

# MODELO SARIMAX APLICADO AOS CASOS DE NOTIFICAÇÕES DE DENGUE EM RECIFE, PE

**Elias Silva de Medeiros**

Universidade Federal de Lavras

[eliasestatistica@gmail.com](mailto:eliasestatistica@gmail.com)

**Carolina Cristina Bicalho**

Universidade Federal de Lavras

[carolinabicalho@gmail.com](mailto:carolinabicalho@gmail.com)

**Thelma Sáfadi**

Universidade Federal de Lavras

[safadi@des.ufla.br](mailto:safadi@des.ufla.br)

**Kelly Pereira de Lima**

Universidade Federal de Lavras

[kelly\\_limaadm@hotmail.com](mailto:kelly_limaadm@hotmail.com)

## RESUMO

As mudanças climáticas têm impactado de maneira negativa na saúde humana, e isto tem influenciado no aumento da incidência de doenças infecciosas transmitidas por mosquitos, em especial a dengue que pode causar sérios riscos a saúde humana, provocando doença febril aguda. Os modelos estatísticos baseados em análises de séries temporais têm sido amplamente utilizados para identificar fatores climáticos que estão relacionados com as notificações de casos de dengue, bem como, para previsões destas notificações. O objetivo deste estudo é analisar o comportamento temporal das notificações de dengue na cidade de Recife - Pernambuco (PE), considerando os efeitos climáticos. Os resultados indicaram uma relação entre os casos de dengue com fatores climáticos. De acordo com os critérios REQM e MAPE o modelo Autorregressivo Integrado de Média Móvel Sazonal (SARIMA) mais o efeito atual, com uma e com duas defasagens da precipitação foi adequado para prever a quantidade de notificações de dengue em Recife.

## ABSTRACT

Climate change has had a negative impact on human health, and this has influenced the increased incidence of mosquito-borne infectious diseases, especially dengue, which can cause serious human health risks, leading to an acute febrile illness. Statistical models based on time series analysis have been widely used to identify climatic factors that are related to reports of dengue cases as well as for predictions of these reports. The objective of this study is to analyze the temporal behavior of dengue notifications in the city of Recife - Pernambuco (PE), considering the climatic effects. The results indicated a relationship between dengue cases and climatic factors. According to the RMSE and MAPE criteria the Seasonal Autoregressive Integrated Moving Average (SARIMA) model plus the current effect, with one and two lags of rainfall was adequate to predict the number of dengue notifications in Recife.

**Palavras-chave:** Dengue, Doenças infecciosas, Modelo SARIMAX.

## 1 INTRODUÇÃO

As mudanças climáticas têm impactado de maneira negativa na saúde humana, e isto tem influenciado no aumento da incidência de doenças infecciosas transmitidas por mosquitos. A dengue transmitida pelo mosquito *Aedes aegypti* é a arbovirose mais grave e prevalente nos dias atuais. De acordo com Brunkard et al. (2008) [1] estima-se de 50 a 100 milhões de novos casos de dengue a cada ano, embora esta seja uma subestimativa da verdadeira incidência, uma vez que, muitos casos provavelmente não são notificados porque os sintomas da dengue são semelhantes aos da gripe.

As ações tomadas pela saúde pública têm por finalidade monitorar as regiões de maior incidência, realizando a coleta, análise e interpretação dos dados sobre as doenças transmitidas por este mosquito. Com este tipo de procedimento pretende-se acompanhar e analisar as tendências na incidência de doenças infecciosas, fazendo com que as políticas de saúde pública possam propor formas de minimizar os riscos destas doenças infecciosas (Luz et al., 2008 [2]). A dengue pode causar sérios riscos à saúde humana, como por exemplo, provocando doença febril aguda. No início da fase febril, ela pode imitar um largo espectro de doenças febris, incluindo: mononucleose infecciosa, chikungunya, coxsackie e outras infecções por enterovírus, parvovírus B19, rubéola, sarampo, malária, riquetsia e leptospirose e sepsis bacteriana (Singhi et al., 2007 [3]).

Alguns trabalhos apontam uma forte associação entre a incidência da dengue e fatores climáticos, tais como, precipitação, temperaturas e umidade (Donalísio e Glasser, 2002 [4]; Brunkard et al., 2008 [1]; Souza et al., 2010 [5]; Hii et al. 2012 [6]). Por exemplo, Brunkard et al. (2008) [1] mencionaram em seu trabalho que a incidência de dengue apresentou um incremento quando aumentada a temperatura máxima semanal e o mesmo comportamento ocorrendo quando observado o aumento no índice pluviométrico. Eventos climáticos extremos aumentam a incidência de doenças transmitidas por mosquitos. Áreas que recebem um aumento na frequência ou magnitude de eventos climáticos extremos, provavelmente, irão proporcionar uma expansão de criadouros do vetor e habitat larval. Os eventos climáticos extremos, além de aumentar a incidência de dengue, resultam em longos períodos de seca, inundações, vendavais e tempestades, entre outros (Brunkard et al., 2008 [1]).

Os modelos estatísticos baseados em análises de séries temporais têm sido amplamente utilizados para identificar fatores climáticos que estão relacionados com as notificações de casos de dengue, bem como, para previsões destas notificações (Brunkard et al., 2008 [1]; Luz et al., 2008 [2]; Hii et al. 2012 [6]; Wongkoon et al., 2012 [7]). O objetivo deste estudo é realizar uma investigação da dinâmica climática que colabora com o crescimento de notificações da dengue na cidade de Recife.

