



IV-010 - ANÁLISE DE PRINCIPAIS COMPONENTES APLICADA AOS PARÂMETROS DE QUALIDADE DE ÁGUA NA BACIA DO RIO DAS VELHAS

Gerson José de Mattos Freire⁽¹⁾

Engenheiro Civil pela Pontifícia Universidade Católica de Minas Gerais. Especialista em Geoprocessamento pelo Instituto de Geociências da Universidade Federal de Minas Gerais (IGC-UFMG). Mestre em Análise e Modelagem de Sistemas Ambientais pelo IGC-UFMG. Consultor da Aluvial Engenharia e Meio Ambiente Ltda. Doutorando em Arquitetura e Urbanismo na Escola de Arquitetura da Universidade Federal de Minas Gerais (EA-UFMG).

Ana Clara Mourão Moura⁽²⁾

Arquiteta e Urbanista pela Universidade Federal de Minas Gerais. Especialista em Planejamento Territorial e Urbano pela PUC-MG e Universidade de Bologna. Mestre em Geografia (Organização Humana do Espaço) pela Universidade Federal de Minas Gerais. Doutora em Geografia (Geoprocessamento) pela Universidade Federal do Rio de Janeiro (2002). Professor Adjunto da Universidade Federal de Minas Gerais - Departamento de Urbanismo. Coordenadora do Laboratório de Geoprocessamento da Escola de Arquitetura da Universidade Federal de Minas Gerais (EA-UFMG).

Endereço^(1,2): Rua Paraíba, 697 - 30.130-140 - Bairro Funcionários - Belo Horizonte, MG – Brasil
gersonj@ufmg.br

RESUMO

As águas da Bacia do Rio das Velhas vêm merecendo a atenção de pesquisadores e estudiosos ao longo dos tempos. O monitoramento sistemático, que consiste no acompanhamento dos aspectos qualitativos das águas, visando a produção de informações é efetuado ao longo da bacia pelo Projeto Águas de Minas. Este é responsável pelo monitoramento da qualidade das águas superficiais desde 1997. O programa disponibiliza uma série histórica da qualidade das águas no Estado, consistindo-se em dados indispensáveis ao gerenciamento dos recursos hídricos. Ações em curso, implementadas pelo IGAM através do programa Águas de Minas, visam a melhoria da qualidade das Águas na Bacia do Rio das Velhas.

O presente trabalho pretende avaliar a variabilidade dos dados encontrados na série histórica de dados disponíveis na Bacia do Rio das Velhas – Minas Gerais, através da Análise de Principais Componentes sobre dados de 35 pontos de monitoramento na bacia. Destes 35 pontos, 18 se encontram na calha principal do rio das Velhas. A análise dos principais componentes efetuada consiste no cálculo dos autovalores e correspondentes autovetores de uma matriz de variâncias-covariâncias ou de uma matriz de coeficientes de correlação entre variáveis, no caso presente os dados provenientes da série histórica do monitoramento. O trabalho descreve brevemente as operações efetuadas sobre a base de dados para a análise das principais componentes, que foi efetuada com o cálculo dos autovalores e correspondentes autovetores de uma matriz de variâncias-covariâncias ou de correlações entre variáveis, através do software TANAGRA, desenvolvido por pesquisadores da Universidade de Lyon.

A análise permite concluir que as variáveis analisadas apresentam um elevado grau independência, com correlações significativas em casos isolados.

PALAVRAS-CHAVE: Qualidade das águas, monitoramento, estatística, Análise de Principais Componentes.

INTRODUÇÃO

As águas da Bacia do Rio das Velhas vêm merecendo a atenção de pesquisadores e estudiosos ao longo dos tempos. Martius e Spix (1817) creditavam a malária, endêmica em toda a Bacia do São Francisco no início do século XIX, à água “enjoativa, embora fresca”; E se utilizaram, em diversas ocasiões, dos locais de desova dos jacarés como marcadores dos níveis atingidos pelo rio na época de cheia.

