



Predição do número de notificações de dengue em Maringá via modelo SARIMA

Aline Edlaine de Medeiros¹, Eniuce Menezes de Souza², Isolde Terezinha Santos Previdelli³, Marcos Vinícius Oliveira Peres⁴ e Guilherme Zulbatch da Cunha⁵

¹ Programa de Pós-graduação em Bioestatística PBE-UEM

² Programa de Pós-graduação em Bioestatística PBE-UEM

³ Programa de Pós-graduação em Bioestatística PBE-UEM

⁴ Departamento de Engenharia Agrícola DEA-UEM

⁵ Programa de Pós-Graduação em Saúde na Comunidade PGCS-USP

RESUMO

Este estudo tem por objetivo desenvolver um modelo para a predição do número de casos de dengue em Maringá, Paraná, Brasil, por técnicas de análise de séries temporais, em particular, o modelo SARIMA. Considerando as notificações no período de janeiro de 2005 à março de 2016, realizou-se o ajuste do modelo, obteve-se os valores preditos para os meses de abril à setembro de 2016, e efetuou-se a validação do modelo. Os resultados obtidos mostram que o modelo SARIMA $(0, 0, 2) \times (0, 1, 1)_{12}$ é eficiente em prever o número de notificações de dengue, mostrando-se um aliado no controle da doença.

Palavras chave: Dengue; SARIMA; Previsão.

1 INTRODUÇÃO

A dengue é causada pelo vírus RNA, que é um Arbovírus do gênero Flavivírus, pertencente à família Flaviviridae, que em geral, é transmitida ao ser humano através da picada do mosquito fêmea da família Aedes Aegypti.

Uma vez que Aedes Aegypti se adapta muito bem a climas subtropicais, países como o Brasil, tem lidado com epidemias da doença. Em 2016 o país registrou 1.438.624 casos prováveis de dengue no país até 17/09, dos quais, 72.048 ocorreram na região Sul (MINISTÉRIO DA SAÚDE, 2016)[5]. Em particular, a cidade de

Maringá-PR registrou 8467 notificações da doença de janeiro a outubro de 2016, e esteve em situação epidêmica em vários anos anteriores, como 2007 e 2010. Neste contexto, descrever e prever o número de notificações de dengue mostra-se um problema desafiador, e de grande utilidade para as autoridades de controle de endemias, sendo este o objeto de estudo deste trabalho.

2 METODOLOGIA

A cidade de Maringá localiza-se na região noroeste do estado do Paraná, Brasil (23°25' latitude Sul e 51°57' longitude Oeste), com extensão territorial de 487,052km² com uma população estimada de 403.063 mil habitantes (IBGE,2016)[1]. Conforme descrito por Ely(2013)[3] aspectos ambientais da cidade de Maringá, tais, como precipitação e temperatura, tornam a cidade favorável à eclosão dos ovos, ação e proliferação dos mosquitos *Aedes aegypti*. Além disso, a cidade possui um elevado número de notificações da doença, em anos como, 2007, 2008 e 2010.

O número de notificações da doença foi disponibilizado pelo setor de Vigilância Sanitária da Secretária Municipal de Saúde de Maringá, e compreendem o período de janeiro de 2005 à outubro de 2016. Para realizar o ajuste da série utilizou-se o período de janeiro de 2005 à março de 2016, e os meses subsequentes foram empregados na validação do modelo.

Dentre as técnicas empregadas para a modelagem em séries temporais, adotou-se o modelo *Seasonal Autoregressive Integrated Moving Average* (SARIMA), devido a sua capacidade de modelar dados com as características descritas, em detrimento de modelos como o autoregressivos integrado de médias móveis (ARIMA) e o autoregressivo de médias móveis(ARMA).

Seja X_t uma série temporal, um modelo SARIMA com s observações por período, denotado por $SARIMA(p, d, q) \times (P, D, Q)_s$, é dado por $\phi(B)\Phi(B^s)W_t = \theta(B)\Theta(B^s)\epsilon_t$, em que, $\phi(B) = (1 - \alpha_1 B - \dots - \alpha_p B^p)$, $\Phi(B^s) = (1 - \phi_s B^s - \dots - \phi_P B^{Ps})$, $\theta(B) = (1 - \beta_1 B - \dots - \beta_q B^q)$, $W_t = \nabla^d \nabla_s^D X_t$ e $\Theta(B^s) = (1 - \theta_s B^s - \dots - \theta_Q B^{Qs})$ [2]. Em outras palavras, $\phi(B)$ é o operador autoregressivo, $\Phi(B^s)$ é operador autoregressivo sazonal, $\theta(B)$ é operador de médias móveis, $\Theta(B^s)$ é operador de médias móveis sazonal, W_t são as diferenciações necessárias para tornar a série estacionária, e ϵ_t é o termo do erro, que admiti-se ser independente e identicamente distribuído e é denominado ruído branco. O ruído branco, representa fenômenos externos aos dados observados e que o modelo não é capaz de explicar, no caso da dengue, o ruído branco poderia ser o efeito da temperatura, ou da precipitação, entre outras informações que possam influenciar o número de notificações da doença.

Para realizar o ajuste do modelo utilizou-se os pacotes *astsa* e *forecast*, disponíveis no software *R*[4], versão 3.3.1. Este software estatístico permitiu realizar a análise gráfica da série, os testes de Ljung Box para análise dos resíduos, e a predição do número de notificações nos meses de abril à dezembro de 2016. Como critério para seleção do melhor modelo utilizou-se o critério de informação de Akaike (AIC) e a análise da variância do modelo estimado.

