

Predicting the number of cases of dengue infection in Ribeirão Preto, São Paulo State, Brazil, using a SARIMA model

Previsão do número de casos de dengue em Ribeirão Preto, São Paulo, Brasil, por um modelo SARIMA

Edson Zangiacomi Martinez ¹
Elisângela Aparecida Soares da Silva ¹

¹ Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, Brazil.

Correspondence

E. Z. Martinez
Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.
Av. Bandeirantes 3900, Ribeirão Preto, SP 14049-900, Brasil.
edson@fmrp.usp.br

Abstract

This study aimed to develop a forecasting model for the incidence of dengue in Ribeirão Preto, São Paulo State, Brazil, using time series analysis. The model was performed using the Seasonal Autoregressive Integrated Moving Average (SARIMA). Firstly, we fitted a model considering monthly notifications of cases of dengue recorded from 2000 to 2008 in Ribeirão Preto. We then extracted predicted values for 2009 from the adjusted model and compared them with the number of cases observed for that year. The SARIMA (2,1,3) (1,1,1)₁₂ model offered best fit for the dengue incidence data. The results showed that the seasonal ARIMA model predicts the number of dengue cases very effectively and reliably, and is a useful tool for disease control and prevention.

Dengue; Forecasting; Brostatistics

Introduction

Dengue is a disease of great importance to public health in tropical nations, particularly in South-east Asia and Central and South America. It is caused by four serotypes of a flavivirus – DENV1, DENV2, DENV3 and DENV4 – classified on biological and immunological criteria. Dengue is transmitted between human hosts by several species of day-feeding mosquitoes, such as the *Aedes aegypti*. Infection can be asymptomatic or it can manifest as an undifferentiated febrile illness, known as dengue fever, characterized by symptoms including fever, headaches, myalgia and retro-orbital pain ¹. Some infections result in dengue hemorrhagic fever (DHF), a syndrome that, in its most severe form, can be life-threatening ².

By the final decade of the twentieth century, *Ae. aegypti* and the four dengue viruses had spread to nearly all countries of the tropical world, and tens of millions are infected annually ³. In the 21st century Brazil became the country with the most reported cases of dengue fever in the world ⁴: more than three million cases were reported there from 2000 to 2005; that is approximately 70% of reported dengue fever cases in the Americas ⁵. The Southeast region of Brazil – and, as a special case, the city of Ribeirão Preto – has been most affected by dengue ^{6,7,8,9}.

The first reported dengue outbreak in Brazil occurred in 1922, in Niterói, Rio de Janeiro Sta-

te¹⁰, and the first laboratory-confirmed dengue outbreak was reported in 1981-1982, in Roraima State¹¹. After a period with no cases being diagnosed, the disease reappeared in Rio de Janeiro State in 1986, when the DENV1 virus was introduced¹². In 1990, during a period of high DENV1 activity, the DENV2 virus serotype was isolated in Niterói¹³, and the first cases of DHF were documented there¹⁴. Thus, a major dengue epidemic in Rio de Janeiro State was caused by the simultaneous circulation of DENV1 and DENV2, with a total of 140,000 reported cases¹⁵. In the following years, the DENV2 serotype spread to other regions of Brazil, with more severe clinical presentations¹⁶. In Ceará State, northeastern Brazil, DENV2 was first identified in 1994, at which time the first cases of DHF were notified¹⁷. The DENV3 serotype was first isolated in December 2000 in Nova Iguaçu, Rio de Janeiro State¹⁸, marking the start of a period of co-circulation of DENV1, DENV2 and DENV3. In 2002, the number of dengue cases increased in susceptible populations that had only experienced DENV1 and DENV2 epidemics¹⁹, and DENV3 virus later spread more broadly in Brazil. In 2008, DENV4-positive samples were obtained from patients in Amazonas State²⁰, the first time this serotype was isolated in Brazil in 25 years. By the end of March 2010, the São Paulo State health authorities reported more than 34,000 cases of dengue²¹. In 2010, approximately 30,000 confirmed cases of dengue were reported in Ribeirão Preto, the largest outbreak to date in that municipality (see <http://www.cve.saude.sp.gov.br/>).

Statistical tools used in epidemiology to monitor and predict dengue and other infectious diseases have included time series analysis techniques²², such as autoregressive integrated moving average (ARIMA) models^{23,24,25}. In the epidemiological literature, recent articles have used ARIMA models to describe the temporal pattern of diseases such as influenza²⁶, malaria²⁷ and dengue^{28,29,30,31,32}.

In this paper, we study the performance of the seasonal ARIMA model (SARIMA) in describing and predicting the monthly number of notified cases of dengue in Ribeirão Preto (São Paulo). Using dengue incidence data from 2000 to 2008, and the Box-Jenkins modeling approach³³, we fit a SARIMA³⁴ model to dengue incidence, and then used the fitted model to out-of-sample predict dengue incidence for the year 2009.

Methods

Ribeirão Preto is a municipality in northeastern São Paulo State, Brazil (21°10'42" South latitude

and 47°48'24" West longitude), with an economy based on agribusiness. The frequency of confirmed cases of dengue in Ribeirão Preto was obtained from the Divisão de Vigilância Epidemiológica of the Secretaria Municipal de Saúde de Ribeirão Preto (available at <http://www.ribeiraopreto.sp.gov.br>). The dataset includes the monthly number of cases, and the study period was from 2000 to 2009.