Já em 1867, Richard Burton, após sua notável Viagem de Canoa de Sabará ao Oceano Atlântico, destacava a forma encaixada do rio, a composição do fundo em cascalho, “com a presença de material arenoso mais fino; Quase não apresenta lama, exceto a proveniente dos afluentes.” O autor admirou-se com o valor potencial dos



grandes cardumes de peixes que percorriam as águas. Depois de uma descrição de diversos tipos de armadilhas para a captura de peixes, o autor recomenda que os pescadores devam estar equipados “com os maiores anzóis de água doce e com o sistema de enrolamento mais resistente; do contrário os peixes que pesam mais de cinquenta quilos o surpreenderão”. Este era o cenário que cativou o explorador inglês, há cento e quarenta anos atrás, período dilatado para o avanço tecnológico, porém muito curto sob a perspectiva da vida do curso d’água.

Nos dias atuais, a gestão ambiental de bacias tem no monitoramento de qualidade das águas um dos seus mais importantes instrumentos para o estabelecimento de políticas e metas. O monitoramento consiste no acompanhamento sistemático dos aspectos qualitativos das águas, visando a produção de informações. Os dados são destinados à comunidade científica, ao público em geral e, principalmente, às diversas instâncias decisórias. Assim, o monitoramento é um dos fatores determinantes no processo de gestão ambiental, uma vez que propicia uma percepção sistemática e integrada da realidade ambiental.

O esforço pelo conhecimento da qualidade das águas redundava em ganho social, na economia e correta aplicação de recursos para o saneamento e ações de revitalização; Pela identificação de fontes poluidoras e a possibilidade de se tomarem medidas que impeçam o lançamento de poluentes nos corpos hídricos.

Em Minas Gerais, o Projeto Águas de Minas é responsável pelo monitoramento da qualidade das águas superficiais desde 1997. O programa disponibiliza uma série histórica da qualidade das águas no Estado, que será comentada adiante neste trabalho, e gera dados indispensáveis ao gerenciamento correto dos recursos hídricos.

Em todo o Estado, o Programa de monitoramento das águas superficiais conta com 260 pontos de amostragem em 2009, nas oito principais bacias hidrográficas de Minas Gerais, com uma cobertura de 98,3% da área estadual. As coletas são feitas a cada trimestre, com um total de 4 campanhas anuais. As amostras coletadas e as respectivas análises físico-químicas, bacteriológicas e ecotoxicológicas são realizadas pela Fundação Centro Tecnológico de Minas Gerais (CETEC), órgão vinculado à Secretaria de Estado de Ciência e Tecnologia.

Já o monitoramento das Águas Subterrâneas vem sendo efetuado pelo IGAM desde 2005, quando foi implantada uma Rede Piloto de Monitoramento da Qualidade das Águas Subterrâneas nas bacias dos rios Verde Grande, Riachão e Jequitaiá, todas inseridas na bacia do Rio São Francisco. A rede piloto é constituída por 39 poços tubulares profundos, instalada numa área aproximada de 36.000 km².

O presente trabalho pretende avaliar a variabilidade dos dados encontrados na série histórica de dados disponíveis na Bacia do Rio das Velhas – Minas Gerais, através da Análise de Principais Componentes.

MATERIAIS E MÉTODOS

A Rede básica de monitoramento de qualidade de água na Bacia do Rio das Velhas é composta de 35 pontos, monitorados desde Setembro de 1997. Por sua característica de localização, é a rede que fornece dados mais abrangentes da situação da bacia hidrográfica. Destes 35 pontos de monitoramento, 18 se encontram na calha principal do rio das Velhas. A localização das estações, bem como o ano desde o qual vem sendo monitorados os parâmetros de qualidade de água no local, encontram-se descritos na tabela a seguir.

