

ANÁLISE DE AGRUPAMENTOS COMO FERRAMENTA DE AVALIAÇÃO DE QUALIDADE DE ÁGUA: ESTUDO DE CASO RIO ARRAIAL, NA REGIÃO USINA HIDRELÉTRICA (UHE) GUARICANA

Leonardo Tourinho Carazzai^{1} & Nicole M. Brassac de Arruda²*

Resumo – O objetivo deste estudo foi avaliar a qualidade da água do rio Arraial, na região da UHE Guaricana. Foram selecionadas três estações de monitoramento: montante e jusante do empreendimento (ambientes lóticos), e reservatório (ambiente intermediário). Os dados utilizados são pretéritos, oriundos do programa de monitoramento, desenvolvido pela COPEL. Coletas trimestrais ocorridas entre 2005 e 2016, totalizaram 44 observações por estação de monitoramento. Os dados de cada estação de amostragem foram divididos de acordo com a estação do ano em que foram coletados, sendo calculadas as medianas para cada uma das variáveis avaliadas. Com base nas medianas, a análise de agrupamentos hierárquicos revelou a formação de cinco grupos distintos. O grupo com maior similaridade entre os elementos foi o G1, onde agruparam-se coletas de inverno e verão, das estações lóticas de monitoramento. No grupo G2, reuniram-se somente as coletas da estação de reservatório. Nos grupos G3, G4 e G5, novamente as estações lóticas foram listadas, em coletas de primavera e outono. Desta forma, observa-se que os grupos de maior similaridade (G1 e G2) foram formados e separados entre si pela velocidade do fluxo da água. Já os demais grupos foram discriminados pela variável coliformes termotolerantes, que apresentou-se elevada, além da sazonalidade.

Palavras-Chave – Bacia litorânea, Paraná, Análise multivariada.

CLUSTER ANALYSIS AS TOOL TO EVALUATE WATER QUALITY FROM ARRAIAL RIVER, IN THE REGION OF GUARICANA HPP

Abstract – This study aimed to evaluate the water quality of the Arraial river, in the Guaricana HPP region. Three monitoring stations were selected: upstream and downstream (lotic environments), and at Guaricana's reservoir (intermediate environment). The data used are from the monitoring program developed by COPEL. Quarterly samplings were collected between 2005 and 2016, summarizing 44 observations, per monitoring station. Data of each sampling station were divided according to the season of the year in which they were collected, and the medians for each one of the evaluated variables were calculated. Based on the medians, the cluster analysis revealed the formation of five distinct groups. The group with the greatest similarity among the elements was the G1, where winter and summer samplings of the lotic monitoring stations were grouped. In group G2, only the samplings of the reservoir station were grouped. In groups G3, G4 and G5, the lotic stations were listed again, in spring and fall samplings. It was observed that the groups of greater similarity (G1 and G2) were formed and separated from each other by the velocity of the water flow. The other groups were discriminated by the variable thermotolerant coliforms, which was elevated, besides the seasonality.

Keywords – Costal basin, Paraná state, Multivariate analysis.

¹ Graduando em Ciências Biológicas, Universidade Positivo, leonardotourinho98@hotmail.com. * Autor correspondente

² Professora Titular Universidade Positivo; Pesquisadora Institutos LACTEC, n.brassac@lactec.org.br

INTRODUÇÃO

De acordo com Mendonça e Leitão (2008), dados do Banco Mundial indicam que oitenta países passam, nos dias de hoje, por problemas relacionados à escassez de água. Soma-se a esta informação também o fato de que a população mundial cresce de forma acelerada em áreas onde a água já é um recurso escasso.

Além disto, o crescimento da população mundial, e sua consequente aglomeração próxima a corpos hídricos, formando centros urbanos, causa uma grande pressão ambiental sobre rios e lagos, além de gerar necessidade maior de energia e reservação de água para o consumo (TUNDISI; TUNDISI, 2008). Assim, observa-se a relevância do tema frente aos demais problemas da atualidade. Segundo Mendonça e Leitão (2008), na gestão de recursos hídricos, e sua problemática, está um dos grandes desafios da sociedade moderna.