Este artigo descreve na seção 2.1 o conjunto de dados utilizado na pesquisa. Na seção 2.2 são apresentados os modelos de séries temporais, em especial, o modelo SARIMAX. Na seção 3 tem-se a aplicação do método aos casos de notificações de dengue na cidade de Recife. Na seção 4 é realizada uma discussão acerca da necessidade de monitoramento e predição dos casos de dengue. Na seção 5 têm-se as considerações finais da pesquisa.

## 2 MATERIAL E MÉTODOS

### 2.1 MATERIAL

O município de Recife é a capital do estado de Pernambuco, localizado na Região Nordeste do Brasil. A capital possui um área territorial de aproximadamente  $218 \text{ km}^2$  e com uma população de 1,6 milhões de habitantes. A cidade fica localizada entre a latitude  $8^\circ$

3' 15" S e longitude 34° 52' 53" W. A sua área territorial é composta por 67,43% de morros, 23,26% de planícies, 9,31% de aquáticas e, 5,58% de Zonas Especiais de Preservação Ambiental (Bitoun et al., 2010 [8]).

A série histórica relacionada à dengue é obtida pelas notificações de indivíduos residentes na cidade de Recife, independente de sua confirmação, exceto os descartados, pois em situações de epidemia nem sempre é possível confirmar todos os casos. A série foi constituída de dados mensais, referente ao período de 2001 a 2012, obtida pelo Sistema de Informação de Agravos de Notificação (SINAN).

As variáveis climáticas utilizadas neste artigo foram a Precipitação Total Mensal (mm) e a Umidade Relativa Média (%). Estas variáveis foram obtidas pela estação meteorológica convencional de Recife (Curado), Pernambuco (OMM: 82900), que fica localizada a  $-8,05^\circ$  de latitude,  $-34,95^\circ$  de longitude e 10 metros de altitude. As séries climáticas estão disponíveis no endereço eletrônico do INMET - CSC - BDMEP - Banco de Dados Meteorológicos para Ensino e Pesquisa.

## 2.2 MÉTODOS

Em um modelo auto-regressivo de médias móveis (ARMA), o valor futuro de uma variável é assumido como sendo uma função linear de várias observações passadas e de erros aleatórios. O processo ARMA possui a seguinte forma:

$$y_t = \theta_0 + \sum_{i=1}^p \phi_i y_{t-i} - \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t. \quad (1)$$

Na equação (1),  $y_t$  é a série observada no instante  $t$ ,  $t = 1, 2, \dots, n$ ;  $\varepsilon_t$  é o erro aleatório associado à observação  $y_t$ ;  $\phi_i$  ( $i = 1, 2, \dots, p$ ) são os parâmetros autorregressivos e  $\theta_j$  ( $j = 1, 2, \dots, q$ ) são os parâmetros de médias móveis do modelo;  $p$  e  $q$  são ordens do mesmo. Admite-se que os erros aleatórios ( $\varepsilon_t$ ) sejam ruído branco (RB), com média zero, variância constante  $\sigma^2$  e não autocorrelacionados,  $\varepsilon_t \sim RB(0, \sigma^2)$ . A equação 1 implica uma série de importantes casos especiais. Se  $q = 0$ , torna-se um modelo AR de ordem  $p$ . Quando  $p = 0$ , o modelo se reduz ao MA de ordem  $q$ . O processo ARIMA visa determinar a ordem do modelo apropriado ( $p, q$ ) e o número de diferenças  $d$  (Zhang, 2003 [9]). Uma vez identificada uma série não estacionária, tem-se a necessidade de aplicar certa quantidade de diferenciação com o intuito de obter-se uma série estacionária. A equação (1) pode ser representada da forma reduzida:

$$\phi(L) \Delta^d y_t = \theta(L) \varepsilon_t. \quad (2)$$

Na equação (2),  $L$  representa o operador de retardo,  $\Delta^d$  é a diferença  $(1 - L)^d$ ;  $\phi$  e  $\theta$  são os operadores polinomiais. Assim, tem-se o ARIMA( $p, d, q$ ), sendo  $p$  e  $q$  definidos como na equação (1) e  $\Delta^d$  é a parte de integração não sazonal de ordem  $d$ .

Dentre a família de modelos ARIMA, existe um modelo que é utilizado quando se tem uma série sazonal, ou com ciclos sazonais, conhecido como o modelo SARIMA. Adicionalmente, quando se tem um modelo não estacionário (média não constante ao longo do tempo), fazem-se necessárias uma ou duas diferenças, para que se tenha uma série estacionária, possibilitando o ajuste do modelo. Um modelo adequado que expressa este processo é denotado por SARIMA ( $p, d, q$ )  $\times$  ( $P, D, Q$ ) $_s$ , e pode ser escrito como:

$$\Phi_P(L^s) \phi(L) \Delta_s^D \Delta^d y_t = \Theta_Q(L^s) \theta(L) \varepsilon_t. \quad (3)$$

Na equação (3),  $\Delta_s^D$  é a diferença sazonal  $(1 - L^s)^D$ ,  $\Delta^d$  como a diferença não-sazonal  $(1 - L)^d$  e  $\Phi, \Theta$  são os operadores polinomiais. Na parte sazonal,  $P$  representa a ordem sazonal autorregressiva,  $D$  a ordem de integração sazonal,  $Q$  a ordem de média móvel sazonal e  $s$  o período ou duração da sazonalidade. Por exemplo: se a sazonalidade for anual,  $s = 12$ , se trimestral,  $s = 4$  (Saz, 2011 [10]).

