

Modelagem da incidência do dengue na Paraíba, Brasil, por modelos de defasagem distribuída

Modeling of dengue incidence in Paraíba State, Brazil, using distributed lag models

Izabel Cristina Alcantara de Souza ¹
Rodrigo Pinheiro de Toledo Vianna ²
Ronei Marcos de Moraes ¹

Abstract

The literature provides several statistical models to explain dengue incidence. However, there is disagreement as to the real contribution of models based on climate factors or models based on relative variables to combat the vector, because their variability fails to provide a satisfactory statistical explanation for the behavior of dengue incidence. Distributed lag models (DLM) posit that a variable response Y will be explained by the presence of a variable X at the same moment in time t and in the previous moments ($t-1, \dots$). The current study presents the use of DLM in dengue modeling. Among the several models tested, two displayed apparently interesting results. A DLM using rainfall was not validated statistically. Another model using the number of counties (municipalities) with dengue presented valid and satisfactory results. This DLM was also feasible from the perspective of State Health Departments, allowing the development of a model with interesting statistical results and good accuracy, using a single data source.

Statistical Models; Dengue; Incidence

¹ Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, João Pessoa, Brasil.

² Centro de Ciências da Saúde, Universidade Federal da Paraíba, João Pessoa, Brasil.

Correspondência

R. M. Moraes
Departamento de Estatística,
Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba.
Cidade Universitária s/n,
João Pessoa, PB
58051-900, Brasil.
ronei@de.ufpb.br

Introdução

Tem crescido nos últimos anos, o investimento em pesquisas com a finalidade de explicar o comportamento epidemiológico do dengue ¹. Por ser uma doença transmitida por vetores, cujo principal é o *Aedes aegypti*, a busca de relações entre fatores ambientais e a propagação da doença tem sido investigada com a finalidade de instituir mecanismos de controle e prevenção de epidemias. Tais mecanismos são baseados no controle biológico do vetor, evitando a proliferação do mosquito baseando-se no conhecimento de fatores ambientais, tais como, temperatura, umidade e chuvas ².

Estudos realizados no Brasil e em outras partes do mundo apontam para uma relação entre fatores climáticos e a proliferação do dengue ^{3,4}. Hales et al. ⁵ afirmam a existência de uma relação entre fatores climáticos e incidência mensal do dengue em 22 ilhas do sul do Pacífico. Por outro lado, Hay et al. ⁶ analisaram a série temporal mensal de febre hemorrágica do dengue em Bangucoque, correlacionando-a com as variáveis temperatura e precipitação mensal, e observaram que o ciclo periódico anual destas variáveis não possui similaridade com a periodicidade anual do dengue.

O clima afeta o meio ambiente onde o *Ae. aegypti* está presente, mas não é o único fator que caracteriza a sua existência ². Os criadouros deste vetor podem ser reproduzidos artificialmente

devido a problemas urbanos, tais como: armazenamento de água em recipientes inadequados, falta de fiscalização apropriada em estabelecimentos e residências que armazenam pneus, vasos de plantas, assim como nos depósitos de lixo^{1,3}. Esta realidade possibilita a existência do dengue, como doença endêmica, nas áreas mais áridas do país, onde os períodos de seca podem se prolongar por um longo período.

No Estado da Paraíba, localizado no Nordeste do Brasil, o clima é marcado pela presença de apenas duas estações climáticas: a chuvosa e a seca, de modo que a pluviometria é o principal fator de modificação climática da região ao longo do ano⁷. O padrão espaço-temporal da temperatura média é caracterizada pela pequena variação anual, quase sempre inferior a 5°C⁸. Essa homogeneidade térmica se contrasta com o alto grau de heterogeneidade espacial e temporal do regime pluviométrico, no qual a distância do litoral define um gradiente de umidade, sendo as áreas mais afastadas daquelas mais secas^{8,9}. Nesse estado, o período chuvoso geralmente é curto, ocorrendo no verão nas regiões mais áridas e no inverno nas áreas mais úmidas⁸. Na estação do outono as chuvas são pouco frequentes, mas no entanto, este período é o de maior incidência do dengue no estado⁹. Esse fato não é o mesmo observado em outros estados do Brasil onde a sazonalidade do dengue coincide com a da chuva, segundo o Ministério da Saúde¹⁰.

