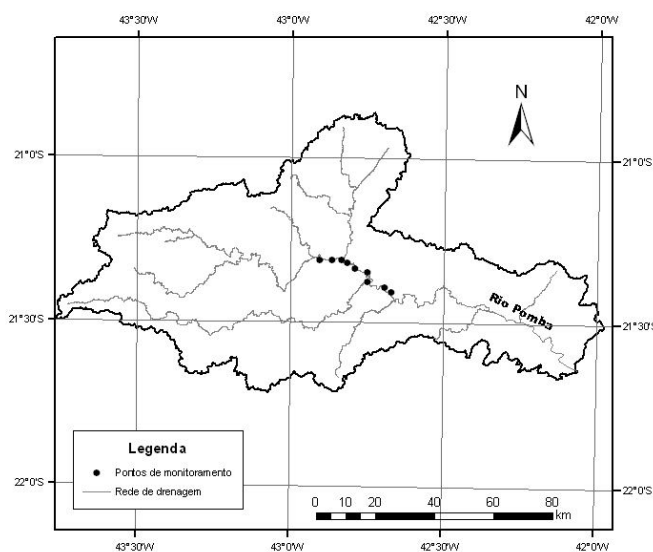


Figura 1. Bacia hidrográfica do Rio Pomba

O Rio Pomba nasce nas proximidades de Barbacena, MG, na Serra da Mantiqueira, em trecho denominado Serra da Conceição, a 1.100 m de altitude, percorrendo aproximadamente 290 km no sentido sudeste até sua confluência com o Rio Paraíba do Sul. Ao longo deste percurso recebe vários afluentes, sendo os mais relevantes os da margem esquerda, isto é, os rios Paraopeba e Xopotó e, da margem direita, os rios Formoso, Novo, Pardo e o Ribeirão dos Monos.

O clima da bacia é o tipo Tropical Quente e Úmido, com verões quentes e chuvosos e invernos com estiagem de 4 a 5 meses. Nos pontos de altitude mais elevada os verões são brandos e o clima é classificado como Tropical Superúmido, sem seca ou com subseca. A precipitação média anual dessa região é de 1.200 a 1.600 mm e a temperatura média anual varia de 17,4 a 24,7 °C.

O trecho denominado médio Rio Pomba se inicia no município de Astolfo Dutra percorrendo o município de Dona Euzébia e o vilarejo de Sinimbu, finalizando no município de Cataguases; todos esses municípios pertencem ao estado de Minas Gerais; neste percurso se encontram os dois maiores afluentes na bacia os rios Xopotó e Novo.

Monitoramento da qualidade da água

Visando avaliar a qualidade das águas do médio Rio Pomba, realizaram-se duas coletas de amostras de água, em nove pontos de monitoramento (Figura 1) durante o período de out/2008 a jan/2009.

Tabela 1. Características dos pontos de monitoramento no médio Rio Pomba

Identificadores	Estações amostrais	Comprimento (km)*	Altitude (m)	Coordenadas geográficas (UTM) - Zona 23	
P1	Início	0,0	215	717413	7641968
P2	Astolfo Dutra	6,0	209	721781	7642055
P3	Rio Xopotó	9,6	208	724833	7641946
P4	Dona Euzébia	13,0	207	726948	7641029
P5	Faz. Cachoeira do Pomba	17,4	206	729459	7639024
P6	Sinimbu	22,0	205	733524	7637561
P7	Rio Novo	29,9	198	733537	7634439
P8	Cataguases	38,4	165	739454	7632657
P9	Taquara Preta	42,8	163	741655	7630991

* Comprimento em relação ao primeiro ponto de coleta ("Início")

As coletas de amostras de água para caracterização físico-química foram realizadas no dia 16 de outubro de 2008, representando o período seco e, no dia 19 de janeiro de 2009, sinalizando o período chuvoso. O trecho de estudo, totalizando 42,8 km de extensão, se inicia a seis quilômetros à montante do município de Astolfo Dutra, seção de monitoramento denominada "Início", percorrendo os municípios de Astolfo Dutra e Dona Euzébia e finalizando no município de Cataguases, no bairro Taquara Preta; mencionados municípios pertencem ao estado de Minas Gerais (Tabela 1). O critério adotado para se eleger os pontos de monitoramento foi a existência de áreas rurais, aglomerados urbanos e de indústrias.

