

## Algoritmo GARMA-Scan para previsão de clusters espaço-temporais de doenças.

Almeriane da Silva Almeida<sup>1</sup>(PQ), Max Sousa de Lima<sup>2</sup>(PQ).

<sup>1</sup>Universidade Federal do Amazonas, Departamento de Estatística, Av, Rodrigo Otávio Jordão Ramos 1200, Coroado 1, 69067-005, Manaus AM, Brasil.

\* almerianealmeida@gmail.com

**Palavras Chave:** Séries Temporais; Modelos GARMA; Scan Espacial; Cluster Espaciais, Previsão.

### Introdução

O aumento de doenças relacionado às mudanças climáticas e à intervenção humana é uma preocupação crescente para a saúde global, evidenciada por surtos de dengue no Brasil. Este trabalho tem como objetivo prever mapas de surtos utilizando técnicas estatísticas de detecção de clusters espaço-temporais, como DBSCAN e Scan Espacial.

A metodologia desenvolvida combina modelos de séries temporais, como o Modelo Linear GARMA (Modelos Lineares Generalizados Auto-Regressivos de Médias Móveis<sup>1</sup> que podem ser utilizadas para previsão do número e intensidade de casos para um tempo futuro  $t+k$  ( $k$  dias,  $k$  meses à frente), e técnicas espaciais (Scan Espacial) que estima com alta precisão a localização dos clusters. Dessa forma, desenvolveu-se o algoritmo GARMA-Scan, que é explicável e permite que analistas de saúde pública entendam os fatores que contribuem para a disseminação das doenças. As principais contribuições incluem a construção de um modelo de previsão que integra dados ambientais, climáticos, demográficos e sociais, além de um algoritmo automático para prever o número e risco de casos com até três meses de antecedência. As previsões utilizam informações do IBGE, INMET e DATASUS, demonstrando a aplicabilidade do modelo na estimativa de clusters futuros de dengue nos municípios do Amazonas.

### Material e Métodos

Nesse trabalho, em nossas aplicações do algoritmo proposto usamos os dados mensais de contagem dos casos de dengue disponíveis no (DATASUS) para o período de 2009 a 2022, dados ambientais e climáticos (número de dias com chuva, precipitação total, temperatura máxima, média e mínima, e proporção de dias em que a temperatura mínima ultrapassou 26°C) disponíveis no (INMET) e dados demográficos de população disponíveis no (IBGE), além dos pacotes, integrados ao ambiente do software R.

O modelo GARMA com distribuição de Poisson é dado através da seguinte relação<sup>2</sup>:

$$\log(\mu(t, s)) = x'(t, s)\beta + \sum_{j=1}^p \hat{\phi}_j \{ \log(y(t-j, s)) + x'(t-j, s)\beta \} + \sum_{i=1}^q \hat{\theta}_i \log \left( \frac{\hat{y}(t+k-j, s)}{\hat{\mu}(t+k-j, s)} \right)$$

Onde  $\mu(t, s)$  é a média da distribuição Poisson no tempo  $t$  em uma área  $s$ , que é modelada como função das observações anteriores  $y(t-j, s)$  e  $x'(t, s)\beta$ , que denota o vetor transposto das variáveis explicativas pelo vetor  $\beta' = (\beta_1, \dots, \beta_s)$  que quantifica as distribuições de cada variável explicativa para o total de casos,  $\phi' = (\phi_1, \dots, \phi_p)$  e  $\theta' = (\theta_1, \dots, \theta_q)$  são vetores de parâmetros autorregressivos e de média móvel respectivamente. Essa estimação é realizada usando a biblioteca *glarma* do software. Após essa estimação, as previsões de casos da doença  $k$ -meses à frente em cada município  $s$  serão realizadas através de  $\hat{y}(t+k, s) = \hat{\mu}(t+k, s)$ .