Given a stationary time series of data $Y' = (Y_1, Y_2, \dots, Y_n)$, an autoregressive moving average (ARMA) model, denoted by ARMA (p, q), consists of two parts, an autoregressive (AR) part of order p and a moving average (MA) part of order q . Thus, the ARMA model of order p and q , denoted by ARMA (p, q), is given by

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

where μ is a constant, $\phi' = (\phi_1, \phi_2, \dots, \phi_p)$ is a vector of autoregressive coefficients, $\theta' = (\theta_1, \theta_2, \dots, \theta_q)$ is a vector of moving average coefficients, and ε_t are error terms assumed to be independent, identically-distributed random variables sampled from a distribution with mean equal to zero and variance σ_ε^2 . In time series analyses, the variables ε_t are commonly referred to as white noise, and they are interpreted as an exogenous effect that the model is not able to explain. Considering the time series of monthly dengue incidence, these white noises can be, for example, an effect of climatic variables, prevention and education campaigns, introduction/reintroduction of a dengue serotype in a susceptible population, or random factors.

If the time series show evidence of non-stationarity, the data can be stationarized by introducing difference operators in the model. The first difference operator is given by $\Delta Y_t = Y_t - Y_{t-1}$. The k 'th difference operator is given by $\Delta^k Y_t = (1 - B)^k Y_t$, where B is the lag operator given by $B^k = Y_{t-k} / Y_t$. Thus, we obtain the autoregressive integrated moving average (ARIMA), denoted by ARIMA (p, d, q), where d is the number of differencing passes. The mathematical form of the ARIMA (p, d, q) model is

$$\phi(B) (1 - B)^d Y_t = \theta(B) \varepsilon_t$$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$. Thus, an important issue in fitting an ARIMA model is to identify the appropriate order of differencing needed to stationarize the series.

A seasonal ARIMA model (SARIMA) with S observations per period, denoted by SARIMA (p, d, q)(P, D, Q)_S, is given by

$$\Phi(B^S) \phi(B) (1 - B)^d (1 - B^S)^D Y_t = \Theta(B^S) \theta(B) \varepsilon_t$$

where $\Phi(B^S) = 1 - \phi_{S,1}B^S - \phi_{S,2}B^{2S} - \dots - \phi_{S,p}B^{pS}$ and $\Theta(B^S) = 1 + \theta_{S,1}B^S + \theta_{S,2}B^{2S} + \dots + \theta_{S,Q}B^{QS}$ are seasonal polynomial functions of order P and Q , respectively, which satisfy the stationarity and invertibility conditions.

In order to analyze the time series for dengue incidence in Ribeirão Preto over the years 2000 to 2008, we defined $S = 12$, given that we have 12 observations per year. Our first step used plots of the autocorrelation and partial autocorrelation functions^{35,36} (correlograms of the time series) to identify possible values for the autoregressive or moving average components. The second step was to obtain maximum likelihood estimates for the parameters of the SARIMA models, according to the different values of p , d , q , P , D and Q . Thirdly, we verified the goodness of fit of each model by plotting the autocorrelation and partial autocorrelation of residuals, and by using the Ljung-Box test³⁷. The fourth step compared the models by the Akaike information criterion (AIC)³⁸, where the preferred model is the one with the lowest AIC value. Finally, we extracted predicted values for 2009 from the best SARIMA model, and compared them with the number of new cases observed in that year.

As a criterion for comparing the predictive ability of the models, let K be a measure defined by the sum of squared differences between predicted and observed values at each month, each divided by its respective predicted value. Thus, the preferred model is the one with the lowest K value.

All analyses were performed using R software (The R Foundation for Statistical Computing, Vienna, Austria; <http://www.r-project.org>)³⁹.

Results

The monthly numbers of cases of dengue notified in Ribeirão Preto from 2000 to 2009 are shown in Table 1. Note that 2001, 2006 and 2007 are years with large numbers of individuals with the disease, and a graphical description of monthly cases of the disease (Figure 1) identifies a peak of cases in March, April, and May.

Let $X' = (X_1, X_2, \dots, X_n) = (8, 22, 31, 73, \dots, 3, 16)$ be the vector containing the monthly cases of dengue between the years 2000 and 2008, as shown in Table 1. In order to obtain a more stationary time series, let $Y' = (\log(X_1+1), \log(X_2+1), \dots, \log(X_n+1))$, instead of the original values X . In order to avoid taking natural logarithms of zero values, we added 1 to the number of dengue cases reported in each month, given that there are months with no cases recorded. In the absence of mathematical formalisms, a time series is said to be station-

ary if it oscillates around a constant mean value, and with a constant variance. However, a graph of the series Y_1, Y_2, \dots, Y_n against time (not shown in this article) describes a rising trend, but we ascertained that this series can be stationary after one difference operation. This suggests that it is appropriate to consider an order $d = 1$ in fitting the model to the data.

Figure 2 shows graphs of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the transformed series. The shape of the ACF describes a seasonal effect with a period of $S = 12$ months. The PACF suggests that p should be equal to 2 or 3, given that partial autocorrelations are near to zero at all lags that exceed 3, and the ACF suggests a moving average of order q equal to 2 or 3, given that its autocovariances are close to zero at all lags that exceed 3.

