



XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

OTIMIZAÇÃO DO MONITORAMENTO DA QUALIDADE DA ÁGUA NA LAGOA MIRIM/RS UTILIZANDO A ANÁLISE DE AGRUPAMENTO

Mylena Feitosa Tormam^{1}; Carina Krüger Bork²; Julia Manzke³; Hugo Alexandre Soares Guedes⁴*

Resumo – A otimização do monitoramento da qualidade da água é essencial para uma gestão eficaz dos recursos hídricos. O objetivo deste estudo foi identificar e determinar os locais de monitoramento homogêneos na Lagoa Mirim/RS/Brasil. As metodologias estatísticas foram aplicadas a dados obtidos com a análise de amostras de água coletadas no período compreendido entre 2013 e 2016, provenientes de sete pontos de monitoramento. A Análise de Agrupamento foi utilizada para verificar se os pontos de monitoramento poderiam ser agrupados em padrões espacialmente coerentes. Os resultados encontrados auxiliam na otimização das redes de monitoramento da qualidade da água, por meio da identificação de pontos homogêneos de monitoramento da qualidade da água, possibilitando futuras estratégias na escolha do número e da localização dos pontos de monitoramento que impactem na redução da quantidade de análises e do custo do procedimento de avaliação da qualidade da água.

Palavras-Chave – análise de agrupamento, pontos de monitoramento, Lagoa Mirim

OPTIMIZATION OF WATER QUALITY MONITORING IN THE MIRIM LAKE USING CLUSTER ANALYSIS

Abstract – Optimizing water quality monitoring is essential for effective management of water resources. The objective of this study was to identify and determine the homogeneous monitoring sites in the Mirim lake (RS/Brazil). The statistical methodologies were applied to data obtained from the analysis of water samples collected in the period between 2013 and 2016, from seven monitoring points. Cluster Analysis was performed to verify whether the monitoring points could be grouped into spatially coherent patterns. The results help optimize the water quality monitoring networks by identifying homogeneous points of water quality monitoring, enabling future strategies in choosing the number and location of monitoring points that impact on the reduction of the number of analyzes and the cost of the water quality assessment procedure.

Keywords – cluster analysis, monitoring points, Mirim lake

¹ Aluno de Graduação, Centro de Engenharias/Engenharia Civil, Universidade Federal de Pelotas. tormam.mylena@gmail.com.

² Aluno de Pós-Graduação, Centro de Desenvolvimento Tecnológico/Programa de Pós-Graduação em Recursos Hídricos, Universidade Federal de Pelotas. borkcarina@gmail.com.

³ Aluno de Graduação, Centro de Engenharias/Engenharia Civil, Universidade Federal de Pelotas. julia-manzke@gmail.com

⁴ Professor Adjunto, Centro de Engenharias/ Engenharia Civil, Universidade Federal de Pelotas. hugo.hydro@gmail.com.



XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

1. INTRODUÇÃO

A possível variabilidade na qualidade da água pode ser devida a atividades antropogênicas, variações naturais durante as estações ou decorrente de vários processos bioquímicos ou químicos (AWADALLAH; YOUSRY, 2012). Para o estudo da qualidade das águas em um corpo hídrico é necessário o conhecimento de diferentes variáveis, as quais devem ter uma abrangência temporal e espacial. A fim de obter uma boa compreensão das mudanças espaciais e temporais nas características dos recursos hídricos, é necessário ter dados indicando a qualidade da água, devendo ser recolhidos em uma base contínua (BILGIN, 2015).

A variação espacial e temporal da qualidade da água é, muitas vezes, exigida a longo prazo em programas de monitoramento, com o objetivo de avaliar os aspectos físicos, químicos e biológicos das águas receptoras (CHOW, 2015). Por conseguinte, é comum gerar uma base de dados demasiadamente longa e de difícil interpretação dos resultados, principalmente em se tratando da inter-relação entre variáveis qualitativas e pontos de monitoramento.

