

Avaliação da proliferação da dengue nos estados brasileiros por meio da Análise de Cluster

Assessment of dengue proliferation in Brazilian states using Cluster Analysis

Evaluación de la proliferación del dengue en los estados brasileños mediante Análisis de Cluster

DOI:10.34119/bjhrv7n2-371

Originals received: 03/15/2024

Acceptance for publication: 04/01/2024

Priscila Neves Faria

Doutora em Ciências (Estatística e Experimentação Agronômica) pela Escola Superior de Agricultura "Luiz de Queiroz" - Universidade de São Paulo (ESALQ - USP)

Instituição: Universidade Federal de Uberlândia

Endereço: Uberlândia, Minas Gerais, Brasil

E-mail: priscilaneves@ufu.br

Bruna Queiroz de Melo Prado

Graduada em Estatística

Instituição: Universidade Federal de Uberlândia

Endereço: Uberlândia, Minas Gerais, Brasil

E-mail: bruna.est.ufu@hotmail.com

RESUMO

A cada ano a dengue vem mostrando que se tornou um grave problema de saúde pública no Brasil e em 2024, o rápido surto de dengue no Brasil já se transformou em epidemia em seis estados. Epidemias requerem estudo das possíveis causas de cada região, bem como a discussão de estratégias que visem à prevenção. Com o objetivo de analisar o número absoluto de casos prováveis e a incidência da dengue em 26 estados brasileiros, nos anos de 2020, 2021 e 2023, o presente estudo utilizou a análise de cluster, técnica do aprendizado de máquina não supervisionado. Dentre os resultados, foi obtido que o padrão de formação dos clusters se altera ao avaliar número absoluto de casos prováveis e a incidência da dengue, concluindo que o impacto que informar e detalhar sobre a incidência de dengue é de suma importância no registro de informações associadas à proliferação da doença, bem como identificação dos estados cujo políticas de combate precisam ser direcionadas.

Palavras-chave: similaridade, agrupamentos, epidemiologia, *Aedes aegypti*.

ABSTRACT

Each year dengue has shown that it has become a serious public health problem in Brazil and in 2024, the rapid outbreak of dengue in Brazil has already turned into an epidemic in six states. Epidemics require study of the possible causes in each region, as well as the discussion of strategies aimed at prevention. With the objective of analyzing the absolute number of probable cases and the incidence of dengue in 26 Brazilian states, in the years 2020, 2021 and 2023, the present study used cluster analysis, an unsupervised machine learning technique. Among the

results, it was obtained that the pattern of cluster formation changes when evaluating the absolute number of probable cases and the incidence of dengue, concluding that the impact of informing and detailing the incidence of dengue is of paramount importance in recording associated information the spread of the disease, as well as identification of states whose combat policies need to be targeted.

Keywords: similarity, groupings, epidemiology, *Aedes aegypti*.

RESUMEN

Cada año, el dengue ha demostrado que se ha convertido en un grave problema de salud pública en Brasil y, en 2024, el rápido brote de dengue en Brasil ya se ha convertido en una epidemia en seis estados. Las epidemias requieren el estudio de las posibles causas en cada región, así como la discusión de estrategias encaminadas a su prevención. Con el objetivo de analizar el número absoluto de casos probables y la incidencia de dengue en 26 estados brasileños, en los años 2020, 2021 y 2023, el presente estudio utilizó análisis de conglomerados, una técnica de aprendizaje automático no supervisado. Entre los resultados se obtuvo que el patrón de formación de conglomerados cambia al evaluar el número absoluto de casos probables y la incidencia de dengue, concluyendo que el impacto de informar y detallar la incidencia de dengue es de suma importancia en el registro de información asociada a la propagación de la enfermedad, así como la identificación de los Estados cuyas políticas de combate deben dirigirse.

Palabras clave: similitud, agrupaciones, epidemiología, *Aedes aegypti*.

1 INTRODUÇÃO

A dengue é uma doença que se tornou um grave problema de saúde pública no Brasil e possui um padrão sazonal, com maior taxa de incidência nos cinco primeiros meses do ano, período mais quente e úmido, característico dos climas tropicais (Braga; Valle, 2007). Essa arbovirose é causada por quatro tipos de vírus do gênero *Flavivirus* e transmitida por mosquitos do gênero *Aedes*.

