

# Influência do clima urbano da cidade de Campo Grande, MS, na quantidade de casos registrados de dengue: um estudo de caso via modelo de regressão Poisson

Influence of urban climate in the city of Campo Grande, MS, on the number of registered cases of dengue: a case study via a Poisson regression model

*Influencia del clima urbano en la ciudad de Campo Grande, MS, en el número de casos de dengue registrados: un estudio de caso mediante el modelo de regresión de Poisson*

Erlandson Ferreira Saraiva<sup>1</sup>

Leandro Sauer<sup>1</sup>

Mariana Villela Flesh<sup>1</sup>

Recebido em: 23/02/2022; aceito em: 17/04/2023

DOI: <http://dx.doi.org/10.20435/inter.v24i3.3653>

**Resumo:** No Brasil, o primeiro caso de dengue foi registrado na década de 1980 e, desde então, os casos vêm ocorrendo de forma continuada, sendo atualmente um dos principais problemas de saúde do país. Como a proliferação do mosquito transmissor da doença depende de variáveis ambientais, tais como temperatura e chuvas, para completar seu ciclo de vida, é de interesse entender as relações existentes entre o clima e os casos de dengue. Com o objetivo de contribuir com o sistema de vigilância da dengue na cidade de Campo Grande, MS, este artigo propõe um modelo estatístico para identificar as variáveis climáticas que podem estar relacionadas com o número de casos de dengue. Identificadas as variáveis, o modelo ajustado permite fazer projeções e simular diferentes cenários. Dessa forma, pode auxiliar na tomada de decisões com relação à implementação de medidas de combate e/ou controle do mosquito transmissor. Além disso, desenvolvemos um estudo para verificar a existência de períodos de defasagem, *i.e.*, se o número de casos de dengue registrado em um mês depende dos valores registrados para as variáveis ambientais no mês anterior ou do mês atual.

**Palavras-chave:** dengue; distribuição de Poisson; modelo de regressão de Poisson; predição.

**Abstract:** In the 1980s, the first dengue fever cases were recorded in Brazil. Since then, the cases have been occurring continuously, and currently, dengue fever is one of the country's main health problems. Since the proliferation of the mosquito that transmits the disease depends on environmental variables, such as temperature and rainfall, to complete its life cycle, it is of interest to understand the relationships between climate and dengue cases. In order to contribute to the dengue surveillance system in the city of Campo Grande, MS, Brazil, this article proposes a statistical model to identify the climate variables that may be related to the number of dengue fever cases. Once the variables have been identified, the fitted model allows us to make projections and develop simulations of different sceneries. In this way, it can assist in decision-making regarding the implementation of measures to combat and/or control the transmitting mosquito. In addition, we developed a study to verify the existence of lag periods, *i.e.*, if the number of dengue cases recorded in a month depends on the values recorded for the environmental variables in the previous month or of the current month.

**Keywords:** dengue fever; Poisson distribution; Poisson regression model; prediction.

**Resumen:** En Brasil, el primer caso de dengue se registró en la década de 1980 y, desde entonces, los casos se vienen presentando de manera continua, siendo actualmente uno de los principales problemas de salud del país. Como la proliferación del mosquito transmisor de la enfermedad depende de variables ambientales, como la temperatura y las precipitaciones, para completar su ciclo de vida, es de interés comprender las relaciones entre el clima y los casos de dengue. Con el fin de contribuir al sistema de vigilancia del dengue en la ciudad de Campo Grande, MS, Brasil, este artículo propone un modelo estadístico para identificar las variables climáticas que pueden estar relacionadas con el número de casos de dengue. Una vez identificadas las variables, el modelo ajustado permite realizar proyecciones y simular diferentes escenarios. De esta forma, puede ayudar en la toma de decisiones sobre la implementación de medidas para combatir y/o controlar el

<sup>1</sup> Universidade Federal de Mato Grosso do Sul (UFMS), Campo Grande, Mato Grosso do Sul, Brasil.



mosquito transmissor. Además, desarrollamos un estudio para verificar la existencia de períodos de rezago, es decir, si el número de casos de dengue registrados en un mes depende de los valores registrados para las variables ambientales en el mes anterior o en el mes actual.

**Palabras clave:** dengue; distribución de Poisson; modelo de regresión de Poisson; predicción.

---

## 1 INTRODUÇÃO

De acordo com o Ministério da Saúde, a dengue “é uma doença infecciosa febril aguda, que pode se apresentar de forma benigna ou grave” (Biblioteca Virtual em Saúde do Ministério Da Saúde, 2007). Conforme Gubler (1998), Gubler (2002) e Azil *et al.*, (2010), esta doença é causada por um vírus pertencente à família *Flaviviridae*, com quatro sorotipos, denominados de DENV-1, DENV-2, DENV-3 e DENV-4, e é transmitida principalmente pelas fêmeas do mosquito *Aedes aegypti* e *Aedes albopictus*. A dengue ocorre especialmente em países de climas tropicais, pois o clima quente e úmido favorece o desenvolvimento e a proliferação do mosquito *Aedes aegypti*, sendo este bem adaptado a climas com temperaturas entre 20°C e 46°C (Costa, 2001).

No Brasil, o primeiro caso de dengue foi registrado na década de 1980 e, desde então, os casos vêm ocorrendo de forma continuada. Atualmente, é um dos principais problemas de saúde pública do país. Na ausência de uma vacina para prevenção e controle da dengue, a eliminação dos criadouros do *Aedes aegypti* ainda é a única estratégia eficaz (Wan *et al.*, 2010). Na cidade de Campo Grande, MS, uma estratégia de controle da dengue é desenvolvida pela prefeitura com o auxílio dos agentes de saúde comunitários e de campanhas de conscientização e limpeza em bairros. Este programa de vigilância, que é prático e econômico, também serve como um programa de alerta precoce para que ações corretivas possam ser tomadas antes que ocorra um surto.

Como a proliferação do mosquito transmissor da doença depende da temperatura, da disponibilidade de água e de alguns outros fatores climáticos para completar seu ciclo de vida, é de interesse entender as relações existentes entre variáveis climáticas e os casos de dengue. Com o objetivo de contribuir com o sistema de vigilância da dengue na cidade de Campo Grande, MS, este artigo propõe um modelo estatístico para identificar as variáveis climáticas que podem influenciar no número de casos de dengue. Identificadas as variáveis, o modelo ajustado permite fazer projeções e simular diferentes cenários de surtos da doença. Portanto pode auxiliar na tomada de decisões com relação à implementação de medidas de combate e/ou controle do vetor transmissor da doença.

Como a variável resposta de interesse, número de casos de dengue registrados por mês, é uma variável discreta (uma contagem), o modelo considerado é o de regressão de Poisson. Como variáveis explicativas, consideramos: o mês, a temperatura média do mês, a média das temperaturas do mês, a precipitação pluviométrica do mês, os dias com chuvas no mês e se o mês faz parte do período denominado de época de chuvas. Para estimação dos parâmetros do modelo, consideramos o método de máxima verossimilhança, em que as estimativas foram obtidas utilizando o comando *glm* do *software* R (CRAN R, 2021). Além disso, desenvolvemos um estudo para verificar a existência de períodos de defasagem, *i.e.*, se o número de casos de dengue registrados em um mês depende dos valores registrados para as variáveis ambientais no mês anterior ou do mês atual.