Os modelos apresentados relacionam a dependência temporal em estudo com o seu próprio passado e não utilizam explicitamente as informações contidas em outras séries temporais pertinentes. Como já mencionado anteriormente, além das variações sazonais, a incidência de dengue pode estar relacionada por vários fatores exógenos, principalmente com as condições climáticas. Para analisar com maior precisão a relação existente entre a proliferação do mosquito causador da dengue com as variáveis meteorológicas, serão utilizados modelos de séries temporais com covariáveis.

Na literatura existem outras metodologias que incorporam covariáveis no ajuste da série temporal, tais como, função de transferência, regressão dinâmica e modelos de intervenção (Weron, 2007 [11]). A seguir será detalhado o modelo ARMA com variáveis exógenas, sendo a extensão análoga para as demais classes de modelo citada no parágrafo anterior.

Seja o modelo ARMA definido na equação (2) desprezando o operador de diferença, o processo  $y_t$ , sem especificações adicionais, tem média zero. Uma das formas de modelar os processos em que a série  $y_t$  apresenta média não-nula, é incluir termos ao modelo. Existem duas formas possíveis de representar processos com média não-nula.

A primeira forma é definir  $\mu_t$  como a média incondicional de  $y_t$ , ou seja, o valor central de sua distribuição marginal. Portanto, a série  $\tilde{y}_t = y_t - \mu_t$  tem a média 0, e a equação (2) aplica-se em  $\tilde{y}_t$ . Na prática, assumindo que  $\mu_t$  é uma função linear de algumas variáveis observáveis  $x_t$ , tem-se:

$$\phi(L)(y_t - x_t\beta) = \theta(L)\varepsilon_t. \quad (4)$$

A equação (4) pode ser entendida como um “modelo de regressão com erros ARMA” e sua estrutura pode ser representada por duas equações:  $y_t = x_t\beta + \mu_t$  e  $\phi(L)\mu_t = \theta(L)\varepsilon_t$ . Esse modelo que foi apresentado também é conhecido como “ARMAX” (ARMA + variáveis exógenas). No modelo ARMAX, o valor atual da série temporal  $y_t$  é expresso linearmente em termos de seus valores passados, em termos de valores anteriores do erro e, adicionalmente, em termos de valores presentes e passados das variáveis exógenas. Entretanto nessa modelagem não foi levado em consideração a estacionariedade das séries dessas covariáveis, sendo isto um objeto para estudos futuros.

A segunda maneira de incluir o termo  $\mu_t$  e basear a representação na média condicional de  $y_t$ , que é o valor central da distribuição de  $y_t$  dado seu próprio passado. Assumindo, novamente, que isto pode ser representado como uma combinação linear de algumas variáveis observáveis  $z_t$ , o modelo é então escrito da forma:

$$\phi(L)y_t = z_t\gamma + \theta(L)\varepsilon_t. \quad (5)$$

A equação (5) tem a vantagem de que  $\gamma$  pode ser imediatamente interpretado como o vetor de efeitos marginais das variáveis  $z_t$  na média condicional de  $y_t$ . E, adicionando defasagens de  $z_t$  a esta especificação, pode-se estimar modelos de Função de Transferência (que generalizam ARMA adicionando os efeitos das variáveis exógenas distribuídas ao longo do tempo).

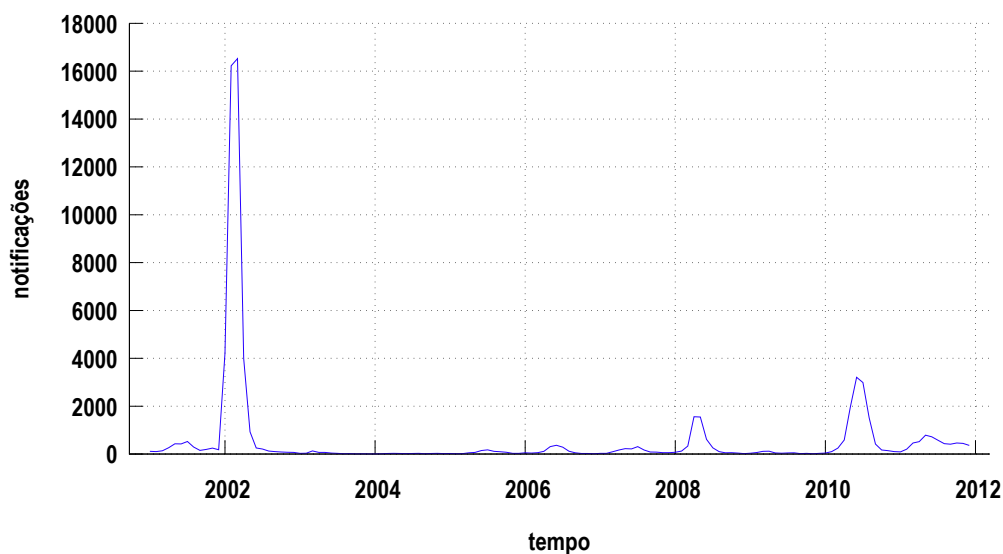
Para os ajustes dos modelos SARIMA e SARIMAX foi utilizado o programa Gretl 2016.d (Gnu Regression, Econometrics and Time-series Library [12]) é um programa livre e escrito em linguagem C, podendo ser obtido pelo endereço, <http://gretl.sourceforge.net/>. O Gretl fornece uma maneira de estimar os modelos escritos em 4 e 5 utilizando máxima verossimilhança (MV) e máxima verossimilhança condicional (MVC), respectivamente. Nesse artigo, para obter as estimativas dos estimadores dos parâmetros, foi utilizado o método MV. As estimativas pelo método MV utiliza todas as  $n$  observações. Entretanto, o método MVC é ponderado pelas quantidades de parâmetros ( $k$ ), utilizando para obtenção das estimativas  $n - k$  observações.