Table 2 shows AIC values for the SARIMA models corresponding to different choices of p and q . Problems with convergence of the computational algorithm used for estimating the model parameters are encountered when using $D = 0$. Therefore, we defined $D = 1$ in all models in Table 2. The SARIMA model with order $(2,1,3)(1,1,1)_{12}$ has the lowest AIC of the models described in Table 2. For that model, the autoregressive parameters ϕ_1 and ϕ_2 are estimated as 1.40 (standard error - SE = 0.186) and -0.533 (SE = 0.182), respectively, and the moving-average parameters θ_1 , θ_2 and θ_3 are estimated as -1.820 (SE = 0.211), 1.314 (SE = 0.303) and -0.494 (SE = 0.138), respectively. The seasonal components $\phi_{12,1}$ and $\theta_{12,1}$ are estimated as 0.085 (SE = 0.119) and -1.0 (SE = 0.169), respectively.

Also for the SARIMA $(2,1,3)(1,1,1)_{12}$ model, Figure 3 shows the standardized residuals, their histogram and ACF graphs and p-values for the Ljung-Box³⁷ statistic. Panel (a) of Figure 3 suggests that the standardized residuals estimated from this model should behave as an independent and identically distributed sequence with mean zero and constant variance. Panel (b) suggests that the residuals are normally distributed (in addition, the p-value Kolmogorov-Smirnov test is 0.26, and we thus do not reject the null hypothesis of normality). The ACF of the residuals shown in Panel (c) suggests autocorrelations close to zero. This means that the residuals did not deviate significantly from a zero-mean white-noise process. Panel (d) shows p-values for the Ljung-Box statistic. Given the high p-values associated with the statistics, we cannot reject the null hypothesis of independence in this residual series. We can therefore say that the model identified fits the data well.

The graph in Figure 4 compares the number of dengue cases observed between 2000 and 2008

Table 1

Number of recorded cases of dengue between 2000 and 2009 in Ribeirão Preto, São Paulo State, Brazil.

Month	Year									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
January	8	30	53	69	0	27	292	36	88	41
February	22	80	81	86	5	66	554	151	111	101
March	31	457	73	246	6	105	1,312	656	219	353
April	73	1,308	55	240	6	175	2,366	1,019	373	441
May	48	960	56	112	7	82	1,222	664	127	372
June	19	273	7	30	0	28	175	106	55	88
July	1	60	0	10	2	12	40	46	25	23
August	6	8	0	2	0	6	19	12	16	3
September	0	6	0	2	6	3	2	4	17	13
October	0	3	0	0	0	17	6	13	8	14
November	1	1	7	0	4	59	6	5	3	46
December	1	4	14	0	10	57	3	10	16	55
Total	210	3,190	346	797	46	637	5,997	2,722	1,058	1,550

Source: Divisão de Vigilância Epidemiológica, Ribeirão Preto City Health Department (Secretaria Municipal de Saúde de Ribeirão Preto), São Paulo State, Brazil.

Figure 1

Number of notified cases of dengue between 2000 and 2009 in Ribeirão Preto, São Paulo State, Brazil.

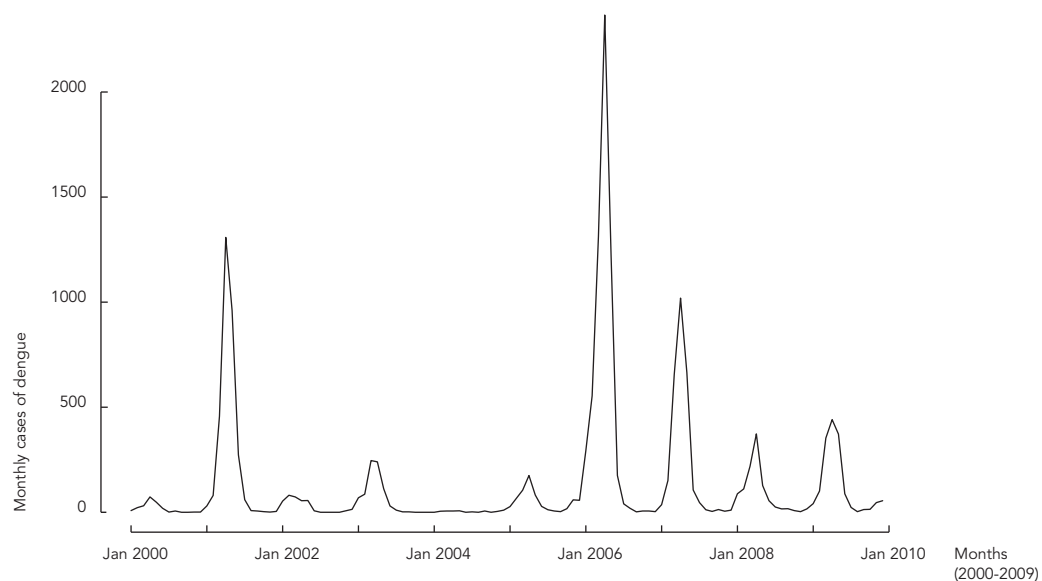


Figure 2

Autocorrelation function and partial autocorrelation function plotted against time lag.