Para o ajuste do modelo, realizamos uma coleta de dados referente ao número de casos de dengue registrados no período de 2007 a 2019 no Brasil, por região, e depois com foco no

Centro-Oeste, no estado de Mato Grosso do Sul e, mais precisamente, em Campo Grande (foco do estudo). Estes dados estão disponíveis gratuitamente no *website*: <http://tabnet.datasus.gov.br>. Contudo, o *site* não disponibilizava os dados referentes aos anos de 2020 e 2021. Devido a isto, não foi possível incluir os dados desses últimos dois anos no estudo. Os valores para as variáveis ambientais consideradas no estudo foram obtidos no *website* [www.cemtec.ms.gov.br/boletins-meteorologicos](http://www.cemtec.ms.gov.br/boletins-meteorologicos).

Obtido o conjunto de dados, nossa primeira análise consistiu em fazer uma análise descritiva dos dados ao longo dos anos. Para isto, primeiramente, consideramos uma padronização do número de casos por 100.000 habitantes. Apresentamos os resultados em gráficos que mostram a quantidade de casos no decorrer dos anos. Nesse panorama, o Centro-Oeste apresentou as maiores quantidades de casos em quase todos os anos analisados. Após esta análise descritiva dos dados, desenvolvemos a modelagem e o procedimento de estimação dos parâmetros do modelo de regressão de Poisson. O modelo ajustado indicou que as variáveis mês, média das temperaturas máxima do mês, precipitação pluviométrica do mês, dias com chuva no mês e época de chuvas têm relação com a quantidade de casos de dengue.

O restante do artigo está organizado da seguinte maneira. Na seção 2, apresentamos a análise descritiva dos dados. Na seção 3, apresentamos o modelo de regressão de Poisson com e sem defasagem. Na seção 4, apresentamos os resultados obtidos com o ajuste do modelo aos dados de Campo Grande, MS. A seção 5 apresenta as considerações finais.

## **2 DADOS SOBRE A DENGUE NO BRASIL E EM CAMPO GRANDE**

Nesta seção, apresentamos um panorama sobre as quantidades de casos de dengue registrados no Brasil entre os anos de 2007 e 2019, por região. Em seguida, apresentamos o panorama com foco nos estados da região Centro-Oeste, especialmente Mato Grosso do Sul. Finalizamos a seção descrevendo as quantidades de casos registrados na cidade Campo Grande.

### **2.1 Dengue no Brasil**

Entre os anos de 2007 e 2019, foram registrados, no Brasil, onze milhões trinta e oito mil e duzentos e quarenta e nove (11.038.249) casos de dengue. A Tabela 1 mostra a distribuição de frequências do número de casos de dengue registrados por ano e região. Para cada um dos anos, é apresentado o total de casos, o percentual em relação ao total registrado naquele ano (total por linha, entre parênteses) e o número de casos para cada 100.000 habitantes (valor entre colchetes).

Como pode ser observado na Tabela 1, nos anos de 2013, 2015, 2016 e 2019, foram registrados mais de um milhão de casos de dengue em todo o Brasil. Com exceção dos anos de 2009 e 2017, os percentuais de casos da região Sudeste são maiores do que todas as outras regiões, o que é esperado, em virtude de esta região apresentar a maior população. Nos anos de 2009 e 2017, os maiores percentuais de casos foram registrados na região Nordeste.

Por facilidade de interpretação, o ideal é padronizar o número de casos e fazer a interpretação dos valores para cada 100.000 habitantes, por exemplo. Neste cenário, a região Centro-Oeste apresenta a maior quantidade de casos por 100.000 habitantes, com exceção dos anos de 2008 e 2011. Nos últimos oito anos (de 2012 a 2019), a região Centro-Oeste sempre apresentou as maiores quantidades. Uma explicação para este fato é que esta região apresenta um clima

muito favorável à proliferação do transmissor da doença. No ano de 2008, a região Nordeste apresentou a maior quantidade de casos por 100.000 habitantes, e, no ano de 2011, a maior quantidade foi registrada na região Norte.

Por outro lado, a região Sul apresentou os menores valores percentuais por 100.000 habitantes em todos os anos. Isto ocorreu devido ao clima mais ameno, com temperaturas médias mais baixas do que as outras regiões do Brasil, o que não favorece o desenvolvimento do mosquito *Aedes aegypti*. Este fato é discutido por Donalísio e Glasser (2002) e Silva, Mariano e Scopel, (2017), que argumentam que “as fêmeas do mosquito, infectadas e submetidas a temperaturas de aproximadamente 32°C, teriam 2,64 vezes mais chance de completar o período de incubação extrínseco do que aquelas submetidas a temperaturas amenas”.

Tabela 1 – Quantidade de casos de dengue por região e ano

Ano	Região					Total
	Norte	Nordeste	Sudeste	Sul	Centro-Oeste	
2007	36.939 (7,38%) [233]	125.629 (25, 11%) [237]	207.064 (41, 38%) [258]	27.798 (5,55%) [101]	103.002 (20, 58%) [733]	500.531 (100%) [262]
2008	47.505 (8, 46%) [299]	180.726 (32, 18%) [340]	286.926 (51, 08%) [357]	2.074 (0, 37%) [8]	44.428 (7, 91%) [316]	561.659 (100%) [294]
2009	56.603 (13, 75%) [357]	127.363 (30, 93%) [240]	115.454 (28, 04%) [144]	1.737 (0, 42%) [6]	110.576 (26, 86%) [786]	411.733 (100%) [216]
2010	97.738 (9, 92%) [616]	169.602 (17, 21%) [319]	460.714 (46, 75%) [573]	41.590 (4, 22%) [152]	215.844 (21, 90%) [1.535]	985.488 (100%) [517]
2011	114.251 (16, 58%) [720]	174.842 (25, 36%) [329]	334.770 (48, 57%) [416]	29.360 (4, 26%) [107]	36.046 (5, 23%) [256]	689.277 (100%) [361]
2012	42.306 (7, 26%) [267]	219.237 (37, 65%) [413]	248.175 (42, 62%) [309]	4.808 (0, 83%) [17]	67.793 (11, 64%) [482]	582.365 (100%) [305]
2013	48.803 (3, 41%) [308]	147.207 (10, 28%) [277]	906.487 (63, 29%) [1.128]	66.555 (4, 65%) [243]	263.080 (18, 37%) [1.871]	1.432.234 (100%) [750]
2014	48.302 (8, 17%) [304]	90.489 (15, 31%) [170]	312.181 (52, 81%) [388]	23.062 (3, 90%) [84]	117.094 (19, 81%) [833]	591.128 (100%) [310]
2015	32.347 (1, 90%) [204]	328.951 (19, 38%) [620]	1.051.700 (61, 95%) [1.309]	52.110 (3, 07%) [190]	232.693 (13, 70%) [1.655]	1.697.801 (100%) [890]
2016	38.621 (2, 54%) [243]	326.071 (21, 47%) [614]	864.899 (56, 94%) [1.076]	71.325 (4, 70%) [260]	217.928 (14, 35%) [1.550]	1.518.858 (100%) [796]
2017	21.987 (9, 04%) [139]	84.845 (34, 86%) [160]	53.848 (22, 13%) [67]	2.604 (1, 07%) [9]	80.052 (32, 90%) [569]	243.336 (100%) [128]
2018	17.789 (6, 68%) [112]	66.561 (24, 99%) [125]	73.144 (27, 46%) [91]	1.739 (0, 65%) [6]	107.154 (40, 22%) [762]	266.387 (100%) [140]

Ano	Região					Total
	Norte	Nordeste	Sudeste	Sul	Centro-Oeste	
2019	36.961 (2, 37%) [233]	216.795 (13, 92%) [408]	1.025.980 (65, 88%) [1.277]	49.860 (3, 20%) [182]	227.856 (14, 63%) [1.621]	1.557.452 (100%) [816]

Fonte: Elaborada pelos autores.