O método Scan Espacial<sup>3</sup>, é uma abordagem específica para detectar clusters espaciais de incidência de doenças ou eventos de saúde pública em um mapa. Neste trabalho o GARMA Poisson é o modelo ajustado. Assim, pode-se mostrar que a função de varredura do algoritmo GARMA-Scan, chamada de LR, para a estimação de clusters futuros  $k$ -meses à frente do tempo  $t$  da forma:

$$LR_{t+k(z)} = \begin{cases} I(z)^{C_{t+k,z}} O(z)^{C_{t+k}-C_{t+k,z}}, & \text{se } I(z) > 1 \\ 1, & \text{caso contrário.} \end{cases}$$

Onde  $C_{t+k}$  é o total de casos previstos para todo o mapa no tempo  $(t+k)$  e  $C_{t+k,z}$  é o total de casos previsto para a zona  $z$  do mapa fixada no tempo  $(t+k)$  com  $I(z) = \frac{C_{t+k,z}}{\hat{\mu}_{t,z}}$  e  $O(z) = \frac{C_{t+k}-C_{t+k,z}}{C_{t+k}-\hat{\mu}_{t,z}}$ . Quando  $LR_{t+k(z)}$  ultrapassa um valor de referência,  $v_0$  dizemos que a zona  $z$  é um possível clusters espaço-temporal de base  $z$  a altura no intervalo  $(t, t+k)$ .

Portanto, o algoritmo GARMA-Scan proposto é descrito da seguinte forma:

**Passo 1:** Dado a série de casos  $Y(t, s)$  observados no espaço-tempo  $[s \times t]$ . Para um município  $s$  fixado ajustamos um modelo GARMA para a série.

**Passo 2:** Obtemos os valores previstos  $\hat{y}(t+k, s)$  com  $k > 0$  e  $s = 1, 2, \dots, L$  (onde  $L = 62$ , representado os municípios do estado do Amazonas).

**Passo 3:** Para cada possível zona  $z$  formada no mapa do Amazonas (note que nesse caso podemos ter até  $2^{62} - 1$  zonas formadas por municípios) e tempo  $t+k$ , definimos  $v_0$  e calculamos  $LR_{t+k(z)}$ .

**Passo 4:** se  $LR_{t+k(z)} > v_0$ , estimamos a zona  $z$  como a possível parte espacial do cluster no mapa. Este Passo;4 do algoritmo foi implementado da biblioteca *scanstatistics* do software R, onde a quantidade  $v_0$  é ajustada automaticamente dentro do algoritmo Scan.

## Resultados e Discussão

Apresentamos o ajuste do modelo GARMA para os 62 municípios do Estado do Amazonas, aplicado ao número mensal de casos de dengue notificados entre janeiro de 2009 e dezembro de 2021 usados para ajuste do modelo e 2022 para comparação das previsões. Os ajustes GARMA foram realizados utilizando a biblioteca *glarma* do software R para encontrar o modelo mais adequado para cada município. O objetivo foi avaliar a capacidade do modelo GARMA em ajustar-se adequadamente e realizar boas previsões para os casos futuros de dengue, utilizando o conjunto de covariáveis citadas anteriormente, 11 modelos foram testados para cada município. A seleção dos melhores modelos foi baseada no menor valor do Erro Quadrático Médio de Predição (EQMP).

A análise revelou que fatores climáticos e temporais impactam a incidência de dengue. O número de casos do mês anterior ( $Y_{t-1}$ ,  $Y_{t-2}$ ) tende a ter um efeito positivo significativo sobre os casos atuais, sugerindo uma continuidade na transmissão da doença, que está relacionada ao ciclo de vida do mosquito transmissor, em média de 45 dias. A influência da precipitação variou entre os municípios, sugerindo que as condições urbanas locais afetam a proliferação do mosquito. A temperatura também teve papel relevante, com temperaturas máximas e médias tendendo a inibir a incidência de dengue, sendo a faixa ideal para a proliferação do mosquito entre 30 e 32 graus. Esses resultados destacam a complexidade das dinâmicas da dengue na região, mostrando que os fatores climáticos e temporais não afetam uniformemente todos os municípios. A análise sugere que as estratégias de controle da dengue precisam ser adaptadas às condições específicas de cada localidade, considerando as variáveis climáticas predominantes e os padrões de transmissão ao longo do tempo.