A obtenção de dados climáticos limitou-se à variável precipitação pluviométrica, uma vez que apenas esta variável é medida em todo o Estado da Paraíba¹¹. As variáveis relativas ao controle do dengue, armazenadas pelo banco de dados do FAD (Febre Amarela e Dengue), tais como o índice de Breteau, predial, número de imóveis trabalhados, agentes, consumo de larvicida, aduicida etc., não foram utilizados no processo de modelagem¹². As justificativas para tal exclusão são relativas à qualidade desses dados, sendo encontrados os seguintes problemas: (1) longos períodos sem coletas de informações para alguns municípios paraibanos; (2) ausência de informação nos meses de novembro e dezembro de 1999 em 220 dos 223 municípios paraibanos; (3) curto período de informação, de janeiro de 1999 (início do sistema FAD na Paraíba) até dezembro de 2003. Nos anos de 2004 ocorreram problemas no armazenamento das informações da base de dados do FAD em grande parte dos municípios paraibanos.

Por ser uma doença de notificação compulsória, todos os casos de dengue devem ser notificados ao Sistema de Informação de Agravos de Notificação (SINAN)¹³. Por conta disso, o SINAN reúne informações sobre o dengue dos órgãos

de saúde pública e privada, sendo, portanto, a principal fonte de informação sobre o dengue no Brasil. No entanto, a base do SINAN apresenta problemas de subnotificação do dengue, já que grande parte da população costuma se automedicar quando acometida pelos sintomas característicos do dengue, buscando orientação médica somente nos casos mais graves da doença.

Este estudo tem o objetivo de estabelecer modelos para incidência do dengue com base nas variáveis explicativas: precipitação pluviométrica e número de municípios com dengue. O período de estudo compreende os meses de janeiro de 1998 a junho de 2005, com base nos dados oficiais de notificação do SINAN, e refere-se ao Estado da Paraíba.

Método

Os Modelos de Defasagem Distribuída (MDD) partem do princípio que a dependência da variável resposta Y em relação a variável explicativa X não é instantânea. Acredita-se que a reação de Y_t no instante t não é explicada somente pelo ocorrido em X_t no tempo t , mas também pelo histórico de observação de X nos instantes $t-1$, $t-2$,... Supondo que Y_t é afetada por X com uma defasagem de tempo finito, isto é, a influência de X sobre Y_t se esgota após k períodos de tempo anteriores a t , pode-se definir o modelo de defasagem finita de k períodos de tempo por (1)¹⁴.

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k} + \varepsilon_t \quad (1)$$

O modelo MDD é amplamente utilizado em estudos econométricos empíricos, pois "permite tornar a teoria econômica estática em dinâmica ao levar em conta explicitamente o papel do tempo"¹⁴ (p. 630). Na área de saúde, considerando Y_t como sendo a incidência de uma determinada doença no instante t , se existir razões para crer que os impactos da variável explicativa X sobre Y_t sejam prolongados por um período de tempo k , o modelo (1) pode ser adaptado para estudos epidemiológicos. Para estimação do MDD, Almon¹⁵ propôs estabelecer restrições aos coeficientes de (1) com base no teorema de Weierstrass, o qual afirma que qualquer função contínua num intervalo fechado e finito pode ser aproximada por um polinômio de grau $g < k$ em d , onde d é o tamanho da defasagem $d = 1, 2, \dots, k$. Seja $f(d) = \theta_0 + \theta_1 d + \theta_2 d^2 + \dots + \theta_g d^g$ uma função polinomial em d de grau g , tal que $\beta_d = f(d)$, e fazendo $w_{it} = \sum_d d^i x_{t-d}$, obtém-se um novo modelo, denominado *modelo construído*, dado por (2)¹⁶.

$$y_t = \alpha + \theta_0 w_{0t} + \theta_1 w_{1t} + \theta_2 w_{2t} + \dots + \theta_g w_{gt} + \varepsilon_t \quad (2)$$

Estima-se o modelo construído, dado pela equação (2), usando-se o método dos mínimos quadrados, em que o erro ε_t é admitido ser uma

variável aleatória não correlacionada e normalmente distribuída com média zero e variância constante igual a σ^2 , isto é, $\varepsilon_t \sim N(0, \sigma^2)$ ¹⁷. A autocorrelação refere-se à existência de correlação não nula entre os erros. Quando o modelo de regressão é multivariado, costuma-se adotar a notação matricial para modelo (2), expressa por (3), e utiliza-se o método de mínimos quadrados para estimar o vetor $\Theta = [\alpha, \theta_0, \theta_1, \dots, \theta_g]'$ que define o modelo.