## 2.2 Dengue no Centro-Oeste

Com relação à região Centro-Oeste, entre os anos de 2007 e 2019, foram registrados um milhão oitocentos e vinte e seis mil e seiscentos e trinta e um (1.826.631) casos de dengue. A Tabela 2 mostra a quantidade de casos de dengue na região Centro-Oeste por ano e estado.

Nos últimos sete anos (de 2013 a 2019), o estado de Goiás apresentou os maiores valores percentuais em relação ao total de casos registrados na região Centro-Oeste. Com destaque para os anos de 2017 e 2018, em que mais de 80% dos casos registrados de dengue na região Centro-Oeste foram neste estado. Com base no número de casos por 100.000 habitantes, o estado de Goiás também apresentou os maiores valores entre os anos de 2011 e 2018. No ano de 2019, o estado de MS foi o que apresentou maior número de casos por 100.000 habitantes, sendo 2.264 casos para cada 100.000 habitantes.

Tabela 2 – Quantidade de casos de dengue na região Centro-Oeste

Ano	Estado				Total
	MS	MT	GO	DF	
2007	70.322 (68,25%) [2.871]	16.598 (16,11%) [547]	14.853 (14,42%) [247]	1.254 (1,22%) [49]	103.027 (100%) [733]
2008	759 (1,71%) [31]	6.783 (15,27%) [223]	35.533 (79,98%) [592]	1.351 (3,04%) [52]	44.426 (100%) [316]
2009	13.482 (11,88%) [550]	52.464 (46,20%) [1.728]	46.690 (41,12%) [778]	913 (0,8%) [35]	113.549 (100%) [808]
2010	62.856 (29,13%) [2.566]	36.848 (16,94%) [1.205]	100.864 (46,75%) [1.680]	15.471 (7,17%) [602]	215.770 (100%) [1.535]
2011	6.263 (17,38%) [256]	5.008 (13,91%) [165]	23.145 (64,25%) [385]	1.607 (4,46%) [62]	36.023 (100%) [256]
2012	9.277 (13,69%) [3.188]	32.848 (48,45%) [1.150]	23.856 (35,19%) [2.174]	1.814 (2,67%) [768]	67.795 (100%) [1.873]
2013	78.072 (29,66%) [3.188]	34.912 (13,26%) [1.150]	130.542 (49,58%) [2.174]	19.738 (7,50%) [768]	263.264 (100%) [1.873]
2014	3.397 (2,90%) [139]	7.183 (6,14%) [237]	92.905 (79,34%) [1.547]	13.609 (11,62%) [529]	117.094 (100%) [833]
2015	31.990 (13,75%) [1.306]	21.524 (9,25%) [709]	168.934 (72,60%) [2.814]	10.245 (4,40%) [399]	232.693 (100%) [1.655]

Ano	Estado				Total
	MS	MT	GO	DF	
2016	47.150 (21,64%) [1.925]	21.377 (9,79%) [704]	129.798 (59,57%) [2.162]	19.603 (9%) [763]	217.928 (100%) [1.550]
2017	2.322 (2,90%) [95]	9.399 (11,74%) [310]	64.080 (80,05%) [1.067]	4.251 (5,31%) [165]	80.052 (100%) [569]
2018	5.650 (5,27%) [231]	7.294 (6,81%) [240]	91.728 (85,60%) [1.528]	2.482 (2,32%) [96]	107.154 (100%) [762]
2019	55.459 (24,34%) [2.264]	11.230 (4,93%) [370]	121.709 (53,41%) [2.027]	39.458 (17,32%) [1.535]	227.856 (100%) [1.621]

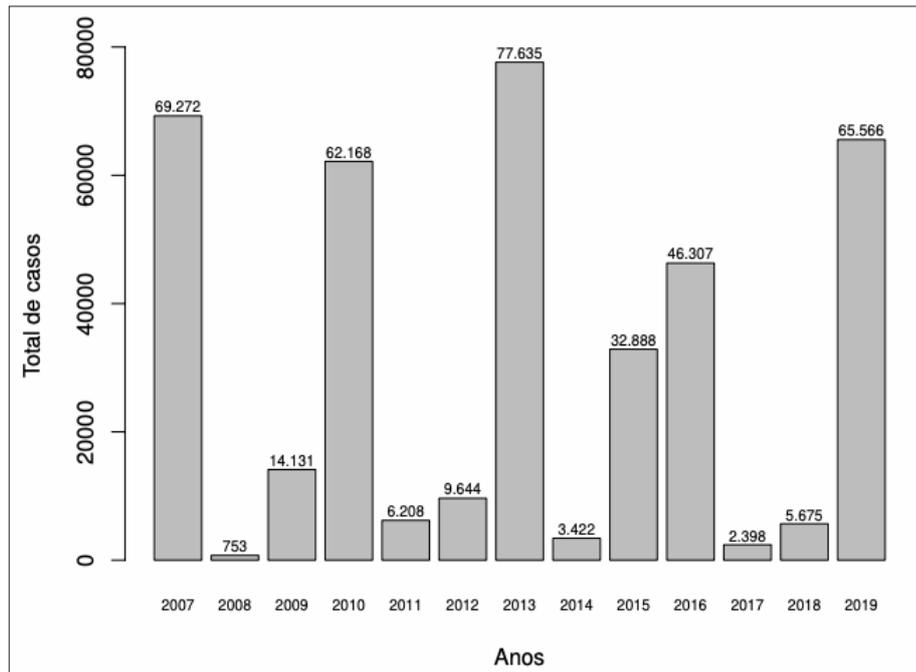
Fonte: Elaborada pelos autores.

### 2.3 Dengue em MS

Entre os anos de 2007 e 2019, foram registrados, em todo o estado de MS, trezentos e noventa e seis mil e sessenta e sete (396.067) casos da dengue. O Gráfico 1 mostra, em barras, a quantidade de casos registrados em cada um dos anos em Mato Grosso do Sul. Nota-se que, nos anos de  $A_1 = \{2007, 2010, 2013, 2016, 2019\}$ , houve um aumento significativo do número de casos da dengue no estado. Em anos seguintes aos anos  $A_1$ , no caso, os anos de  $A_2 = \{2008, 2011, 2014, 2017\}$ , ocorreram quedas significativas na quantidade de casos registrados. Uma explicação plausível para esta queda é a de que, após um ano com grandes números de casos, os agentes governamentais e a população agem em conjunto para conter a proliferação do mosquito transmissor e, dessa forma, fazem com que ocorra uma queda do número de casos.

Porém o cuidado em conter a proliferação do transmissor parece diminuir no ano seguinte à queda do número de casos (anos  $A_3 = \{2009, 2012, 2015, 2018\}$ ), ocasionando um aumento dos casos confirmados da dengue. Então, voltamos a um ano de crescimento significativo do número de casos. De 2009 para 2010, houve um aumento de 339,94% na quantidade de casos confirmados; de 2012 para 2013, o aumento foi de 705,01%; de 2015 para 2016, o aumento foi de 40,80%, e, de 2018 para 2019, o aumento foi de 1.055,35%. Ou seja, o gráfico indica o seguinte padrão: a cada três anos, ocorre um aumento significativo na quantidade de casos confirmados da dengue.

