

## ADEQUABILIDADE DE ANÁLISES MULTIVARIADAS PARA AVALIAÇÃO DE DADOS DE PEGADA HÍDRICA

Patrícia da Silva Costa (1); Rener Luciano de Souza Ferraz (2); Newcélia Paiva Barreto (1); Suely de Lima Santos (1); José Dantas Neto (2)\*

(1) Centro de Saúde e Tecnologia Rural – CSTR da Universidade Federal de Campina Grande – UFCG. [paty\\_16costa@hotmail.com](mailto:paty_16costa@hotmail.com), [newcelia.barreto@bol.com.br](mailto:newcelia.barreto@bol.com.br), [suely126@hotmail.com](mailto:suely126@hotmail.com). (2) Centro de Tecnologia e Recursos Naturais – CTRN da UFCG, [ferraz340@gmail.com](mailto:ferraz340@gmail.com), [zedantas@deag.ufcg.edu.br](mailto:zedantas@deag.ufcg.edu.br). \*Orientador

### Introdução

Diagnosticar a pegada hídrica de pessoas em ambientes acadêmicos, por exemplo, pode gerar indicadores preponderantes para massificação de ações educativas com alto índice de sucesso, sobretudo em virtude de esse público ser potencialmente multiplicador de informações relevantes acerca do tema. Entretanto, cabe a observância de que os dados gerados a partir da subjetividade das pessoas integrantes dos grupos amostrais podem apresentar elevada dispersão e invalidar o método de análise estatístico empregado para análise destes dados, fazendo-se necessário que se busque técnicas que melhor descrevam dados subjetivos (SOUZA et al., 2014).

Tendo em vista a necessidade de soluções estatísticas para melhor representatividade dos dados obtidos em pesquisas científicas, postula-se que o emprego de análise de componentes principais (PCA - *Principal Component Analysis*) mostra-se eficiente para análise de dados, tornando a explicação do fenômeno estudado mais simples (MUSINGARABWI et al., 2016). Em sentido complementar, a análise de agrupamento (*Cluster Analysis*) auxilia na visualização e interpretação dos resultados a partir da estrutura de grupos (WANG R et al., 2016). Estas técnicas foram aplicadas com êxito em pesquisas quali-quantitativas a fim de averiguar opinião pública (FULCO et al., 2016), sazonalidade da produção de culturas agroindustriais (FONSECA et al., 2016) e discussões sobre as melhores técnicas de análise de dados (FIGUEIREDO FILHO et al., 2016).

Assim, objetivou-se com este trabalho verificar a adequabilidade de análises multivariadas para avaliação de dados da pegada hídrica obtidos de grupos amostrais da comunidade acadêmica do Centro de Saúde e Tecnologia Rural da Universidade Federal de Campina Grande.

### Metodologia

Para testar a adequabilidade de análises multivariadas, foram coletadas informações, entre os dias 26 e 29 de abril de 2016, no Centro de Saúde e Tecnologia Rural da UFCG localizado no município de Patos – PB, inserido na Mesorregião do Sertão do estado da Paraíba, no Bioma Caatinga.

A pesquisa exploratória (GIL, 2010) consistiu na aplicação de 140 questionários para estimativa da pegada hídrica, conforme metodologia adaptada de Hoekstra et al. (2011). Foram avaliados sete grupos amostrais representados pelos cursos de graduação (Ciências Biológicas, Engenharia Florestal, Medicina Veterinária e Odontologia) e de pós-graduação (Mestrado em Ciências Florestais, Zootecnia e Medicina Veterinária) do CSTR. Em cada grupo, foi tomada amostra de tamanho  $n=20$ . As variáveis consideradas neste estudo foram representadas pelos componentes da pegada hídrica total (PHTot), fracionados em pegada hídrica doméstica (PHDom), industrial (PHInd) e de alimentos (PHAli), sendo esta última fracionada em pegada hídrica de cereais (PHCer), carnes (PHCar), vegetais (PHVeg), frutas (PHFru), laticínios (PHLat), bebidas (PHBeb), gorduras (PHGor), açúcares (PHAçu), ovos (PHOvo) e outros (PHOut).