Durante as coletas utilizou-se um frasco de PVC, com volume aproximado de 1.000 mL de água, a qual foi transferida para garrafas plásticas e as amostras armazenadas em caixa de isopor, a uma temperatura média de 4 °C e conduzidas ao Laboratório de Qualidade da Água, do Departamento de Engenharia Agrícola (DEA), na Universidade Federal de Viçosa (UFV), onde foram determinadas as concentrações de 18 variáveis: oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO₅), pH, temperatura, condutividade elétrica, alcalinidade, demanda química de oxigênio (DQO), fósforo total, nitrogênio total, nitrogênio orgânico, nitrogênio amoniacal, nitrato, nitrogênio total Kjeldahl, sólidos totais, sólidos suspensos totais, sólidos dissolvidos totais, sólidos inorgânicos suspensos e detritos. A variável temperatura foi medida in loco, com auxílio de um termômetro com sensibilidade de 0,1 °C; as demais análises laboratoriais seguiram a metodologia descrita por APHA (2005).

Análises estatísticas

A associação entre as variáveis que influenciam a qualidade da água do médio Rio Pomba foi identificada mediante o uso de análise estatística multivariada - Análise Fatorial/Análise de Componente Principal (AF/ACP). Para a análise dos dados utilizou-se o software Statistica®, versão 7.0.

A AF/ACP demanda três etapas: a primeira consiste na preparação da matriz de correlação [R]; a segunda na extração dos fatores comuns e a possível redução de variáveis explicativas e, por último, na rotação dos eixos relativos aos fatores comuns, com o objetivo de tornar a solução mais simples e facilmente interpretável (Toledo & Nicoletta, 2002).

Segundo Wilks (2006), após a construção da matriz original de dados deve-se analisar a ordem de grandeza das variáveis uma vez que variáveis com ordem de grandeza discrepante

devem ser padronizadas, para evitar que o fator de escala influencie nas decisões finais de estudo, dando maior ênfase a uma variável de escala maior (com maior módulo da variância em seus dados) em relação a uma variável de escala menor.

Deste modo, a análise dos componentes principais consistiu no cálculo dos autovalores e correspondentes autovetores de uma matriz de correlações [R]. Nesta matriz simétrica os elementos da diagonal principal foram as variâncias das frequências absolutas padronizadas de cada classe (variáveis padronizadas) e os demais elementos os valores de covariância entre pares de variáveis padronizadas (Melo Júnior et al., 2006).

A padronização das variáveis foi feita dividindo-se os desvios em relação à média de cada variável ($X_{ij} - \bar{X}_j$) por seu desvio padrão $S(X_j)$ como é apresentado na Eq. 1.

$$Y_{ij} = \frac{X_{ij} - \bar{X}_j}{S(\bar{X}_j)} \quad (1)$$

em que Y_{ij} , $S(X_j)$ e X_j são, respectivamente, a variável padronizada, o desvio-padrão e a média da j-ésima variável original.

Os dados padronizados em forma de matriz foram expressos por $X = (x_{ij})$, em que $i = 1 \dots n$ amostragens (18) e $j = 1 \dots p$ variáveis (18). De posse dessa matriz de dados determinou-se a matriz de correlação [R] ($p \times p$), para p igual a 18 variáveis de qualidade de água, analisadas nesta pesquisa. A principal razão para se usar a matriz de correlação é se eliminar o problema de escalas e unidades diferenciadas em que as variáveis são medidas. A eficiência da descrição dos dados por meio das componentes vai depender da porcentagem de variação total que cada componente contém, a qual é representada pelos autovalores (Hair et al., 2005).

Do ponto de vista estatístico-matemático, a análise dos componentes principais busca transformar o conjunto original das variáveis observadas em um novo conjunto de variáveis, denominadas componentes principais. A principal característica dessas componentes, além da ortogonalidade, é que são obtidas em ordem decrescente de máxima variância. Portanto, são calculadas de forma que a primeira componente principal explique o máximo da variabilidade total dos dados; a segunda explique o máximo da variabilidade restante dos dados, sendo não-correlacionada com a primeira; a terceira explique o máximo da variabilidade total restante dos dados, sendo não-correlacionada com a primeira e a segunda componentes e assim

Tabela 2. Matriz de correlação das variáveis de qualidade de água no médio Rio Pomba

