

## **AVALIAÇÃO DO IMPACTO DO SANEAMENTO BÁSICO NO DESENVOLVIMENTO SÓCIOECONÔMICO DO SEMIÁRIDO BRASILEIRO ATRAVÉS DE ANÁLISE DE COMPONENTES PRINCIPAIS**

Eckhard Mozart Bezerra da Costa (1); Fabiana de Oliveira Araújo (2)

(<sup>1</sup>Universidade Federal do Rio Grande do Norte, mozartbezerra@hotmail.com; <sup>2</sup>Universidade Federal do Rio Grande do Norte, fabianabio@gmail.com)

**Resumo:** O semiárido brasileiro ocupa uma extensa área no território nacional e abriga mais de 20 milhões de pessoas. Suas características climáticas, muitas vezes são tomadas como justificativa para o baixo desenvolvimento socioeconômico das cidades da região. Entretanto, o semiárido brasileiro é o mais chuvoso do mundo e outras peculiaridades colaboram para essa realidade. Em meio a isso, essa região recebe baixos investimentos no setor de saneamento e por isso possui índices sanitários desfavoráveis. Dessa forma, esse artigo buscou evidenciar que as cidades com melhores índices de desenvolvimento socioeconômico (representado por IDHM e PIB) são as que possuem as melhores taxas de abrangência de atendimento desses serviços à população. A ferramenta utilizada para isso foi a análise de componentes principais que é uma técnica da estatística multivariada que consiste em transformar um conjunto de variáveis originais em outro conjunto de variáveis de mesma dimensão denominadas de componentes principais. Os componentes principais apresentam propriedades importantes: cada componente principal é uma combinação linear de todas as variáveis originais, são independentes entre si e estimados com o propósito de reter, em ordem de estimação, o máximo de informação, em termos da variação total contida nos dados. Dessa forma, o trabalho também indicou quais medidas de saneamento básico estão mais associadas com desenvolvimento socioeconômico e exibiu um ranking de desempenho, das cidades do semiárido norterio-grandense, com base nos índices sanitários e de desenvolvimento socioeconômico. Os resultados servirão de base para a elaboração de planos de saneamento e para melhorar a distribuição de investimentos na área.

**Palavras-chave:** saneamento básico, desenvolvimento socioeconômico, PCA.

**Palavras-Chave:** saneamento básico, desenvolvimento socioeconômico, PCA.

### **INTRODUÇÃO**

O semiárido brasileiro se estende por oito estados da região Nordeste (Alagoas, Bahia, Ceará, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe) mais o norte de Minas Gerais. A região ocupa quase um quinto do território nacional, abrange mais de 20% dos municípios brasileiros e abriga mais de 23 milhões de pessoas (BRASIL, 2010). Segundo os dados populacionais, a grande maioria (93,39%) dos 1.135 municípios do semiárido brasileiro é considerada de pequeno porte, seguida de 5,02% médio e 1,59% de grande porte, com contingentes populacionais variando entre 0 a 50.000, 50.001 a 100.000 e 100.001 a 900.000 habitantes, respectivamente. Os municípios de pequeno porte residem 65,23% da população total do semiárido, enquanto que os de médio porte 16,48% e nos grandes, 18,30%. (BRASIL, 2014).

(83) 3322.3222

contato@conidis.com.br

[www.conidis.com.br](http://www.conidis.com.br)

Em relação ao clima, a escassez de chuva e o limitado sistema de armazenamento de água pluvial tornam a região com características que se aproximam da aridez. Dessa forma, geralmente se afirma que não chove o suficiente, o que é uma verdade relativa, pois existem diferenças marcantes do ponto de vista da precipitação pluviométrica anual de uma região para outra. O Semiárido brasileiro é o mais chuvoso do mundo, porém, as chuvas são concentradas em poucos meses e mais de 90% de suas águas não são aproveitadas em virtude da evaporação e do escoamento superficial. (BRASIL, 2017).