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & w_{01} & w_{11} & \dots & w_{g1} \\ 1 & w_{02} & w_{12} & \dots & w_{g2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w_{0n} & w_{1n} & \dots & w_{gn} \end{bmatrix} \cdot \begin{bmatrix} \alpha \\ \theta_0 \\ \vdots \\ \theta_g \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$Y = W\Theta + \varepsilon$ (3)

Seja Φ o vetor estimado de Θ , o método de mínimos quadrados permite estimar Θ selecionando valores para Φ que minimize a soma de quadrados do erro $\varepsilon = Y - W\Theta$, ou seja, minimizar $SQE = (Y - W\Theta)'(Y - W\Theta)$, em que o símbolo $(Y - W\Theta)'$ representa a matriz transposta de $(Y - W\Theta)$. A solução obtida ao minimizar SQE é dada em (4).

$$\Phi = (W'W)^{-1}W'Y$$
 (4)

Baseando-se nas estimativas dos coeficientes do modelo construído obtém-se os coeficientes do modelo MDD, por meio da relação (5):

$$\begin{aligned} \alpha &= \alpha \\ \beta_0 &= \theta_0 \\ \beta_1 &= \theta_0 + \theta_1 + \theta_2 + \dots + \theta_g \\ \beta_2 &= \theta_0 + 2\theta_1 + 2^2\theta_2 + \dots + 2^g\theta_g \\ &\vdots \\ \beta_g &= \theta_0 + d\theta_1 + d^2\theta_2 + \dots + d^g\theta_g \end{aligned}$$
 (5)

Três observações são importantes quanto ao modelo MDD, no qual a primeira refere-se ao fato de que o modelo (2) pode apresentar problemas de multicolinearidade, isto é, exata relação linear entre as variáveis independentes. No entanto, se for por possível estimar os seus coeficientes, o problema se resume a uma maior dispersão dos coeficientes, o que não invalida o modelo ¹⁴. O segundo problema consiste na autocorrelação dos resíduos, ou seja, a existência de uma correlação não nula entre os resíduos ε_t e ε_{t+h} tomados em um intervalo de tempo h , o que conduz à subestimação dos erros padrões. Para verificar a existência de autocorrelação pode-se utilizar o teste de Durbin Watson (no caso de modelos com intercepto) ou o teste de aleatoriedade de Geary ¹⁴. Finalmente, o terceiro problema consiste em determinar o tamanho da defasagem k e o grau do polinômio g . Fixado o tamanho da defasagem k para obter o grau do polinômio, caso este não seja conhecido *a priori*, pode-se estimar o modelo (2) com um polinômio de grau $g = k - 1$ e ir diminuindo este grau até encontrar um modelo que apresente um

bom ajuste. Nessa fase de escolha do grau g pode-se utilizar testes de comparação de modelos e o método *stepwise backward* ¹⁸, que consiste na obtenção das equações com polinômios de grau $0, 1, 2, \dots, g$, escolhendo iterativamente a equação que apresenta o melhor ajuste. Para estabelecer o tamanho da defasagem, caso este também não seja conhecido, utiliza-se o critério de Schwarz (CS) ¹⁹, que consiste em determinar um valor para k que minimize a função (6). Na prática, estima-se o modelo (2) para vários valores de defasagem k e escolhe-se o que apresentar o menor valor para CS.

$$CS = k \ln(n) + n \ln(SQRes / n)$$
 (6)

Fonte dos dados

Utilizou-se três variáveis oriundas de fontes distintas, são elas: (1) número de casos notificados do dengue na Paraíba por mês, período de janeiro de 1998 a junho de 2005 (fonte: dados do SINAN obtidos na Secretaria de Estado da Saúde do Estado da Paraíba – SES/PB ¹²); (2) população residente estimada da Paraíba por mês, período de janeiro de 1998 a junho de 2005 – foram realizadas projeções mensais para o dia 15 de cada mês com base nas estimativas anuais da população divulgadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) e disponibilizado pelo Departamento de Informática do SUS (DATASUS) ²⁰; (3) precipitação média na Paraíba por mês, período de janeiro de 1998 a junho de 2005, obtida na Agência Executiva de Gestão das Águas do Estado da Paraíba ¹¹.