Variável	OD	DBO	Temp	CE	Alcalin	DQO	PT	Norg	NH ₄	NTK	NT	ST	SST	SDT	SIS
OD	1,00														
DBO	-0,58	1,00													
Temp	-0,57	0,18	1,00												
CE	0,02	-0,24	0,61	1,00											
Alcalin	-0,23	-0,34	0,63	0,77	1,00										
DQO	-0,70	0,32	0,64	0,12	0,38	1,00									
PT	-0,09	0,31	0,03	0,21	-0,11	-0,10	1,00								
Norg	0,16	0,23	-0,62	-0,49	-0,71	-0,49	0,28	1,00							
NH ₄	-0,76	0,53	0,38	-0,01	0,12	0,31	0,48	0,08	1,00						
NTK	0,08	0,28	-0,57	-0,49	-0,69	-0,45	0,33	0,99	0,18	1,00					
NT	0,15	0,27	-0,55	-0,34	-0,62	-0,50	0,35	0,97	0,13	0,97	1,00				
ST	0,24	-0,06	-0,23	0,33	-0,02	-0,61	0,47	0,29	0,08	0,29	0,37	1,00			
SST	0,01	0,35	-0,39	-0,14	-0,26	-0,43	0,60	0,61	0,37	0,64	0,63	0,63	1,00		
SDT	0,30	-0,38	0,04	0,55	0,20	-0,42	0,09	-0,15	-0,22	-0,17	-0,06	0,75	-0,04	1,00	
SIS	0,01	0,26	-0,47	-0,19	-0,37	-0,43	0,54	0,61	0,32	0,63	0,63	0,75	0,84	0,24	1,00

OD (mg L⁻¹) – oxigênio dissolvido; DBO (mg L⁻¹) – demanda bioquímica de oxigênio; Temp (°C) – temperatura; CE (μS cm⁻¹) – condutividade elétrica; Alcalin (mg L⁻¹ CaCO₃) – alcalinidade; DQO (mg L⁻¹) – demanda química de oxigênio; PT (mg L⁻¹ - PO₄) – fósforo total; Norg (mg L⁻¹ - N) – nitrogênio orgânico; NH₄ (mg L⁻¹ - N) – nitrogênio amoniacal; NTK (mg L⁻¹ - N) – nitrogênio Kjeldahl; NT (mg L⁻¹ - N) – nitrogênio total; ST (mg L⁻¹) – sólidos totais; SST ((mg L⁻¹) – sólidos suspensos totais; SDT (mg L⁻¹) – sólidos dissolvidos totais; SIS (mg L⁻¹) – sólidos inorgânicos suspensos

sucessivamente, até que o número de componentes principais seja no máximo igual ao número de variáveis.

De acordo com Andrade et al. (2007a), nem sempre a matriz das componentes principais obtida na fase de extração é de fácil interpretação. Este fato ocorre quando as cargas dos fatores estão próximas do valor médio. Para suplantar tal limitação, utilizou-se do procedimento de rotação ortogonal da matriz das cargas fatoriais, gerando uma nova matriz de cargas fatoriais que apresenta melhor significado interpretativo aos fatores. Nesta pesquisa empregou-se a rotação ortogonal utilizando-se o método Varimax. Outras informações sobre o assunto podem ser encontradas em Dillon & Goldstein (1984), Rencher (2002) e Hair et al. (2005).

RESULTADOS E DISCUSSÃO

A matriz de correlação [R] composta das variáveis estudadas no médio Rio Pomba, pode ser vista na Tabela 2. A análise de componentes principais resultou na redução de 3 das 18 variáveis analisadas. As variáveis limnológicas nitrato, pH e detrito, não se mostraram significativas na explicação da variância total, podendo ser desconsideradas no estudo de qualidade de água. Neste tipo de análise é interessante diminuir o espaço amostral uma vez que, reduzindo o número de variáveis de qualidade de água, reduz-se também o número de análises laboratoriais, economizando tempo e recursos, sendo que esses recursos poderiam ser alocados em outras fases do monitoramento dos recursos hídricos (Zeng & Rasmussen, 2005).

Observa-se, na Tabela 2, que todas as variáveis apresentam correlação superior a 0,5 com pelo menos duas outras variáveis de qualidade de água. Segundo Helena et al. (2000), coeficientes de correlação superiores a 0,5 expressam uma forte relação entre as variáveis de qualidade de água. Pode-se observar, também, que NT, Norg e NTK apresentaram correlações superiores a 0,9. As variáveis Temp, NT e SST apresentaram maiores coeficientes de correlação com maior número de variáveis.

Constata-se correlação muito forte e positiva entre as variáveis Norg e NTK ($r = 0,99$), Norg e NT ($r = 0,97$), NTK e NT

($r = 0,97$) e SST e SIS ($r = 0,84$). As correlações muito fortes entre as frações de nitrogênio eram esperadas, uma vez que estão correlacionadas, assim como as frações de sólidos.