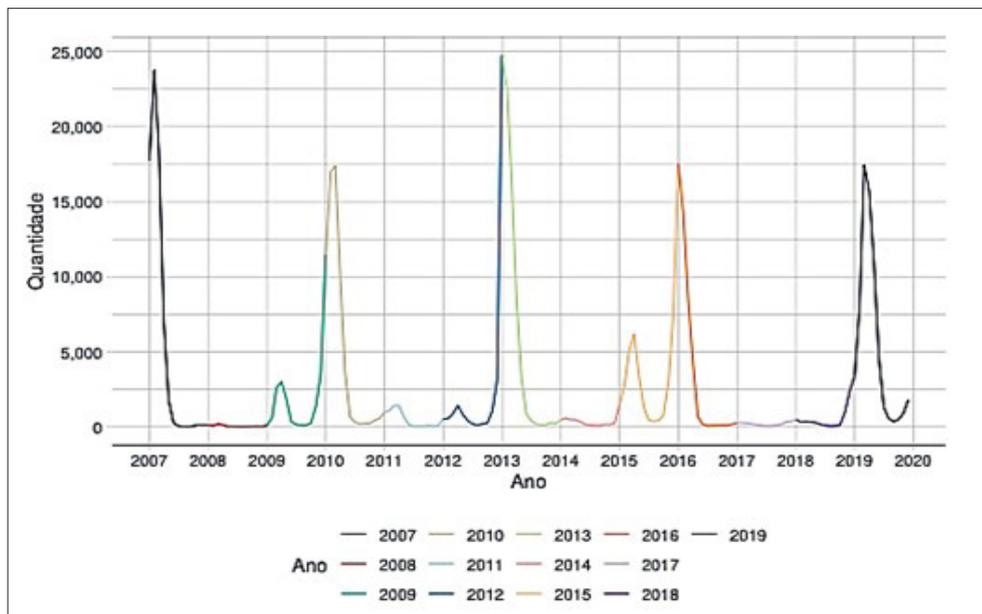
Gráfico 1 – Quantidade de casos de dengue em MS



Fonte: Elaborado pelos autores.

O Gráfico 2 mostra, em linhas, a evolução dos casos de dengue em MS, por mês, entre os anos de 2014 e 2019. Nota-se que a quantidade de casos começa a aumentar a partir do quarto trimestre de cada ano e começa a diminuir somente após o primeiro trimestre do ano seguinte.

Gráfico 2 – Evolução dos casos de dengue em MS, de 2007 a 2019



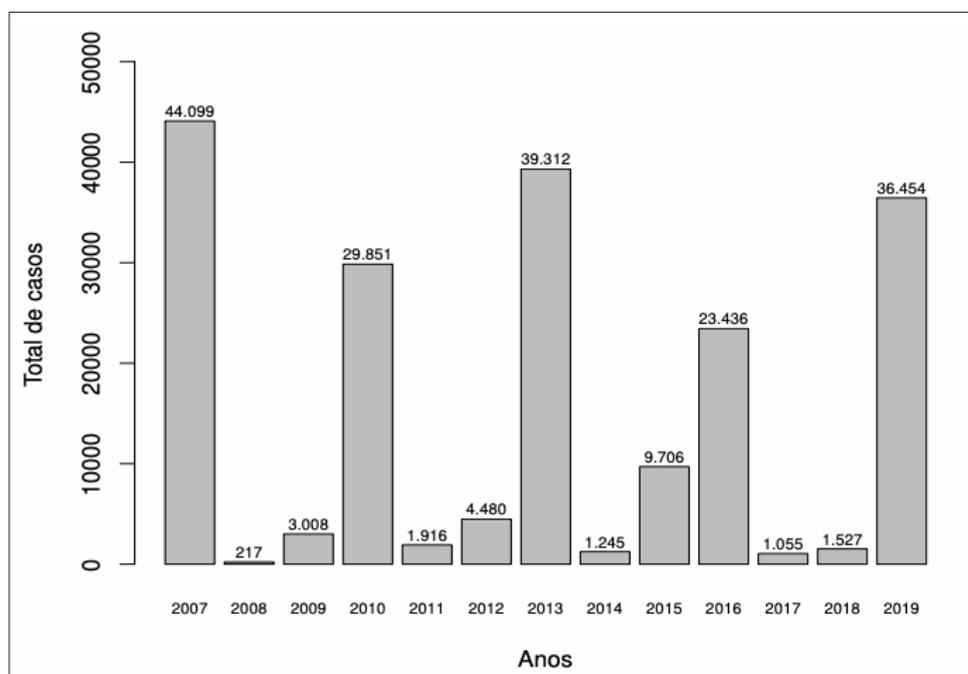
Fonte: Elaborado pelos autores.

## 2.4 Dengue em Campo Grande

Entre os anos de 2007 e 2019, foi registrado, em Campo Grande, um total de cento e noventa e seis mil e trezentos e seis casos de dengue (196.306). O Gráfico 3 mostra, em barras, as quantidades registradas por ano. Nota-se que os valores registrados em Campo Grande apresentam um comportamento cíclico. A cada três anos, ocorre um aumento significativo na quantidade de casos confirmados da dengue (um surto).

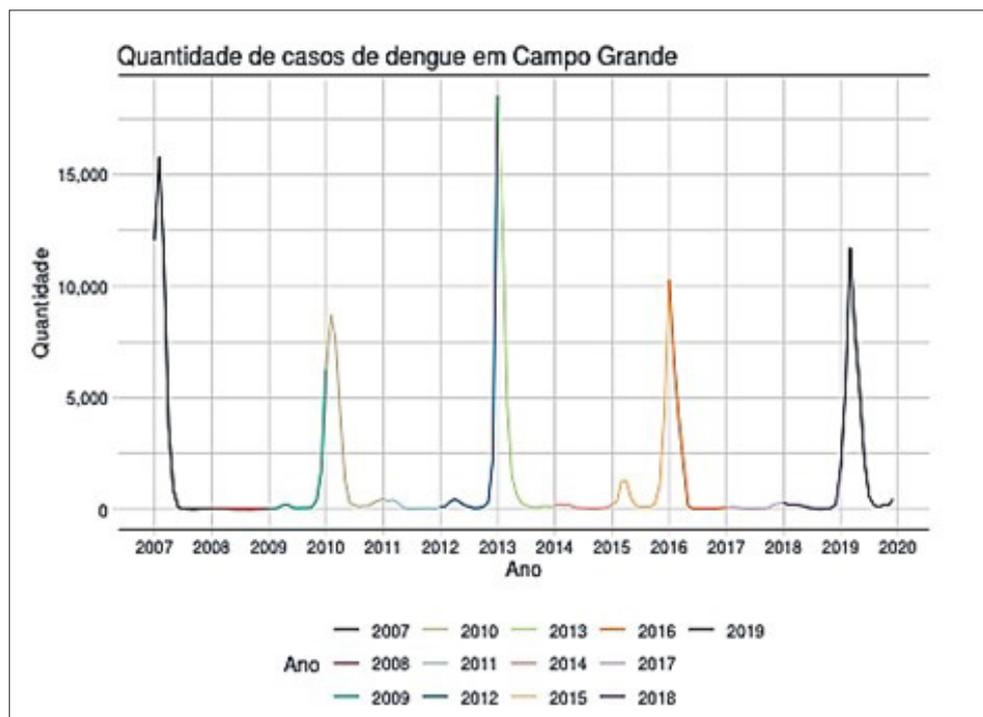
O Gráfico 4 mostra, em linhas, a evolução da quantidade de casos de dengue em Campo Grande por mês, entre os anos de 2007 e 2019. Os dados registrados mostram que o período de grande crescimento do número de casos de dengue, na capital, ocorre a partir do mês de outubro de cada ano e vai até o mês de abril do ano seguinte. Este aumento de casos neste período é devido, principalmente, ao clima da cidade nesta época, que é caracterizado por muito calor e fortes chuvas, o que favorece a proliferação do transmissor da doença, o mosquito *Aedes aegypti*.