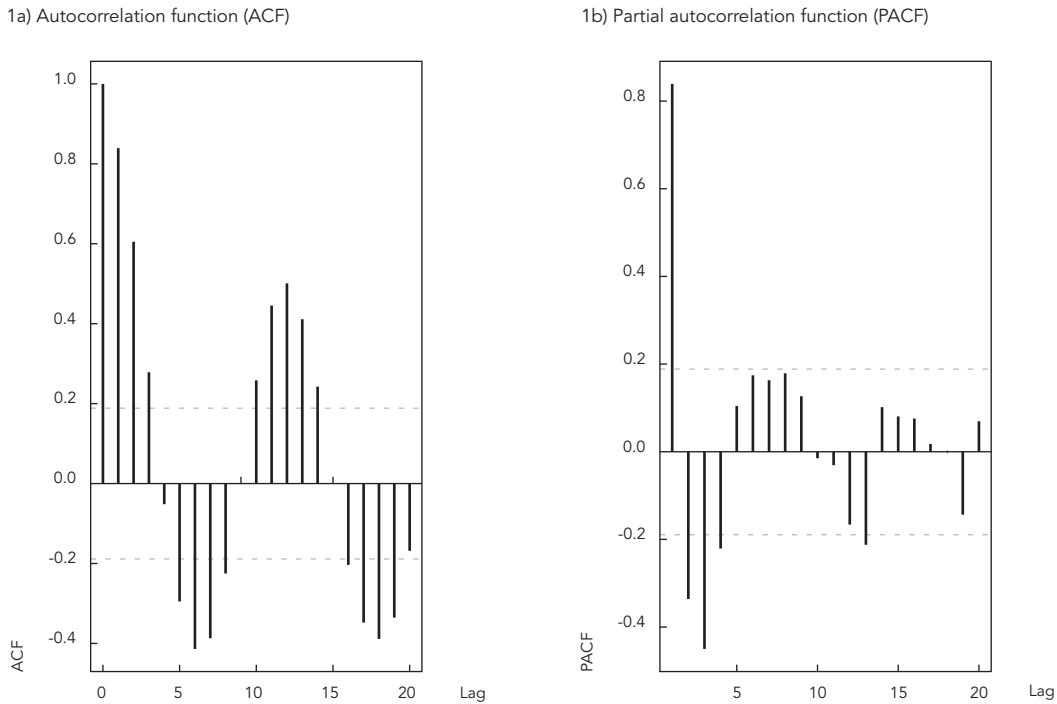


Table 2

Akaike Information Criterion (AIC) values for different seasonal autoregressive integrated moving average (SARIMA) $(p,1,q)(1,1,1)_{12}$ models and estimates for σ^2_ε .

Model	AIC value	Estimate for σ^2_ε
$(2,1,3)(1,1,1)_{12}$	253.01	0.5193
$(2,1,2)(1,1,1)_{12}$	256.65	0.5773
$(2,1,1)(1,1,1)_{12}$	257.40	0.5956
$(1,1,3)(1,1,1)_{12}$	256.19	0.5742
$(1,1,2)(1,1,1)_{12}$	254.97	0.5793
$(1,1,1)(1,1,1)_{12}$	256.49	0.6015
$(3,1,3)(1,1,1)_{12}$	260.18	0.5741
$(3,1,2)(1,1,1)_{12}$	259.30	0.5810
$(3,1,1)(1,1,1)_{12}$	257.37	0.5813

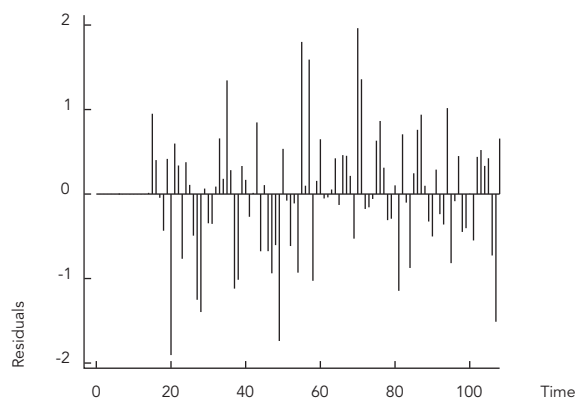
with the values obtained by the SARIMA $(2,1,3)(1,1,1)_{12}$ model. The observed and estimated values agreed very closely, except for the two highest peaks in the time series (corresponding to April 2001 and April 2006), and for March 2007.

Predicted 2009 values are shown in Table 3, where they are compared with the observed number of dengue cases. Table 3 presents out-of-sample predicted values obtained from the SARIMA $(2,1,3)(1,1,1)_{12}$ model and other more parsimonious models with lower AIC values: $(1,1,2)(1,1,1)_{12}$ and $(1,1,3)(2,1,1)_{12}$. Note the reasonably good match between observed and predicted dengue cases obtained by all these models for 2009. The graph in Figure 5 compares the observed number of dengue cases and the corresponding out-of-sample predicted values for 2009 obtained from the SARIMA $(2,1,3)(1,1,1)_{12}$ model, where the vertical lines are 95% prediction intervals. As the figure shows, the observed and predicted values are very close to each other, despite the relatively large range of the intervals associated with the out-of-sample predictions for March to June. K (the sum of squared differences between predicted and observed values at each month) equals 1039.4, 1380.4 and 1379.5, respectively, for the SARIMA models of orders $(2,1,3)(1,1,1)_{12}$, $(1,1,2)(1,1,1)_{12}$ and $(1,1,3)(2,1,1)_{12}$. On that criterion, we identified the SARIMA $(2,1,3)(1,1,1)_{12}$ as the model that best predicts