A segunda etapa da análise fatorial concerne à decomposição da matriz de correlação tendo como objetivo identificar os grupos de poluição no médio Rio Pomba representados pelos fatores comuns e independentes (componente principal). O melhor comportamento das variáveis de qualidade de água no médio Rio Pomba foi aquele composto por três componentes (Tabela 3). A seleção do número de componentes teve como base os princípios sugeridos por Jolliffe (2002), ou seja, que a porcentagem acumulada da variância total entre 70 e 90% oferece uma idéia razoável da representação da variância original. Adotou-se, como critério para a seleção dos componentes principais, a porcentagem acumulada da variância total dos dados originais de 70%.

Nota-se, na Tabela 3, que as três componentes (CP1, CP2 e CP3), determinadas pela AF/ACP, explicam 74,30% da variância total das variáveis originais, componentes que expressam a

Tabela 3. Matriz de peso fatorial das variáveis da qualidade de água nos três componentes principais selecionados

Variável	CP1	CP2	CP3
Nitrogênio Kjeldahl	0,93	0,13	0,09
Nitrogênio Orgânico	0,93	0,05	0,14
Nitrogênio Total	0,91	0,04	0,01
Sólidos Inorgânicos Suspensos	0,80	-0,03	-0,39
Sólidos Suspensos Totais	0,79	0,10	-0,47
Alcalinidade	-0,73	-0,09	-0,51
Temperatura	-0,70	0,37	-0,44
OD	0,20	-0,85	0,24
DBO	0,29	0,77	-0,13
Nitrogênio Amoniacal	0,14	0,71	-0,52
Sólidos Dissolvidos Totais	-0,04	-0,69	-0,42
DQO	-0,61	0,67	0,00
CE	-0,49	-0,31	-0,70
Fósforo Totais	0,42	0,14	-0,68
Sólidos Totais	0,49	-0,47	-0,64
Autovalor	6,18	3,45	3,00
% variância explicada	36,36	20,29	17,65
% variância acumulada	36,36	56,65	74,30

CP – Componente principal

relação entre fatores e variáveis e permitem identificar as variáveis com maiores inter-relações em cada componente. Os valores elevados dos pesos fatoriais sugerem quais são as variáveis mais significativas em cada fator.

A primeira componente explica 36,36% da variabilidade dos dados e está associada a variáveis indicativas de cargas de nutrientes provenientes de esgoto doméstico e uso de compostos nitrogenados. Na elaboração de um índice de qualidade de água em uma microbacia de uso agrícola e urbana no município de Guaíra, SP, Toledo & Nicolella (2002) também encontraram, na primeira componente da análise fatorial, um número maior de variáveis, sendo as de maior peso o fósforo e o nitrogênio amoniacal.

As componentes CP2 e CP3 explicaram 20,29 e 17,65% da variância total da matriz de variáveis padronizadas, respectivamente. Pode-se observar que as componentes CP1 e CP2 apresentaram variáveis comuns na explicação da variabilidade das variáveis de qualidade de água no médio Rio Pomba, o que torna a matriz de fatores de difícil interpretação. Para superar esta limitação aplicou-se a transformação ortogonal dos eixos pelo emprego do algoritmo Varimax. Este processo maximiza a variância entre os fatores, alterando a raiz característica sem afetar a proporção da variância total explicada pelo conjunto (Andrade et al., 2007a). Pesquisadores como Silva & Sacomani (2001) e Melo Júnior et al. (2006) obtiveram uma matriz de interpretação mais fácil com a aplicação do algoritmo Varimax (Tabela 4).

Observa-se, na Tabela 4, uma redistribuição melhor da variância total entre as três componentes. Comparando-se as Tabelas 3 e 4, verifica-se redução do percentual da variância total explicada pelas componentes CP1 e CP2 além de um consequente aumento do percentual da variância, explicado pela componente CP3, sem ocorrer variação do total explicado pelas variáveis. A primeira componente rotacionada (FV1) passou a explicar 31,89% da variância enquanto a terceira componente rotacionada (FV3) passou a contribuir com 23,11 em substituição aos 17,65% registrados antes da transformação ortogonal.