Tabela 1: Descrição e Cronologia de instalação das Estações

Código	Descrição do local das estações	Instalação
BV013	Rio das Velhas logo a montante da foz do Rio Itabirito	1997
BV035	Rio ITABIRITO a jusante do córrego Cata Branca	1997
BV037	Rio das Velhas logo a jusante da foz do rio Itabirito	1997
BV062	Ribeirão ÁGUA SUJA próximo de sua foz no rio das Velhas	1997
BV063	Rio das Velhas logo a jusante do ribeirão Água Suja	1997
BV067	Rio das Velhas a montante do ribeirão Sabará	1997
BV076	Ribeirão SABARÁ próximo de sua foz no rio das Velhas	1997
BV083	Rio das Velhas logo a jusante do ribeirão Arrudas	1997
BV105	Rio das Velhas logo a jusante do ribeirão do Onça	1997
BV130	Ribeirão da MATA próximo de sua foz no rio das Velhas	1997
BV135	Rio TAQUARAÇU próximo de sua foz no rio das Velhas	1997
BV137	Rio das Velhas na Ponte Raul Soares	1997
BV139	Rio das Velhas a montante da ETA/COPASA em Bela Fama	1997
BV140	Ribeirão JEQUITIBÁ próximo de sua foz no rio das Velhas	1997
BV141	Rio das Velhas na cidade de Santana de Pirapama	1997
BV142	Rio das Velhas a montante da foz do rio Paraúna	1997
BV143	Rio PARAÚNA próximo de sua foz no rio das Velhas	1997
BV146	Rio das Velhas a jusante do rio Pardo Grande	1997
BV147	Rio BICUDO próximo de sua foz no rio das Velhas	1997
BV148	Rio das Velhas na cidade de Várzea da Palma	1997
BV149	Rio das Velhas – Guaicuí	1997
BV152	Rio das Velhas entre os rios Paraúna e Pardo Grande	1997
BV153	Rio das Velhas a jusante do ribeirão da Mata	1997
BV154	Ribeirão do ONÇA próximo de sua foz no rio das Velhas	1997
BV155	Ribeirão ARRUDAS próximo de sua foz no rio das Velhas	1997
BV156	Rio das Velhas logo a jusante do rio Jaboticatubas	1998
BV160	Ribeirão das NEVES próximo de sua foz no ribeirão da Mata	2000
BV161	Ribeirão SANTO ANTÔNIO próximo de sua foz no rio das Velhas	2000
BV162	Rio CIPÓ a montante da foz do rio Paraúna	2000
BV133	Rio Jaboticatubas a montante de Jaboticatubas	2005
BV136	Rio Jaboticatubas a jusante de Jaboticatubas	2005
BV144	Ribeirão da Onça a jusante da ETE de Cordisburgo	2005
BV145	Rio Pardo Pequeno a Jusante de Monjolos	2005
BV150	Rio das Velhas a jusante do Rio Paraúna, em Senhora do Glória	2008
BV151	Rio das Velhas a jusante do córrego do vinho - Lassance	2008

A figura a seguir mostra a distribuição espacial dos pontos de monitoramento na calha principal do Rio das Velhas.