Calcularam-se os coeficientes de incidência do dengue por 10 mil habitantes, variável denotada por V_p para o período de janeiro de 1998 a junho de 2005, segundo a metodologia de taxas de incidências de doenças, que pode ser encontrada em Laurenti et al. ²¹. Em seguida obteve-se a contagem de municípios que apresentaram casos de dengue em cada mês do período de estudo. Dessa forma, as variáveis trabalhadas neste estudo são: incidência do dengue, número de municípios com dengue e precipitação pluviométrica, sendo todas referentes ao período de janeiro de 1998 a junho de 2005.

Resultados

O estudo do modelo explicativo a partir de variáveis climáticas foi limitado a apenas a precipitação pluviométrica, visto que é a única coletada em todo o estado. Porém, é a variável mais importante e o principal fator de modificação climática da Região Nordeste como um

todo ^{7,8}. No caso do modelo explicativo a partir de variáveis relativas ao combate do vetor (como: número de depósitos e imóveis infectados, número de agentes trabalhando, tempo de trabalho, gastos com larvicidas e adulticidas etc.), vários foram os problemas encontrados: em muitas dessas variáveis houve falta sistemática de informação e diversas vezes, quando ela estava presente, a falta de confiabilidade nos seus registros, apresentando variações inesperadas e sem uma explicação plausível para isso. Dessa maneira, a modelagem do dengue baseada nessas variáveis teoricamente relevantes não puderam ser utilizadas na explicação da incidência do dengue na Paraíba. Modelos testados a partir dessas variáveis não resultaram em modelos com explicação estatística significativa. Todos foram ajustados usando-se o pacote estatístico R (R Development Core Team; <http://www.r-project.org>) e os gráficos foram confeccionados usando-se a planilha eletrônica BrOffice (BrOffice.org; <http://www.openoffice.org.br>).

Assim, neste estudo, foram testados dois modelos: o primeiro é baseado em variáveis climáticas (precipitação pluviométrica) e o segundo em uma variável da base de dados do SINAN (número de municípios com dengue). O primeiro modelo apresentou um bom coeficiente de determinação, mas a dispersão dos resíduos mostrou que este modelo não responde à incidência do dengue de forma satisfatória. O segundo apresentou melhores resultados na modelagem da incidência do dengue e possui a vantagem de não necessitar de outra fonte externa ao SINAN.

Os modelos obtidos têm como variável resposta a incidência do dengue V_t que deve atender às hipóteses clássicas do modelo de regressão linear para que se possa aplicar a metodologia de MDD. Para isso, testou-se a normalidade da série de incidência do dengue e como a distribuição de probabilidade de V_t não é normal, de acordo com o teste de Lilliefors (Kolmogorov-Smirnov), com $D = 0,2696$ e valor $p < 2,0.10^{-16}$, aplicou-se a transformação raiz quinta à variável V_t , isto é, obteve-se $Y_t = V_t^{0,2}$ com distribuição de probabilidade normal ($D = 0,0672$, com valor $p = 0,4042$), atendendo assim à hipótese de normalidade do modelo de regressão.

Modelo baseado na precipitação

Foram obtidos os modelos com k defasagens de 1 a 15 e em cada modelo estimado utilizou-se o método *stepwise* para decidir qual o grau g do polinômio para todo $g < k$. Para decidir qual é o tamanho adequado da defasagem, calculou-se o CS escolhendo-se o modelo com o menor CS. O modelo com defasagem $k = 1$ foi o que

apresentou menor CS, igual a -207,84. O modelo construído estimado foi $y_t = 0,214w_{0t}$ o que corresponde ao MDD igual a $y_t = 0,214x_t + 0,214x_{t-1}$. Nesses modelos a constante não foi significativa e o coeficiente foi estimado por um polinômio de grau zero, isto é, $g = 0$ e $\theta_0 = \beta_0 = \beta$. Sendo assim, ambos os modelos estimados (construído e MDD) relacionam a raiz quinta da incidência do dengue com a soma das precipitações médias nos tempos t e $t-1$.