Segundo Silva et al. (2022), a dengue constitui um sério problema de saúde pública, devido a gravidade de sua infecção, podendo evoluir para o óbito. A necessidade de estudo e acompanhamento dessa doença está relacionada à sua morbidade, mortalidade e necessidade de várias estratégias para o seu controle (Ferreira et al., 2009).

A dengue ocorre em climas tropicais e subtropicais, principalmente em áreas urbanas e semiurbanas e o mosquito *Aedes aegypti*, o principal vetor urbano da dengue, possui uma grande capacidade de adaptação ao ambiente urbano. São vários os fatores que contribuem para a proliferação deste mosquito e, de acordo com Ribeiro et al. (2020), as áreas urbanas possuem as condições favoráveis para sua proliferação já que o processo de urbanização está cada vez

mais rápido, gerando alteração climática e aumento da circulação de pessoas, ocasionando muitas vezes condições ambientais precárias dos grandes centros como aumento de lixo, deficiência no abastecimento de água e na limpeza urbana. Além disso, impactos ambientais nas paisagens e nos ecossistemas, a globalização, as condições climáticas favoráveis à reprodução dos vetores como ambiente quente e úmido, além da falta de implementação de ações eficientes de combate e controle vetorial são fatores importantes que contribuem para a proliferação do mosquito.

A incidência da dengue vem crescendo substancialmente no mundo nas últimas duas décadas, com casos relatados à OMS que vão de meio milhão no ano 2000 para acima de 4,2 milhões no ano de 2022 (Nações Unidas, 2023). Segundo a Organização Mundial de saúde (OMS), aproximadamente metade da população mundial está em risco de contrair a doença. Segundo o Ministério da Saúde, o país concentra quase 70% dos casos notificados da doença nas américas desde 2000.

A importância da produção científica em torno dessa temática é visível, no intuito de descrever o perfil dos casos notificados de dengue em território brasileiro, contribuindo, assim, para o redirecionamento das estratégias dos programas de controle dessa endemia e fomentar ações de prevenção à doença e promoção à saúde, visto que ainda não há controle efetivo da dengue no país.

No intuito de identificar o perfil dos casos de dengue nos estados, a análise de cluster é uma metodologia que vem para somar. Constitui um dos métodos multivariados mais utilizados para se classificar elementos em categorias de similaridade e vem ganhando cada vez mais espaço na área da epidemiologia uma vez que além de avaliar semelhanças entre localidades em relação às variáveis analisadas, também tem sido frequentemente utilizada avaliando a similaridade em relação a um período de tempo. Esse método considera um conjunto inicial de elementos, aos quais são associadas medidas de várias grandezas, utilizadas para se obter grupos de objetos assemelhados em relação aos valores assumidos por essas variáveis (Everitt, 1993). Além de possibilitar a construção de grupos de acordo com as similaridades dos indivíduos, a análise de cluster possibilita também representá-los de maneira bidimensional, por meio de um dendrograma (Moita Neto e Moita, 1998).

Na área de saúde pública, a análise de cluster é aquela que procura obter subgrupos de indivíduos na população com riscos em relação a determinados agravos para possíveis prevenções (Frei, 2006). A aplicação da análise de cluster em estudos envolvendo saúde pública vem compor o arsenal de métodos da epidemiologia na identificação de regiões preocupantes no que tange a proliferação de doenças.

Silva e Machado (2018) aplicaram a análise de cluster nas capitais do Nordeste brasileiro a fim de obter o nível de similaridade entre dengue e variáveis climáticas, sociodemográficas e de saneamento no período de 2001 a 2012. Dentre os resultados obtidos, uma alta correlação da dengue com os índices de tratamento de água e de esgoto, sugerindo estar relacionado à manutenção do ciclo de vida do mosquito-vetor devido à disponibilidade de água.

Da Silva et al. (2020) analisaram a distribuição espacial da incidência da dengue no estado da Paraíba entre 2007 e 2016, obtendo a autocorrelação espacial da dengue e a associação com variáveis socioambientais. Como resultado identificou-se que a doença vem avançando e acometendo maior número de cidades a cada ano.

Assim, com base no exposto, pretende-se com este trabalho analisar tanto o número absoluto de casos prováveis quanto a incidência de dengue em 26 estados brasileiros, nos anos de 2020, 2021 e 2023 por meio da técnica de análise de cluster. Pretende-se com esta análise identificar os estados com comportamentos similares em relação a incidência da dengue, além de observar se o padrão dos clusters se altera quando analisado apenas o número absoluto de casos prováveis.