Quanto ao desenvolvimento socioeconômico, no semiárido brasileiro 60,09% dos seus municípios apresentaram Índice de Desenvolvimento Humano Municipal (IDHM) variando de Muito Baixo a Baixo envolvendo mais de 9 milhões de habitantes. Além disso, todos os municípios do semiárido brasileiro apresentaram IDHM inferior ao do Brasil (0,727). Apenas uma pequena parcela dos habitantes do semiárido reside em municípios com IDHM considerado alto. O Produto Interno Bruto *per capita* (PIB) do Semiárido brasileiro (R\$ 6.520,35) apresenta-se quase 70% menor que a média brasileira e quase 32% menor que o da região Nordeste. A grande maioria dos municípios do Semiárido apresentam PIB inferior a R\$ 6.520,35 (BRASIL, 2014).

Contudo, a maioria dos problemas ligados ao semiárido brasileiro não são de ordem climática, mas sim decorrentes da ausência de políticas que promovem o desenvolvimento da região, como a implantação de medidas de saneamento básico (BRASIL, 2017). É tanto que, a universalização do serviço de abastecimento de água não é uma realidade em todos os municípios da região; atualmente mais de 34000 habitantes não contam com fornecimento de água através de rede de distribuição e utilizam-se de soluções alternativas de acesso a água como chafariz público, bica, carro pipa e/ou cisterna. Estudos realizados apontaram que para o cenário do ano de 2015, menos de 30% das sedes municipais contariam com abastecimento de água satisfatório e em mais de 15% das sedes haveria necessidade de investimentos em novos mananciais para garantir a oferta de água. O ponto crítico dos sistemas de abastecimento de água do semiárido brasileiro está mais associado a sua capacidade de produção do que na insuficiência de água nos atuais mananciais existentes. Com respeito ao esgotamento sanitário, apenas 243 sedes municipais possuem sistema de coleta de esgoto; contudo, a existência deste serviço não constitui garantia de atendimento a toda a população urbana dessas sedes, visto que menos de 45% desta população é efetivamente atendida. Neste contexto estima-se que na região semiárida cerca de 10,9 milhões de habitantes das áreas urbanas não dispõem do serviço de coleta de esgoto, sendo as fossas, sumidouros, valas a céu aberto

e/ou lançamento direto nos corpos hídricos, os principais destinos dos dejetos gerados evidenciando a ausência de investimentos e a exposição dos habitantes a diversas doenças infecciosas e parasitárias (BRASIL, 2014).

Em vista disso, apesar da baixa atenção reservada ao setor, saneamento básico é um fator essencial para o desenvolvimento econômico e social de um país. Os serviços de água tratada, coleta e tratamento dos esgotos levam à melhoria da qualidade de vida das pessoas, sobretudo na saúde infantil com redução da mortalidade infantil, melhorias na educação, na expansão do turismo, na valorização dos imóveis, na renda do trabalhador, na despoluição dos rios e na preservação dos recursos hídricos (BRASIL, 2017). Esta conclusão é a mesma que se obtém quando se faz a análise da história da humanidade e também a mesma que atualmente pode-se encontrar em publicações referentes a estudos e pesquisas científicas. Resultados de diversas pesquisas na área do saneamento ambiental têm comprovado que comunidades ou cidades com características semelhantes, exceto na situação do saneamento do meio, encontram-se em diferentes condições de desenvolvimento econômico e, sobretudo, de saúde pública, sendo notadamente melhores os níveis de saúde e qualidade de vida das populações que dispõem de eficientes serviços de saneamento básico. Do mesmo modo que, uma análise da história da humanidade sob a ótica do desenvolvimento e da qualidade de vida, mostra que toda civilização considerada desenvolvida em certa época era também a que detinha maior domínio sobre as boas práticas de saneamento ambiental naquele tempo (ANDRADE NETO, 2017).

Diante do exposto e com base nos dados de IDHM, PIB e de Serviços de Saneamento Básico do semiárido, este trabalho teve como objetivo geral avaliar o efeito das medidas saneamento básico no desenvolvimento socioeconômico da região, apontando as medidas de maior influência. Também foi elaborado um ranking de desempenho das cidades do semiárido norterriograndense em relação aos dados coletados. A partir dos resultados os gestores poderão direcionar políticas públicas diferenciadas de acordo com os níveis de desenvolvimento socioeconômico e deficiências sanitárias de cada cidade.