3 RESULTADOS E DISCUSSÕES

Pode-se observar que a série temporal apresenta alguns picos, por exemplo, no anos de 2007 e 2015, anos em que a cidade de Maringá sofreu com a epidemia da doença, além disso, o comportamento da função de autocorrelação (FAC) indica sazonalidade ($s = 12$). A FAC também forneceu possíveis valores para q , que podem ser 1 ou 2, e função de autocorrelação parcial (FACP) sugere que p é igual a 1 ou 2, conforme descrito na figura 2.

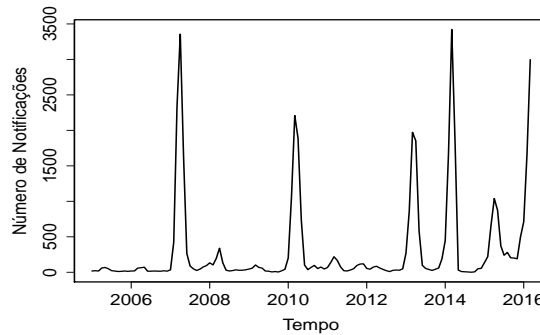


Figura 1: Série temporal do log das notificações de dengue

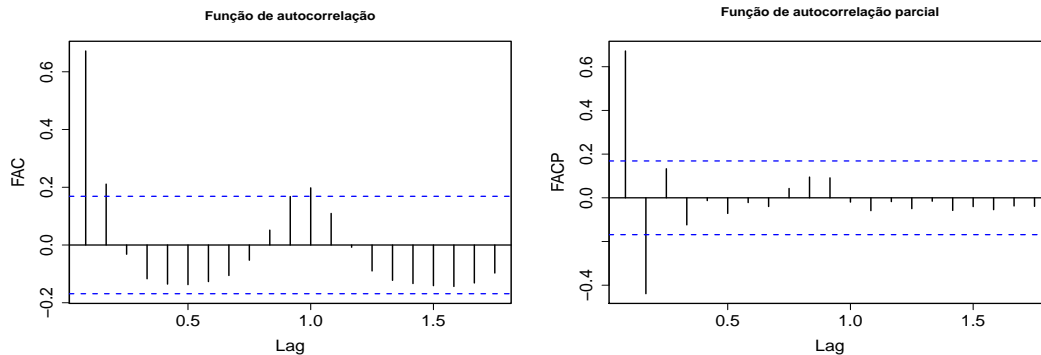


Figura 2: FAC e FACP

Após realizar o ajuste do modelo SARIMA, conforme os valores de médias móveis e processos autoregressivos sugeridos, concluiu-se que o modelo SARIMA com ordem $(0, 0, 2) \times (0, 1, 1)_{12}$ foi o que forneceu o melhor ajuste, visto que, obteve o menor AIC e variância. A tabela 1 apresenta alguns dos modelos analisados e respectivas variâncias e AIC. O teste de Ljung-Box forneceu o valor 7.275 para a estatística χ^2 , com valor-p de 0.839, e consequentemente, a hipótese de que os resíduos são aleatórios não é rejeitada.

A partir do modelo SARIMA $(0, 0, 2) \times (0, 1, 1)_{12}$, efetuou-se a predição dos meses de abril a setembro de 2016 e realizou-se a comparação entre os valores preditos e os valores observados neste período descrita na tabela 2, visando a validação do modelo. Nota-se que os valores preditos são próximos dos valores observados, por

Tabela 1: Akaike e variância para alguns modelos ajustados

Modelo	Variância	AIC
ARIMA(1, 1, 1)(0, 1, 1) ₁₂	108.205	1783, 78
ARIMA(1, 2, 2)(0, 1, 1) ₁₂	110.578	1783, 79
ARIMA(2, 1, 1)(0, 1, 1) ₁₂	100.745	1774, 66
ARIMA(0, 0, 2)(0, 1, 1) ₁₂	89.777	1771, 75

exemplo, no mês de abril, em que existe uma diferença de apenas dez notificações, sugerindo que este modelo pode ser utilizado para a estimação do número futuro de notificações. Com base nos valores observados e preditos pelo modelo, a um nível

Tabela 2: Valores observados e valores preditos das notificações

Mês	Observado	Predito
Abril	2163	2173
Maio	474	628
Junho	102	122
Julho	79	72
Agosto	124	78
Setembro	153	65

de 95% de confiança, espera-se que ocorram aproximadamente 8890 notificações de dengue em 2016, o que indica um aumento em comparação ao ano de 2015, e que pode ter ocorrido, devido aos elevados registros de chuvas e altas temperaturas no município, o que sugere uma análise futura destas covariáveis. Para 2017 espera-se obter 5132 notificações da doença, um número menor que o ano de 2016, e que pode estar associado a vacinação inserida na cidade neste ano e as recorrentes campanhas e fiscalizações realizadas pela Secretaria de Saúde de Maringá.

Referências

- [1] DEMOGRÁFICO, I. C. Disponível em: <http://www.ibge.gov.br/>. Acesso em 1 (2005).
- [2] EHLERS, R. S. Análise de séries temporais. *Universidade Federal do Paraná* (2007).
- [3] ELY, D. F. Tendências climáticas e a incidência da dengue em cidades do sul do Brasil: Estudo de caso de Londrina, Maringá (PR) e Florianópolis (SC). *Revista Brasileira de Climatologia* 13 (2014).
- [4] R CORE TEAM. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [5] SAÚDE, M. D. Monitoramento dos casos de dengue, febre de chikungunya e febre pelo vírus Zika até a semana epidemiológica 37, 2016.