Gráfico 3 – Quantidade de casos de dengue em Campo Grande, por ano



Fonte: Elaborado pelos autores.

Gráfico 4 – Evolução dos casos de dengue em Campo Grande, de 2007 a 2019



Fonte: Elaborado pelos autores.

Estes resultados mostram a importância de se ter um modelo de projeção do número de casos a partir de variáveis ambientais, para auxiliar e embasar ações de combate à proliferação do mosquito e, conseqüentemente, a redução do número de casos.

Para a modelagem descrita na próxima seção, consideramos as seguintes variáveis:

- $X_1$ : Mês;
  - $X_2$ : Temperatura média do mês;
  - $X_3$ : Precipitação pluviométrica do mês;
  - $X_4$ : Dias com chuva;
  - $X_5$ : Época das chuvas.
- (1)

Os valores das variáveis  $X_2$  a  $X_5$  estão disponíveis gratuitamente no *website* do Centro de Monitoramento do Tempo e do Clima de MS (CEMTEC): <https://www.cemtec.ms.gov.br>. Porém o *site* não disponibiliza os dados referentes ao ano de 2007. Assim, para o ajuste do modelo descrito na próxima seção, não consideramos as quantidades de casos de dengue registrados no ano de 2007.

### 3 MODELO

Nesta seção, apresentamos uma modelagem estatística para a identificação das variáveis ambientais relacionadas às quantidades de casos de dengue registrados na cidade de Campo Grande. O modelo considerado é baseado no ajuste de um modelo de regressão de Poisson, em que os parâmetros de interesse são estimados com base nas quantidades de casos de dengue registrados entre os anos de 2008 e 2019 e nos valores registrados das variáveis ambientais descritas em (1).

Para iniciarmos a modelagem, considera-se que  $Y_t$  é a quantidade de casos de dengue registrados no mês  $t$ , para  $t = 1, \dots, n$ , em que  $t = 1$  representa o mês de janeiro do ano de 2008 e  $n = 144$  representa o mês de dezembro do ano de 2019, para  $Y_t \in \{0, 1, 2, 3, \dots\}$ . Assume-se que:

$$Y_t | \mu_t \sim \text{Poisson}(\mu_t), \quad (2)$$

em que  $\mu_t = E(Y_t)$  é o número esperado de casos de dengue no mês  $t$ , para  $t = 1, \dots, n$ .

Do modelo em (2), a probabilidade de observarmos  $y_t$  casos de dengue no mês  $t$  é dado por

$$P(Y_t = y_t | \mu_t) = \frac{\mu_t^{y_t} e^{-\mu_t}}{y_t!}, \quad (3)$$

para  $y_t \in \{0, 1, 2, 3, \dots\}$ . Ou seja, fixado o valor de  $\mu_t$ , a equação em (3) nos dá a probabilidade de observarmos  $Y_t = y_t$ , para  $y_t = 0, 1, \dots$  e  $t \in \{1, \dots, n\}$ . Por exemplo, se  $\mu_{72} = 100$ , isto é, se no mês de dezembro do ano de 2019 é esperado ter 100 casos de dengue, então a probabilidade de se observar 120 casos de dengue é

$$P(Y_{72} = 120 | \mu_{72} = 100) = \frac{100^{120} e^{-100}}{120!} = 0,0056. \quad (4)$$

Porém, na prática, o valor de  $\mu_t$  é uma quantidade desconhecida, e os valores  $y_t$ 's são quantidades observáveis. Assim, nosso objetivo é estimar  $\mu_t$  a partir dos valores observados das quantidades de casos de dengue,  $y_t$ 's, e dos valores observados das variáveis  $X_j$ 's, para  $j = 1, 2, 3, 4, 5$ , descritas em (1).

Para relacionar o número esperado de casos de dengue no mês  $t$  com as variáveis ambientais  $X_j$ 's, considera-se que o logaritmo de  $\mu_t$  é dado pelo seguinte modelo de regressão:

$$\log(\mu_t) = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + \beta_4 x_{4t} + \beta_5 x_{5t}, \quad (4)$$

em que  $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$  são os parâmetros do modelo de regressão e  $x_t = (1, x_{t1}, x_{t2}, x_{t3}, x_{t4}, x_{t5})$  é a  $t$ -ésima linha da matriz observada  $\mathbf{x}$ , de dimensão  $n \times 6$ , com a primeira coluna sendo composta de valores 1's e as colunas 2 a 6 sendo compostas pelos valores observados das variáveis  $X_j$ 's, respectivamente. Para maiores detalhes sobre o uso da função de ligação logarítmica da equação (4), sugere-se consultar os trabalhos de McCullagh e Nelder (1989) e Cameron e Trivedi (1998).

### 3.1 MODELO COM DEFASAGEM

O modelo em (4) assume que o número de casos de dengue no mês  $t$ , denotado por  $y_t$ , depende dos valores registrados para as variáveis ambientais no mesmo mês, para  $t = 1, \dots, n$ . Por exemplo, o número de casos de dengue no mês de janeiro de 2020 depende dos valores registrados para as variáveis  $X_j$ 's no mês de janeiro de 2020. Porém isto nem sempre é verdade, pois o reflexo dos valores registrados para as variáveis ambientais  $X_j$ 's em um mês  $t$  qualquer pode ser observado apenas no mês seguinte,  $t + 1$ . Assim, para incorporar este fato ao modelo, considera-se que o logaritmo de  $\mu_t$  é dado pelo seguinte modelo de regressão:

$$\log(\mu_t) = \beta_0 + \beta_1 x_{1(t-d)} + \beta_2 x_{2(t-d)} + \beta_3 x_{3(t-d)} + \beta_4 x_{4(t-d)} + \beta_5 x_{5(t-d)}, \quad (5)$$

em que  $d$  representa a defasagem que pode ocorrer. Para as análises, assumimos  $d \in \{0, 1\}$ , em que,  $d = 0$  indica a não ocorrência de defasagem, ou seja, o número de casos de dengue depende dos valores registrados para  $X_j$ 's no mesmo mês  $t$ , e  $d = 1$  indica uma defasagem de

um mês, isto é, o número de casos de dengue em um mês  $t$  depende dos valores registrados para as variáveis  $X_j$ 's no mês anterior, para  $t = 2, \dots, n$ .