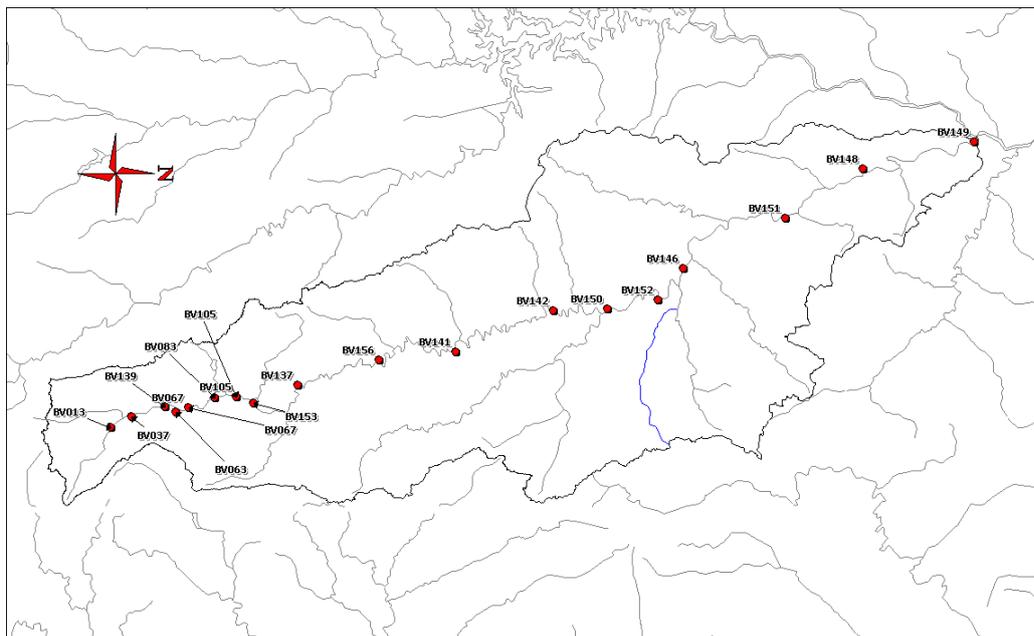


Figura 1: das Velhas - rede de monitoramento na calha principal.

A análise dos principais componentes é o cálculo dos autovalores e correspondentes autovetores de uma matriz de variâncias-covariâncias ou de uma matriz de coeficientes de correlação entre variáveis, no caso presente os dados provenientes da série histórica do monitoramento “Águas de Minas”.

A análise das principais componentes foi empregada como uma técnica de transformação de variáveis. Cada variável medida foi considerada como um eixo de variabilidade, estando correlacionada com variáveis de controle; Assim, os dados foram transformados de tal modo a descrever a mesma variabilidade total existente, com o mesmo número de eixos originais, porém não mais correlacionados entre si. A análise pode ser descrita, de forma gráfica, como a rotação de pontos existentes num espaço multidimensional originando eixos, ou componentes principais, que dispostos num espaço a duas dimensões representem variabilidade suficiente que possa indicar algum padrão a ser interpretado.

BASE DE DADOS DISPONÍVEL

A série histórica de monitoramento compreende dados obtidos no período de setembro de 1997 a março de 2009, em quatro campanhas anuais. No monitoramento são analisados parâmetros físicos, químicos, microbiológicos e bioensaios ecotoxicológicos de qualidade de água, levando em conta os mais representativos, os quais são relatados a seguir:

Parâmetros Físicos: temperatura do ar e da água, condutividade elétrica, sólidos totais, sólidos dissolvidos, cor, turbidez, sólidos em suspensão, sólidos totais, sólidos dissolvidos, alcalinidade total, alcalinidade bicarbonato, dureza total, dureza de cálcio, dureza de magnésio;

Parâmetros Químicos: pH, oxigênio dissolvido, %OD saturação, demanda bioquímica de oxigênio (DBO), demanda química de oxigênio (DQO), série de nitrogênio (orgânico, amoniacal, nitrato e nitrito), fósforo total, surfactantes aniônicos, óleos e graxas, cianetos, fenóis, cloretos, ferro, potássio, sódio, sulfetos, magnésio, manganês, alumínio, zinco, bário, cádmio, boro, arsênio, níquel, chumbo, cobre, cromo (III), cromo (IV), selênio, mercúrio, cálcio, amônia, substâncias tensoativas;

Parâmetros microbiológicos: coliformes fecais, coliformes totais e estreptococos totais; clorofila, cianobactérias e feofitina;

Além destes parâmetros armazenam-se dados referentes a Bioensaios Ecotoxicológicos, que são ensaios de toxicidade crônica visando aprimorar as informações referentes à toxicidade causada pelos lançamentos de substâncias tóxicas nos cursos d'água.