## **METODOLOGIA**

### **Local de estudo**

O presente estudo foi realizado em 67 municípios pertencentes à região semiárida do estado do Rio Grande do Norte, Brasil. Os dados foram fornecidos pelo IBGE (Instituto Brasileiro de Geografia e Estatística) e SNIS (Sistema Nacional de Informações sobre Saneamento).

Foram coletados, de cada município, dados de IDHM e PIB per capita (representando o desenvolvimento socioeconômico) e taxas de atendimento à população dos serviços de abastecimento de água (AA), esgotamento sanitário (ES) e coleta de resíduos sólidos (RS) (representando o saneamento básico). Para a pesquisa atual, essa abordagem foi suficiente, pois as variáveis IDHM e PIB são índices que englobam juntas três dimensões do desenvolvimento humano: longevidade, educação e renda. Além do mais, as outras variáveis compreendem três das quatro atividades que compõem o sistema de saneamento básico no Brasil.

### **Análise de dados**

Uma das técnicas estatísticas mais utilizadas em diversas áreas do conhecimento científico é a análise multivariada, na qual um dos objetivos é a redução do número de variáveis utilizadas, ou ainda a possibilidade de se investigar o comportamento espacial e temporal das variáveis consideradas. Dentre as técnicas de análise multivariada, será utilizada neste artigo a análise de componentes principais (PCA).

A PCA é uma técnica de modelagem da estrutura de covariância. Essa técnica estatística transforma linearmente um conjunto original de variáveis, inicialmente correlacionadas entre si, num conjunto substancialmente menor de variáveis não correlacionadas que contém a maior parte da informação do conjunto original (HONGYU, 2016). Procura-se redistribuir a variação observada nos eixos originais de forma a se obter um conjunto de eixos ortogonais não correlacionados (MANLY, 1986; HONGYU, 2015). O objetivo principal da PCA é o de explicar a estrutura da variância e covariância de um vetor aleatório, composto de  $p$ -variáveis aleatórias, por meio de combinações lineares das variáveis originais. Essas combinações lineares são chamadas de componentes principais e são não correlacionadas entre si. Os componentes principais retidos devem ter autovalores (variâncias) maiores que um (Regra de Kaiser). Esta técnica pode ser utilizada para geração de índices e agrupamento de indivíduos. A análise agrupa os indivíduos de acordo com sua variação, isto é, os indivíduos são agrupados segundo suas variâncias, ou seja, segundo seu comportamento dentro da população.

Para se estabelecer um índice que possibilite ordenar um conjunto de  $n$  objetos, segundo critério definido por um conjunto de  $m$  variáveis adequadas, é necessário escolher os pesos ou ponderações das variáveis de tal forma que traduzam a informação contida na variável. Logo, é

desejável que o índice contenha o máximo de informação fornecida pelo conjunto de variáveis selecionadas.

### Validação da análise

Devido a variância não constante dos resíduos, os dados sofreram transformação  $\log_x+1$ . Além disso, para realizar a análise, os dados da matriz de correlação precisaram ser correlacionados, porém não independentes. Essa condição foi verificada através da matriz de correlação dos dados e pelo Teste de Esfericidade de Bartlett. Esse teste avalia a hipótese de que as variáveis sejam não correlacionadas na população. A hipótese nula do teste diz que a matriz de correlação da população não é uma matriz identidade, o que indica que a PCA é apropriada.

## RESULTADOS E DISCUSSÃO

### Validação da análise

Os dados da matriz de correlação (Tab. 1) são suficientemente correlacionados, com exceção das variáveis ES e RS. Contudo, a baixa correlação entre essas variáveis não influenciou na qualidade dessa análise, já que o objetivo foi avaliar a associação entre variáveis de grupos distintos (desenvolvimento socioeconômico e saneamento básico). Da mesma forma, o resultado do Teste de Esfericidade de Bartlett (Tab. 1) também atendeu aos pressupostos de utilização da PCA.