### 3 RESULTADOS

A Figura 1 apresenta a série histórica de notificações de dengue na cidade de Recife, na qual se percebe um comportamento heterogêneo nos dados. Adicionalmente, observou-se

nessa figura um aumento significativo das notificações de dengue nos anos de 2008 e 2010. Nota-se que o ano de 2002 apresentou meses com mais de 16 mil casos de notificações da doença. Montenegro et al. (2006) [13] relatam o caso de epidemia de dengue ocorrida na cidade de Recife no ano de 2002, o que resultou em 214 casos de dengue hemorrágica, sendo 14 óbitos. Lima et al. (2016) [14] mencionam que, durante a estação chuvosa, seguida de elevadas temperaturas, os casos de dengue na cidade de Recife apresentaram um aumento significativo entre o período de 2008 a 2010.

Para fins de previsão, o ano de 2012 será retirado da base de dados, para que, após o ajuste de um modelo, possa-se comparar o valor observado com o seu respectivo valor previsto. Com o intuito de verificar a necessidade de transformação na série de notificação de dengue foi realizado um gráfico de amplitude versus média.



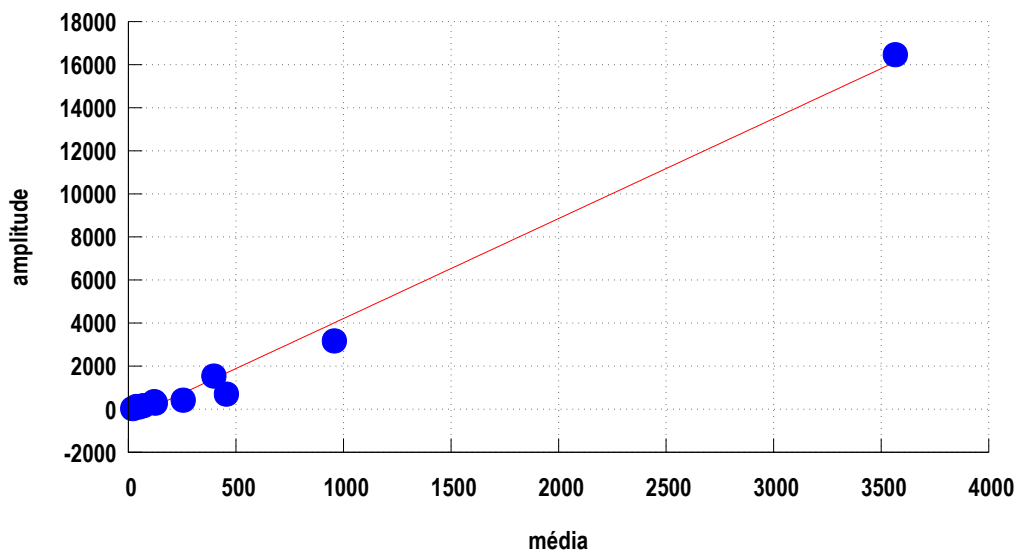
**FIGURA 1:** Quantidade de notificações de dengue entre 2001 e 2011 em Recife

Na Figura 2 tem-se o gráfico da amplitude versus média, a qual observa-se um comportamento linear. Assim, há fortes indícios da necessidade de uma transformação logarítmica ( $\ln$ ) nos dados em estudo. Para verificar a necessidade da transformação foi aplicado o teste de hipótese para o coeficiente de inclinação, que foi de 4,64, no qual a sua hipótese nula é de este coeficiente é estatisticamente nulo. O valor  $p$  do teste foi inferior a 0,01, indicando a rejeição dessa hipótese, confirmando assim, a necessidade da transformação  $\ln$ . Assim, foi aplicado a transformação  $\ln$  na série original das notificações de dengue, ver Figura 3.

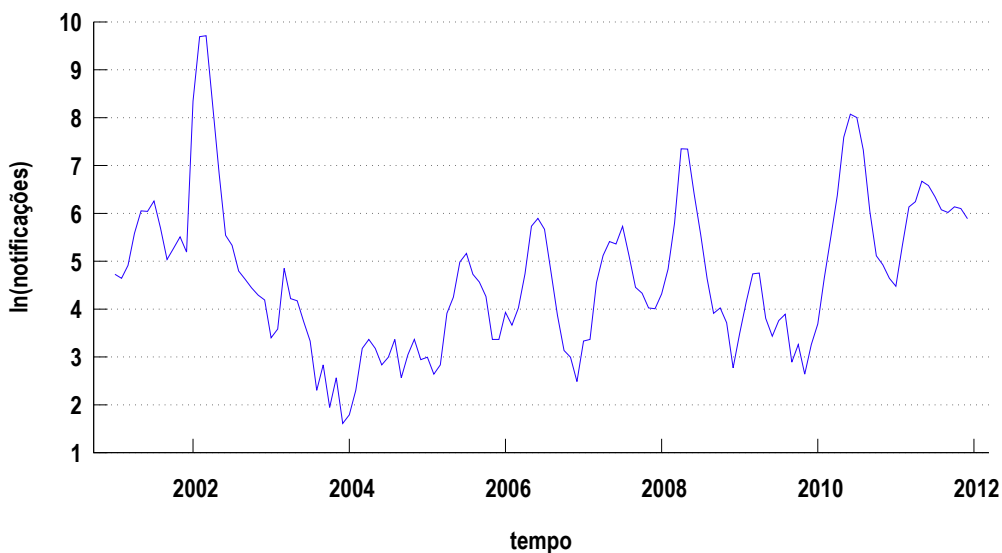
Na Figura 3 mostra que os casos de notificações de dengue, aparentemente, apresentam um comportamento constante ao longo do tempo. Assim, para analisar se o componente de tendência é significativo na série foi aplicado o Teste de Cox-Stuart (Sprenst e Smeeton, 2007 [15]). Após a aplicação desse teste foi obtido um valor  $p$  de 0,11, indicando, ao nível de 0,05 de significância, que não existe o componente de tendência na série, sendo assim, não é necessário aplicar uma diferença, na parte não sazonal, nesta série. Na Figura 4 tem o periodograma (Brockwell e Davis, 2013 [16]), indicando um possível ciclo sazonal de 12 meses ou 1 ano.