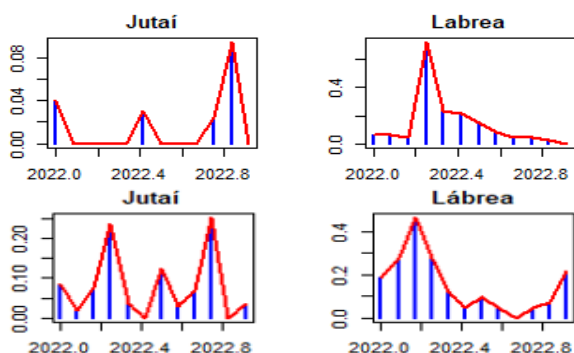


Figura 1. Série temporal real versus a série prevista da taxa de casos de dengue para 2022.

A Fig 1 apresenta a série temporal real versus a série prevista da taxa de casos de dengue para todos os meses de 2022 dos municípios de Jutai e Lábrea no Estado do Amazonas. Essas séries foram elaboradas para comparar as previsões com os dados observados e avaliar a eficácia dos modelos GARMA em prever o comportamento futuro da série temporal, essencial para a aplicação do método Scan e o planejamento de intervenções de saúde pública.

A análise espaço-temporal dos casos de dengue nos municípios do Amazonas foi realizada usando o método GARMA-Scan. O estudo identificou áreas com alta concentração de casos de dengue (clusters espaço-temporais) entre janeiro e maio de 2023, avaliando 1326 zonas espaciais com janelas temporais de 1 a 5 meses. Os resultados principais indicam que a zona  $z_1$ , composta por Atalaia do Norte, Guajará e Ipixuna, apresentou uma quantidade de casos aproximadamente 12 vezes maior que o esperado ( $I(z) = 12,36$ ) com uma LLR de 801,24 e um valor-p de 0,001, indicando um surto significativo previsto para começar em fevereiro e durar cerca de 4 meses. Comparado com dados históricos, houve uma mudança no padrão de distribuição de casos, destacando-se uma gravidade maior na zona  $z_1$  em 2022/2023 em contraste com 2021/2022. A análise também revelou outros focos significativos, como a zona  $z_2$  (Caapiranga, Anamá, Manacapuru), evidenciando surtos não mencionados no boletim da FVS<sup>4</sup>, indicando uma distribuição geográfica preocupante.

## Conclusões

No presente trabalho, foi desenvolvido o algoritmo GARMA-Scan para previsão espaço-temporal de clusters de doenças, com foco na dengue. A aplicação do GARMA-Scan mostrou-se eficaz para identificar clusters de casos de dengue no Amazonas, destacando a influência das variáveis climáticas e do ciclo de vida do mosquito na disseminação da doença. A análise apontou Atalaia do Norte como o principal centro de clusters de dengue, com um valor-p de 0.001, além de identificar outros municípios como Guajará, Ipixuna, Caapiranga, Anamá e Manacapuru. Esses achados permitem que as autoridades priorizem recursos e intervenções de forma mais eficaz. O desenvolvimento do GARMA-Scan representa um avanço significativo na gestão de surtos de doenças, apesar da limitação de ainda não ter sido avaliado por simulação para prever clusters futuros, algo que será abordado em estudos futuros, incluindo sua aplicação a outras doenças transmitidas por vetores.

## Agradecimentos

Agradecemos à Fundação de Amparo à Pesquisa do Estado do Amazonas pelo suporte financeiro.

## Referências

- <sup>1</sup>MICHAEL, M. A.; RIGBY, R. A.; STASINOPOULOS, D. M. *Modelos Lineares Generalizados Auto-Regressivos de Médias Móveis*. *Journal of the American Statistical Association*, v. 98, n. 461, p. 214-223, 2003.
- <sup>2</sup>BENJAMIN, M. A.; RIGBY, R. A.; STASINOPOULOS, D. M. *Generalized autoregressive moving average models*. *Journal of the American Statistical Association*, v. 98, n. 461, p. 214-223, 2003.
- <sup>3</sup>KULLDORFF, M. *A spatial scan statistic*. *Communications in Statistics: Theory and Methods*, v. 26, n. 6, p. 1481-1496, 1997.
- <sup>4</sup>FUNDAÇÃO DE VIGILÂNCIA EM SAÚDE DO AMAZONAS - DRA. ROSEMARY COSTA PINTO (FVS-RCP). *Boletim Epidemiológico*, Ano 2, nº 02, Janeiro de 2023. Governo do Estado do Amazonas, Secretaria de estado de Saúde do Amazonas (SES-AM)