Os dados foram submetidos à padronização, de modo a tornar a média nula e variância unitária. A estrutura multivariada dos resultados foi avaliada por meio de análise de componentes principais

(ACP), de modo a condensar a quantidade de informação relevante contida no conjunto de dados originais em um menor número de dimensões (componentes principais), resultantes de combinações lineares das variáveis originais geradas a partir dos autovalores mais elevados na matriz de covariância. Para cada componente principal (CP), procedeu-se análise de *cluster* por método hierárquico, *Ward's minimum variance*, considerando-se para tanto as variáveis relevantes na composição de cada componente principal (HAIR et al., 2009).

## Resultados e discussão

Com base na análise de componentes principais (ACP), foi possível condensar o número de variáveis originais em três componentes principais (CP1, CP2 e CP3), os quais juntos retêm 88,67% da variância total acumulada. A escolha destes CPs foi baseada nos autovalores  $\geq 1,0$ . Para a seleção das variáveis, foram adotados valores  $\geq 0,60$  (em módulo), conforme o critério de Kaiser (1958) (Tabela 1).

Tabela 1. Variância relativa e cargas das variáveis associadas aos três primeiros componentes principais formados a partir de 12 variáveis estimadas em sete grupos amostrais da comunidade acadêmica do Centro de Saúde e Tecnologia Rural da UFCG. Patos - PB, 2016.

| CPs | $\sigma^2$ | Cargas das variáveis de Pegada Hídrica |      |       |       |       |       |       |       |       |       |       |       |
|-----|------------|--|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |            | Dom                                    | Ind  | Cer   | Car   | Veg   | Fru   | Lat   | Beb   | Gor   | Açu   | Ovo   | Out   |
| CP1 | 43,77      | -0,71                                  | 0,17 | -0,16 | -0,87 | -0,28 | -0,77 | -0,93 | -0,81 | -0,64 | -0,53 | -0,44 | -0,92 |
| CP2 | 29,53      | -0,61                                  | 0,05 | 0,91  | -0,21 | 0,74  | 0,57  | -0,23 | -0,47 | 0,63  | 0,67  | -0,52 | -0,11 |
| CP3 | 15,37      | -0,13                                  | 0,84 | -0,28 | -0,13 | 0,52  | 0,16  | 0,09  | -0,17 | 0,28  | -0,41 | 0,66  | -0,08 |

CP: componente principal;  $\sigma^2$ : variância relativa; Dom: pegada hídrica domiciliar; Ind: industrial; Cer: cereais; Car: carnes; Veg: vegetais; Fru: frutas; Lat: laticínios; Beb: bebidas; Gor: gorduras; Açu: açúcares e Out: outros componentes da pegada hídrica.

O primeiro componente principal (CP1) reteve 43,77% da variância total relevante. Neste componente, o grupo amostral representado pelo curso de mestrado em Medicina Veterinária foi divergente dos demais por apresentar propriedades específicas, notadamente em relação aos componentes da pegada hídrica domiciliar, de carnes, de frutas, de laticínios, de bebidas, de gorduras e outros. O segundo componente principal (CP2) responde por 29,53% da variância total, sendo este CP importante para discriminar os cursos de graduação em Medicina Veterinária (MVE) e mestrado em Ciências Florestais (PCF), onde este último possui maior PH de cereais, vegetais e açúcares em detrimento destas variáveis no curso de MVE (Figura 1).