Assim, dados os valores observados  $(y, x)$ , a função de verossimilhança para  $\beta$  é dada por

$$L(\beta|y, x) = \prod_{t=1}^n \frac{\mu_t^{y_t} e^{-\mu_t}}{y_t!},$$

em que, da Equação (5), o valor esperado  $\mu_t$  é definido em termos dos parâmetros  $\beta$  e das covariáveis  $x$  por meio da seguinte expressão:  $\mu_t = \exp\{\beta X_{(t-d)}\} = \exp\{\beta_0 + \beta_1 x_{1(t-d)} + \beta_2 x_{2(t-d)} + \beta_3 x_{3(t-d)} + \beta_4 x_{4(t-d)} + \beta_5 x_{5(t-d)}\}$ , para  $t = 1, \dots, n$ .

A função log-verossimilhança é dada por

$$l(\beta|y, x) = \sum_{t=1}^n y_t \log(\mu_t) - \mu_t - \log(y_t) \propto \sum_{t=1}^n y_t \beta x_{(t-d)} - \exp\{\beta x_{(t-d)}\}. \quad (6)$$

em que  $\log(\mu_t) = \beta x_{(t-d)}$  dado na equação (5), para  $t = 1, \dots, n$ .

O estimador de máxima verossimilhança é dada pelas soluções das seguintes equações

$$\frac{\partial l(\beta|y, x)}{\partial \beta_j} = \sum_{t=1}^n x_{j(t-d)} (y_t - \exp\{\beta_0 + \beta_1 x_{1(t-d)} + \beta_2 x_{2(t-d)} + \beta_3 x_{3(t-d)} + \beta_4 x_{4(t-d)} + \beta_5 x_{5(t-d)}\}) = 0,$$

para  $j = 0, \dots, 5$  com  $x_{j(t-d)} = 1$  para todo  $j = 0$ . Porém estas equações não possuem solução analítica. Devido a isto, devem ser resolvidas numericamente, utilizando, por exemplo, o algoritmo de Newton-Raphson. Neste texto, obtemos as estimativas de máxima verossimilhança utilizando o comando *glm* do *software* R com a opção *family = Poisson*.

#### 4 RESULTADOS

Nesta seção, apresentamos o ajuste do modelo de regressão de Poisson às quantidades de casos de dengue registrados na cidade de Campo Grande, entre os anos de 2008 e 2019. Assim, seja  $D = [y, x]$  uma matriz de dados de dimensão  $144 \times 7$ , em que a coluna 1 contém os valores  $y$ 's (quantidade de casos de dengues) e as colunas 2 a 6 contêm os valores das variáveis  $X_j$ 's, para  $j = 1, \dots, 5$ , descritas em (1).

Nossa primeira análise consiste em verificar se pelo menos uma das cinco variáveis é importante para explicar a resposta  $Y$ . Isso leva ao seguinte teste de hipótese

$$H_0: \beta_j = 0 \text{ para todo } j \text{ versus } H_1: \beta_j \neq 0$$

para pelo menos um  $j$ , para  $j = 1, 2, 3, 4, 5$ .

Utilizando o teste da razão de verossimilhanças, a estatística de teste é dada por

$$S = -2 \log \frac{L(\beta_0 | y)}{L(\beta | x, y)} = -2 \log\{L(\beta_0 | y) + 2 \log L(\beta | x, y)\},$$

em que  $L(\beta_0|D)$  é a função de verossimilhança para um modelo composto apenas pelo intercepto, e  $L(\beta|D)$  é a função de verossimilhança para o modelo completo composto pelas variáveis,  $X_1, X_2, X_3, X_4$  e  $X_5$  e, em que,  $\beta = (\beta_1, \beta_2, \beta_3, \beta_4$  e  $\beta_5)$ . A partir deste ponto, vamos denotar o modelo completo sem defasagem  $d = 0$  e com defasagem ( $d = 1$ ) por  $M_0(0)$  e  $M_0(1)$ , respectivamente. Sob a hipótese nula  $H_0$ , a estatística de teste  $S$  segue uma distribuição qui-quadrado com  $k$

graus de liberdade,  $S \sim X_k^2$ , em que  $k$  é o número de variáveis explicativas; ver, por exemplo, McCullagh e Nelder (1989) e Casella e Berger (2002). Para aplicação do teste de hipótese, consideramos um nível de significância  $\alpha = 0,05$ .

A Tabela 3 mostra o resumo dos resultados do teste para o modelo  $M_0(0)$ . Como o p-valor é menor do que o nível de significância  $\alpha$ ,  $p\text{-valor} < \alpha$ , então rejeitamos a hipótese nula  $H_0$ . O resultado do teste para o modelo  $M_0(1)$  é similar, *i. e.*,  $p\text{-valor} < \alpha$ . Este resultado mostra que pelo menos uma das variáveis explicativas é importante para explicação da variável respostas  $Y$ , quantidade de casos de dengue no mês  $t$ , para  $t = 1, \dots, n$ .

Tabela 3 – Teste da razão de verossimilhanças

Modelo	-2 Log Probabilidade	Estatística $S$	$p$ -valor	Graus de liberdade
Intercepto	467.350,40			
$M_0$	304.420,00	162.930,40	3	< 0.00001

Fonte: Elaborado pelos autores.

Além do modelo completo  $M_0(d)$ , ajustamos cinco outros modelos conforme descritos na Tabela 4, para  $d = 0, 1$ . Ou seja, a partir dos modelos  $M_0(0)$  e  $M_0(1)$ , obtemos os modelos  $M_j(0)$  e  $M_j(1)$  retirando a variável  $X_j$ , respectivamente, para  $j = 1, 2, 3, 4, 5$ . Para selecionar o melhor modelo, consideramos os critérios de seleção de modelos *Akaike information criterion*, denotado por AIC (Akaike 1974, Bozdogan, 1987) e *Bayesian Information criterion*, denotado por BIC (Schwarz, 1978). O melhor modelo é aquele que possui o menor valor AIC e BIC.

Para obtenção dos valores AIC, utilizamos o comando *step* do *software* R. Este comando fornece os valores AIC dos modelos considerados a partir da aplicação de um algoritmo *stepwise*. Para obter os valores BIC, implementamos os modelos utilizando o comando *glm* com a opção *family = Poisson*. Então, utilizamos o comando *BIC (modelo)*

Tabela 4 – Variáveis usadas para ajustar os modelos  $M_j, j = 1, 2, 3, 4, 5$

Modelo	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$
Variável retirada	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$

Fonte: Elaborada pelos autores.

A Tabela 5 mostra os valores de AIC e BIC dos seis modelos considerados com  $d = 0$  e  $d = 1$ . Nota-se que o modelo  $M_0(0)$  apresenta o menor valor AIC e BIC (destacado em negrito), ou seja, os dois critérios indicam o modelo  $M_0(0)$  como sendo o melhor dentre os considerados. Além disso, temos que o melhor modelo é composto pelas cinco variáveis e não apresenta defasagem. Isto mostra que os valores das variáveis ambientais de um determinado mês influenciam a quantidade de casos registrados no próprio mês.

Tabela 5 – Valores AIC e BIC para os modelos  $M_j, j = 0, \dots, 5$

Modelo	AIC	BIC	Modelo	AIC	BIC
$M_0(0)$	<b>304.432</b>	304.450	$M_0(1)$	314.636	314.653
$M_1(0)$	393.863	393.878	$M_1(1)$	319.718	319.733
$M_2(0)$	314.035	314.050	$M_2(1)$	335.160	335.175

Modelo	AIC	BIC	Modelo	AIC	BIC
$M_3(0)$	313.257	313.272	$M_3(1)$	323.907	323.922
$M_4(0)$	306.934	306.949	$M_4(1)$	316.342	316.357
$M_5(0)$	305.717	305.731	$M_5(1)$	318.975	318.990

Fonte: Elaborada pelos autores.