Assim, para verificar a existência de sazonalidade, foi aplicado o teste de Kruskal-Wallis (Sprenst e Smeeton, 2007 [15]). A hipótese nula é de que não existe sazonalidade. Após a aplicação desse teste, foi obtido o valor  $p$  inferior a  $< 0,001$ , indicando, ao nível de 0,05 de significância, uma necessidade de aplicar uma diferença sazonal.

Na Tabela 1 tem-se os valores do Critério de Informação Bayesiano (BIC), Critério de Informação de Akaike (AIC), Raiz do Erro Quadrático Médio (REQM) e o Erro Absoluto Médio Percentual (MAPE) [Wang et al., 2009 [17]] para os diferentes modelos SARIMA, cujos resíduos eram ruído branco, considerando com e sem variáveis exógenas. As siglas

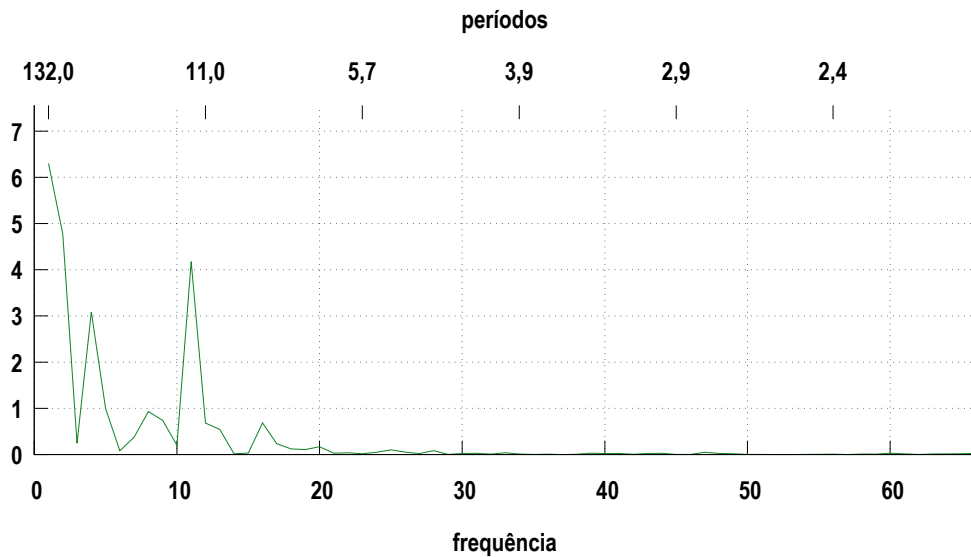


**FIGURA 2:** Gráfico de amplitude versus média da quantidade de notificações de dengue, entre os anos de 2001 a 2011, Recife, Pe. A linha vermelha indica o ajuste da regressão linear



**FIGURA 3:** Quantidade de notificações de dengue entre 2001 e 2011 em Recife da série transformada,  $\ln(dengue)$





**FIGURA 4:** Periodograma para a série transformada

Prec e Umid significam precipitação e umidade, respectivamente.

Modelo	BIC	AIC	REQM	MAPE
$(1, 0, 0)(0, 1, 0)_{12}$	303, 2586	302, 3759	1, 4841	25, 4120%
$(1, 0, 1)(0, 1, 0)_{12}$	299, 1981	297, 7433	1, 3934	23, 6470%
$(1, 0, 1)(0, 1, 1)_{12}$	254, 1556	242, 6244	0, 7134	9, 0410%
$(1, 0, 1)(0, 1, 1)_{12} + Umid_{t-1}$	252, 6336	238, 2197	0, 6989	8, 6197%
$(1, 0, 1)(0, 1, 1)_{12} + Prec_t + Prec_{t-1} + Prec_{t-2}$	254, 7108	239, 4656	0, 6309	7, 391%

**TABELA 1:** Modelos adequados aos dados e seus respectivos valores do Critério de Informação Bayesiano (BIC), Critério de Informação de Akaike (AIC), Raiz do Erro Quadrático Médio (REQM) e o Erro Absoluto Médio Percentual (MAPE)

Pela Tabela 1 o modelo que apresentou o menor valor de AIC e BIC foi o *SARIMA*  $(1, 0, 1)(0, 1, 1)_{12}$  mais o efeito da Umidade com uma defasagem (Modelo 1). Porém, o modelo *SARIMA*  $(1, 0, 1)(0, 1, 1)_{12}$  mais o efeito da Precipitação atual e com uma e duas defasagens (Modelo 2) apresentou o menores *REQM* e o *MAPE*.

