





CLUSTERIZAÇÃO PARA REDUÇÃO DO NÚMERO DE VARIÁVEIS DE QUALIDADE DA ÁGUA

Luana Nunes Centeno – <u>luananunescenteno@gmail.com</u>

Universidade Federal de Pelotas, Doutoranda no PPG em Recursos Hídricos. Endereço: R. Gomes Carneiro, 01. 96010-610. Balsa, Pelotas - RS, Brasil.

Samanta Tolentino Cecconello – <u>satolentino@gmail.com</u>

Instituto Federal de Educação, Ciência e Tecnologia Sul-rio grandense, Câmpus Pelotas.

Alexssandra Daiane Soares de Campos- <u>alexssandra1_sc@yahoo.com.br</u> Universidade Federal de Pelotas.

Thais Palumbo Silva - <u>thaispalumbosilva@hotmail.com</u> Universidade Federal de Santa Maria.

Miguel David Fuentes Guevara - <u>miguelfuge@hotmail.com</u> Universidade Federal de Pelotas.

Resumo: A análise da qualidade da água de um manancial é de extrema relevância quando se trata de agua de consumo humano. Todavia necessita de investimentos elevado para a caracterização de diversos parâmetros de qualidade da água. Cientes disto este estudo teve como objetivo reduzir o número de variáveis de qualidade da água, bem como identificar as possíveis fontes de poluição do Rio Capivari/RS – Brasil. Para isso, utilizou-se dados de qualidade da água secundários, neles empregou-se a estatística multivariada através da análise de agrupamento hierárquico, na qual a medida de similaridade adotada foi a distância euclidiana. O método Ward's foi utilizado para ligação entre os grupos. Por meio da Análise de Agrupamento, foi possível retirar cinco variáveis de qualidade da água, sendo elas: T_{AR}, NH₃, ST e Cl⁻ e CE. Ademais após o corte do dendrograma, em 70% soi possível identificar as possíveis fontes de poluição. Concluindo-se assim que através da técnica estatística multivariada, por meio da estatística multivariada em função da Análise de Agrupamento, foi possível retirar variáveis similares, contribuindo para a redução dos custos com monitoramento da qualidade da água no Rio Capivari/RS.

Palavras-chave: Monitoramento Ambiental, Análise de Cluster e Qualidade da Àgua.

















CLUSTERIZATION TO REDUCE THE NUMBER OF WATER QUALITY VARIABLES

Abstract: The analysis of the water quality of a spring is of extreme relevance when it comes to water for human consumption. However, there is a need for high investments in the characterization of several parameters of water quality. Acknowledgments This study aimed to reduce the number of water quality variables, as well as to identify the possible sources of pollution of the Capivari River / RS - Brazil. For this, secondary water quality data were used, in which multivariate statistics were employed through the hierarchical grouping analysis, in which the measure of similarity adopted was the Euclidean distance. The Ward's method was used to link the groups. By means of the Grouping Analysis, it was possible to obtain five water quality variables: ART, NH3, TS and Cl⁻ and EC. In addition, after cutting the dendrogram, 70% of the possible sources of pollution can be identified. It was concluded that, through the multivariate statistical technique, through the multivariate statistic as a function of the Grouping Analysis, it was possible to obtain similar variables, contributing to the reduction of costs with monitoring of water quality in the Rio Capivari / RS.

Keywords: Environmental Monitoring, Cluster Analysis and Water Quality.

1. INTRODUÇÃO

A poluição hídrica decorrente das ações antrópicas podem tornar os mananciais impróprios para usos mais restritivos, como por exemplo, ao abastecimento de água (BORA; GOSWAMI, 2016; NAZIR et al., 2016; HELLER; PÁDUA, 2010; BRAGA 2015), tendo ciência que estas decorrem de atividades antrópicas que ocorrem de maneira desordenada, é de extrema importância o monitoramento da qualidade da água (CENTENO, 2017).

È pertinente destacar que quando se referindo a poluição esta pode ser entendida como a entrada de substâncias químicas, físicas e biológicas, que alteram de forma direta ou indireta a natureza dos sistemas hídricos, prejudicando os diferentes usos da água a que se destinam (SPERLING, 2005; ABREU; CUNHA, 2015; TAVARES, 2014). Porquanto estas substâncias quando integradas sem o devido tratamento no ambiente aquático, sejam eles, através de fontes pontuais ou difusas, e dependendo da estrutura hidráulica do manancial em questão, podem impedir a autodepuração dos corpos hídricos e, portanto, alterar o equilíbrio e a dinâmica do manancial (GOMES et al., 2014; SOUZA, 2015).