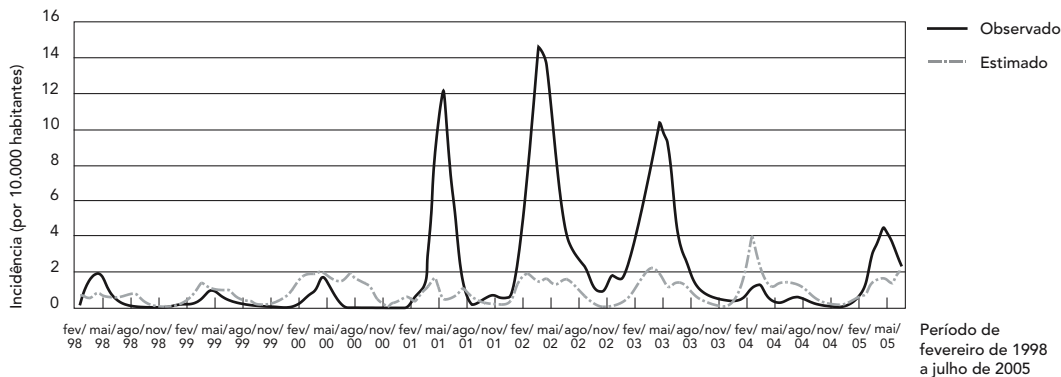
O modelo construído $y_t = 0,214w_{0t}$ apresenta um bom ajuste, $R^2 = 0,9077$, mas no entanto, subestimou a incidência do dengue. A análise de resíduos, o teste de Kolmogorov-Smirnov e o teste de Geary foram utilizados para verificar qualidade de ajuste do modelo. A hipótese de normalidade dos resíduos foi rejeitada (valor $p = 4,3.10^{-7}$) no nível de 1% de significância. Além de não atender à hipótese de normalidade, os resíduos são também autocorrelacionados. Como o modelo construído não tem a constante, não é recomendado utilizar o teste de Durbin Watson, utilizando-se em seu lugar o teste de aleatoriedade de Geary para os resíduos. Com base no teste de Geary, a estatística z calculada da normal padrão é 2,232 que corresponde a um valor $p = 0,0128$. Assim, rejeitaram-se ambas as hipóteses: de normalidade e aleatoriedade dos resíduos a 1% de insignificância. Esses resultados demonstram que o modelo construído $y_t = 0,214w_{0t}$ e conseqüentemente o MDD $y_t = 0,214x_t + 0,214x_{t-1}$ não atende às hipóteses do modelo de regressão e por estes motivos não é válido em termos estatísticos.

Elevando-se à quinta potência os dados ajustados pelo modelo MDD, isto é, realizando-se a transformação inversa $z_t = y_t^5$, obteve-se a variável z_t na mesma unidade da incidência do dengue por 10 mil habitantes. Pode-se visualizar a qualidade do ajuste no gráfico da Figura 1. Vê-se claramente que o modelo estimado subestimou a incidência do dengue nos anos de maior incidência e superestimou nos de baixa incidência, apresentando um comportamento diferente da série observada.

O comportamento anual da incidência do dengue é apresentar uma curva de incidência no primeiro semestre do ano, mas o modelo ajustado obteve em cada ano mais de uma curva. Esse comportamento assemelha-se ao da precipitação pluviométrica, que na Paraíba caracteriza-se por dois períodos de chuva: um no verão, nas áreas mais úmidas, e um no inverno, nas áreas mais áridas. Graficamente, a Figura 1 mostra o que já foi constatado usando-se o teste de Geary: o modelo não atende às hipóteses estatísticas de aleatoriedade, necessárias para validar um modelo de regressão pretendido.

Figura 1

Série do modelo ajustado B (estimada) e da raiz quinta da incidência do dengue (observada).



Modelo baseado no número de municípios com dengue

No caso do modelo da incidência do dengue em função do número de municípios com a doença o tamanho da defasagem, $k = 5$, é o que apresentou menor CS, igual a -364,73. O modelo construído estimado foi $y_t = 0,502w_{0t} - 0,292w_{1t} + 0,033w_{2t}$, o que corresponde ao MDD igual a $y_t = 0,502x_t - 0,243x_{t-1} + 0,051x_{t-2} - 0,074x_{t-3} - 0,132x_{t-4} - 0,124x_{t-5}$. Esses modelos não possuem constantes e os coeficientes foram estimados por um polinômio de grau dois. A partir do modelo MDD pode-se verificar que os coeficientes decrescem até a quarta defasagem, voltando a crescer na quinta. Esses coeficientes informam o grau de importância de cada defasagem da variável explicativa no modelo.