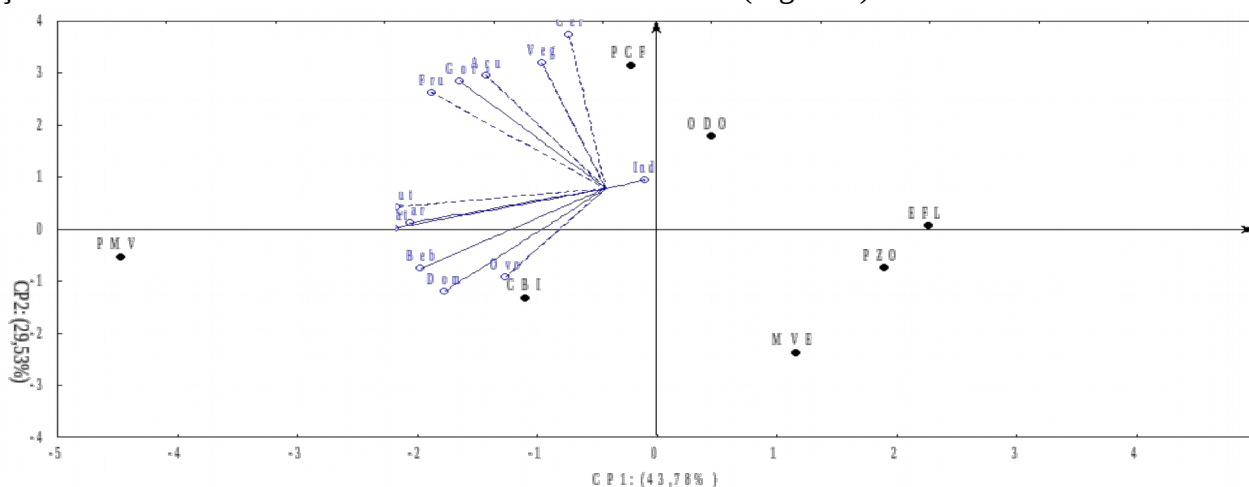


Figura 1. Projeção bidimensional (Biplot) da posição relativa dos grupos amostrais e as respectivas variáveis nos dois primeiros componentes principais (CP1 e CP2). CBI: Ciências Biológicas; EFL: Engenharia Florestal; MVE: Medicina Veterinária; ODO: Odontologia; PCF: Ciências Florestais; PZO: Zootecnia; PMV: Medicina Veterinária. Dom: pegada

hídrica domiciliar; Ind: industrial; Cer: cereais; Car: carnes; Veg: vegetais; Fru: frutas; Lat: laticínios; Beb: bebidas; Gor: gorduras; Açú: açúcares e Out: outros componentes da pegada hídrica. Patos - PB, 2016.

No terceiro componente principal (CP3) ficaram retidos 15,37% da variância total, onde foram separados os grupos amostrais representados pelos cursos de graduação em Engenharia Florestal (EFL) e Odontologia (ODO), de modo que este último curso mencionado possui pegada hídrica industrial e de ovos superiores quando comparado ao curso de Engenharia Florestal, conforme pode ser observado nas projeções dos vetores destas variáveis sobrepostas à posição relativa dos grupos amostrais (Figura 2).