O Brasil é um país cuja matriz elétrica é fortemente embasada na hidreletricidade e, desta forma, o país detém grande número de reservatórios vocacionados para este fim (GOLDEMBERG; LUCON, 2007; ARRUDA, 2014). Segundo Arruda (2014), trabalhos sobre a qualidade de água de reservatórios tem sido mais frequentes nos últimos 20 anos, em especial, pelo aumento das exigências ambientais relacionadas à construção usinas hidrelétricas e monitoramento dos corpos hídricos.

Segundo Trindade (2013), para a adequada gestão dos recursos hídricos, o monitoramento e a avaliação da qualidade de água são de grande importância. A autora complementa que estes procedimentos auxiliam na análise de tendências no gerenciamento de bacias hidrográficas, sendo essenciais para o funcionamento para os instrumentos de gestão de recursos hídricos indicados na legislação brasileira.

Para uma avaliação criteriosa, o monitoramento dos dados deve ser realizado de forma contínua, fornecendo uma base histórica consistente para trabalhos futuros que desejem obter as estimativas da variação da qualidade da água superficial do local (TRINDADE, 2013; ARRUDA, 2014).

Dentre as ferramentas de interpretação de dados de qualidade de água estão as análises estatísticas multivariadas, como a análise de agrupamentos, a análise de discriminantes e a análise fatorial (ARRUDA, 2014).

A Análise de Agrupamentos (AA) é um método estatístico que tem se mostrado adequado para avaliação de dados de qualidade de água, pois visa à simplificação de grandes matrizes de dados, aproximando e afastando variáveis ou observações de acordo com sua similaridade ou dissimilaridade (SINGH et al., 2004; MINGOTI, 2005; KNUPP, 2007; SIECIECHOWICZ et al., 2011; TRINDADE, 2013).

Exemplos da aplicação de AA para estudos de qualidade de água são encontrados em todo o Brasil (FRANÇA, 2009; BERTOSSI et al., 2013; SHEELA et al., 2012). Podem ser citados os trabalhos de Guedes et al. (2011), que utilizaram AA para identificar os grupos de poluição presentes no rio Pomba (MG) e de Palácio et al. (2009), que realizaram um trabalho de interpretação da qualidade de água do rio Curu, Ceará, com o objetivo de identificar as variáveis determinantes da qualidade de água local.

O presente estudo visou analisar a qualidade de água no rio Arraial, na região da Usina Hidrelétrica de Guaricana e seu reservatório, a montante e jusante do barramento, com base em séries históricas (dados pretéritos), usando a análise de agrupamentos como técnica estatística. Visou também estabelecer diferenças espaciais e sazonais na qualidade da água estudada.

METODOLOGIA

O presente estudo avaliou dados de qualidade de água advindos do rio Arraial, que se encontra inserido na bacia Litorânea no estado do Paraná, na região da Usina Hidrelétrica de Guaricana (COPEL, 1999).

Os dados utilizados foram advindos do programa de monitoramento da qualidade da água de reservatórios e corpos hídricos relacionados, desenvolvido pela COPEL – Companhia Paranaense de Energia, que é a concessionária responsável pelo empreendimento. Neste programa, foram estabelecidas três estações de amostragem para a área de estudo, conforme a Figura 1. Os dados gerados neste monitoramento entre os anos de 2005 e 2016 foram utilizados para a aplicação da análise estatística multivariada selecionada.

A estação de monitoramento E1 localiza-se no rio Arraial, em sua porção lótica, a montante do reservatório da UHE Guaricana. A estação E2 localiza-se no reservatório de Guaricana, em sua porção mais profunda, nas proximidades do barramento. O reservatório de Guaricana tem tempo de residência médio de 12 dias. A estação E3 está locada no rio Arraial, a jusante do reservatório e da casa de força da usina (COPEL, 1999).