O modelo construído $y_t = 0,502w_{0t} - 0,292w_{1t} + 0,033w_{2t}$, apresenta um bom ajuste, $R^2 = 0,9894$, e os resíduos atendem às hipóteses da regressão linear. Novamente, a análise de resíduos, o teste de Kolmogorov-Smirnov e o teste de Geary foram utilizados para verificar a qualidade de ajuste do novo modelo. Realizando-se o teste de Kolmogorov-Smirnov para verificar a hipótese H_0 : os resíduos têm distribuição normal com média zero, obteve-se a estatística $D = 0,1215$ o que corresponde a um valor $p = 0,1499$. Dessa maneira, não rejeita-se a hipótese H_0 no nível de 1% de significância e conclui-se que os resíduos são normais com média zero. A partir do teste de aleatoriedade de Geary para os resíduos obteve-se a estatística calculada igual a -0,87465 que corresponde a um valor $p = 0,19088$. Sendo assim, não se rejeita

a hipótese de aleatoriedade dos resíduos a 1% de significância. Sob o ponto de vista estatístico, esse modelo obedece às restrições de normalidade das variáveis e normalidade e aleatoriedade dos resíduos, portanto, é um modelo estatístico válido para o estudo do dengue. Além disso, sob o ponto de vista das Secretarias Estaduais de Saúde, é um modelo viável que permite com uma única fonte de informação estabelecer um modelo com resultados estatísticos interessantes e de boa acurácia.

As séries observadas e a quinta potência dos dados ajustados, série estimada, podem ser visualizadas no gráfico da Figura 2. Pode-se observar que a série estimada está bem próxima da observada, apresentando apenas casos de subestimação nos anos de 2002 e 2003.

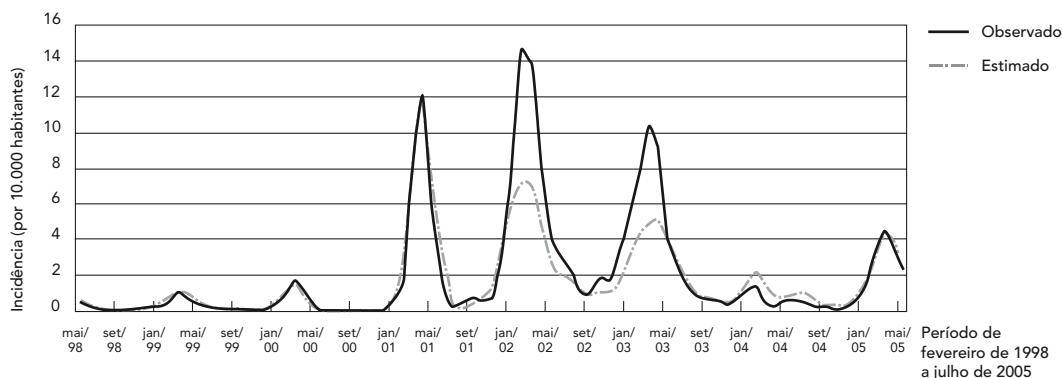
O comportamento do modelo ajustado descreve a característica de sazonalidade da incidência do dengue com valores máximos oscilando entre os meses de março a maio. Nota-se que o início das curvas de crescimento anual se dá com cinco meses de antecedência, o que corresponde à defasagem de cinco meses descrita pelo modelo ($k = 5$).

Discussão

Na literatura, geralmente obtém-se modelos do dengue baseados em fatores climáticos, em que as características geográficas e a qualidade de tais dados possibilitam tal modelagem^{3,4,5}. Já nos modelos que se baseiam nas variáveis relativas ao vetor como resposta à incidência do den-

Figura 2

Séries de incidência do dengue (observada) e a série ajustada (estimada) para o modelo explicativo de modelos de defasagem distribuída (MDD) com defasagem de cinco meses.



gue, a dificuldade geralmente está na qualidade com que estes dados são produzidos.

Os dados oficiais do Estado da Paraíba apresentam limitações quanto às informações climáticas e às relativas ao vetor do dengue. Nesse estado, apenas a precipitação pluviométrica, principal fator de modificação climática da região, é coletada. As informações relativas ao combate do vetor (tais como: número de depósitos e imóveis infectados, número de agentes trabalhando, tempo de trabalho, gastos com larvicidas e adulcidades etc.) apresentam problemas de falta de informação e quando ela está presente, a falta de sua confiabilidade. Dessa maneira, modelos do dengue baseados nessas variáveis teoricamente relevantes não se mostraram significativos na explicação da incidência da doença no caso do Estado da Paraíba, onde há particularidades sobre as características climáticas e problemas de coletas de informações.

Este estudo apresenta uma proposta de utilização do MDD na modelagem da incidência do dengue, utilizando como variáveis explicativas a precipitação pluviométrica, num primeiro modelo, e o número de municípios com dengue em outro. O modelo MDD baseado na precipitação média como variável explicativa não atendeu aos critérios de validade estatística, pois as hipóteses de normalidade e ausência de autocorrelação dos resíduos foram rejeitadas. O modelo também mostrou-se inadequado pois superestima a incidência de dengue e portanto, os dados pluviométricos não confirmaram ser uma boa variável explicativa para a ocorrência do dengue na Paraíba.