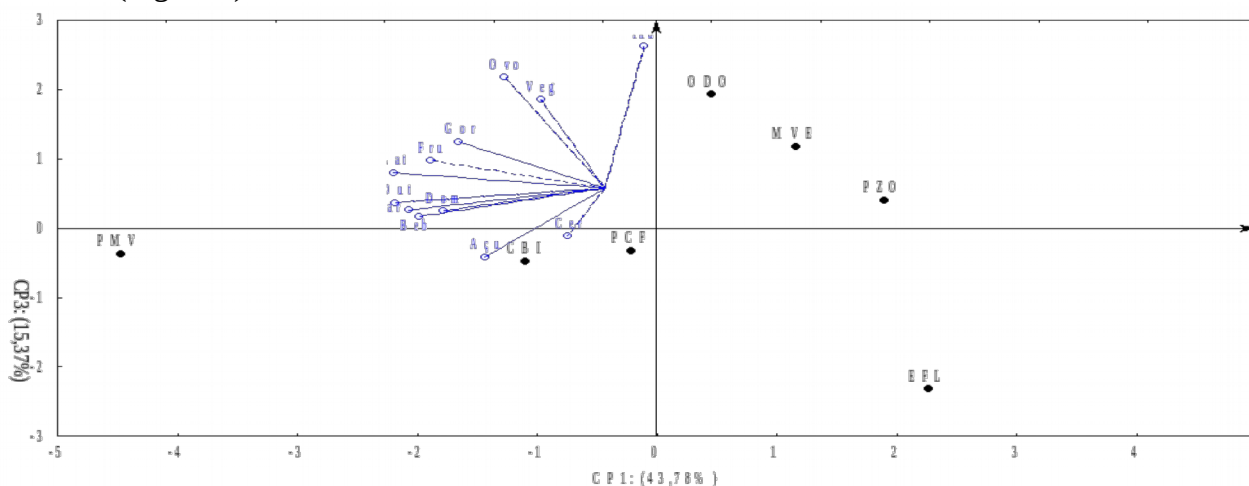


Figura 2. Projeção bidimensional (Biplot) da posição relativa dos grupos amostrais e as respectivas variáveis no primeiro e terceiro componentes principais (CP1 e CP3). CBI: Ciências Biológicas; EFL: Engenharia Florestal; MVE: Medicina Veterinária; ODO: Odontologia; PZF: Ciências Florestais; PZO: Zootecnia; PMV: Medicina Veterinária. Dom: pegada hídrica domiciliar; Ind: industrial; Cer: cereais; Car: carnes; Veg: vegetais; Fru: frutas; Lat: laticínios; Beb: bebidas; Gor: gorduras; Açú: açúcares e Out: outros componentes da pegada hídrica. Patos - PB, 2016.

O cálculo da pegada hídrica é distribuído em três dimensões: uso industrial, doméstico e alimentação. Para cada categoria de entrevistados essas dimensões variaram, o que pode ser atribuído a estilo de dieta, nível de conscientização e renda (MOREIRA e BARROS, 2015). Realmente, estas três dimensões foram evidenciadas em virtude da formação de três CPs a partir das 12 variáveis originais, embora a variável PH doméstica tenha se correlacionado com alguns componentes da PH de alimentos e, também a variável PH industrial tenha se correlacionado com a componente PH de ovos. Estas constatações evidenciam que o emprego de análise de componentes principais (ACP) pode ser empregada com eficiência para discriminar grupos amostrais quanto aos componentes de suas pegadas hídricas totais nas três dimensões. De fato, a ACP tem sido providencial para facilitar o entendimento de fenômenos complexos relacionados à qualidade de água (ROCHA e PEREIRA, 2016) e opinião de pessoas (SILVA et al., 2015; FULCO et al., 2016).

A partir dos três CPs formados, foi possível ilustrar a estrutura de grupos contida nos cursos com base nos componentes da pegada hídrica com expressividade em cada componente principal. Assim, foi possível a formação de três grupos a partir das variáveis (PH: domiciliar, carnes, frutas, laticínios, bebidas, gorduras e outros) com cargas superiores a 0,6 no CP1, sendo o primeiro grupo (Grupo I) formado pelos cursos de mestrado em Ciências Florestais e graduação em Odontologia, o segundo grupo (Grupo II) formado pelos cursos de graduação em Medicina Veterinária e Engenharia Florestal, além do curso de mestrado em Zootecnia, enquanto que um terceiro grupo (Grupo III) foi formado a partir da dissimilaridade entre os cursos de mestrado em Medicina Veterinária e graduação em Ciências Biológicas (Figura 3).