De acordo com Simeonov (2003), na grande maioria dos casos, a análise de grandes bases de dados exige uma interpretação complexa. Para contornar esse problema, muitos estudos vêm utilizando métodos multivariados de análise (GOMES, 2014; AWADALLAH; YOUSRY, 2012; YANG, 2015). Esses estudos comprovam que a abordagem multivariada em conjunto de dados ambientais altamente complexos traz melhores informações, propiciando resultados mais eficazes na gestão qualitativa dos recursos hídricos.

Sendo assim, os métodos estatísticos multivariados têm se mostrado úteis para a interpretação, modelagem e classificação de grandes conjuntos de dados, avaliações da qualidade da água e o gerenciamento dos fluxos superficiais (SIMEONOV, 2003). Conforme estudo de Mahmood (2011), a aplicação de tais métodos oferece uma melhor compreensão da qualidade da água para interpretar os conjuntos de dados.

Dentre os diferentes métodos multivariados, a análise de agrupamento é usada para agregar pontos de monitoramento em grupos coerentes, com base nos valores registrados ou nas correlações entre eles (AWADALLAH; YOUSRY, 2012). Gomes (2014) enfatiza em seu estudo que a análise de agrupamento permite a interpretação de matrizes de dados complexos para melhor compreender a qualidade da água e o estado ecológico do sistema estudado. A análise de agrupamento classifica objetos, de modo que cada objeto é semelhante aos outros no grupo em função de um critério de seleção predeterminado. Os grupos resultantes de objetos devem exibir alta homogeneidade interna, dentro do grupo, e alta heterogeneidade externa, entre os grupos (MUANGTHONG, 2016). Um processo de agrupamento hierárquico pode ser representado como uma sequência aninhada ou árvore, chamada dendrograma, que mostra como os grupos que são formados nas várias etapas do processo estão relacionados.

Neste sentido, o presente estudo buscou interpretar e avaliar as semelhanças e diferenças da qualidade da água entre os pontos de monitoramento ao longo da Bacia Hidrográfica da Lagoa Mirim, no sul do Estado do Rio Grande do Sul, durante os anos de 2013 e 2016, com o objetivo de verificar se os pontos de monitoramento podem ser agrupados de uma forma espacialmente coerente e se são recomendados mais (ou menos) locais de amostragem para melhor descrever a variabilidade espacial.



XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

2. METODOLOGIA

2.1 Área de Estudo

A Bacia Hidrográfica da Lagoa Mirim está localizada no sul do Estado do Rio Grande do Sul entre os paralelos 31°30' e 34°30' S e entre os meridianos 52° e 56°O, e possui uma área de aproximadamente 62.250 km². Trata-se de um espaço geográfico binacional sob um regime de águas compartilhadas entre o Brasil e o Uruguai, em que 47% do território da bacia pertencem ao Brasil e 53% são de domínio uruguaio (MACHADO, 2012).

A Lagoa Mirim é o principal recurso hídrico da bacia hidrográfica da Lagoa Mirim e compõe um dos principais corpos hídricos do sistema lagunar meridional da América do Sul. O seu espelho d'água apresenta uma área aproximada de 3.749 Km², 185 km de comprimento e largura média de 20 km (MACHADO, 2012).

O clima da região, segundo classificação de Köppen (ALVARES, 2013), é subtropical de verão úmido e quente, do tipo Cfa, com temperaturas médias dos meses mais quentes superiores a 22°C e dos mais frios inferiores a 18°C e precipitação pluvial média anual de 1.450 mm. Azambuja (2005) destaca que a bacia possui uma distribuição bastante irregular ao longo da sua extensão, o que resulta tanto na ocorrência de grandes inundações quanto em períodos de seca prolongada. Na região a área do complexo de terras baixas, destaca-se a produção de arroz irrigado por gravidade ocupado praticamente toda a área, compartilhando-a com a pecuária em sistema rotacional e silvicultura nas terras altas. (STEINKE; SAITO, 2010).