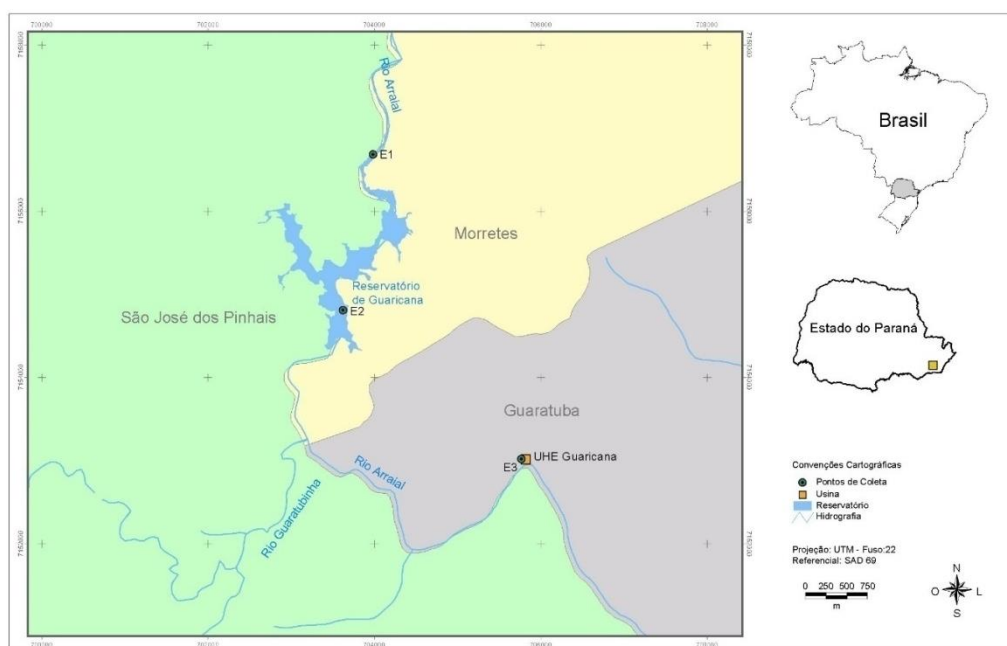


Figura 1. Mapa da área de estudo (Fonte: LACTEC, 2016)

A metodologia empregada para as coletas realizadas pela equipe de campo da COPEL foi a proposta por Santos et al. (2001). No total, foram realizadas 44 amostragens, por estação de monitoramento, para cada uma das variáveis de qualidade de água selecionadas.

Do total de variáveis analisadas pela concessionária, foram selecionadas 11 para análise da qualidade de água neste estudo, a saber: transparência da água (m), temperatura da água (°C), oxigênio dissolvido - OD (mg.L^{-1}), pH, condutividade ($\mu\text{S.cm}^{-1}$), fósforo total - PT (mg.L^{-1}), nitrogênio total - NT (mg.L^{-1}), sólidos totais - ST (mg.L^{-1}), turbidez (NTU), demanda bioquímica de oxigênio - DBO (mg.L^{-1}) e coliformes termotolerantes (NMP. 100 mL^{-1}).

Os dados de resultados analíticos (laudos) e de fichas de campos foram tabelados em planilhas eletrônicas no software Microsoft Excel, consolidados e validados, a fim de buscar eventuais erros de digitação ou laboratoriais, ou ainda, falhas nas séries de dados.

Para a avaliação inicial, os dados foram separados de acordo com a estação do ano em que foram coletados, por estação de amostragem e, a partir disto, foram calculadas as medianas das variáveis, por estação do ano.

A técnica de estatística multivariada que foi utilizada foi Análise de Agrupamentos Hierárquicos (Cluster), que separa os dados através da similaridade, formando grupos com elementos de um mesmo padrão e separando-os através das diferenças dos padrões para as mesmas características (MINGOTI, 2005; SIECIECHOWICZ et al., 2011).

Com o objetivo de se distinguir os dados das estações de monitoramento de maneira sazonal, a letra correspondente à estação do ano foi adicionada após a sigla de estação de monitoramento, como por exemplo: E1_V, representa dados da estação E1, coletados nos verões entre 2005 e 2016, e assim sucessivamente.

Para a realização da análise de agrupamentos, foi primeiramente realizado o cálculo da correlação cofenética da matriz. As distâncias testadas foram: Distância Euclidiana, quadrado da distância Euclidiana, Mahalanobis e Cityblock, com ligações do tipo simples, completa, média, centróide e ward. Esta análise foi realizada através de função programada no Software Matlab (FRANÇA, 2009).

Em seguida, os dados foram tratados no software Statistica 11 pela função Análise de Agrupamento, sendo a interpretação baseada nos gráficos gerados.

RESULTADOS E DISCUSSÃO

Os valores medianos para as variáveis avaliadas, por estação de monitoramento, por estação do ano, encontram-se na Tabela 1.

Tabela 1. Resultados da mediana das variáveis avaliadas, por estação de monitoramento e por estação do ano.