Na Tabela 2 é mostrada a comparação entre os valores reais e os preditos, na escala original, da quantidade de notificações de dengue nos meses do ano de 2012 na cidade de Recife, indicando que os dois modelos fornecem um ajuste aceitável para a quantidade de notificações de dengue. Como base nos valores desta tabela, o Modelo 1 apresentou a *REQM* e o *MAPE* de 927 e 44%, respectivamente. Adicionalmente, o Modelo 2 apresentou a *REQM* e o *MAPE* de 895 e 37%, respectivamente. Com o intuito de utilizar o modelo com melhor capacidade de previsão, optou por utilizar o Modelo 2.

Valor	Mês												Total
	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez	
Observado	1642	2545	3127	1892	870	425	372	190	100	88	112	81	11444
Modelo 1	499	656	964	1058	828	562	559	292	163	103	90	56	5871
Modelo 2	572	731	1087	911	673	465	423	258	140	108	91	62	5521

**TABELA 2:** Valores para os meses do ano de 2012 e os respectivos valores previstos obtidos por meio do Modelo 1 e do Modelo 2

O Modelo 2 é expresso a seguir e as suas estimativas encontram-se na Tabela 3.

$$(1 - \phi_1 B)(1 - B^{12})Z_t = (1 - \theta B)(1 - \Theta B)\varepsilon_t + \beta_1 Prec_t + \beta_e Prec_{t-1} + \beta_2 Prec_{t-2} \quad (6)$$

Parâmetro	Estimativa	Erro padrão	Z	valor p
$\phi_1$	0,9023	0,0453	19,9100	< 0,01
$\theta_1$	0,2166	0,0822	2,6360	0,0084
$\Theta_1$	-0,9990	0,2112	-4,7350	< 0,01
$Prec_t$	0,0009	0,0005	1,9870	0,0469
$Prec_{t-1}$	0,0015	0,0006	2,5770	0,0100
$Prec_{t-2}$	0,0008	0,0005	1,6250	0,1042

**TABELA 3:** Estimativas, Erro Padrão e Valor p para os parâmetros do modelo 6

Na Tabela 3 estão as estimativas dos parâmetros para o Modelo 2 bem como o valor p. Muitas vezes em um modelo estatístico pode ocorrer o fato de uma variável não ser significativa, mas sua presença no modelo é de fundamental importância, uma vez que esta pode estar influenciando, significativamente, o comportamento de outras variáveis. Quando inserida a precipitação no modelo, foram testados modelos com diferentes defasagens para esta variável. Quando considerada a precipitação mais o seu efeito com um mês de defasagem não foram estatisticamente significativas. Porém, quando inserida a precipitação com até duas defasagens foi visto que a precipitação até com uma defasagem era significativa. Assim, notou-se que o valor p para a precipitação com a defasagem de dois meses é superior a 0,05, mesmo assim foi mantido no modelo devido esta variável influenciar, significativamente, nas estimativas das defasagens anteriores.

O próximo passo agora é a validação do modelo por meio de uma análise de resíduo. Na Figura 5 tem-se os gráficos da autocorrelação e autocorrelação parcial dos resíduos do modelo (6). Visualmente, os resíduos apresentam ruído branco. Para verificar esta hipótese será aplicado o teste de Ljung-Box.

O teste de Ljung-Box (Box e Pierce, 1970 [18]) para autocorrelação apresentou uma estatística do teste  $Q_{48} = 40,4014$  que é inferior ao valor de uma qui-quadrado com 45 graus de liberdade,  $\chi_{45}^2 = 61,6562$ . Assim, não se rejeita a hipótese, ao nível de 0,05 de significância, de que os resíduos são ruído branco.

Na Figura 6 têm-se os valores observados e previstos na escala logarítmica e os intervalos de confiança para a quantidade de notificação de dengue no período analisado, inclusive para o ano de 2012, o qual foi retirado do conjunto de dados para que posteriormente ao ajuste viesse a ser comparado com os valores previsto pelo modelo.

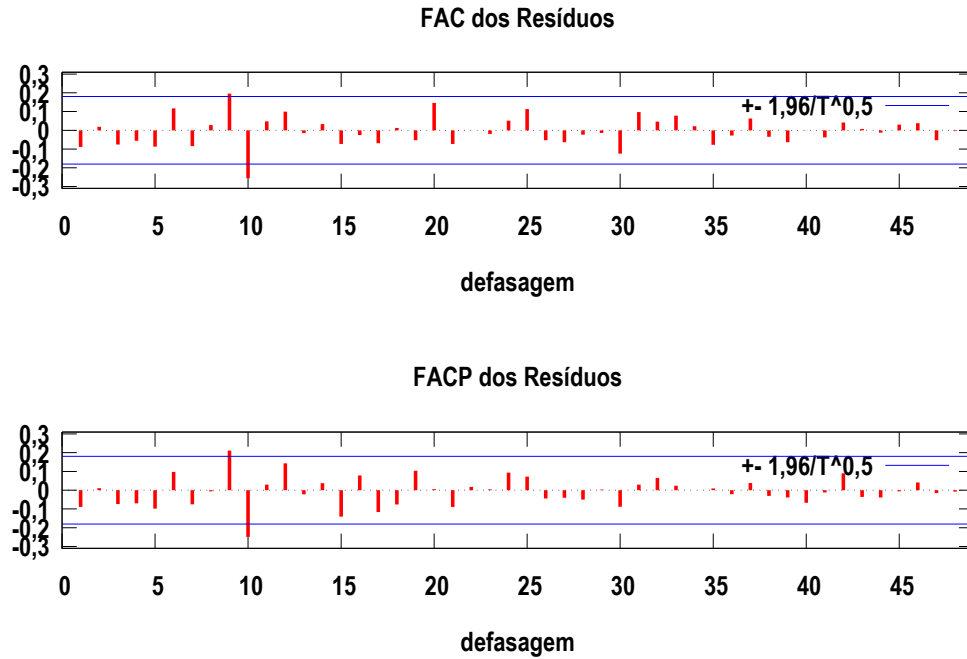
As barras indicam o intervalo à 95% de confiança para cada valor mensal previsto pelo modelo durante o ano de 2012. Nota-se que o ajuste foi satisfatório, conseguindo “capturar” o pico de incidência que ocorreu em Junho de 2010. Vale ressaltar que no ano de 2012 foi um ano em que ocorreu uma epidemia de dengue na cidade de Recife, PE.