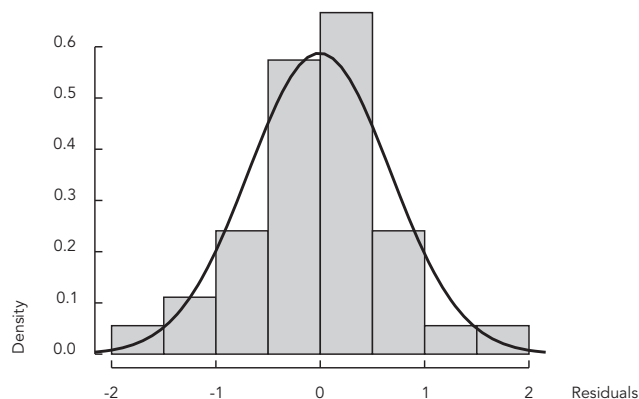
Figure 3

Diagnostic plots for time series fit, seasonal autoregressive integrated moving average (SARIMA) (2,1,3)(1,1,1)₁₂.

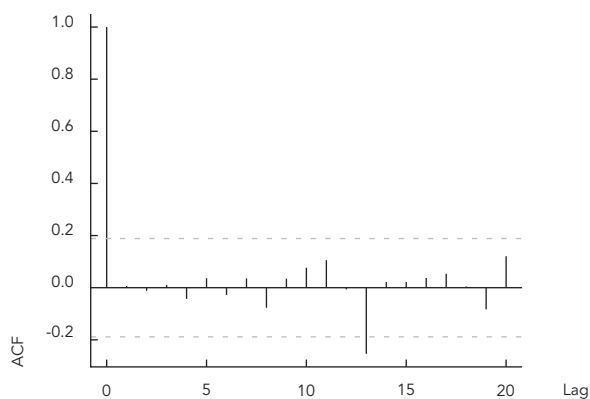
1a) Graph of standardized residuals versus observation order



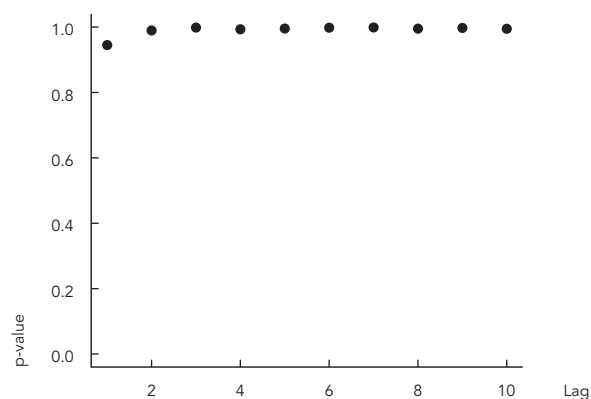
1b) Histogram of residuals compared with a normal curve



c) Autocorrelation function (ACF) of residuals



d) p values for Ljung-Box statistic



the number of dengue cases recorded in 2009 in Ribeirão Preto.

Discussion

Efforts to model dengue incidence in various parts of the world have used statistical approaches for time series analysis. Wongkoon et al.³⁰ developed a SARIMA model on the monthly data collected between 2003 and 2006 in Northern Thailand, and found that the SARIMA (2,0,1)(0,2,0)₁₂ was appropriate to predict the number of cases of dengue hemorrhagic fever for 2007. Promprou et al.²⁹ forecasted the monthly number of dengue hemorrhagic fever cases in South-

ern Thailand by an ARIMA (1,0,1) model. Silawan et al.³¹ showed that a SARIMA (2,1,0)(0,1,1)₁₂ model was suitable to determine temporal patterns and forecast dengue incidence in Northeastern Thailand. Choudhury et al.³² showed that a SARIMA (1,0,0)(1,1,1)₁₂ model was suitable for forecasting dengue incidence in Dhaka, Bangladesh. Luz et al.²⁸ developed a SARIMA (2,0,0)(1,0,0)₁₂ model for monitoring dengue incidence in Rio de Janeiro, Brazil, from 1997 to 2004. All these studies showed that the number of dengue cases in a given month can be estimated by the number of dengue cases occurring one (when $p = 1$) or two (when $p = 2$) months prior. In the present article considering data from Ribeirão Preto, we also showed that the number of dengue

Table 3

Observed number of dengue cases in 2009 and corresponding out-of-sample predicted values obtained from seasonal autoregressive integrated moving average (SARIMA) models.

Month	Observed values, 2009	Out-of-sample predicted values for 2009		
		SARIMA (2,1,3)(1,1,1) ₁₂	SARIMA (1,1,2)(1,1,1) ₁₂	SARIMA (1,1,3)(1,1,1) ₁₂
January	41	44.7	46.0	47.8
February	101	99.0	115.7	112.5
March	353	224.9	245.8	280.0
April	441	386.6	374.8	373.2
May	372	256.7	225.1	245.3
June	88	70.4	54.8	56.0
July	23	26.4	18.9	20.3
August	3	14.8	9.9	10.3
September	13	10.5	6.8	7.2
October	14	9.1	5.9	6.1
November	46	11.7	7.8	8.3
December	55	19.4	12.6	13.1
Total	1,550	1,174.3	1,124.1	1,180.2

Figure 4

Observed number of notified dengue cases between 2000 and 2008 in Ribeirão Preto, São Paulo State, Brazil, and number of cases estimated by seasonal autoregressive integrated moving average (SARIMA) (2,1,3)(1,1,1)₁₂ model (represented by points).