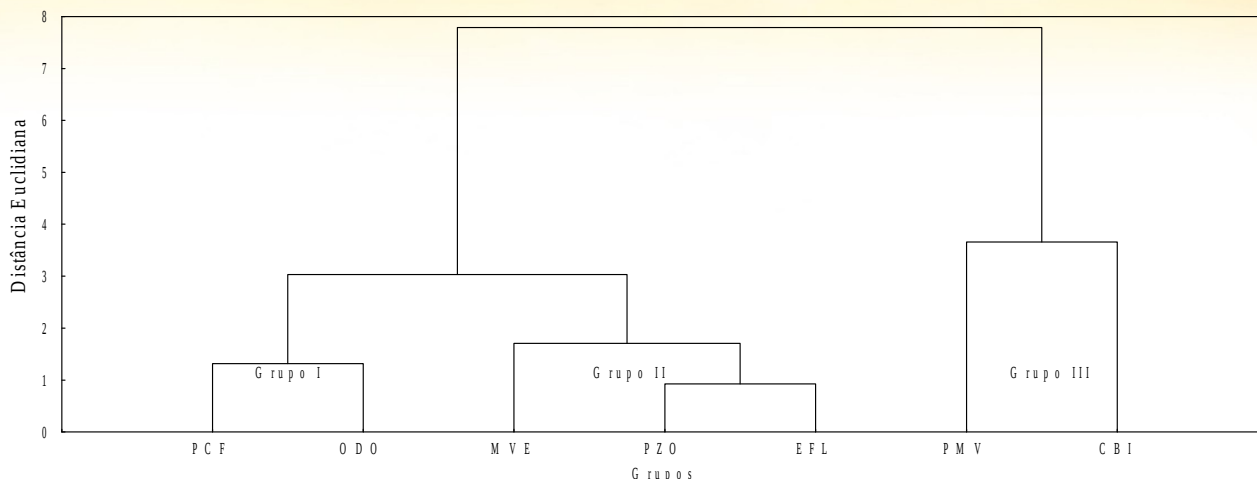


Figura 3. Dendrograma de agrupamento dos grupos amostrais, construído a partir das variáveis com contribuição relevante no primeiro componente principal (CP1). CBI: Ciências Biológicas; EFL: Engenharia Florestal; MVE: Medicina Veterinária; ODO: Odontologia; PCF: Ciências Florestais; PZO: Zootecnia; PMV: Medicina Veterinária.

Com base no segundo componente principal (CP2), formado a partir dos componentes da PH de cereais, vegetais e açúcares, foi possível a formação de dois grupos, sendo o primeiro formado pelos cursos de mestrado em Ciências Florestais e graduação em Odontologia, enquanto que o segundo grupo foi formado pelos cursos de graduação em Medicina Veterinária, Engenharia Florestal e Odontologia, além dos cursos de mestrado em Medicina Veterinária e Zootecnia (Figura 4).

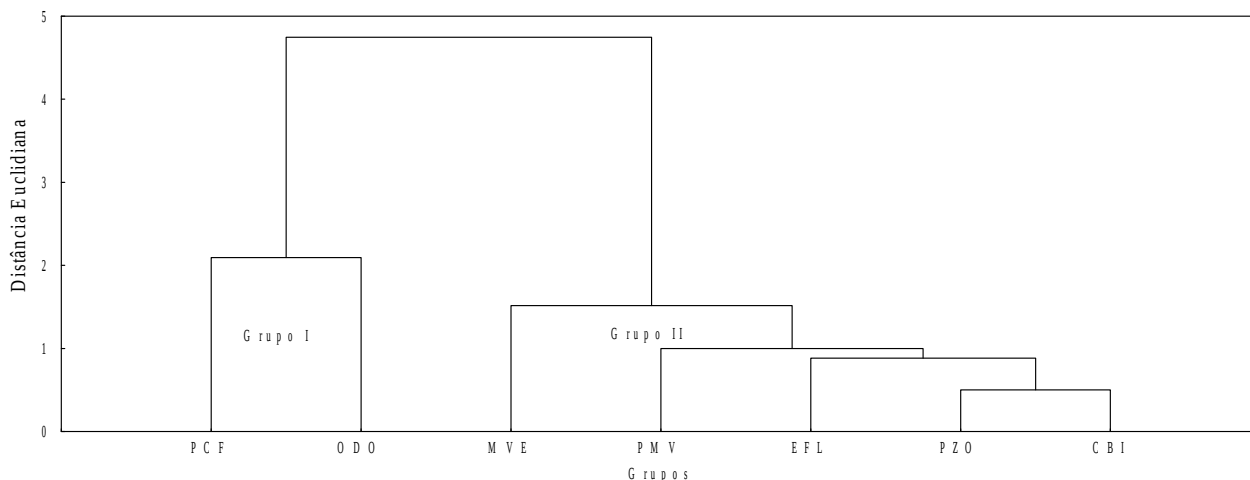


Figura 4. Dendrograma de agrupamento dos grupos amostrais, construído a partir das variáveis com contribuição relevante no segundo componente principal (CP2). CBI: Ciências Biológicas; EFL: Engenharia Florestal; MVE: Medicina Veterinária; ODO: Odontologia; PCF: Ciências Florestais; PZO: Zootecnia; PMV: Medicina Veterinária.