Sendo assim, faz-se necessário o monitoramento das variáveis limnológias de acordo com os usos pretendidos. Todavia para avaliar os impactos gerados sobre a qualidade da água é necessário analisar as suas variações ambientais, bem como os processos físicos, químicos e biológicos, que ocorrem de maneira dinâmica em uma bacia hidrográfica (BILGIN; KONANÇ, 2016; LOBATO et al., 2015; MONICA; CHOI, 2016).

Ademais quando os parâmetros físicos, químicos e biológicos são analisados separadamente não é possível obter uma compreensão global do corpo hídrico, principalmente por profissionais de outras áreas do conhecimento, necessitando assim de ferramentas capazes de analisálos simultaneamente (GOMES et al., 2014; SOUZA, 2015).

Sendo assim se ter uma visão sistêmica do corpo hídrico é necessário analisar inúmeras variáveis que denotam a qualidade da água, simultaneamento, ou pelo menos que estas estejam de alguma maneira correlacionadas entre si. Todavia este processo torna-se complexo e necessita de altos investimentos econômicos, o que muitas vezes acaba por inviabilizar o processo principalmente em

Realização



Correalização



Informações:







cidades com poucos recursos. Diante disso, é extremamente importante tentar reduzir o número de variáveis limnológias empregadas em estudos ambientais, utilizando, assim, apenas variáveis essenciais para cada corpo d'água estudado, identificando as possíveis fontes de poluição decorrentes de atividades antrópicas.

A Análise de Agrupamento – AA vem sendo muito empregada em estudos de qualidade da água com a finalidade de identificar as principais fontes de poluição e reduzir o número de variáveis limnológicas (HASHEMI et al., 2014; HAIR et al., 2009; MINGOTI, 2013). Ao contrario da Análise de Componentes Principais – ACP, a qual exige que o número de amostras seja 10 vezes maior que o número de variáveis limnológicas, a AA não necessita. Apresentando, assim, fácil empregabilidade no monitoramento hidrológico, em que a periodicidade nas coletas muitas vezes é difícil (VICINI, 2005; CORRAR; PAULO; FILHO DIAS, 2014; BODRUD-DOZA et al., 2016; GOMES et al., 2014; GOMÉZ et al., 2014).

Cabe destacar que o foco principal desta técnica é separa as variáveis de qualidade da água envolvidas no estudo, em grupos, na qual as variáveis limnológias pertencentes ao mesmo grupo sejam similares entre si, porém, diferentes dos demais grupos formados. Além de simplificar a interpretação dos dados de qualidade da água, a Análise de Agrupamento pode, segundo Hair et al. (2009) e Ferreira (2010), identificar as principais fontes de poluição, sendo possível também retirar as variáveis de qualidade da água desde que apresentem a mesma representação no grupo formado (HAIR et al., 2009).

Diante do exposto, este estudo teve como objetivo reduzir o número de variáveis de qualidade da água, bem como identificar as possíveis fontes de poluição do Rio Capivari/RS – Brasil

2. MATERIAIS E MÉTODOS

2.1. Caracterização da Área de Estudo

O Rio Capivari pertence à Bacia Litoral Médio é localizado dentro da Região da Bacia Litorânea (FUNDAÇÃO ESTADUAL DE PROTEÇÃO AMBIENTAL HENRIQUE LUIZ ROESSLE - FEPAM, 2018) e é monitorado pelo Programa de Gerenciamento Costeiro da FEPAM (GERCO/RS). A série temporal das variáveis limnológicas empregadas neste estudo foram referentes ao ponto denominado GER 26.

Sendo que o ponto empregado neste estudo denominado pela FEPAM de GER 26, esta localizado na Longitude -50,5487316; Latitude -30,1440847. Este manancial no que tange a economia, suas disponibilidades estão destinadas para o abastecimento público, plantio, indústria e comércio de arroz irrigado; além de ser utilizado para criação e comércio de ovinos, bovinos, equinos e piscicultura. (FEPAM, 2018).