O programa Águas de Minas disponibilizou os dados trimestrais de monitoramento, compreendendo um total de 1369 resultados de 57 dos parâmetros já citados, desde a data de instalação do ponto de monitoramento até o 1º. Trimestre de 2009. Os dados das diversas séries anuais obtidos foram consolidados em tabela única para a sua análise conjuntural, preservadas as características das tabelas originais e incluídos indicadores da estação do ano (Seca ou chuvosa).

OPERAÇÕES SOBRE A BASE DE DADOS

A análise das principais componentes efetuou-se com o cálculo dos autovalores e correspondentes autovetores de uma matriz de variâncias-covariâncias ou de correlações entre variáveis, através do software TANAGRA, desenvolvido por pesquisadores da Universidade de Lyon. O TANAGRA é um software livre para mineração de dados, para finalidades acadêmicas e de pesquisa. Propõe diversos métodos para mineração de dados, análise exploratória, classificação estatística, de computação e de pesquisa em bases de dados. TANAGRA conta com processos de classificação supervisionada e não-supervisionada, tais como clusterização, análise fatorial, estatísticas parametrizadas e não parametrizadas, regras de associação, seleção de características e algoritmos da construção.

Para a operação, os dados existentes foram consolidados em tabela única e submetidos ao algoritmo de blocos descritos na figura a seguir.

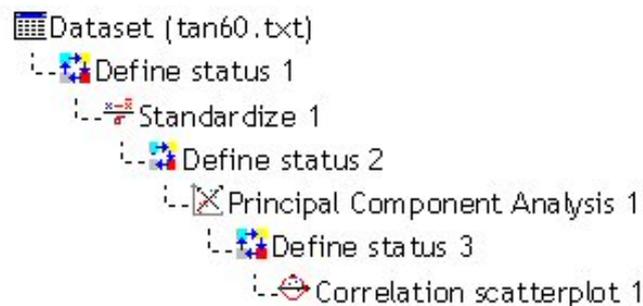


Figura 2: Algoritmo de Blocos – Tanagra.

Os dados foram analisados para as campanhas de seca e chuva separadamente, com entradas e os parâmetros-alvos selecionados entre todos os dados existentes. Como as variáveis de análise, devido a escalas diferentes de mensurações empregadas, não podem ser diretamente comparadas, tornou-se necessário preliminarmente a padronização, de modo que as variáveis transformadas passaram a ter média zero e variância unitária, através da transformação "z", expressa como se segue.

$$\frac{(x - \underline{x})}{\rho x} \quad (1)$$

Assim, com variáveis padronizadas, a matriz de variâncias-covariâncias e a de coeficientes de correlação tornaram-se idênticas.

RESULTADOS ENCONTRADOS

Define-se a variância total existente no conjunto de dados multivariados como a soma das variâncias de cada uma das variáveis. Na matriz de variâncias-covariâncias essas variâncias individuais constituem os elementos da