Estação de Monitoramento	Disco de Secchi (m)	Temperatura da água (°C)	OD (mg.L ⁻¹)	pH	Condutividade (µS.cm ⁻¹)	Fósforo Total (mg.L ⁻¹)	Nitrogênio Total (mg.L ⁻¹)	Sólidos Totais (mg.L ⁻¹)	Turbidez (NTU)	Coliformes Termo. (NMP.100mL ⁻¹)	DBO (mg.L ⁻¹)
E1_I	0,70	15,30	8,80	7,40	32,00	0,020	0,70	33,00	2,98	26	2,00
E1_O	0,50	16,00	8,75	7,20	26,00	0,020	0,82	31,50	4,00	485	2,00
E1_P	0,70	19,00	8,54	7,00	23,00	0,030	0,80	39,00	4,23	100	2,00
E1_V	0,73	19,80	8,29	7,25	22,50	0,020	0,80	33,50	3,50	41	2,00
E2_I	1,40	17,30	9,30	7,80	27,00	0,020	0,70	35,00	4,00	1	2,00
E2_O	1,35	17,60	8,06	7,30	28,00	0,027	0,95	34,50	4,50	4	2,00
E2_P	1,10	21,50	8,60	7,30	26,00	0,030	0,50	37,00	4,00	2	2,00
E2_V	0,98	24,60	8,15	7,90	21,50	0,030	0,65	38,00	5,00	1	2,45
E3_I	0,50	15,50	9,10	7,30	29,00	0,020	0,60	34,00	4,00	42	2,00
E3_O	0,50	16,50	8,70	7,30	26,50	0,020	0,65	34,00	5,00	31	2,00
E3_P	0,50	18,75	8,75	7,00	25,00	0,023	0,85	33,50	5,00	200	2,00
E3_V	0,50	21,10	8,46	6,95	24,00	0,023	0,60	39,50	5,16	55	2,05

Após a realização do cálculo da correlação cofenética, a distância que obteve valor mais próximo de 1,0 foi a distância Euclidiana, com ligação média (0,9838). O elevado valor para a relação cofenética aumenta a confiabilidade nas conclusões obtidas a partir da análise do dendrograma.

Com a realização da análise de agrupamentos, um gráfico (dendrograma) das ligações entre as estações de monitoramento x estações do ano foi gerado. O uso da AA permitiu separar as estações de monitoramento e estações do ano em cinco grupos distintos, conforme a Figura 2. Nesta figura, os grupos (G) foram marcados de um a cinco.

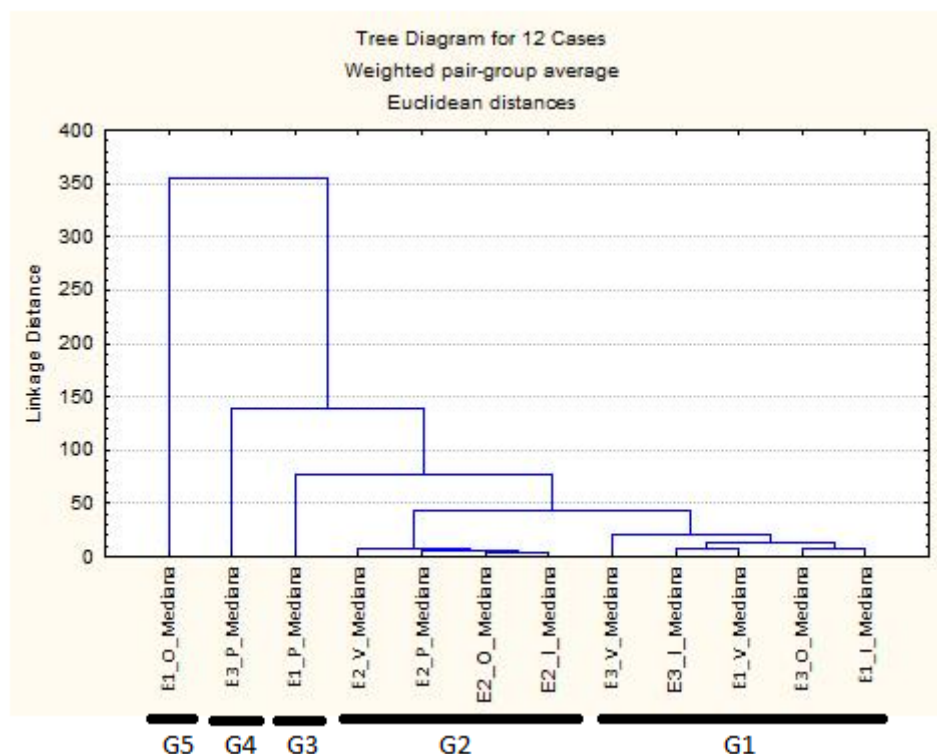


Figura 2. Dendrograma representando os cinco agrupamentos principais.