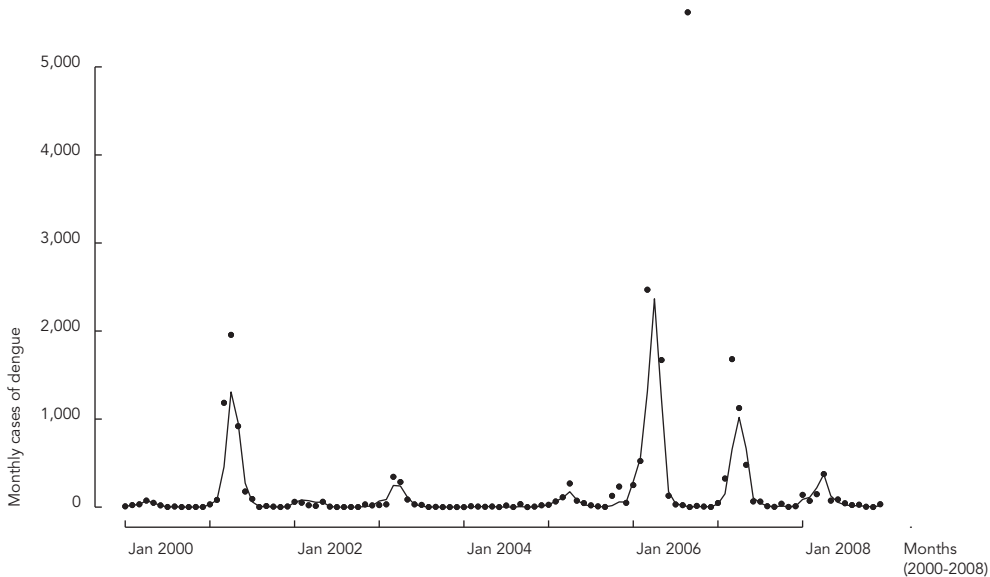
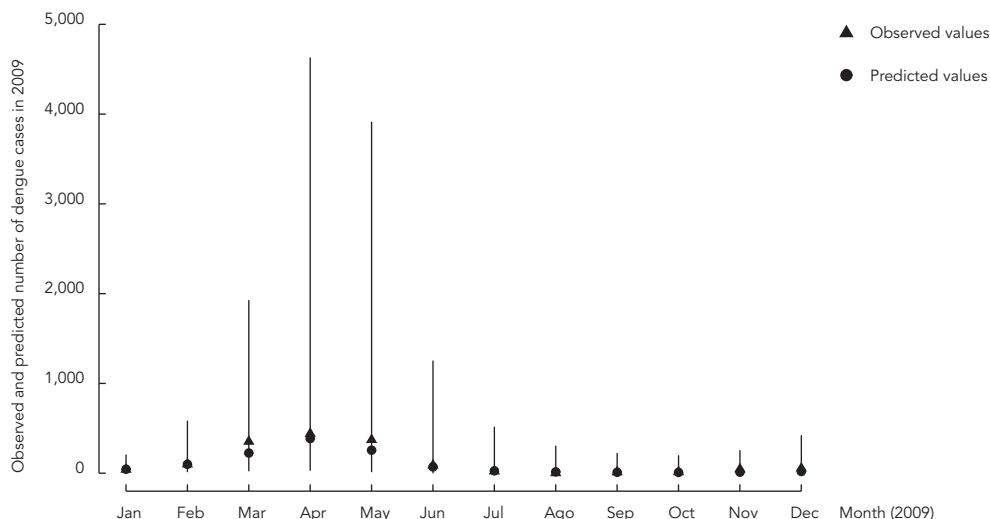


Figure 5

Observed number of dengue cases in 2009 (represented by triangles) and the respective out-of-sample predicted values (represented by circles) obtained from seasonal autoregressive integrated moving average (SARIMA) $(2,1,3)(1,1,1)_{12}$ model. The vertical lines are 95% prediction intervals.



cases in a given month can be estimated from the number of dengue cases occurring one and two ($p = 2$) months prior, and twelve ($S = 12$ and $P = 1$) months prior, but we found that a moving-average model of order q equal to 3 is suitable for the data between 2000 and 2008.

The results from this study show that the seasonal ARIMA model is a very effective and reliable predictive model for determining the number of dengue cases in a population, and is a useful tool for disease control and prevention. Allard⁴⁰ claims that ARIMA models are a useful tool for interpreting surveillance data, and that the usefulness of forecasting expected numbers of infectious disease reports consists not so much in detecting outbreaks or providing probability statements, but in giving decision makers a clearer idea of the variability to be expected among future observations. We found that the SARIMA model's predictions for 2009 agree reasonably well with the observed incidence of dengue. However, these out-of-sample predictions may not be credible for forecasting the number of dengue cases in epidemic years, when the observed monthly incidence is significantly higher than the expected number of new cases for the period. This large number of cases may be a consequence of a lack of immunity in a population exposed for the first time to a given dengue viral serotype. A dengue

virus type 1 (DENV1) outbreak is known to have started in Ribeirão Preto in November 1990⁷, and DENV2 and DENV3 were introduced in São Paulo State in 1997 and 2003, respectively (see <http://www.cve.saude.sp.gov.br/>). Thus, the highest peaks in the time series shown in Figure 1 may be a direct consequence of the introduction or reintroduction of different serotypes. Note, however, in Figure 4, that the SARIMA $(2,1,3)(1,1,1)_{12}$ model produced good estimates at each month, even though the time series contains periods with large numbers of dengue cases. Figure 4 shows that the model failed to estimate the number of dengue cases in April 2006, but in the following months the model again provided estimates with good precision. These results suggest that the out-of-sample forecast values for 2009 obtained from the SARIMA model are not subject to an effect due to the introduction or reintroduction of different dengue serotypes.