Tabela 1-Matriz de correlação de Pearson e resultado do Teste de Bartlett

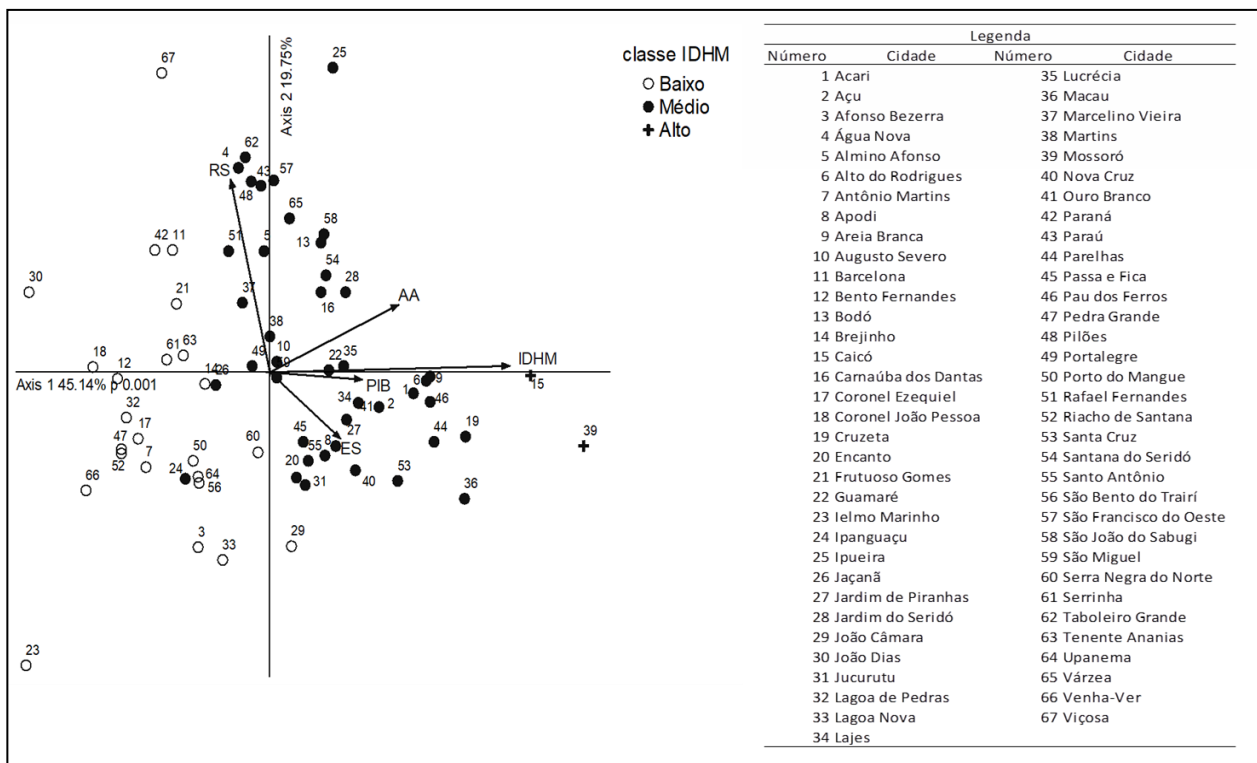
	Matriz de correlação de Pearson					Teste de Bartlett	
	IDHM	PIB	AA	ES	RS		
IDHM	1					Qui-quadrado	67,74
PIB	0,38	1				Graus de liberdade	10
AA	0,62	0,17	1			p-valor	< 0,001
ES	0,31	0,13	0,21	1			
RS	0,37	0,17	0,49	-0,01	1		

### Análise de componentes principais

A análise de componentes principais (PCA) explicou 64,9% da variabilidade dos dados nos dois primeiros eixos (eixo1 = 45,15%; eixo 2= 19,75%; Fig. 1). As variáveis mais importantes para o eixo 1 foram, em ordem decrescente, IDHM (0,93), PIB (0,58), AA (0,68) e ES (0,50). Em relação ao eixo 2, RS (0,80) foi a variável mais importante. Os resultados da PCA indicaram que o

primeiro componente principal refletiu a classificação do IDHM. No lado negativo do eixo 1, as unidades de amostragem das cidades com baixo IDHM não apresentaram alta correlação com nenhuma variável. No lado positivo do eixo 1, as unidades de amostragem das cidades com IDH médio e alto correlacionaram-se com AA, ES e PIB.