Tabela 4. Fatores dos componentes da matriz transformada pelo algoritmo Varimax

Variável	FV1	FV2	FV3
Alcalinidade	-0,89	0,04	0,05
Nitrogênio Orgânico	0,88	-0,01	0,27
Nitrogênio Kjeldahl	0,87	0,08	0,29
Nitrogênio Total	0,80	0,03	0,38
CE	-0,80	-0,07	0,42
Temperatura	-0,77	0,48	-0,09
OD	0,16	-0,93	0,11
Nitrogênio Amoniacal	-0,02	0,84	0,24
DBO	0,31	0,78	0,00
DQO	-0,43	0,63	-0,48
Sólidos Totais	0,07	-0,15	0,96
Sólidos Inorgânicos Suspensos	0,42	0,17	0,75
Sólidos Dissolvidos Totais	-0,31	-0,36	0,72
Fósforo Total	0,06	0,31	0,61
Sólidos Suspensos Totais	0,47	0,19	0,60
Autovalor	5,42	3,28	3,54
% variância explicada	31,89	19,30	23,11
% variância acumulada	31,89	51,19	74,30

FV – Fator Varimax

No primeiro fator as variáveis alcalinidade, nitrogênio orgânico, Kjeldahl e total, CE e temperatura, apresentaram peso superior a 0,75, indicando que essas variáveis são as mais significativas na definição da qualidade da água do médio Rio Pomba. De acordo com Liu et al. (2003), fatores que possuem peso superior a 0,75 são indicativos de forte carga entre as variáveis de qualidade de água; assim, este fator representa o indicativo de poluição difusa no médio Rio Pomba, causada por fertilizantes orgânicos oriundos de áreas agrícolas, principalmente entre as estações amostrais de Astolfo Dutra e Sinimbu e, também, por nutrientes presentes no esgoto doméstico.

O segundo fator é explicado pelas variáveis OD, nitrogênio amoniacal, DBO e DQO, as quais indicam a poluição por matéria orgânica no eixo longitudinal no médio Rio Pomba. No Brasil e de maneira geral, assim como na bacia do Rio Pomba, apenas 47% dos municípios têm rede coletora de esgotos e somente 18% dos esgotos recebem algum tratamento.

O FV3 é explicado pelas variáveis sólidos totais, sólidos inorgânicos suspensos, sólidos dissolvidos totais, fósforo total e sólidos suspensos totais as quais refletem, basicamente, o escoamento superficial oriundo de áreas agrícolas e a contribuição de esgotos domésticos e resíduos sólidos dispostos inadequadamente nas margens do curso de água. Segundo informações do Censo 2000 do Instituto Brasileiro de Geografia e Estatística (IBGE, 2002), 40% dos domicílios da área urbana e 99% da área rural realizam a queima, enterram ou dispõem o lixo em terreno baldio, rio, lago e mar. A influência da erosão hídrica também foi identificada por Singh et al. (2005) em estudos de qualidade de água e aporte de poluição do Rio Gomti, na Índia. O comportamento das variáveis mostrou que, para o primeiro componente, foram significantes as variáveis relacionadas com os sais dissolvidos na água enquanto para o segundo componente os elementos relacionados com a poluição agrícola apresentaram maior significância e, para o terceiro, os maiores pesos foram atribuídos ao escoamento superficial.

CONCLUSÕES

1. A técnica estatística multivariada por meio da aplicação da análise fatorial/análise de componentes principais promoveu redução no número de variáveis de qualidade de água, uma vez que o melhor comportamento das variáveis ocorreu com a inclusão de 15 das 18 variáveis analisadas.

2. A análise fatorial/análise de componentes principais permitiu a seleção de três componentes indicadores da qualidade das águas do médio Rio Pomba, explicando 74,30% da variância total.

3. As variações na qualidade da água no médio Rio Pomba foram definidas por um grupo de nutrientes (influência de fontes de poluição difusa, como drenagem das áreas agrícolas e esgoto doméstico), por um grupo orgânico (indicativo de lançamento de esgotos domésticos in natura no curso de água) e por um grupo de sólidos (representando o processo de erosão hídrica).

AGRADECIMENTOS

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro a esta pesquisa.