In addition, climate changes have potential impact on dengue transmission, and in a future study more accurate predictions should be made by introducing meteorological variables such as temperature, pressure, humidity and rainfall into the model. These variables are known to be associated with an increase in the number of available breeding places for *Ae. aegypti* and, accordingly, in the risk of transmission of dengue.

Resumo

Este estudo tem por objetivo desenvolver um modelo para a predição do número de casos de dengue em Ribeirão Preto, São Paulo, Brasil, por técnicas de análise de séries temporais. Para isto, foi utilizado o modelo SARIMA (Seasonal Autoregressive Integrated Moving Average). Inicialmente, ajustamos um modelo considerando o número mensal de casos notificados de dengue entre os anos 2000 e 2008 em Ribeirão Preto. Em uma etapa seguinte, obtivemos, com base nesse modelo, valores preditos para 2009, os quais comparamos com os valores observados neste ano. O modelo SARIMA $(2,1,3)(1,1,1)_{12}$ trouxe o melhor ajuste para os dados de incidência de dengue. Os resultados obtidos neste artigo mostram que o modelo SARIMA é bastante eficiente em descrever o número de casos de dengue no período em estudo e em prever valores em meses futuros, mostrando-se uma útil ferramenta para estratégias de controle e prevenção da doença.

Dengue; Previsões; Bioestatística

Contributors

Both authors contributed equally to preparing this manuscript.

Acknowledgments

The second author is supported by a grant from the Fundação de Amparo ao Ensino, Pesquisa e Assistência do Hospital das Clínicas, Faculdade de Medicina de Ribeirão Preto (FAEPA), São Paulo.

References

1. Nunes-Araújo FRF, Ferreira MS, Nishioka SA. Dengue fever in Brazilian adults and children: assessment of clinical findings and their validity for diagnosis. *Ann Trop Med Parasitol* 2003; 97:415-9.
2. Rigau-Pérez JG, Clark GG, Gubler DJ, Reiter P, Sanders EJ, Vorndam AV. Dengue and dengue haemorrhagic fever. *Lancet* 1998; 352:971-7.
3. Halstead SB. The XXth century dengue pandemic: need for surveillance and research. *World Health Stat Q* 1992; 45:292-8.
4. Teixeira MG, Costa MC, Barreto F, Barreto ML. Dengue: twenty-five years since reemergence in Brazil. *Cad Saúde Pública* 2009; 25 Suppl 1:S7-18.
5. Siqueira Jr. JB, Martelli CM, Coelho GE, SImplicio AC, Hatch DL. Dengue and dengue hemorrhagic fever, Brazil, 1981-2002. *Emerg Infect Dis* 2005; 11:48-53.
6. Hino P, Santos CC, Santos MO, Cunha TN, Santos CB. Temporal evolution of dengue fever in Ribeirão Preto, São Paulo State, 1994-2003. *Ciênc Saúde Coletiva* 2010; 15:233-8.
7. Figueiredo LTM, Owa MA, Carlucci RH, Oliveira L. Estudo sobre o diagnóstico laboratorial e sintomas do dengue, durante epidemia ocorrida na região de Ribeirão Preto, SP, Brasil. *Rev Inst Med Trop São Paulo* 1992; 34:121-30.
8. Pontes RJS, Dal Fabbro AL, Rocha GM, Santiago RC, Figueiredo LTM, Silva AAMCC, et al. Dengue in the Southeast Region of Brazil: a preliminary note. *Rev Saúde Pública* 1991; 25:315-7.
9. Passos ADC, Rodrigues EMS, Dal-Fabbro AL. Dengue control in Ribeirão Preto, São Paulo, Brazil. *Cad Saúde Pública* 1998; 14 Suppl 2:123-8.
10. Dietz VJ, Gubler DJ, Rigau-Perez JG, Pinheiro F, Schatzmayr HG, Bailey R, et al. Epidemic dengue 1 in Brazil, 1986: evaluation of a clinically based dengue surveillance system. *Am J Epidemiol* 1990; 131:693-701.
11. Osanai CH, Travassos-Rosa AP, Tang AT, Amaral AS, Passos AD, Tauil PL. Outbreak of dengue in Boa Vista, Roraima. Preliminary report. *Rev Inst Med Trop São Paulo* 1983; 25:53-4.