## 4 DISCUSSÕES

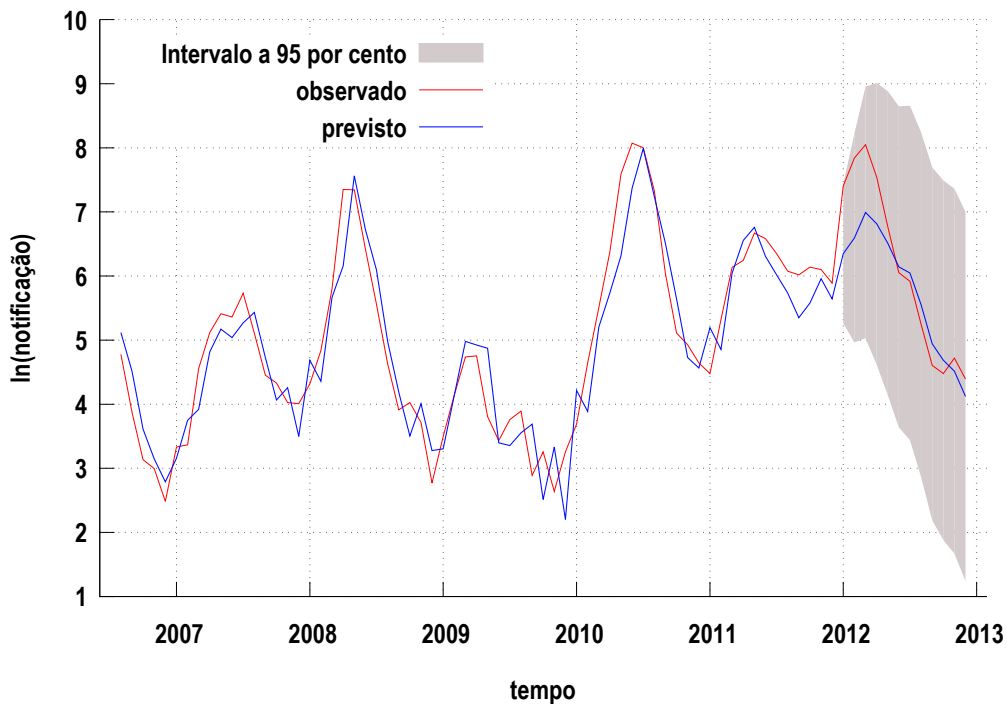
Na literatura diversos estudos têm indicado uma relação existente entre os casos de dengue e os fatores climáticos de uma região. Wu et al. (2007) [19] utilizaram uma análise de séries temporais para estudar os impactos que as mudanças climáticas tiveram sobre a incidência de dengue em Kaohsiung, no sul de Taiwan. No trabalho supracitado os autores verificaram que a temperatura média mensal estava negativamente associada com a incidência de dengue.

Neste artigo para os casos de dengue em Recife foi visto que as notificações de dengue estavam relacionadas com a umidade e a precipitação, sendo a precipitação ocasionando um maior impacto na modelagem dessas notificações. Hales et al. (2002) [20] realizaram





**FIGURA 5:** Função de autocorrelação (FAC) e autocorrelação parcial (FACP) dos resíduos do modelo ajustado



**FIGURA 6:** Valores observados e previstos para a série transformada das notificações de dengue em Recife, Pe

um estudo, no âmbito global, sobre a relação que as mudanças climáticas têm sobre os casos de dengue, concluindo que a umidade, precipitação e temperatura proporcionam condições favoráveis à reprodução dos mosquitos. Bicalho et al. (2014) [21] utilizaram da metodologia de séries temporais para estudar os efeitos que algumas variáveis climáticas têm sobre a quantidade de casos notificados de dengue na cidade de Lavras, Brasil. Estes

autores identificaram que séries de temperatura máxima e de precipitação influenciam a incidência de dengue na região em estudo.

Em termos de comparações, foram ajustados neste trabalho diversos modelos com presença e também ausência dos fatores climáticos como variáveis explicativas. Durante a modelagem foram detectados dois modelos no ajuste. Um modelo com a presença da variável umidade com uma defasagem de um mês e outro com a precipitação com um e dois meses de defasagens. O modelo quando considerada precipitação apresentou melhores resultados, considerando as estatísticas REQM e MAPE.

O estudo da relação entre os fatores climáticos de uma região e os casos de notificações de dengue permite obter resultados que auxiliem nas detecções de futuras epidemias, fazendo com que os setores governamentais responsáveis pela saúde pública possam realizar estratégias e planos para o combate e prevenção das doenças causadas pelo mosquito (Bicalho et al., 2014 [21]).