2.2 Obtenção dos dados

Os dados utilizados no estudo são provenientes de sete pontos de monitoramento na bacia: 1: Praia do Pontal; 2: Fazenda Bretanha; 3: Fazendo São Francisco; 4: Capilha; 5: Curral Alto; 6: Vila Anselmi; e 7: Porto Santa Vitória (Figura 1). As coletas de amostras foram realizadas no período compreendido entre junho de 2013 a dezembro de 2016, totalizando 22 coletas em cada local de amostragem. Foram avaliadas nove variáveis de qualidade de água: temperatura, turbidez, pH, condutividade elétrica, oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO), nitrogênio total, fósforo total, sólidos totais. A coleta das amostras e as análises laboratoriais foram realizadas pela equipe técnica da Agência da Lagoa Mirim, seguindo a metodologia descrita por APHA (2005).

XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

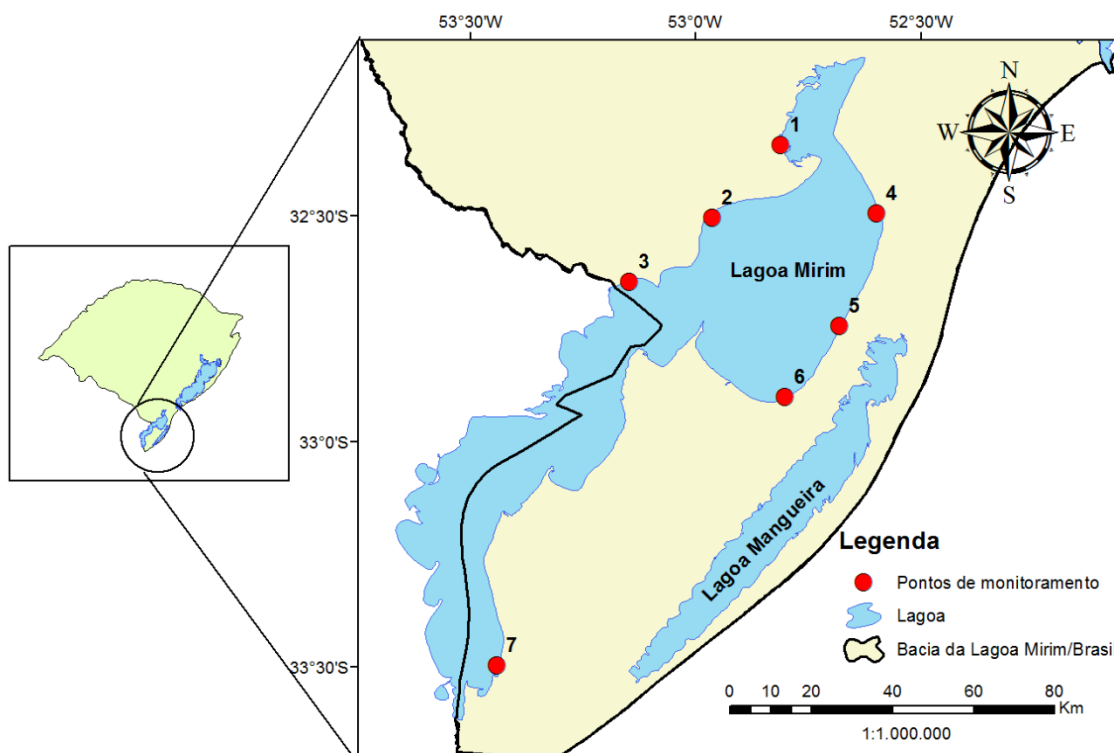


Figura 1 – Área de estudo e pontos de monitoramento.

2.3 Análise dos dados

Inicialmente, buscou-se preparar os dados para a análise posicionando-os em matrizes, onde nas linhas ficaram os pontos amostrais e nas colunas ficaram as variáveis de qualidade de água. Os dados foram padronizados e aplicou-se o teste não paramétrico de Kolmogorov-Smirnov, com nível de significância de 5% ($p < 0,05$), para verificar se as variáveis seguiam a distribuição normal. Em seguida, realizou-se a análise descritiva para investigar as diferenças numéricas entre as médias das variáveis.