No G1, de forma geral, ficaram agrupadas as estações de montante e jusante, em coletas realizadas no inverno e no verão. Observa-se assim, uma maior influência da característica lítica das estações, em detrimento da caracterização sazonal da qualidade de água. Situação similar foi observada por Sheela et al. (2012), em que a análise de agrupamentos pareou as estações de monitoramento em função da posição das mesmas, em detrimento da avaliação dos diferentes períodos de coleta.

No G2, agruparam-se as coletas realizadas na estação de reservatório, nas quatro estações do ano. Novamente a característica de velocidade da água no ecossistema ficou evidente, uma vez que os elementos pertencentes à G2 são todos do reservatório de Guaricana. O reservatório de Guaricana tem tempo de residência de 12 dias e desta forma, apresenta característica de ambiente intermediário (entre ecossistemas lóticos e lêntico). Nesta situação, a qualidade de água é, geralmente, beneficiada, em função dos processos físicos do reservatório, como a sedimentação, com valores mais baixos para turbidez, sólidos e valores mais elevados de transparência e oxigênio dissolvido. Também, na estação de reservatório, foram registrados os menores valores de coliformes termotolerantes.

Os agrupamentos G3, G4 e G5 reuniram, individualmente, elementos da estação de montante (G3) e jusante (G4) na primavera, e de montante no outono (G5). Nestes três grupos a sazonalidade aparece de forma mais clara, visto que os elementos, apesar de em grupos distintos, são mais semelhantes entre si, pela proximidade. Nesta situação, são agrupadas coletas em que a concentração

da variável coliformes termotolerantes apresentou valores medianos mais elevados. Sieciechowicz et al. (2011), avaliando dados de qualidade de água do reservatório de Pitangui por meio de análise de agrupamentos, também observou que a concentração da variável coliformes termotolerantes foi a principal caracterizadora dos grupos formados, além dos sólidos totais e turbidez.

CONCLUSÕES

Os resultados do trabalho permitiram a separação dos dados de qualidade de água das estações de monitoramento, separadas de acordo com as estações do ano, em cinco grupos.

Observou-se que o G1 foi formado por estações de monitoramento com característica de velocidade da água rápida (ambientes lóticos), nos meses de verão e inverno, enquanto o G2, apresentou elementos exclusivos da estação de reservatório. Ou seja, avalia-se que na formação dos dois primeiros grupos, o posicionamento das estações, bem como a velocidade do fluxo da água (lótico para os elementos do G1 e lântico para os elementos do G2) teve maior peso na definição da similaridade entre elementos dos grupos, em detrimento da sazonalidade.

Além disso, na formação dos demais grupos, foram registrados elementos de estações de monitoramento de fluxo rápido (lóticos), observou-se a relevância da variável coliformes termotolerantes. Esta variável, indicadora da contaminação das águas por material de origem fecal, esteve mais elevada nos elementos formadores do G3, G4 e G5. Nestes grupos registraram-se duas coletas de primavera e uma de outono, mostrando que, mesmo secundariamente, a sazonalidade também foi relevante no arranjo dos elementos no dendrograma.