2.2. Compilação dos Dados

Os dados utilizados neste estudo foram secundários, disponibilizados pela FEPAM/RS, que monitora desde 1992, a cada seis meses, a qualidade da água da região litorânea (FEPAM, 2018). Foram utilizados dados compreendidos entre o período de 2005 a 2013, totalizando 11 amostras, sendo que, a escolha dos parâmetros de qualidade da água foi realizada levando em consideração a menor percentagem de falhas amostrais, decorrentes da inexistencia de dados em determinados períodos.

Os parâmetros de qualidade de água analisados foram: Demanda Bioquímica de Oxigênio (DBO₅), Coliformes Totais (CT), Fósforo Total (PT), Nitrogênio Amoniacal (NH₃), Nitrogênio Total (NT),Oxigênio Dissolvido (OD), Turbidez (TH), Potencial Hidrogeniônico (pH), Temperatura da água (T_{H2O}), Temperatura do Ar (T_{AR}), Condutividade Elétrica (CE), Cloretos (Cl⁻) e Sólidos Totais (ST).

As coletas das amostras foram realizadas pela FEPAM de acordo com a metodologia descrita pela Companhia Ambiental do Estado de São Paulo. Os métodos analíticos, bem como a preservação das amostras, seguem os procedimentos definidos por APHA (1998).

Realização



Correalização





Informações:







2.3. Escalonamento dos dados

O escalonamento dos dados foi realizada para cada variável em cada ponto separadamente através do método do desvio padrão, pois os mesmos apresentam unidades de medidas distintas havendo necessidade de ponderação de cada indicador de qualidade dentro de sua respectiva massa de dados. Na análise fatorial quando se escolhe obter a matriz através do método de correlação os dados devem apresentar-se escalonados.

O escalonamento dos parâmetros de qualidade da água seguindo a metodologia da média foi realizado subtraindo-se os desvios em relação à média de cada variável, por seu desvio padrão, conforme a equação 1 (HAIR, 2009).

$$Y_{ij} = \frac{X_{IJ} - \overline{X_J}}{S(\overline{X}_I)}$$
 Equação (1)

Sendo:

Y_{ij}: variável escalonada; S (X_i): desvio padrão e;

X_{ij}: concentração da variável original;

 \bar{X}_i : média da variável.

Os dados após serem escalonados foram expressos em forma de matriz por $X=(X_{ij})$, onde i=1,2,...,n número de amostras e j=1,2,...p variáveis.

2.4. Análise de Cluster

A Análise de Agrupamento - AA visa a dividir os parâmetros de qualidade da água da amostra em grupos homogêneos, sendo que, para isso, as variáveis similares devem pertencer a um mesmo grupo.

Com este intuito neste estudo, optou-se por utilizar a distância Euclidiana, denominada como distância em linha reta, para medir a similaridade entre as variáveis estudadas, ou seja, medir a semelhança entre os objetos a serem agrupados (HAIR et al., 2009). O método de ligação adotado foi o método da variância mínima ou método de Ward's; de acordo com Ferreira (2010), neste procedimento a escolha de qualquer par de agrupamento é realizada através da combinação de agregados que juntos diminuem a soma interna de quadrados no conjunto completo de agrupamentos separados.

O corte do dendograma no presente estudo foi estabelecido em relação às maiores distâncias em que os grupos foram formados, porém, levando em conta a não heterogeneidade dos dados dentro de cada grupo formado (BERTOSSI et al, 2013; FERREIRA, 2010; MINGOTTI, 2013).

3. RESULTADOS E DISCUSSÃO

A Figura 2 apresenta o comportamento do dendograma com as 13 variáveis linminológicas monitoradas no ponto GER 26.



Realização













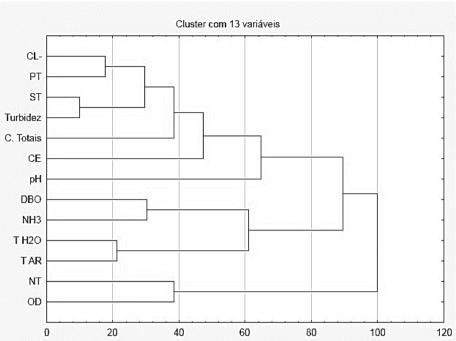


Figura 1: Dendograma do ponto GER 26 com os 13 parâmetros de qualidade da água de um trecho do Rio Capivari

Conforme podemos observar na Figura 1 o dendrograma do ponto GER 26 foi compostos pelas variáveis: Demanda Bioquímica de Oxigênio (DBO₅), Coliformes Totais (CT), Fósforo Total (PT), Nitrogênio Amoniacal (NH₃), Nitrogênio Total (NT), Oxigênio Dissolvido (OD), Turbidez (TH), Potencial Hidrogeniônico (pH), Temperatura da água (T_{H2O}), Temperatura do Ar (T_{AR}), Condutividade Elétrica (CE), Cloretos (Cl⁻) e Sólidos Totais (ST).