12. Schatzmayr HG, Nogueira RM, Travassos-Rosa AP. An outbreak of dengue virus at Rio de Janeiro. *Mem Inst Oswaldo Cruz* 1986; 81:245-6.
13. Nogueira RMR, Miagostovich MP, Lampe E, Schatzmayr HG. Isolation of dengue virus type 2 in Rio de Janeiro. *Mem Inst Oswaldo Cruz* 1990; 85:253.
14. Nogueira RM, Miagostovich MP, Lampe E, Souza RW, Zagne SMO, Schatzmayr HG. Dengue epidemic in the state of Rio de Janeiro, Brazil, 1990-1: co-circulation of dengue 1 and dengue 2 serotypes. *Epidemiol Infect* 1993; 111:163-70.
15. Nogueira RM, Miagostovich MP, Schatzmayr HG, Santos FB, Araújo ES, Filippis AM, et al. Dengue in the State of Rio de Janeiro, Brazil, 1986-1998. *Mem Inst Oswaldo Cruz* 1999; 94:297-304.
16. Costa-Vasconcelos PF, Menezes DB, Melo LP, Pessoa P, Rodrigues SG, Travassos-Rosa E, et al. A large epidemic of dengue fever with dengue hemorrhagic cases in Ceará State, Brazil, 1994. *Rev Inst Med Trop São Paulo* 1995; 37:253-5.
17. Cavalcanti LP, Coelho IC, Vilar DC, Holanda SG, Escóssia KN, Souza-Santos R. Clinical and epidemiological characterization of dengue hemorrhagic fever cases in northeastern, Brazil. *Rev Soc Bras Med Trop* 2010; 43:355-8.
18. Nogueira RMR, Miagostovich MP, Filippis AMB, Pereira MAS, Schatzmayr HG. Dengue type 3 in Rio de Janeiro, Brazil. *Mem Inst Oswaldo Cruz* 2001; 96:925-6.
19. Nogueira RM, Schatzmayr HG, Filippis AM, Santos FB, Cunha RV, Coelho JO, et al. Dengue virus type 3, Brazil, 2002. *Emerg Infect Dis* 2005; 11:1376-81.
20. Figueiredo RM, Naveca FG, Bastos MS, Melo MN, Viana SS, Mourão MP, et al. Dengue virus type 4, Manaus, Brazil. *Emerg Infect Dis* 2008; 14:667-9.
21. Romano CM, Matos AM, Araújo ES, Villas-Boas LS, Silva WC, Oliveira OM, et al. Characterization of Dengue virus type 2: new insights on the 2010 Brazilian epidemic. *PLoS One* 2010; 5:e11811.
22. Latorre MRDO, Cardoso MRA. Time series analysis in epidemiology: an introduction to methodological aspects. *Rev Bras Epidemiol* 2001; 4:145-52.
23. Hamilton JD. Time series analysis. Princeton: Princeton University Press; 1994.
24. Tsay RS. Analysis of financial time series. 2nd Ed. New York: John Wiley & Sons; 2005.
25. Cordeiro DM, Cordeiro GM. A combined forecast model: an application to the monthly series of dengue reports in Pernambuco state - Brazil. *Revista de Matemática e Estatística* 2004; 22:57-80.
26. Soebiyanto RP, Adimi F, Kiang RK. Modeling and predicting seasonal influenza transmission in warm regions using climatological parameters. *PLoS One* 2010; 5:e9450.
27. Tian L, Bi Y, Ho SC, Liu W, Liang S, Goggins WB, et al. One-year delayed effect of fog on malaria transmission: a time-series analysis in the rain forest area of Mengla County, south-west China. *Malar J* 2008; 7:110.
28. Luz PM, Mendes BV, Codeço CT, Struchiner CJ, Galvani AP. Time series analysis of dengue incidence in Rio de Janeiro, Brazil. *Am J Trop Med Hyg* 2008; 79:933-9.
29. Promprou S, Jaroensutasinee M, Jaroensutasinee K. Forecasting dengue haemorrhagic fever cases in Southern Thailand using ARIMA Models. *Dengue Bull* 2006; 30:99-106.
30. Wongkoon S, Pollar M, Jaroensutasinee M, Jaroensutasinee K. Predicting DHF incidence in Northern Thailand using time series analysis technique. *World Academy of Science, Engineering and Technology* 2007; 26:216-20.
31. Silawan T, Singhasivanon P, Kaewkungwal J, Nimmamitya S, Suwonkerd W. Temporal patterns and forecast of dengue infection in Northeastern Thailand. *Southeast Asian J Trop Med Public Health* 2008; 39:90-8.
32. Choudhury MAHZ, Banu S, Islam MA. Forecasting dengue incidence in Dhaka, Bangladesh: a time series analysis. *Dengue Bull* 2008; 32:99-106.
33. Box G, Jenkins G. Time series analysis: forecasting and control. 4th Ed. New York: John Wiley & Sons; 2008.
34. Nobre FF, Monteiro AB, Telles PR, Williamson GD. Dynamic linear model and SARIMA: a comparison of their forecasting performance in epidemiology. *Stat Med* 2001; 20:3051-69.
35. Hamilton DC, Watts DG. Interpreting partial autocorrelation functions of seasonal time series models. *Biometrika* 1978; 65:135-40.
36. Ramsey FL. Characterization of the partial autocorrelation function. *Ann Stat* 1974; 2:1296-301.
37. Ljung GM, Box GEP. On a measure of lack of fit in time series models. *Biometrika* 1978; 65:297-303.
38. Claeskens G, Hjort NL. Model selection and model averaging. Cambridge: Cambridge University Press; 2008.
39. Cowpertwait PSP, Metcalfe AV. Introductory time series with R. New York: Springer; 2009.
40. Allard R. Use of time-series analysis in infectious disease surveillance. *Bull World Health Organ* 1998; 76:327-33.

Submitted on 21/May/2010

Final version resubmitted on 25/Mar/2011

Approved on 11/Apr/2011