2 METODOLOGIA

Os dados coletados para o presente estudo são referentes ao número absoluto de casos prováveis de dengue (clássica e febre hemorrágica) em 26 estados brasileiros, nos anos de 2020, 2022 e 2023, disponibilizadas no Sistema Nacional de Agravos de Notificação (SINAN). Este sistema registra os casos confirmados e suspeitos da doença considerando que o número de casos prováveis são todos os notificados como suspeitos (indivíduo que reside em área onde são registrados casos de dengue ou que tenha viajado nos últimos 14 dias para alguma área com notificação de transmissão ou presença do mosquito *Aedes aegypti*). Dentre as 27 unidades federativas, o estado do Espírito Santo não foi informado no banco de dados coletados do SINAN para os anos de 2022 e 2023 e, portanto, este estado não entrou nas análises estatísticas deste estudo.

A análise de cluster foi aplicada com o objetivo de analisar a similaridade entre os estados em relação a dois tipos de informação: número absoluto de casos prováveis e a incidência dos casos de dengue. As variáveis foram os anos 2020, 2021 e 2023. As análises foram implementadas com o auxílio do software estatístico R.

A análise de cluster se inicia a partir de uma matriz de dissimilaridade, que é obtida por meio da medida de distância mais adequada, visto que os procedimentos em análise de agrupamento são influenciados pela natureza das variáveis ou dos atributos dos objetos. De acordo com Khattree & Naik (2000) e Cruz e Carneiro (2006), a distância euclidiana e a distância de Mahalanobis então entre as medidas que mais se destacam dentre as medidas de dissimilaridades conhecidas. Considerando a natureza dos dados no presente estudo, foi utilizada a distância Euclidiana para a obtenção da matriz de distâncias.

Em seguida foi aplicado o Critério de Ward que, segundo Steiner et al. (2008), é um dos métodos hierárquicos aglomerativos mais utilizados na literatura. A sequência de fusão dos agrupamentos para cada método de ligação utilizado foi representada graficamente pelo dendrograma, que auxiliou na identificação dos agrupamentos dos estados brasileiros.

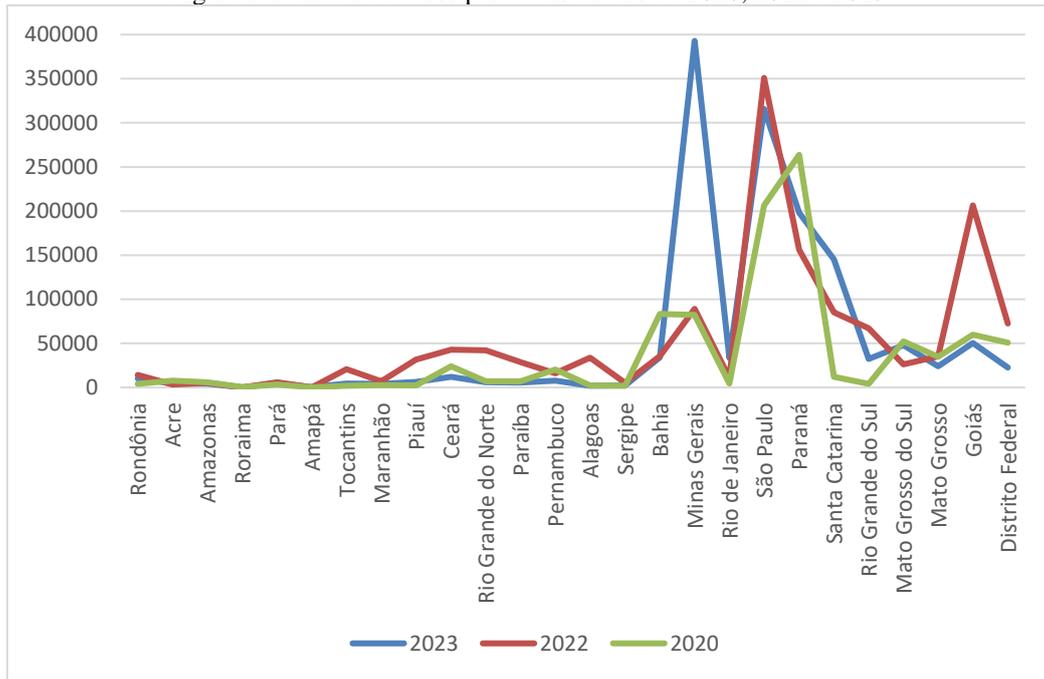
A análise de cluster gerou o gráfico dendrograma no qual o eixo vertical indica o nível de dissimilaridade entre os elementos. No eixo horizontal, são identificadas as unidades da federação (estados). As linhas verticais, partindo dos elementos amostrais agrupados, têm altura correspondente ao nível em que os elementos foram considerados dissimilares, isto é, a distância do agrupamento (Mingoti, 2013). Assim, quanto menor a linha vertical, mais similares são os elementos interligados por tal linha. Por fim, foi obtida a incidência de dengue para cada estado, cálculo feito dividindo os números absolutos de casos prováveis pela população residente de cada estado vezes 100.000 habitantes.