Como a temperatura ideal para proliferação do *Aedes aegypti* fica em torno de 30°C a 32°C, também consideramos o ajuste dos modelos  $M'_j(0)$  e  $M'_j(1)$  obtidos pela substituição da variável  $X_1 =$  “temperatura média do mês” pela variável  $X'_1 =$  “Média das temperaturas máximas do mês”, respectivamente, para  $j = 0, 1, 2, 3, 4, 5$ .

A Tabela 6 mostra os valores de AIC e BIC para os modelos  $M'_j(d)$ , para  $d \in \{0,1\}$  e  $j = 0, 1, 2, 3, 4, 5$ . Nota-se que, para os modelos  $M'_j(d)$ , o melhor modelo também é o que não possui defasagem (destacado em negrito). Porém, como este modelo apresenta valores AIC e BIC maiores do que os do modelo  $M_0(0)$ , isto indica que  $M_0(0)$  explica melhor os dados registrados. Ou seja, o melhor modelo não possui defasagem e é composto pelas cinco variáveis explicativas, sendo a variável  $X_1$  a média das temperaturas do mês.

Tabela 6 – Valores AIC e BIC para os modelos  $M_j(d)$ ,  $J = 0, 1, 2, 3, 4, 5$  e  $d = 0,1$   $M'_0(1)$

Modelo	AIC	BIC	Modelo	AIC	BIC
$M'_0(0)$	<b>307.701</b>	307.719	$M'_0(1)$	323.327	323.345
$M'_1(0)$	399.191	399.205	$M'_1(1)$	329.368	329.383
$M'_2(0)$	–	–	$M'_3(1)$	–	–
$M'_3(0)$	317.292	317.307	$M'_3(1)$	333.117	333.131
$M'_4(0)$	309.265	309.279	$M'_4(1)$	326.579	326.594
$M'_5(0)$	309.451	309.465	$M'_5(1)$	330.173	330.187

Fonte: Elaborada pelos autores.

#### 4.1 MODELO AJUSTADO

A Tabela 7 apresenta as estimativas dos parâmetros do modelo  $M'_0(0)$ , o erro-padrão das estimativas, o valor da estatística  $Z$  e o  $p$ -valor do teste de Wald (Wald, 1943). Nota-se que  $p - \text{valor} < \alpha$  para todas as variáveis, indicando que todas são importantes para explicação da variável resposta. A sexta coluna desta tabela mostra os intervalos de 95% para as estimativas dos parâmetros. Como nenhum dos intervalos contém o valor 0, isto também indica a significância de todas as 5 variáveis na explicação da variável resposta.

Tabela 7 – Estimativas para os parâmetros do modelo

Parâmetro	Estimativa	Erro-padrão	Valor-z-	p-value	C.I. 95%
$\beta_0$	2,6673	5.528e-02	48,25	$< 2 e - 16$	(2,5589;2,7755)
$\beta_1$	-0,2501	1.064e-03	-235,10	$< 2 e - 16$	(-0,2522;-0,2480)
$\beta_2$	0,2161	2.254e-03	95,85	$< 2 e - 16$	(0,2117;0,2205)
$\beta_3$	0,0035	3.657e-05	95,33	$< 2 e - 16$	(0,0034;0,0036)
$\beta_4$	-0.0395	7.950e-04	-49,74	$< 2 e - 16$	(-0,0410;-0,0379)
$\beta_5$	0,3566	1.002e-02	35,58	$< 2 e - 16$	(0,3370;0,3763)

Fonte: Elaborada pelos autores.

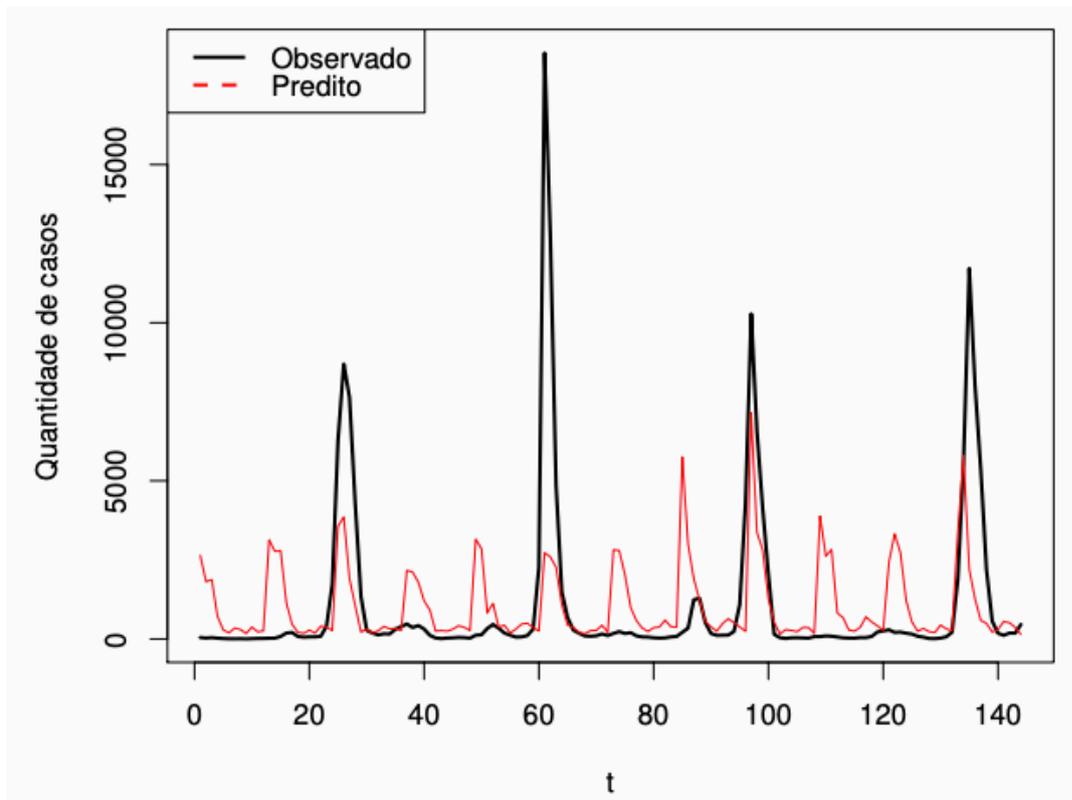
O modelo de regressão linear ajustado é dado por

$$\text{Log}(\mu_t) = 2.6673 - 0.25018x_{t1} + 0.2161x_{t2} + 0.0035x_{t3} - 0.0395x_{t4} + 0.3566x_{t5} \quad (7)$$

Para  $t = 1, \dots, 1 n$ .

O Gráfico 5 mostra as quantidades de casos de dengue registrados no tempo  $t$  e o modelo ajustado (linha em vermelho).

Gráfico 5 – Quantidade registrada e modelo ajustado



Fonte: Elaborado pelos autores.

Nota-se, no Gráfico 5, que o modelo ajustado (linha em vermelho) mostra um comportamento cíclico do número de casos. Em geral, a quantidade de casos da dengue começa a aumentar a partir do mês de outubro de cada ano, seguida de um pico que ocorre entre o final do mês de janeiro e início do mês de fevereiro do próximo ano. A partir da ocorrência do pico, a quantidade começa a diminuir e permanece com esta redução até o mês de setembro, e, em seguida, o ciclo recomeça.