## 5 CONSIDERAÇÕES FINAIS

A proposta desse trabalho foi utilizar uma metodologia de séries temporais, para estudar o comportamento e realizar previsões para as notificações de dengue na cidade de Recife. Por meio da análise de séries temporais foi verificada a relação que as mudanças climáticas têm sobre os casos de dengue. Quando se utilizou de covariáveis no modelo, pode-se perceber um considerável ganho nas previsões, levando a evidências, de que, em casos de dengue quando previstos com variáveis climáticas, há uma precisão significativas nas estimativas. Assim, diante dos resultados obtidos no presente trabalho pode-se observar que o modelo  $SARIMA(1, 0, 1)(0, 1, 1)_{12}$  mais o efeito atual, com uma e com duas defasagens da precipitação foi o que apresentou melhor capacidade de previsão para a quantidade de notificações. As previsões da quantidade de notificações de dengue possibilitam aos órgãos governamentais que possam tomar medidas sobre a incidência de dengue reduzindo a morbidade e mortalidade causada por esta doença.

## REFERÊNCIAS

- [1] J. M. Brunkard, E. Cifuentes, and S. J. Rothenberg, "Assessing the roles of temperature, precipitation, and ENSO in dengue re-emergence on the Texas-Mexico border region," *Salud pública de México*, vol. 50, pp. 227–234, 2008.
- [2] P. M. Luz, B. V. Mendes, C. T. Codeço, C. J. Struchiner, and A. P. Galvani, "Time series analysis of dengue incidence in Rio de Janeiro, Brazil," *The American journal of tropical medicine and hygiene*, vol. 79, pp. 933–939, 2008.
- [3] S. Singhi, N. Kissoon, and A. Bansal, "Dengue and dengue hemorrhagic fever: management issues in an intensive care unit," *Jornal de pediatria*, vol. 83, pp. S22–S35, 2007.
- [4] M. R. Donalísio and C. M. Glasser, "Entomological surveillance and control of dengue fever vectors," *Revista Brasileira de Epidemiologia*, vol. 5, pp. 259–279, 2002.
- [5] S. S. d. Souza, I. G. d. Silva, and H. H. G. d. Silva, "Association between dengue incidence, rainfall and larval density of *Aedes Aegypti*, in the State of Goiás," *Revista da sociedade brasileira de medicina tropical*, vol. 43, pp. 152–155, 2010.
- [6] Y. L. Hii, H. Zhu, N. Ng, L. C. Ng, and J. Rocklöv, "Forecast of dengue incidence using temperature and rainfall," *PLoS neglected tropical diseases*, vol. 6, pp. 1–9, 2012.
- [7] S. Wongkoon, M. Jaroensutasinee, and K. Jaroensutasinee, "Development of temporal modeling for prediction of dengue infection in Northeastern Thailand," *Asian Pacific journal of tropical medicine*, vol. 5, pp. 249–252, 2012.

- [8] J. Bitoun, L. Miranda, M. d. A. Souza, and M. Lyra, “Região Metropolitana do Recife no contexto de Pernambuco no Censo 2010,” tech. rep., 2012.
- [9] G. P. Zhang, “Time series forecasting using a hybrid ARIMA and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [10] G. Saz, “The efficacy of SARIMA models for forecasting inflation rates in developing countries: the case for Turkey,” *International Research Journal of Finance and Economics*, vol. 62, pp. 111–142, 2011.
- [11] R. Weron, *Modeling and forecasting electricity loads and prices: A statistical approach*. John Wiley & Sons, 2007.
- [12] G. Baiocchi and W. Distaso, “Gretl: Econometric software for the gnu generation,” *Journal of applied econometrics*, vol. 18, pp. 105–110, 2003.
- [13] D. Montenegro, H. R. Lacerda, T. M. Lira, D. S. C. d. Oliveira, A. A. F. d. Lima, M. J. B. Guimarães, and P. G. d. Vasconcelos, “Aspectos clínicos e epidemiológicos da epidemia de dengue no Recife, PE, em 2002,” *Rev Soc Bras Med Trop*, vol. 39, no. 1, pp. 9–13, 2006.
- [14] R. C. Lima, E. B. M. Moreira, and R. S. Nóbrega, “Climate Influence on Dengue Epidemic in Recife City by Geographic Information System,” *Revista Brasileira de Geografia Física*, vol. 9, no. 2, pp. 384–398, 2016.
- [15] P. Sprent and N. C. Smeeton, *Applied nonparametric statistical methods*. CRC Press, 2016.
- [16] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*. Springer Science & Business Media, 1991.
- [17] W.-C. Wang, K.-W. Chau, C.-T. Cheng, and L. Qiu, “A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series,” *Journal of hydrology*, vol. 374, no. 3-4, pp. 294–306, 2009.
- [18] G. E. Box and D. A. Pierce, “Distribution of residual autocorrelations in autoregressive-integrated moving average time series models,” *Journal of the American Statistical Association*, vol. 65, no. 332, pp. 1509–1526, 1970.
- [19] P.-C. Wu, H.-R. Guo, S.-C. Lung, C.-Y. Lin, and H.-J. Su, “Weather as an effective predictor for occurrence of dengue fever in Taiwan,” *Acta tropica*, vol. 103, no. 1, pp. 50–57, 2007.
- [20] S. Hales, N. De Wet, J. Maindonald, and A. Woodward, “Potential effect of population and climate changes on global distribution of dengue fever: an empirical model,” *The Lancet*, vol. 360, no. 9336, pp. 830–834, 2002.
- [21] C. C. Bicalho, T. Safadi, and I. Charret, “The influence of climatic factors on dengue epidemics in the cities Cuiaba (Mato Grosso State) and Lavras (Minas Gerais State), Brazil, using statistical methods,” *Rev Bras Biom Sao Paulo*, vol. 32, no. 2, pp. 308–22, 2014.