A partir do terceiro componente principal (CP3), formado pela pegada hídrica industrial e de ovos, foi possível aglomerar os cursos em três grupos, sendo o primeiro formado pelo curso de mestrado em Zootecnia e o curso de graduação em Odontologia. O segundo grupo foi formado pelos cursos de pós-graduação e graduação em Medicina Veterinária, enquanto que o terceiro grupo reúne os cursos de graduação em Engenharia Florestal e Ciências Biológicas e o curso de mestrado em Ciências Florestais (Figura 5).

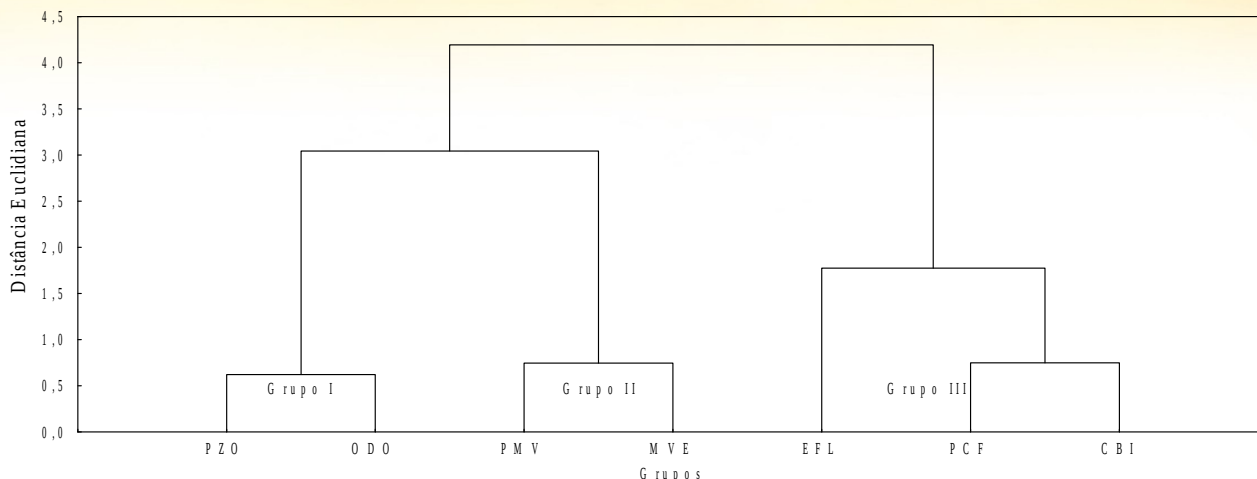


Figura 5. Dendrograma de agrupamento dos grupos amostrais, construído a partir das variáveis com contribuição relevante no terceiro componente principal (CP3). CBI: Ciências Biológicas; EFL: Engenharia Florestal; MVE: Medicina Veterinária; ODO: Odontologia; PCF: Ciências Florestais; PZO: Zootecnia; PMV: Medicina Veterinária.

A formação destes grupos a partir de componentes específicos da pegada hídrica é de grande importância para o planejamento estratégico, notadamente pelo fato de ser possível verificar quais os grupos que possuem maior consumo de água e quais os itens que levam a esse maior consumo, de modo a direcionar ações específicas para cada grupo. Neste sentido, a técnica de agrupamento mostrou-se eficiente para formar aglomerados de cursos em função de suas PHs. Em pesquisas com aplicação de questionários, o emprego de análise de agrupamento mostrou-se eficiente para verificação da satisfação de clientes (ROSÁRIO et al., 2012), da opinião pública sobre potencialidades da mamoneira (FULCO et al., 2016), estratégias de marketing e responsabilidade socioambiental em empresas (SILVA et al., 2015) e percepção das relações no comércio internacional (SILVA et al., 2008), corroborando com os resultados desta pesquisa, denotando que novos estudos podem fazer uso desta técnica com objetividade e facilidade de interpretação de resultados.