3 RESULTADOS E DISCUSSÃO

3.1 ANÁLISES CONSIDERANDO O NÚMERO ABSOLUTO DE CASOS PROVÁVEIS

De acordo com a Figura 1, o estado de Minas Gerais foi o que superou os demais em 2023 contabilizando 392.873 casos prováveis, o que representa 28,6% do total de casos prováveis do mesmo ano. Já São Paulo liderou a lista em 2022 com 350.891 casos prováveis (25,2% do total de casos prováveis do mesmo ano), sendo em 2020 o Paraná foi o que atingiu o maior número de casos (263.769), representando 29,2% do total de casos do mesmo ano.

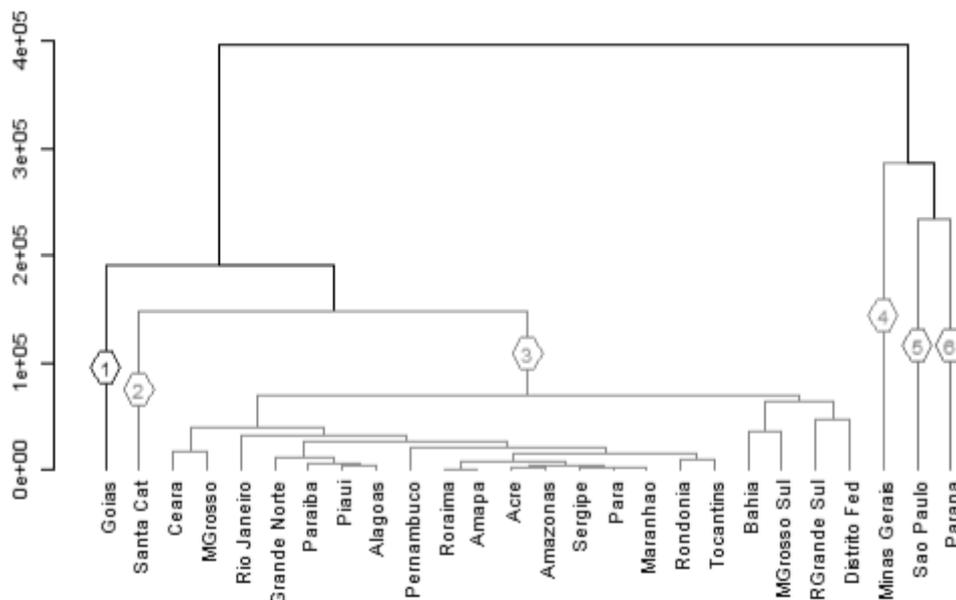
Figura 1. Número de casos prováveis no ano de 2020, 2022 e 2023.



Fonte: Elaborado pelos autores

A análise de cluster utilizando o critério de Ward resultou na formação de seis grupos (Figura 2), em que os estados que constituem o mesmo grupo apresentam comportamento similar em relação ao número absoluto de casos prováveis de dengue no Brasil.

Figura 2. Dendrograma obtido pelo método hierárquico aglomerativo de Ward baseado no número absoluto de casos prováveis de dengue.



Fonte: Elaborado pelos autores

Minas Gerais (segundo estado mais populoso do país), São Paulo (estado mais populoso do país) e Paraná (sexto estado mais populoso do país) foram os estados que compuseram os grupos 4, 5 e 6, respectivamente. Estes estados são também os que apresentaram os maiores números de casos prováveis de dengue nos três anos avaliados.

Estes resultados estão de acordo com estudos que sugerem que ambientes mais populosos tendem a favorecer a difusão espacial de doenças infecciosas, especialmente aquelas transmitidas pelo mosquito *A. aegypti* (Queiroz; Medronho, 2021; Skalinski et al., 2022), uma vez que pode ocorrer adaptação do vetor aos ambientes peri e intradomiciliares, dentre outras causas (Queiroz; Medronho, 2021).