Figura 1-Análise de componentes principais aplicada às variáveis socioeconômicas e sanitárias das cidades do semiárido norterio-grandense. IDHM=índice de desenvolvimento humano municipal; PIB=produto interno bruto per capita; AA=abast. de água; ES=esgot. sanitário.



A Tabela 2 apresenta o ranking de desempenho das cidades do semiárido norterio-grandense por funções sócioeconômicas e sanitárias, no ano de 2010, com base no Índice Geral de Desempenho (IGD). Conforme esclarecimentos de Mingoti e Silva (1997), Vyas e Kumaranayake (2006) e Mingoti (2007), o IGD pode ser realizado apenas com o componente principal 1. Entretanto, o CP1 apresenta um poder de explicação da variância de 45,15% da variância. Desse modo, para proporcionar robustez na avaliação de desempenho foram realizados os cálculos dos componentes 1 a 4, que representam a explicação de 90,13% da variância total ( $45,15+19,75+14,04+11,19 = 90,13\%$ ). No caso, não foi observada a regra de utilizar autovalores somente acima de um (KAISER, 1960), a fim de atender os esclarecimentos de Manly (2008) em utilizar porcentagens altas de variância (80% ou mais) para todos p componentes nas análises



posteriores. Assim, o Índice Geral de Desempenho (IGD), representa a ordenação de desempenho com maior poder de explicação (média ponderada pelos autovalores).

Tabela 2-Ranking das cidades do semiárido norterio-grandense por índices socioeconômicos e sanitários em 2010.

P	Cidade	IGD	P	Cidade	IGD
1°	Mossoró	2,13	35°	São Miguel	0,15
2°	Caicó	2,00	36°	Bodó	0,14
3°	Parelhas	1,17	37°	Apodi	0,13
4°	Acari	1,15	38°	Marcelino Vieira	0,09
5°	Pau dos Ferros	1,14	39°	Portalegre	0,06
6°	Ipueira	1,05	40°	Rafael Fernandes	0,05
7°	Macau	0,92	41°	Augusto Severo	0,05
8°	Santa Cruz	0,86	42°	Serra Negra do Norte	-0,17
9°	Areia Branca	0,81	43°	Jaçanã	-0,25
10°	Cruzeta	0,79	44°	Viçosa	-0,30
11°	São João do Sabugi	0,77	45°	João Câmara	-0,37
12°	Ouro Branco	0,76	46°	Barcelona	-0,50
13°	Jardim do Seridó	0,75	47°	Frutuoso Gomes	-0,56
14°	Lucrécia	0,67	48°	Tenente Ananias	-0,58
15°	Alto do Rodrigues	0,66	49°	Paraná	-0,61
16°	Lajes	0,63	50°	Brejinho	-0,62
17°	Açu	0,62	51°	São Bento do Trairí	-0,63
18°	Várzea	0,60	52°	Afonso Bezerra	-0,71
19°	Santana do Seridó	0,57	53°	Serrinha	-0,74
20°	Carnaúba dos Dantas	0,56	54°	Lagoa Nova	-0,75
21°	Nova Cruz	0,54	55°	Ipanguaçu	-0,77
22°	São Francisco do Oeste	0,48	56°	Upanema	-0,88
23°	Jardim de Piranhas	0,44	57°	Antônio Martins	-1,00
24°	Paraú	0,35	58°	Coronel Ezequiel	-1,05
25°	Almino Afonso	0,34	59°	Bento Fernandes	-1,14
26°	Passa e Fica	0,33	60°	Lagoa de Pedras	-1,15
27°	Santo Antônio	0,30	61°	Riacho de Santana	-1,25
28°	Pilões	0,28	62°	Coronel João Pessoa	-1,31
29°	Taboleiro Grande	0,26	63°	Porto do Mangue	-1,35
30°	Água Nova	0,26	64°	Pedra Grande	-1,41
31°	Encanto	0,25	65°	Venha-Ver	-1,57
32°	Guamaré	0,23	66°	João Dias	-1,68
33°	Martins	0,19	67°	Ielmo Marinho	-2,37
34°	Jucurutu	0,17			