Para as variáveis que apresentaram distribuição normal foi utilizado o teste paramétrico Análise de Variância (ANOVA unidirecional, $p < 0,05$) para verificar se haveria diferenças significativas entre os pontos de monitoramento para cada variável de qualidade de água. Posteriormente, caso houvesse diferença significativa, foi aplicado o teste de Tukey, a um nível de significância de 5% ($p < 0,05$), visando identificar os pontos de monitoramento onde as variáveis de qualidade da água eram significativamente diferentes dos demais pontos (BILGIN, 2015).

Na sequência, foi aplicada a análise de agrupamento utilizando a distância euclidiana como medida de similaridade, a qual visa identificar regiões da Bacia Hidrográfica da Lagoa Mirim homogêneas em relação às variáveis de qualidade da água. O método de Ward, o qual minimiza a



XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

variância entre as variáveis de análise, foi utilizado para realizar as ligações entre os agrupamentos. Em geral, esse método é considerado muito eficiente (AWADALLAH; YOUSRY, 2012). A determinação do número de grupos para o dendrograma é feita de forma visual e pelo critério de inércia entre saltos, em que o maior salto, em relação aos anteriores, é definido como o ponto de corte (MELO JÚNIOR, 2006). Os agrupamentos formados foram avaliados utilizando o Coeficiente de Correlação Cofenética (CCC), o qual estima quão bem a estrutura hierárquica do dendrograma representa em duas dimensões as relações multidimensionais dentro dos dados de entrada (SOKAL; ROHLF, 1962). A concordância entre os dados de entrada e o dendrograma está próxima se o valor do índice estiver perto de 1,0 (RAO; SRINIVAS, 2006).

A análise de agrupamento foi realizada no software R e as demais análises estatísticas foram realizadas no software SPSS 20.0.

3. RESULTADOS E DISCUSSÃO

De acordo com o teste de Kolmogorov Smirnov (K-S), a normalidade de todas variáveis foi confirmada, ou seja, apresentaram nível de significância (p-valor) inferior a 0,05, indicando que as variáveis testadas seguem a distribuição normal. No teste paramétrico ANOVA, apenas a variável condutividade elétrica apresentou p-valor menor que 0,05, ou seja, rejeitou-se a hipótese nula de igualdade e concluiu-se que há diferença estatística significativa entre os pontos amostrais (Tabela 2).

Tabela 2 – ANOVA para as variáveis de qualidade de água nos diferentes pontos amostrais.

Parâmetros	Soma dos Quadrados	df	Quadrados médios	F	p-valor
Temperatura	4,111	6	0,685	0,629	0,706
Turbidez	11,145	6	1,858	2,459	0,058
pH	1,825	6	0,304	0,253	0,952
Condutividade Elétrica	15,079	6	2,513	4,435	0,005
Oxigênio Dissolvido	0,934	6	0,156	0,125	0,992
Demanda Bioquímica de Oxigênio	3,544	6	0,591	0,528	0,781
Nitrogênio Total	0,439	6	0,073	0,058	0,999
Fósforo Total	4,695	6	0,783	0,737	0,625
Sólidos Totais	5,579	6	0,930	0,910	0,507

df: graus de liberdade

F: estatística F de Snedecor

Ao aplicar o teste de Tukey observou-se diferença significativa ($p < 0,05$) entre os locais de amostragem: Praia do Pontal e Vila Ancelmi, Fazenda Bretanha e Vila Ancelmi, Fazenda São

XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

Francisco e Vila Ancelmi. Isso indica que os pontos diferem entre si para a variável condutividade elétrica. Pode-se concluir que a ausência de uma diferença significativa na qualidade da água entre os demais pontos de monitoramento indica uma semelhança na qualidade entre esses pontos (BILGIN, 2015).