## 5 CONSIDERAÇÕES FINAIS

Neste artigo, desenvolvemos uma modelagem estatística para identificar e avaliar a relação existente entre o número de casos de dengue registrados na cidade de Campo Grande, MS, com variáveis climáticas, usando um modelo de regressão de Poisson. Como mostrado pela análise descritiva, o número de casos de dengue apresenta um comportamento sazonal. Além disso, os dados mostram que, a cada dois anos, Campo Grande tem um surto de dengue (Ver Gráfico 3).

Para relacionar os casos de dengue às variáveis explicativas consideradas (ver Eq. (1)), adotamos um modelo de regressão de Poisson. Utilizando este modelo, verificamos a existência ou não de uma defasagem em relação às variáveis explicativas. Além disso, verificamos a adequabilidade de cada variável no modelo, utilizando os critérios de seleção de modelos AIC e BIC. De acordo com estes dois critérios, o melhor modelo não possui defasagem e é composto pelas cinco variáveis explicativas descritas em (1).

Como o modelo ajustado é baseado apenas em valores de variáveis ambientais, das quais não temos controle, e os coeficientes das variáveis  $X_2$  (temperatura média) e  $X_3$  (precipitação pluviométrica) são positivos (ver Equação 7), quanto maior for o valor da temperatura média e/ou a precipitação pluviométrica, maior será o número esperado de casos de dengue. Ou seja, estes resultados mostram que a única maneira de evitarmos a ocorrência dos “picos” de casos de dengue é por algum tipo de intervenção humana, para conter a proliferação do mosquito transmissor e alterar o cenário baseado apenas nas variáveis ambientais. Além disso, a intervenção humana tem de ser permanente, para que não ocorram picos em determinados anos e em outros não. Caso não haja esta intervenção, é esperado que sempre ocorram os picos de casos de dengue na cidade de Campo Grande, como indicado pelo gráfico do modelo ajustado (ver Gráfico 5). Todas as análises estatísticas descritas foram implementadas no *software* R, e os códigos-fonte podem ser obtidos via e-mail aos autores.

## REFERÊNCIAS

- AKAIKE, H. A. *New look at the statistical model identification*. *IEEE Transactions on Automatic Control*, London, UK, v. 19, n. 264, p. 716-23, 1974.
- AZIL, A. H.; LONG, S. A.; RITCHIE, S. A.; WILLIAMS, C. R. The development of predictive tools for pre-emptive dengue vector control: a study of *Aedes aegypti* abundance and meteorological variables in North Queensland, Australia. *Tropical Medicine International Health*, London, v. 15, n. 10, p. 1190-7, 2010.
- BIBLIOTECA VIRTUAL EM SAÚDE DO MINISTÉRIO DA SAÚDE. Dengue. *Portal BVSMS*, Brasília, DF, 2007. Disponível em: <https://bvsms.saude.gov.br/dengue-16>. Acesso em: 9 set. 2021.
- BOZDOGAN, H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, New York, US, v. 52, p. 345-70, 1987.
- CAMERON, A. C.; TRIVEDI, P. K. *Regression Analysis of Count Data*. 2. ed. New York: Cambridge Press, 1998.
- CASELLA, G.; BERGER, R. L. *Statistical inference*. New York, US: Duxbury Pacific Grove, CA, 2002. V. 2.
- COSTA, M. A. R. *A Ocorrência do Aedes aegypti na Região Noroeste do Paraná: um estudo sobre a epidemia da dengue em Paranavaí- 1999, na perspectiva da Geografia Médica*. 2001. 214 f. Dissertação (Mestrado em Institucional em Geografia) – Faculdade Estadual de Educação Ciências e Letras de Paranavaí, Universidade Estadual Paulista, Presidente Prudente, 2001.
- CRAN R. R. *A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2021. Disponível em: <http://R-project.org>. Acesso em: 20 set. 2021.
- DONALISIO, M. R.; GLASSER, C. M. Vigilância entomológica e controle de vetores do dengue. *Revista Brasileira de Epidemiologia*, São Paulo, v. 5 n. 3, p. 259-72, dez. 2002.
- GUBLER, D. J. Dengue and dengue hemorrhagic fever. *Clinical Microbiology Reviews*, Colorado, v. 11, n. 3, p. 480-96, 1998.

GUBLER D. J. (2002). The global emergence/resurgence of arboviral diseases as public health problems. *Archives of medical research*, London, v. 33, n. 4, p. 330-42, 2002.

MCCULLAGH, P.; NELDER J. A. *Generalized linear models*. New York: Chapman / Hall, 1989.

SCHWARZ, G. E. Estimating the dimension of a model. *Annals of Statistics*, New York, n. 6, p. 461-64, 1978.

SILVA, J. S., MARIANO, Z. F., SCOPEL, I. A influência do clima urbano na proliferação do mosquito *Aedes aegypti* em Jataí (GO) na perspectiva da geografia médica. *Hyheia*, Uberlândia, v. 2, n. 5, p. 33-49, dez. 2017.

WALD, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, New York, v. 54, n. 3, p. 426-82, 1943.

WAN, W. Y. F.; WAN, W. H. A.; ALIAS, L. M.; WAH, Y. B. Modelling Dengue Fever (DF) and Dengue Hemorrhagic Fever (DHF) Outbreak Using Poisson and Negative Binomial Model *International Journal of Medical. Medicine and Health Sciences*, London, v. 3, n. 2, p. 1-6, 2010.

### **Sobre os autores:**

**Erlandson Ferreira Saraiva:** Pós-doutor em Estatística pelo Programa Interinstitucional de Pós-graduação em Estatística da UFSCar-USP. Doutor em Estatística pelo departamento de Estatística de Universidade Federal de São Carlos (Des-UFSCar). Mestre em Estatística pelo departamento de Estatística da Universidade Federal de São Carlos (Des-UFSCar). Graduação em Matemática Aplicada e Computacional pela Universidade Católica Dom Bosco (UCDB). Atualmente é professor adjunto do Instituto de Matemática da Universidade Federal de Mato Grosso do Sul (INMA-UFMS). **E-mail:** erlandson.saraiva@ufms.br, **ORCID:** <https://orcid.org/0000-0002-4698-4004>.

**Leandro Sauer:** Pós-doutor em Administração pela Faculdade de Economia, Administração, Contabilidade e Atuária da Universidade de São Paulo (FEA-USP). Doutor em Engenharia Elétrica pela Universidade Estadual de Campinas (UNICAMP). Mestre em Engenharia Elétrica pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Aperfeiçoamento em Estatística Matemática pelo Instituto de Matemática Pura e Aplicada (IMPA), RJ. Graduação em Matemática pela Universidade Federal de Mato Grosso do Sul (UFMS). Atualmente, é professor titular da Escola de Administração e Negócios (ESAN) e professor do Programa de Pós-Graduação em Administração da UFMS, e professor do Programa de Mestrado Profissionalizante em Rede Nacional (PROFIAP). **E-mail:** leandro.sauer@ufms.br, **ORCID:** <https://orcid.org/0000-0003-4882-428X>

**Mariana Villela Flesh:** Graduada em Matemática pelo Instituto de Matemática da Universidade Federal de Mato Grosso do Sul (INMA-UFMS). **E-mail:** mariflesh@hotmail.com, **ORCID:** <https://orcid.org/0000-0003-3821-5034>