Sendo que as variáveis que apresentam a mesma representação dentro de um dendograma gerado podem ser retiradas do estudo (VICINI, 2005). Sendo assim, através do dendograma (Figura 2), observou-se que ocorreu uma similaridade entre as variáveis OD e NT, T_{AR} e T_{H2O} , DBO₅ e NH₃, Turbidez e PT e entre e Cl⁻, portanto optou-se pela retirada das variáveis T_{AR} , NH₃, ST e Cl⁻ e CE. Cabe destacar que entre o OD e NT optou-se por não retirar nenhuma das duas variáveis de qualidade da água tendo em vista a relevância das duas variáveis.

Porquanto é pertinente destacar que dentre a T_{AR} e a T_{H2O} , optou-se por retirar a T_{AR} por ter pouca relevância em estudos ambientais, quando nestes já consta a T_{H2O} sendo um derivado do outro (SPERLING, 2007). Ademais a variável T_{H2O} pode influenciar outras variáveis de qualidade da água, pois a elevação desta variável aumenta a taxa das reações físicas, químicas e biológicas e as taxas de transferência de gases. Já a redução da mesma, diminui a solubilidade dos gases, tornando evidente a importância do monitoramento deste parâmetro (SPERLING, 2005; HELLER, PÁDUA, 2010).

Porquanto a igualdade entre o ST e a Turbidez, fez com que escolhe-se a variável Turbidez para continuar no estudo, tendo em vista que esta representa o grau de interferência da passagem da luz através da água.

Já em relação à NH₃ e DBO₅, optou-se por retirar o NH₃, pois a mesma representa toda a matéria orgânica presente nos corpos hídricos, sendo assim é reponsável pelos microrganismos decompositores, do oxigênio dissolvido na água. Cabe destacar que este é o parâmetro mais importante na caracterização do grau de poluição de um manancial. Além disto, também são representativas dos sólidos totais que foram retirados do estudo (SPERLING, 2007; BRAGA *et al.*, 2015).Todavia este

Realização















manancial em questão possui nos sólidos totais uma maior quantidade de sólidos em suspensão uma vez que esta diretamente ligado a turbidez. (NUVOLARI, 2011).

Entretanto com relação a similaridade existente nas variáveis Cl⁻ e PT, optou-se por retirar os cloretos por ser um sal e ter pouca influência neste curso d'água quando comparado ao PT (SPERLING, 2005). Uma vez que o PT é indispensável para o crescimento de algas e, quando em elevadas concentrações, pode causar eutrofização do manancial, o que pode reduzir a penetração da luz e prejudicar a fotossíntese, que, por consequência, altera a estabilização da matéria orgânica (SPERLING, 2007). Sendo assim na Figura 3 é composta pelo dendograma obtido após a retirada das quatro variáveis.

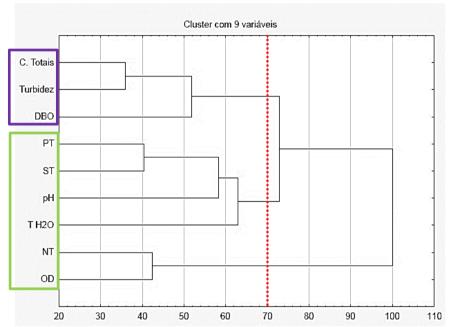


Figura 2: Dendograma do ponto GER 26 com os 9 parâmetros de qualidade da água de um trecho do Rio Capivari

Com base na Figura 2 o dendrograma do ponto GER 26 após a retirada das variáveis similares ficou composto pelas variáveis: Demanda Bioquímica de Oxigênio (DBO₅), Coliformes Totais (CT), Fósforo Total (PT), Nitrogênio Amoniacal (NH₃), Nitrogênio Total (NT), Oxigênio Dissolvido (OD), Turbidez (TH), Potencial Hidrogeniônico (pH), Temperatura da água (T_{H2O}).