O procedimento de agrupamento gerou três grupos de locais de amostragem, indicando independência relativamente alta para cada agrupamento, sendo que cada um possui características semelhantes (Figura 2). O Coeficiente de Correlação Cofenética para os grupos obtidos a partir do algoritmo de ligação de Ward apresentou o valor de 0,78. O agrupamento I é formado por dois pontos (Fazenda São Francisco e Fazenda Bretanha). O agrupamento II também é formado por dois pontos (Praia do Pontal e Curral Alto). E o agrupamento III é formado por três pontos (Vila Ancelmi, Capilha e Porto Santa Vitória). A análise de agrupamento realizada poderia ter agrupado os pontos de monitoramento de forma heterogênea, uma vez que nenhuma variável de localização é introduzida no procedimento. No entanto, a coerência espacial do grupo I é um indicador de zoneamento homogêneo na bacia e da possibilidade de selecionar apenas um local desse grupo para uma rápida avaliação espacial da qualidade da água (AWADALLAH; YOUSRY, 2012). Entretanto, os pontos dos agrupamentos II e III estão distantes uns dos outros, considerando suas localizações geográficas. Isso pode indicar algumas condições pouco usuais da qualidade da água causadas, por exemplo, pela descarga de esgotos de pequenas cidades (KOWALKOWSKI, 2006). Verifica-se, também, que o agrupamento I é caracterizado pela maior distância euclidiana para os outros agrupamentos, ou seja, alta significância de agrupamento, uma vez que a classificação desses grupos varia com o nível de significância (KOWALKOWSKI, 2006).

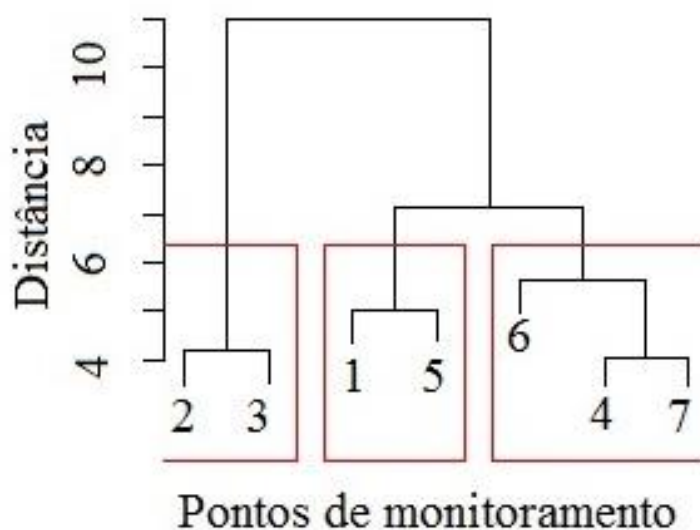


Figura 2 – Dendrograma da Análise de agrupamento dos pontos de monitoramento utilizando a Distância Euclidiana e o Método de Ward.

A aplicação dessa ferramenta nos dados de qualidade da água mostrou que há pontos de monitoramento associados ao mesmo padrão de poluição, o que corresponde a medidas redundantes



XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

e devem ser movimentados para outros locais, otimizando a avaliação da qualidade da água na bacia (GOMES, 2014). Por exemplo, o agrupamento III agregou três pontos de monitoramento, sendo que alguns desses pontos podem ser eliminados e/ou deslocados, visto que apresentaram medições físico-químicas muito semelhantes, ou seja, redundantes. Os resultados indicam que a técnica multivariada é útil para oferecer uma classificação das águas superficiais em toda a região, permitindo melhor planejar a estratégia de amostragem espacial, o que poderá reduzir o número de pontos de monitoramento e os custos associados (MUANGTHONG, 2016).

4. CONCLUSÕES

De acordo com os resultados obtidos foi possível concluir que:

- i) Os pontos de monitoramento diferem estatisticamente entre si em relação à variável condutividade elétrica, indicando uma possível reorganização entre esses pontos.
- ii) Os pontos de monitoramento foram agrupados em três grupos com características similares de qualidade de água. Desta forma, é possível analisar a qualidade de água com menos pontos de coleta, sem perder a qualidade dos resultados. Assim, tem-se uma redução de custos envolvidos com planejamento logístico e análises laboratoriais.