Já o modelo da incidência do dengue pelo número de municípios com dengue apresentou resultados satisfatórios. O MDD apresentou defasagem de cinco meses e os comportamentos dos coeficientes β_p foram descritos por uma função de segundo grau. Mas, qual será o sentido da defasagem de cinco meses? No gráfico da Figura 2 pode-se notar que em cada ano há uma curva de incidência do dengue, na qual os picos oscilam entre os meses de março a maio. Esses meses encontram-se na estação do outono na Paraíba e o início dessas curvas é nos meses do verão. Por exemplo, o verão tem início no dia 21 de dezembro e cinco meses depois tem-se o mês de maio. Comparando esses dois meses, dezembro e maio, na Figura 2, pode-se notar que em dezembro as incidências são pequenas, após o qual a incidência inicia uma curva ascendente. No mês de maio há uma reversão no sentido dessa curva, após o seu ponto máximo. As exceções para tal fato ocorrem apenas nos anos de 2001 e 2004. No ano de 2001 a tendência é semelhante, mas tem um atraso de um mês, como pode ser observado na Figura 2. O ano de 2004 foi atípico em termos climáticos na Paraíba, com índices de precipitação acima da média no verão.

Apesar das exceções, geralmente pode-se verificar que a partir do início do verão, cinco meses depois há uma mudança na tendência de incidência do dengue. Como foi observado anteriormente, o comportamento dos coeficientes β_p no modelo MDD $y_t = 0,502x_t - 0,243x_{t-1} + 0,051x_{t-2} - 0,074x_{t-3} - 0,132x_{t-4} - 0,124x_{t-5}$ é decrescer até a quarta defasagem e na quinta defasagem há uma

reversão na tendência, aumentando o valor dos coeficientes. Diante do exposto, pode-se afirmar que a defasagem cinco corresponde à duração do período necessário para haver mudança na tendência da curva anual da incidência do dengue, a partir do início do verão.

O objetivo principal do trabalho, montar um modelo explicativo que descrevesse a incidência do dengue no Estado da Paraíba, foi atingido e a ocorrência de dengue nos municípios demonstra que o volume da incidência não se dá apenas em poucos municípios com registro

de casos, mas com o aumento do número de municípios com casos registrados, ou seja, o aumento da distribuição espacial da ocorrência foi suficiente para explicar a incidência estadual, com validade estatística e elevada precisão ($R^2 = 0,9894$).

Sob o ponto de vista das Secretarias Estaduais de Saúde, o modelo obtido é sucinto e viável, pois permite com poucas informações estabelecer resultados estatísticos importantes que podem ser utilizados pelos Serviços de Vigilância Epidemiológica das Secretarias Estaduais de Saúde.

Resumo

Existem vários modelos estatísticos na literatura para explicar a incidência do dengue. Porém, há divergências a respeito da real validade de modelos baseados em fatores climáticos e de modelos baseados em variáveis relativas ao combate ao vetor, pois a variabilidade apresentada por estas variáveis não são suficientes para explicar satisfatoriamente o comportamento estatístico da incidência do dengue. Os modelos de defasagem distribuída (MDD) supõem que a variável resposta Y será explicada pela presença de uma variável X no mesmo instante de tempo t e também pelos instantes anteriores (t-1, ...). Este estudo apresenta uma proposta de utilização do MDD na modelagem do dengue. Dentre os vários modelos testados, dois apresentaram resultados aparentemente interessantes. Um modelo MDD usando-se pluviometria não foi validado sob o ponto de vista estatístico. Um outro usando-se o número de municípios com dengue apresentou resultados estatísticos válidos e satisfatórios. Além disso, sob o ponto de vista das Secretarias Estaduais de Saúde, é um modelo viável que permite com uma única fonte de informação estabelecer um modelo com resultados estatísticos interessantes e de boa acurácia.

Modelos Estatísticos; Dengue; Incidência

Colaboradores

I. C. A. Souza participou da organização dos bancos de dados, construção dos modelos, análise estatística, discussão dos resultados sob o ponto de vista estatístico e na elaboração do texto. R. P. T. Vianna contribuiu nas discussões da análise dos modelos sob o ponto de vista epidemiológico e na elaboração do texto final submetido. R. M. Moraes orientou a construção dos modelos e colaborou na análise estatística, discussão dos resultados sob o ponto de vista estatístico e a participação na elaboração do texto.