De acordo com Hair et al (2009) e Vicini (2005), para identificar as possíveis fontes de poluição, é necessário escolher o melhor ponto de corte do dendograma. Corroborando com este estudo, os trabalhos de Bertossi et al. (2013), o corte do dendograma foi realizado a partir do maior distanciamento sem que ocorresse heterogeneidade dentro dos grupos formados e que através dos mesmos fosse possível explicar as fontes de poluição encontradas.

Na Figura 3, observa-se que a maior variação da distância de ligação ocorre em 70%, sendo que, a partir deste ponto ocorre o maior distanciamento da medida de similaridade para a formação dos grupos, ou seja, ocorre uma maior heterogeneidade entre os grupos formados.

Com isso o dendograma obtido neste estudo indicou a formação de dois grupos homogêneos, como pode ser observado na Figura 3, onde o primeiro grupo possui como variáveis de qualidade da água OD, NT, T_{H2O}, pH, ST e PT e o segundo grupo: Coliformes Totais, DBO₅ e Turbidez.

Sendo que possivelmente o grupo 1 pode ter origem nos despejos domésticos e o segundo grupo as atividades agropecuárias (SPERLING 2007; PRADO, TURETTA, ANDRADE, 2010).

Realização



Correalização











4. CONSIDERAÇÕES FINAIS

Conclui-se que através da Análise de Agrupamento, foi possível retirar variáveis similares, contribuindo para a redução dos custos com monitoramento da qualidade da água no Rio Capivari/RS. Ademais estabeleceu-se dois grupos distintos de ações antrópicas, após o corte do dendrograma.

Agradecimentos

O presente trabalho tem o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Fundação Estadual de Proteção Ambiental (FEPAM), Universidade Federal de Pelotas (UFPel) e Programa de Pós-Graduação em Recursos Hídricos (PPG Recursos Hídricos).

5. REFERÊNCIAS

ABREU, C. H. M.; CUNHA, A. C. Qualidade da Água em Ecossistemas Aquáticos Tropicais Sob Impactos Ambientais no Baixo Rio Jari-AP: Revisão Descritiva. **Biota Amazônia**, v. 5, n. 2, p. 119–131, 2015.

AMERICAN PUBLIC HEALTH ASSOCIATION - APHA. **Standard methods for the examination of water and wastewater**. 20^a ed. Washington: American Public Health Association. APHA/AWWA/WEF, 1998.

BERTOSSI, A. P. A. et al. Seleção e agrupamento de indicadores da qualidade de águas utilizando estatística multivariada. **Semina: Ciências Agrárias,** Londrina, v. 34, n. 5, p.2025-2036, set. 2013. Universidade Estadual de Londrina. http://dx.doi.org/10.5433/1679-0359.

BILGIN, A.; KONANÇ, M. U.. Evaluation of surface water quality and heavy metal pollution of Coruh River Basin (Turkey) by multivariate statistical methods. **Environmental Earth Sciences,** [s.l.], v. 75, n. 12, p.1029-1047, jun. 2016. Springer Nature. http://dx.doi.org/10.1007/s12665-016-5821-0.

BODRUD-DOZA, M. D. et al. Characterization of groundwater quality using water evaluation indices, multivariate statistics and geostatistics in central Bangladesh. **Water Science**, [s.l.], v. 30, n. 1, p.19-40, abr. 2016. Elsevier BV. http://dx.doi.org/10.1016/j.wsj.2016.05.001.

BORA, M.; GOSWAMI, D. C.. Water quality assessment in terms of water quality index (WQI): case study of the Kolong River, Assam, India. **Applied Water Science,** [s.l.], v. 6, n. 1, p.1-11, 27 jul. 2016. Springer Nature. http://dx.doi.org/10.1007/s13201-016-0451-y.

BRAGA, B., et al. **Águas doces no Brasil:** capital ecológico, uso e conservação. 4. ed. São Paulo: Escrituras, 2015.

CENTENO, Luana Nunes. **Proposta metodológica para a construção de um índice de qualidade da água na bacia hidrográfica Piratini-São Gonçalo e Mangueira, RS.** 2017. 190 f. Dissertação (Mestrado) - Curso de Ppg em Recursos Hídricos, Universidade Federal de Pelotas, Pelotras, 2017.