REFERÊNCIAS

- ARRUDA, N.M.B. (2014) *Avaliação de variáveis de qualidade de água dos reservatórios das usinas hidrelétricas de Foz da Areia, Segredo e Caxias, como instrumento de gestão da bacia hidrográfica*. 227f. (Tese de Doutorado em Engenharia Florestal) – Universidade Federal do Paraná, Curitiba, 2014.
- BERTOSSI, A.P.A.; MENEZES, J.P.C.; CECÍLIO, R.A.; GARCIA, G.O. ; NEVES, M.A. (2013). Seleção e agrupamento de indicadores da qualidade de águas utilizando estatística multivariada. *Revista Semina: Ciências Agrárias* 34 (5), pp. 2025-2036.
- COPEL - COMPANHIA PARANAENSE DE ENERGIA. *Usina hidrelétrica Guaricana. Relatório Técnico 1999*. Disponível em: <https://goo.gl/ZUpHri>. Acesso em 20 de janeiro de 2017.
- FRANÇA, M.S. (2009). *Análise estatística multivariada dos dados de monitoramento da qualidade de água da Bacia do Alto Iguaçu: uma ferramenta para a gestão de recursos hídricos*. 150 f. Dissertação (Mestrado em Engenharia de Recursos Hídricos e Ambiental) – Universidade Federal do Paraná, Curitiba.
- GOLDEMBERG, J.; LUCON, O. (2007). Energia e meio ambiente no Brasil. *Estudos Avançados* 21 (59), pp. 7-20.
- GUEDES, H.A.S.; SILVA, D.S.; RIBEIRO, C.B.M.; MATOS, A.T.; ELESBON, A.A.A.; SILVA, B.M.B.; GOMES, C.R.; LISBOA, L.; MARTINS, V.S. (2011). Avaliação da qualidade da água do médio Rio Pomba, Minas Gerais, utilizando Análise de Agrupamento. In *Anais do XIX Simpósio Brasileiro de Recursos Hídricos*, Maceió, 2011. pp. 1-14.

KNUPP, E.A.N. (2007). *Uso de métodos estatísticos para dados de qualidade de água: estudo de caso, Rio das Velhas*. 224 f. Tese (Doutorado em Química) – Departamento de Química, Universidade Federal de Minas Gerais, Belo Horizonte.

LACTEC – INSTITUTO DE TECNOLOGIA PARA O DESENVOLVIMENTO (2016). *Relatório anual do automonitoramento da qualidade das águas superficiais do rio Arraial, na região da Usina Hidrelétrica de Guaricana (PR)*. Ciclo 2016: Relatório Técnico. LACTEC: Curitiba. 32p.

MENDONÇA, F.; LEITÃO, S.A.M. (2008). Riscos e vulnerabilidade socioambiental urbana: uma perspectiva a partir dos recursos hídricos. *GeoTextos* 4(1-2), pp. 145-163.

MINGOTTI, S.A. (2005). *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: Editora UFMG, 297 p.

PALACIO, H.A.Q.; ANDRADE, E.M.; LOPES, F.B.; ALEXANDRE, D.M.B.; ARRAES, F.D.D. (2009). Similaridade da qualidade das águas superficiais da bacia do Curu, Ceará. *Revista Ciência Rural* 39 (9): pp. 2494-2500.

SANTOS, I.D.; FILL, H.D.; SUGAI, M.R.V.B.; BUBA, H.; KISHI, R.T.; MARONE, E.; LAUTERT, L.F. (2001). *Hidrometria Aplicada*. Curitiba: CEHPAR. 372p.

SHEELA, A.M.; LETHA, J.; JOSEPH, S.; CHACKO, M.; SANAL KUMAR, S.P.; THOMAS, J. (2012). Water quality assessment of a tropical coastal lake system using multivariate cluster, principal componente and factor analysis. *Lakes & Reservoirs: Research and Management* 17 (2), pp. 143–159.

SIECIECHOWICZ, M.S.F.; ARRUDA, N.M.B.; PIRES, G.D.R.M.; SILVEIRA, L.U.D.; MIRANDA, T.L.G.D.; MÜLLER, I.I. Avaliação da qualidade da água do reservatório da UHE Pitangui (PR) através de técnicas estatísticas multivariadas. In *Anais do XIX Simpósio Brasileiro de Recursos Hídricos*, Maceió, Nov. 2011, pp. 1-14.

SINGH, K.P.; MALIK, A.; MOHAN, D.; SINHA, 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* 38, pp. 3980-3992.

TRINDADE, A.L.C. (2013). *Aplicação de técnicas estatísticas para avaliação de dados de monitoramento de qualidade das águas superficiais da porção mineira da bacia do Rio São Francisco*. 181 f. (Tese de Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) – Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos, Universidade Federal de Minas Gerais, Belo Horizonte.

TUNDISI, J. G.; TUNDISI, T. M. (2008) *Limnologia*. Oficina de Textos: São Paulo.