## Conclusões

Foi possível reduzir as 12 variáveis originais em três variáveis latentes com 88,67% de toda informação relevante neste estudo, possibilitando maior objetividade na tomada de decisões a respeito das ações necessárias junto aos grupos amostrais prioritários evidenciados no resultado da análise de agrupamento.

Dentre os grupos amostrais, a Pós-graduação em Medicina Veterinária teve maior pegada hídrica de alimentos, enquanto que a graduação em Odontologia teve maior pegada industrial, sendo essas duas categorias de pegada as de maior expressividade para separação de grupos. Esses dois grupos são passivos de atividades de educação socioambiental para melhor alinhamento com as políticas de gerenciamento de recursos hídricos no semiárido.

O emprego de estatística multivariada foi adequado para analisar dados de pegada hídrica obtidos de grupos amostrais representados pelos cursos de graduação e pós-graduação do Centro de Saúde e Tecnologia Rural da Universidade Federal de Campina Grande.

## Referências bibliográficas

FIGUEIREDO FILHO, D. B. et al. Precisamos falar sobre métodos quantitativos em Ciência Política. **Revista Latinoamericana de Metodología de la Investigación Social**, n. 11, p. 21 - 39, 2016.

- FONSECA, G. et al. Acompanhamento da produção de mamona nos municípios paulistas entre os anos de 2005 e 2012. **Revista Educação Ambiental em Ação**, v. 14, p. 6-13, 2016.
- FULCO, R. J. et al. Potencialidades da mamoneira sob a ótica dos populares do município de Matão SP. **Revista Educação Ambiental em Ação**, v. 14, p. 1-5, 2016.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 5. Ed. São Paulo: Atlas, 2010.
- HAIR, J. F. et al. **Análise multivariada de dados**. Porto Alegre: Bookman, 6. ed. p. 688, 2009.
- HOEKSTRA, A. Y. et al. **Manual de avaliação da pegada hídrica: estabelecendo o padrão global**. São Paulo: Instituto de Conservação Ambiental, 2011.
- MOREIRA, R. da S.; BARROS, J. D. de S. Pegada hídrica de classes consumidoras que compõem a escola Antônio Landim de Macêdo em Aurora – CE. **Revista Polêmica**, v. 15, n. 1, p. 18 - 27, 2015.
- MUSINGARABWI, D. M. et al. A rapid qualitative and quantitative evaluation of grape berries at various stages of development using Fourier-transform infrared spectroscopy and multivariate data analysis. **Journal Food Chemistry**, v. 190, n. 1, p. 253-262, 2016.
- ROCHA, C. H. B.; PEREIRA, A. M. Análise multivariada para seleção de parâmetros de monitoramento em manancial de Juiz de Fora, Minas Gerais. **Revista Ambiente & Água**, v. 11, n. 1, p. 176 - 187, 2016.
- ROSÁRIO, C. R. do. et al. Modelo para aplicação de análise multivariada através de técnicas estatísticas PCA e HCA sobre questionário de avaliação da satisfação de clientes: Estudo de caso em uma empresa metalúrgica de embalagens metálicas. **Revista Tecnológica**, v.16, n. 1, p. 30 - 39, 2012.
- SILVA, J. C. G. L. da. et al. O uso da análise de correspondência e de cluster para a percepção das relações no comércio internacional: o caso do setor de móveis sul-brasileiro e as barreiras à Alca. **Revista de Administração**, v. 43, n. 1, p. 44 - 58, 2008.
- SILVA, S. C. da. et al. Embalagens ecologicamente corretas: Estratégias de marketing e responsabilidade socioambiental nas empresas. **Revista Educação Ambiental em Ação**, v. 14, p. 1-12, 2015.
- SOUZA, J. L. et al. Pegada hídrica de uma comunidade de consumidores em Fortaleza/CE/Brasil: análise das pegadas rápida e estendida na metodologia ‘waterfootprint network’. **Revista Econômica do Nordeste**, v. 25, n. 3, p. 17 - 32, 2014.
- WANG, R. et al. Flower Pollination Algorithm with Bee Pollinator for cluster analysis. **Journal Information Processing Letters**, v. 116, n. 1, p. 1-14, 2016.