REFERÊNCIAS

- ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift* 22(6), pp.711-728.
- APHA/AWWA/WEF. EATON, A. D. (2005). Standard methods for the examination of water and wastewater. 21^a ed. Washington: American Public Health Association. 1082 p.
- AWADALLAH, A. G.; YOUSRY, M. (2012). Identifying Homogeneous Water Quality Regions in the Nile River Using Multivariate Statistical Analysis. *Water Resources Management* 26(7), pp.2039-2055.
- AZAMBUJA, J. L. F. (2005). Hidrovia da lagoa mirim: um marco de desenvolvimento nos caminhos do Mercosul. 182 f. Dissertação (Mestrado) – Curso de Mestrado Profissionalizante em Engenharia, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.
- BERALDO, A. F. Apostila de Análise Multivariada (2016). Departamento de Estatística, ICE – Universidade Federal de Juiz de Fora, Minas Gerais.
- CHOW, M. F. (2015). Evaluation of surface water quality using multivariate statistical techniques: a case study of Fei-Tsui Reservoir basin, Taiwan. *Environmental Earth Sciences*, 75(1), pp.5-21.
- GOMES, A. I. (2014). Optimization of River Water Quality Surveys by Multivariate Analysis of Physicochemical, Bacteriological and Ecotoxicological Data. *Water Resources Management* 28(5) p.1345-1361.
- KOWALKOWSKI, T. (2006). Application of chemometrics in river water classification. *Water Research* 40(4), p.744-752.
- IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.



XXII SÍMPOSIO BRASILEIRO DE RECURSOS HÍDRICOS

26 de novembro a 01 de dezembro de 2017

Florianópolis- SC

- MACHADO, J. B. (2012). Análise da governança das águas da bacia hidrográfica da lagoa mirim, extremo sul do brasil. 205 f. Dissertação (Mestrado) - Curso de Gerenciamento Costeiro, Universidade Federal do Rio Grande, Rio Grande.
- MELO JÚNIOR, J. C. F.; SEDIYAMA, G. C.; FERREIRA, P. A.; LEAL, B. G. (2006). 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* 10(2), pp.408-416.
- MAHMOOD, A., MUQBOOL, W., MUMTAZ, M. W., & AHMAD, F. (2011). Application of multivariate statistical techniques for the characterization of groundwater quality of Lahore, Gujranwala and Sialkot (Pakistan). *Pakistan Journal of Analytical and Environmental Chemistry* 12, pp.102–112.
- MUANGTHONG, S. (2016). Assessment of surface water quality using multivariate statistical techniques: A case study of the Nampong River Basin, Thailand. *The Journal Of Industrial Technology* 11(1), pp.25-37.
- RAO, A. R.; SRINIVAS, V. (2006). Regionalization of watersheds by hybrid-cluster analysis. *Journal Of Hydrology* 318 (1-4), pp.37-56.
- SIMEONOV, V., STRATIS, J. A., SAMARA, C., ZACHARIADIS, G., VOUTSA, D., ANTHEMIDIS, A., SOFONIOU, M., & KOUIMTZIS, T. (2003). Assessment of the surface water quality in Northern Greece. *Water Research* 37 pp.4119–4124.
- SOKAL R. R.; ROHLF F. J. (1962). The comparison of dendrograms by objective methods. *International Association for Plant Taxonomy (IAPT)* 11(2), pp.33-40.
- STEINKE, V. A.; SAITO C. H. (2010). Avaliação geoambiental do território brasileiro nas Bacias Hidrográficas Transfronteiriças. *Revista Brasileira de Gestão e Desenvolvimento Regional*, 6(1), pp.189-221.
- YANG Q. (2015). Multivariate Statistical Analysis of Hydrochemical Data for Shallow Ground Water Quality Factor Identification in a Coastal Aquifer. *Polish Journal of Environmental Studies* 24(2), pp.769-776.