Agradecimentos

Este trabalho tem suporte do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio de bolsas de pesquisa, e da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), através de apoio financeiro, pelo Projeto CNPq-501584/2003-0. Izabel Cristina Alcantara de Souza é bolsista DTI-7H/CNPq, vinculada a este projeto.

Referência

1. Corrêa PRL, Franca E, Bogutchi TE. Infestação pelo *Aedes aegypti* e ocorrência da dengue em Belo Horizonte. *Rev Saúde Pública* 2005; 39:33-40.
2. World Health Organization. Using Climate to Predict Infectious Disease Outbreaks: a review. Geneva: World Health Organization; 2004.
3. Rosa-Freitas MG, Tsouris P, Sibajev A, Weimann ETS, Marques AU, Ferreira RL, et al. Exploratory temporal and spatial distribution analysis of dengue notifications in Boa Vista, Roraima, Brazilian Amazon, 1999-2001. *Dengue Bull* 2003; 27:63-80.
4. Gubler DJ, Reiter P, Ebi KL, Yap W, Nasci R, Patz JA. Climate variability and change in the United States: potential impacts on vector and rodent-borne diseases. *Environ Health Perspect* 2001; 109 Suppl 2:223-33.
5. Hales S, Weinstein P, Souares Y, Woodward A. El Niño and the dynamics of vectorborne disease transmission. *Environ Health Perspect* 1999; 107:99-102.
6. Hay SI, Myers ME, Burke DS, Vaughn DW, Endy T, Ananda N, et al. Etiology of interepidemic periods of mosquito-borne disease. *Proc Natl Acad Sci USA* 2000; 97:9335-9.
7. Nimer E. Pluviometria e recursos hídricos dos estados de Pernambuco e Paraíba. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística; 1979.
8. Rodrigues JL. Atlas escolar da Paraíba. João Pessoa: Editora Grafset; 2000.
9. Furtado PCH, Souza ICA, Bezerra IS, Moraes RM. Explorando modelos estatísticos de relações entre variáveis climáticas, vetores alados e incidência do dengue. In: *Proceedings of Environmental and Health World Congress 2006*. Santos: Conselho de Pesquisas em Educação e Ciências; 2006. p. 215-9.
10. Farhat R. Dengue: Brasil prepare-se para a batalha. *InformeSaúde* 2003; Ano VII, nº. 196.
11. Agência Executiva de Gestão das Águas do Estado da Paraíba. Monitoramento das chuvas. <http://www.lmrs-semarh.ufcg.edu.br/meteoro/chuvas.shtml> (acessado em 11/Fev/2006).
12. Secretaria de Estado da Saúde do Estado da Paraíba. Base de dados eletrônica dos Sistemas de Informação: Sistema Nacional de Agravos de Notificação – SINAN e o de Febre Amarela e Dengue – FAD. João Pessoa: Secretaria de Estado da Saúde do Estado da Paraíba; 2005.
13. Secretaria de Vigilância em Saúde. Sistema de Informação de Agravos de Notificação. <http://dtr2004.saude.gov.br/sinanweb/> (acessado em 11/Fev/2006).
14. Gujarati DN. *Econometria básica*. São Paulo: Makron Books do Brasil; 2000.
15. Almon S. The distributed lag between capital appropriations and expenditures. *Econometrica* 1965; 33:178-96.
16. Johnston J. *Métodos econométricos*. São Paulo: Editora Atlas; 1974.
17. Cordeiro GM, Paula GA. *Modelos de regressão para análise de dados univariados*. Rio de Janeiro: Instituto Nacional de Matemática Pura e Aplicada; 1989.
18. Johnson RA, Wichern DW. *Applied multivariate statistical analysis*. New Jersey: Prentice Hall; 1992.
19. Schwarz G. Estimating the dimension of a model. *Annals of Statistics* 1978; 6:461-4.
20. Departamento de Informática do SUS. *Informações de saúde: demográficos e sócio-econômicos*. <http://www.datasus.gov.br> (acessado em 11/Fev/2006).
21. Laurenti R, Jorge MHP, Lebrão ML, Gotlieb SLD. *Estatística de saúde*. São Paulo: E.P.U.; 1987.

Recebido em 03/Jul/2006

Versão final reapresentada em 30/Abr/2007

Aprovado em 09/Mai/2007