O estado de Goiás, que formou o grupo 1, foi o estado que ficou em segundo lugar no ano de 2022 no registro de números de casos prováveis (206.548), perdendo apenas para São Paulo. Já Santa Catarina, que formou o grupo 2, em 2023 apresentou 145.328 casos prováveis, ficando em quarto lugar entre os estados. Os demais estados formaram o grupo 3 (Tabela 1).

Tabela 1. Subdivisão dos três primeiros grupos obtidos por meio da Análise de cluster.

Grupo 1	Grupo 2	Grupo 3	
Goiás	Santa Catarina	Ceará	Amapá
		Mato Grosso	Acre
		Rio de Janeiro	Amazonas
		R. Grande do Norte	Sergipe
		Paraíba	Pará
		Piauí	Maranhão
		Alagoas	Rondônia
		Pernambuco	Tocantins
		Roraima	Bahia
			Mato G. do Sul
			R. Grande do Sul
			D. Federal

Fonte: Elaborado pelos autores

Os casos atípicos nos resultados são a presença de estados do Sul do país (Paraná e Santa Catarina), onde a dengue era pouco ativa ou praticamente inexistente. Essa expansão pode estar relacionada às mudanças climáticas onde temperaturas mais altas têm sido observadas nas regiões mais frias do Brasil, além da própria adaptação do mosquito e mobilidade que o vírus da dengue adquiriu ao circular com facilidade por estrada e de avião pelas cidades grandes e médias do Brasil (Mongabay, 2023).

3.2 ANÁLISES CONSIDERANDO A INCIDÊNCIA (NÚMERO DE CASOS POR 100 MIL HABITANTES)

Refazendo as análises a fim de analisar a configuração dos resultados em relação à incidência de dengue nos estados (Tabela 2), é possível observar que no ano de 2023 as maiores incidências foram registradas nos estados de Santa Catarina (1.980,4 casos por 100 mil habitantes), Minas Gerais (1.834,8 casos por 100 mil habitantes), Paraná (1709,0 casos por 100 mil habitantes) e Mato Grosso do Sul (1.683,4 casos por 100 mil habitantes).

Roraima apresentou as menores incidências de dengue nos anos de 2022 e 2023 (14,6 e 9,5 casos por 100 mil habitantes, respectivamente). Já Amapá foi o menor registro de incidência registrado em 2020 (7,8 casos por 100 mil habitantes).

Tabela 2. Incidência da dengue nos estados brasileiros nos anos de 2020, 2022 e 2023.

Estado	Incidência em 2023	Incidência em 2022	Incidência em 2020
Rondônia	570.0	785.1	217.98
Acre	484.1	377.7	852.49
Amazonas	105.3	125.4	141.7
Roraima	14.6	9.5	75.5
Para	56.7	67.4	40.3
Amapá	46.0	33.0	7.8
Tocantins	288.2	1296.9	120.9
Maranhão	59.2	95.8	35.9
Piauí	195.9	963.6	68.0
Ceara	131.9	462.5	261.0
RG do Norte	163.9	1184.4	195.9
Paraíba	133.8	710.3	168.7
Pernambuco	78.0	166.5	209.8
Alagoas	56.0	1002.0	71.7
Sergipe	102.2	224.2	79.8
Bahia	228.6	238.4	555.7
Minas Gerais	1834.8	415.8	384.1
Rio Janeiro	191.5	63.8	25.7
São Paulo	676.4	752.2	442.7
Paraná	1709.0	1346.7	2274.4
Santa Catarina	1980.4	1162.0	161.9
RG do Sul	282.5	586.3	34.6
Mato G.do Sul	1683.4	927.3	1839.8
Mato Grosso	672.2	990.7	981.8