LITERATURA CITADA

- Andrade, E. M. de; Araújo, L. de F. P.; Rosa, M. de F.; Disney, W. Alves, A. B. Seleção dos indicadores da qualidade das águas superficiais pelo emprego da análise multivariada. *Engenharia Agrícola*, v.27, p.683-690, 2007a.
- Andrade, E. M. de; Araújo, L. de F. P.; Rosa, M. de F.; Gomes, R. B.; Lobato, F. A. de O. Fatores determinantes da qualidade das águas superficiais na Bacia do Alto Acaraú, Ceará, Brasil. *Ciência Rural*, v.37, p.1791-1797, 2007b.
- APHA - American Public Health Association. Standard methods for the examination of water and wastewater. 21.ed., New York: AWWA, WPCR. 2005. 937p.
- Dillon, W. R.; Goldstein, M. Multivariate analysis methods and application. New York: John Wiley. 1984. 587p.
- Hair, J. J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. Análise multivariada de dados. 5.ed. Porto Alegre: Bookman, 2005. 593p.
- Helena, B.; Pardo, R.; Vega, M.; Barrado, E.; Fernández, J. M.; Fernández, L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. *Water Research*, v.34, p.807-816, 2000.
- IBGE - Instituto Brasileiro de Geografia e Estatística. Indicadores de desenvolvimento sustentável: Brasil 2002. Rio de Janeiro: IBGE, 2002. 195p.
- Jolliffe, L. T. Principal component analysis. 2.ed. New York: Springer, 2002. 487p.
- Lambrakis, N.; Antonakos, A.; Panagopoulos, G. The use of multicomponent statistical analysis in hydrogeological environmental research. *Water Research*, v.38, p.1862-1872, 2004.
- Liao, S. W.; Gau, H. S.; Lai, W. L.; Chen, J. J.; Lee, C. G. Identification of pollution of Tapeng Lagoon from neighbouring rivers using multivariate statistical method. *Journal of Environmental Management*, v.88, p.286-292, 2008.
- Liu, C. W.; Lin, K. H.; Kuo, Y. M. Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Twain. *Science in the Total Environment*, v.313, p.77-89, 2003.
- Melo Júnior, J. C. F. de; Sedyama, G. C.; Ferreira, P. A.; Leal, B. G. Determinação de regiões homogêneas quanto à distribuição de frequência de chuvas no leste do Estado de Minas Gerais. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.10, p.408-416, 2006.
- Mendiguchía, C.; Moreno, C.; Galindo-Riano, M. D.; García-Vargas, M. Using chemometric tools to assess antropogenic effects in river water a case study: Guadalquivir river (Spain). *Analytica Chimica Acta*, v.515, p.143-149, 2004.
- Merten, G. H.; Minella, J. P. Qualidade da água em bacias hidrográficas rurais: Um desafio atual para a sobrevivência futura. *Agroecologia e Desenvolvimento Rural Sustentável*, v.3, p.33-38, 2002.
- Moraes, L. M.; Paula Júnior, D. R. Avaliação da biodegradabilidade anaeróbia de resíduos da bovinocultura e da suinocultura. *Engenharia Agrícola*, v.24, p.445-454, 2004.
- Papatheodorou, G.; Demopoulou, G.; Lambrakis, N. A long-term study of temporal hydrochemical data in a shallow lake using multivariate statistical techniques. *Ecological Modelling*, v.193, p.759-776, 2006.
- Rencher, A. C. Methods of multivariate analysis. 2.ed. New York: Wiley Interscience, 2002. 738p.
- Silva, A. M. M.; Sacomani, L. B. Using chemical and physical parameters to define the quality of Pardo river water (Botucatu-SP-Brasil). *Water Research*, v.35, p.1609-1616, 2001.
- Simeonov, V.; Stratis, J. A.; Samara, C.; Zachariadis, G.; Voutsas, D.; Anthemidis, A.; Sofoniou, M.; Kouimtzis, T. Assessment of the surface water quality in northern Greece. *Water Research*, v.37, p.119-124, 2003.
- Singh, K. P.; Malik, A.; Mohan, D.; Sinha, S. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques - a case study. *Analytica Chimica Acta*, v.538, p.355-374, 2005.
- Toledo, L. G. de; Nicoletta, G. Índice de qualidade de água em microbacia sob uso agrícola e urbano. *Scientia Agrícola*, v.59, p.181-186, 2002.
- Wilks, D. S. Statistical methods in the atmospheric sciences. 2.ed. San Diego: Elsevier, 2006. 649p.
- Zeng, X.; Rasmussen, T. C. Multivariate statistical characterization of water quality in Lake Lanier, Georgia, USA. *Journal Environmental Quality*, v.34, p.1980-1991, 2005.
- Zhang, X.; Wang, Q.; Liu, Y.; Wu, J.; Yu, M. Application of multivariate statistical techniques in the assessment of water quality in the Southwest New Territories and Kowloon, Hong Kong. *Environmental Monitoring and Assessment*, v.137, p.17-27, 2010.