## CONCLUSÕES

A Análise de Componentes Principais (PCA) permite elaborar um índice de desempenho sócioeconômico e sanitário com base nos indicadores das cidades do semiárido norterio-grandense do ano de 2010. A pesquisa avaliativa amparada em diversos indicadores podem funcionar como um instrumento importante para a melhoria da eficiência do gasto público. Os resultados permitem que os gestores direcionem políticas públicas diferenciadas de acordo com os níveis de desenvolvimento socioeconômico e deficiências sanitárias de cada cidade. Assim, a pesquisa se torna importante, porque desenvolve um modelo empírico de avaliação de políticas públicas que pode ser reaplicado em outras regiões e em diferentes épocas. Por fim, sugere-se a realização de mais pesquisas avaliativas na área, abordando: gastos em infraestrutura de saneamento, índice de escassez hídrica, expectativa de vida, entre outros.

## REFERÊNCIAS

ANDRADE NETO, C. O. de. Desafios na Universalização do Saneamento. p. 51-64. In: João Batista Lucena de Assis. (Org.). Controle Social no Saneamento: Perspectiva para uma Cidade Saudável. Natal: ARSBAN, 2007.

BRASIL. IBGE - Instituto Brasileiro de Geografia e Estatística. Informações sobre os municípios brasileiros, 2010. Disponível em: <<https://cidades.ibge.gov.br/xtras/home.php?lang=>>>. Acesso em: 16 de outubro de 2017.

BRASIL. SNIS – Sistema Nacional de Informações sobre Saneamento. Série histórica, 2010. Disponível em: <<http://app.cidades.gov.br/serieHistorica/>>. Acesso em: 16 de outubro de 2017.

BRASIL. INSA. Abastecimento Urbano de água: panorama para o semiárido brasileiro, 2014. Disponível em: <<http://sigsab.insa.gov.br/welcome>>. Acesso em: 16 de outubro de 2017.

BRASIL. Trata Brasil. O que é saneamento?, 2017. Disponível em: <<http://www.tratabrasil.org.br/o-que-e-saneamento>> . Acesso em: 16 de outubro de 2017.

HONGYU, Kuang; SANDANIELO, Vera Lúcia Martins; DE OLIVEIRA JUNIOR, Gilmar Jorge. Análise de Componentes Principais: Resumo Teórico, Aplicação e Interpretação. E&S Engineering and Science, 2016, 5.1: 83-90.

HONGYU, K. Comparação do GGE-biplot ponderado e AMMI-ponderado com outros modelos de interação genótipo × ambiente. 2015. 155p. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2015.

KAISER, H. F. The application of electronic computers to factor analysis. Educational and Psychological Measurement, v. 20, p. 141 – 151, 1960.



MANLY, B. J. F. Métodos estatísticos multivariados: uma introdução. Tradução de Sara Ianda Carmona. 3. ed. Porto Alegre: Bookman, 2008. 229p.

MANLY, B. F. J. Multivariate statistical methods. New York, Chapman and Hall, 1986. 159 p.

MINGOTI, S. A. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. 1ª reimpressão. Belo Horizonte: Editora UFMG, 2007. 297p.

MINGOTI, S. A.; SILVA, A. F. Um exemplo de aplicação de técnicas de estatística multivariada na construção de índices de preços. Nova Economia, Belo Horizonte, v. 7, n. 2, 1997.

VYAS, S.; KUMARANAYAKE, L. Constructing socio-economic status indices: how to use principal components analysis. Health Policy and Planning, 9 October 2006. 21 (6):459-468. DOI:10.1093/heapol/czl029.