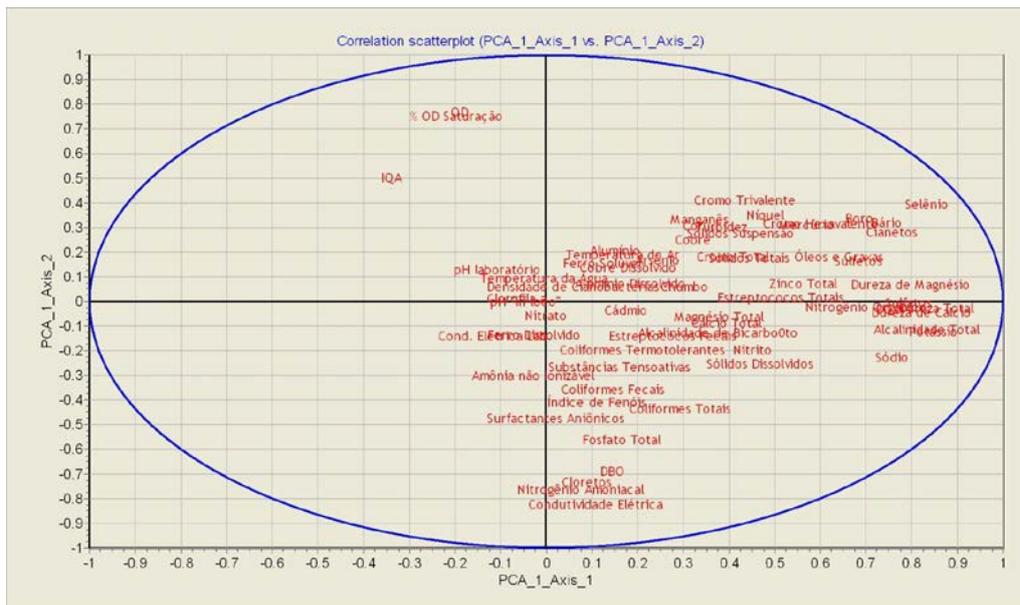


Figura 4: Análise de Principais Componentes – Período de Chuva.

CONCLUSÕES

Utilizando-se da multiplicação da matriz de dados originais pela matriz de autovetores, obteve-se uma matriz de dados transformados que representam projeções dos pontos, num espaço multidimensional, sobre as diversas componentes principais. Como se vê, no período da seca as variáveis com correlações significativas foram:

Turbidez, cor, sólidos totais, sólidos dissolvidos, sólidos em suspensão, Alcalinidade total, Alcalinidade de Bicarbonato, Dureza Total, Dureza de Cálcio, Dureza de magnésio, potássio, sódio, sulfatos, nitrogênio orgânico, DQO, Níquel e Selênio, num total de 17 parâmetros.

Já no período das chuvas as maiores correlações encontradas foram:

Condutividade Elétrica, Alcalinidade total, Alcalinidade de Bicarbonato, Dureza Total, Dureza de Cálcio, Dureza de magnésio, cloretos, potássio, sódio, sulfatos, nitrogênio amoniacal, OD, %ODSaturação, DBO, DQO, Cianetos, Alumínio, Bário, Cobre e Selênio, num total de 20 parâmetros.

Os autovetores encontrados correspondem às componentes principais e são o resultado do carregamento das variáveis originais em cada um deles. Tais carregamentos podem ser considerados como uma medida da relativa importância de cada variável em relação às componentes principais e os respectivos sinais, se positivos ou negativos, indicam relações diretamente e inversamente proporcionais.

A análise revelou traço das matrizes de variância-covariância igual ao número de variáveis analisadas, com a presença de várias delas sem qualquer correlação ou com correlações insignificantes em relação às demais. Esta análise permite concluir que as variáveis analisadas apresentam um elevado grau de independência, com correlações significativas somente em casos isolados e em pontos específicos.

REFERÊNCIAS BIBLIOGRÁFICAS

- SANDERS, T. G. et al. Design of network for monitoring water quality. 4th ed. Michigan: Water Resources Publications, 1983. 328p.
- SHARP, W. E. A topologically optimum water-sampling plan for rivers and streams. Water Resources Research, v. 7, n. 6, p. 1641-1646, 1971.



3. TANAGRA: Ricco Rakotomalala, TANAGRA: Um logiciel de data mining gratuit pour l'enseignement et la recherche. em Actes de EGC'2005, RNTI-E-3, Vol. 2, pp.697-702, 2005. Ricco RAKOTOMALALA Lyon - FRANCE Janeiro 2004. Última modificação: Maio 18o, 2006. Software livre disponível para download em <http://eric.univ-lyon2.fr/~ricco/tanagra/>