CORRAR, L. J; PAULO, E.; DIAS FILHO, J. M.. **Análise multivariada para os cursos de administração, ciências contábeis e economia..** ed. 5 – Reimpressão. São Paulo: Atlas, 2014. 344 p.

FERREIRA, Daniel Furtado. Estatística multivariada. 2. ed. Lavras: UFLA, 2010. 662 p.

Realização















Fundação Estadual de Proteção Ambiental Henrique Roessler – FEPAM. Monitoramento da qualidade da água da região hidrográfica das bacias litorâneas. Site oficial da FEPAM. 2018. Disponível em: http://www.fepam.rs.gov.br/qualidade/monitor_agua_litoral.asp. Acesso em: 01 mar. 2018.

GOMES, A. I. et al. Optimization of river water quality surveys by multivariate analysis of physicochemical, bacteriological and ecotoxicological data. **Water Resources Management**, v. 28, n. 5, p. 1345–1361, 2014.

GOMÉZ, D. M.; Estudio del modelamiento de la calidad del agua del río Sinú, Colombia. **Revista de Ingenierías: Universidad de Medellín**, v. 12, n. 22, p. 33–44, jan./jun.2013.

HAIR JR., J. F. et al. Análise multivariada de dados. 6. ed. Porto Alegre: Bookman, 2009. 688 p.

HASHEMI, S. et al. Channel characterization using multiple-point geostatistics, neural network, and modern analogy: a case study from a carbonate reservoir, southwest Iran. **Journal of Applied Geophysics**, [s.l.], v. 111, p.47-58, dez. 2014.

HELLER, L.; PÁDUA, V. L. de. **Abastecimento de água para consumo humano.** 2. ed. Belo Horizonte: UFMG, 2010. 418 p.

LOBATO, T. C. et al. Construction of a novel water quality index and quality indicator for reservoir water quality evaluation: A case study in the Amazon region. **Journal of Hydrology**, [s.l.], v. 522, p.674-683, mar. 2015. Elsevier BV. http://dx.doi.org/10.1016/j.jhydrol.2015.01.021.

MINGOTI, S. A.. **Análise de dados através de métodos de estatística multivariada:** uma abordagem aplicada. 2. ed. Minas Gerais: UFMG, 2013. 297 p.

MONICA, N.; CHOI, K.. Temporal and spatial analysis of water quality in Saemangeum watershed using multivariate statistical techniques. **Paddy and Water Environment,** [s.l.], v. 14, n. 1, p.3-17, 6 jan. 2015. Springer Nature. http://dx.doi.org/10.1007/s10333-014-0475-6.

NAZIR, H. M.t al. Classification of drinking water quality index and identification of significant factors. **Water Resources Management,** [s.l.], v. 30, n. 12, p.4233-4246, 21 jul. 2016. Springer Nature. http://dx.doi.org/10.1007/s11269-016-1417-4.

NUVOLARI, A.. **Esgoto sanitário:** coleta, transporte, tratamento e reúso agrícola. 2. ed. São Paulo: Blucher, 2011. 565 p.

PRADO, R. B.; TURETTA, A. P. D.; ANDRADE, A. G. de. Manejo e conservação do solo e da água no contexto das mudanças ambientais. Rio de Janeiro: Embrapa Solos, 2010. 419 p.

SOUZA, M. F. Qualidade da água do canal São Gonçalo-RS/Brasil - uma avaliação hidroquímica considerando seus usos múltiplos. 2015. 105 f. Dissertação (Mestrado em Recursos Hídricos) - Universidade Federal de Pelotas. Pelotas, 2015.

SPERLING, M. V. **Introdução à Qualidade das Águas e ao Tratamento de esgotos**. 3. ed. Belo Horizonte: DESA - UFMG, v. 1, 2005. 452 p.

Realização



Correalização











SPERLING, M.. **Estudos e modelagem da qualidade da água de rios.** Belo Horizonte: UFMG, 2007. 7 v. 588 p.

TAVARES, B. S. **Qualidade de água na bacia hidrográfica do rio Una - Pernambuco**. 2014. 110 f. Tese (Doutorado em Engenharia Agrícola) - Universidade Federal Rural de Pernambuco. Recife, 2014.

VICINI, L; SOUZA, A. M. **Análise multivariada da teoria à pratica**. Santa Maria: Biblioteca Central da UFSM, 2005. 215 p.