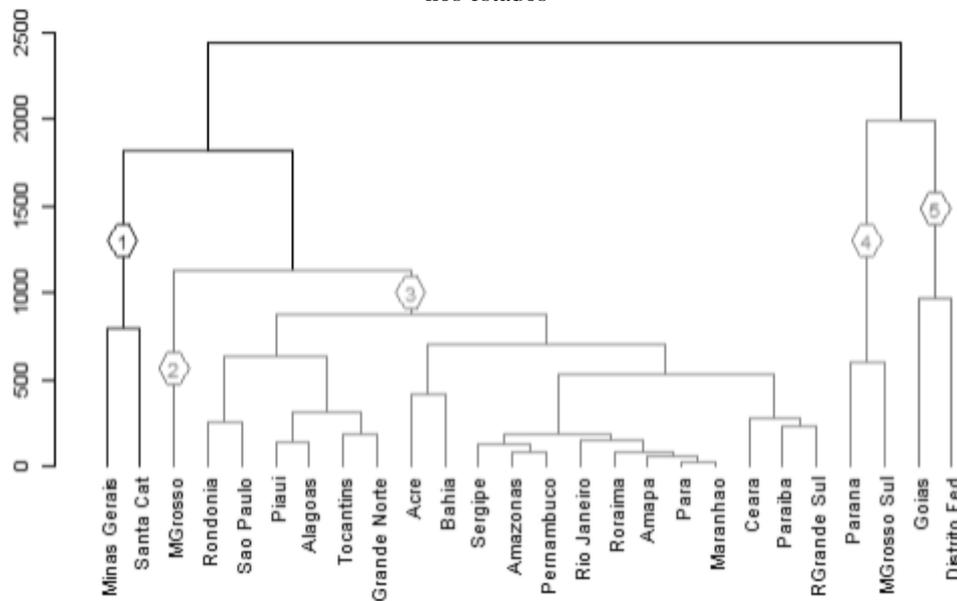
Goiás	699.2	2866.1	830.8
Distrito Federal	732.3	2348.2	1643.5

Fonte: Elaborado pelos autores

Estes resultados estão de acordo com o dendrograma (Figura 3) composto pela formação de 5 grupos, onde estados no mesmo grupo possuem padrão similar nos anos avaliados no que se refere à taxa de incidência da dengue.

É possível notar que o padrão dos grupos foi alterado em comparação com a configuração anterior, sendo que nesta o grupo 1 uniu Minas Gerais e Santa Catarina, o grupo 2 ficou formado apenas por Mato Grosso, o grupo 4 por Paraná e Mato Grosso do Sul, enquanto o grupo 5 uniu Goiás e Distrito Federal. O grupo 3 continha os demais estados.

Figura 3. Dendrograma obtido pelo método hierárquico aglomerativo de Ward baseado na incidência da dengue nos estados



Fonte: Elaborado pelos autores

Dentre vários fatores associados ao aumento da incidência da dengue nos três anos observados neste estudo, estão as ocorrências dos fenômenos El Niño (que ocorreu em 2020 e 2023) e La Niña (que ocorreu em 2022), e as consequências destes ao propiciar mudanças climáticas que levam a temperaturas crescentes (Moraes et al., 2019). Estas ocorrências vêm a promover o aumento dos casos de doenças tropicais, como é o caso da dengue, já que temperaturas mais altas são favoráveis ao desenvolvimento de *A. aegypti* (entre 22°C e 32°C), e para a longevidade e fecundidade dos adultos entre 22°C e 28°C (Ribeiro et al., 2020).

O aumento médio da temperatura e as mudanças nos regimes de chuva no país tem sido observados há vários anos, mas são mais evidentes quando os fenômenos El Niño e La Niña ocorrem. Quanto à dengue, vários fatores estão associados ao aumento do risco de propagação da epidemia, incluindo a mudança na distribuição dos vetores (principalmente *Aedes aegypti* e *Aedes albopictus*), bem como a mudança na circulação de sorotipo do vírus (Brasil, 2023).

4 CONCLUSÃO

A análise de cluster aplicada tanto no número absoluto de casos prováveis quanto nas incidências de dengue nos estados brasileiros resultaram em formações de grupos distintas. A primeira, gerando seis grupos, identificou Minas Gerais agrupado com São Paulo. Já o Paraná, se agrupou com Santa Catarina. A segunda formação foi composta por cinco grupos onde Santa Catarina formou um grupo sem outros estados e Minas Gerais se agrupou com Mato Grosso do Sul. Este resultado mostra que ao avaliar dados de epidemiologia, não apenas o número absoluto de casos prováveis deve ser analisado e reportado em boletins informativos. Adicionar as taxas de incidência da doença às informações associadas à proliferação da doença é de fundamental importância uma vez que produz resultados distintos, principalmente na identificação dos estados cujo políticas de combate precisam ser direcionadas.

REFERÊNCIAS

- BRAGA, Ima Aparecida; VALLE, Denise. Aedes aegypti: histórico do controle no Brasil. *Epidemiol. Serv. Saúde*, Brasília, v. 16, n. 2, p. 113-118, jun. 2007
- CRUZ C.D., CARNEIRO P.C.S. Modelos biométricos aplicados ao melhoramento genético. Second Edition, 2006.
- da Silva, Ellen Tayanne Carla et al. Análise Espacial Da Distribuição Dos Casos De Dengue e Sua Relação Com Fatores Socioambientais No Estado Da Paraíba, Brasil, 2007-2016. *Saúde em Debate*, v. 44, n. 125, p. 465-477, 2020.
- EVERITT, B.S. Cluster analysis. 3rd ed. London: Heinemann Educational Books, 122p. 1993.
- FERREIRA, B. J. et al.. Evolução histórica dos programas de prevenção e controle da dengue no Brasil. *Ciência & Saúde Coletiva*, v. 14, n. 3, p. 961-972, maio 2009.
- FREI, F. Introdução à análise de agrupamentos: teoria e prática. São Paulo: Editora Unesp. 2006.
- Khattree R., Naik D.N.. Multivariate Data Reduction and Discrimination with SAS software. Cary, SAS Institute, 2000.
- MINGOTI, S. A. Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada. Belo Horizonte: UFMG, 2013.
- MOITA NETO J. M.; MOITA G. C. Uma introdução à análise exploratória de dados multivariados. *Química Nova*, v. 21, n. 4, p. 467-469, 1998.
- NAÇÕES UNIDAS [Internet]. OMS: Brasil é o país mais afetado em novo surto de dengue nas Américas, c2023 [cited 2023 Jul 21]. Disponível em: <https://news.un.org/pt/story/2023/07/1817882>. Acesso em: 09 fev. 2024.
- RIBEIRO, A.C.M. et al. Condições socioambientais relacionadas à permanência da dengue no Brasil-2020. *Revista Saúde e Meio Ambiente*, v. 11, n. 2, p. 326-340, 2020.
- SILVA, J. C. B. D.; MACHADO, C. J. S.. Associations between dengue and Socio-Environmental variables in capitals of the brazilian Northeast by Cluster Analysis. *Ambiente & Sociedade*, v. 21, 2018.
- SILVA, T. R. DA. et al.. TENDÊNCIA TEMPORAL E DISTRIBUIÇÃO ESPACIAL DA DENGUE NO BRASIL. *Cogitare Enfermagem*, v. 27, p. e84000, 2022.
- STEINER, M. T. A. et al.. Métodos estatísticos multivariados aplicados à engenharia de avaliações. *Gestão & Produção*, v. 15, n. 1, p. 23-32, jan. 2008.
- Queiroz, E.R.D.S., Medronho, R.A.. Spatial analysis of the incidence of Dengue, Zika and Chikungunya and socioeconomic determinants in the city of Rio de Janeiro, Brazil. *Epidemiol Infect.*, v. 149, e188, 2021.

Skalinski, L.M. et al. The triple epidemics of arboviruses in Feira de Santana, Brazilian Northeast: Epidemiological characteristics and diffusion patterns. *Epidemics*. v. 38:100541, 2022.

MONGABAY [Internet]. Mudanças climáticas devem aumentar transmissões de zika e dengue. c2023 [cited 2023 Jul 17]. Disponível em: <https://brasil.mongabay.com/2023/07/mudancas-climaticas-devem-aumentar-transmissoes-de-zika-e-dengue/>. Acesso em: 18 Jan. 2024.

BRASIL [Internet]. Ministério da Saúde. Minas Gerais registrou mais de 248,5 mil casos prováveis de dengue este ano, c2023 [cited 2023 May 05]. Disponível em: www.gov.br/saude/pt-br/assuntos/noticias-para-os-estados/minas-gerais/2023/maio/minas-gerais-registrou-mais-de-248-5-mil-casos-provaveis-de-dengue-este-ano. Acesso em: 09 fev. 2024.

MORAES, B.C. de, et al. Sazonalidade nas notificações de dengue das capitais da Amazônia e os impactos do El Niño/La Niña. *Cad Saúde Pública*,v. 35, n. 9, e00123417, 2019.

FOCKS, D.A., DANIELS, E., HAILE, D.G., KEESLING, J.E.. A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation and samples of simulation results. *Am J Trop Med Hyg* 1995; 53: 489-506.