Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2020;6(1): 19-29

Original Article

Use of Stochastic Volatility Models in Epidemiological Data: Application to a Dengue Time Series in São Paulo City, Brazil

Jorge Alberto Achcar[1], Ricardo Puziol de Oliveira[2], Emerson Barili[2*]

[1]Medical School, University of São Paulo,Ribeirão Preto, SP, Brazil.
[2]State University of Maringá, Maringá, PR, Brazil.

ARTICLE INFO

Received 15.10.2019
Revised 09.12.2019
Accepted 25.01.2020
Published 10.03.2020

Key words:
Dengue count; Volatility models; Bayesian approach; Markov chain monte carlo methods

ABSTRACT

Background: A study on the dengue daily counting in São Paulo city in a fixed period of time is assumed considering a new regression model approch.
Methods: Under a Bayesian approach, it is introduced a polynomial linear regression model in presence of some covariates which could affect the counts of dengue in São Paulo city considered in the logarithm scale, combined with existing stochastic volatility models usually assumed in financial data analysis. Markov Chain Monte Carlo (MCMC) methods are used to get the posterior summaries of interest.
Results: The new model approach showed some advantages when compared to other existing times series models usually used to model epidemics data.
Conclusion: The use of the polynomial regression model combined with existing volatility models under a Bayesian approach showed that it is possible to get very accurate fit for the counting dengue data in São Paulo city where it is possible to jointly model the means and volatilities (variances) of the epidemiological dengue time series.

Introduction

Dengue is a viral disease transmitted by mosquitoes that spreads most quickly in the world. In the last 50 years, the incidence has increased 30 times with the increase in geographical expansion to new countries. It is estimated that 50 million dengue infections occur annually (Figure 1) and approximately 2.5 billion people live in dengue-endemic countries[1]. About 1.8 billion (more than 70%) of the population at risk for dengue worldwide lives in countries in Southeast Asia and the Western Pacific region. Since 2000, epidemic dengue has spread to new areas and has increased in the already affected areas of the region. In 2003, eight countries - Bangladesh, India, Indonesia, Maldives, Myanmar, Sri Lanka, Thailand and Timor-Leste - were reported dengue cases[1]. Between 2001 and 2008, 1,020,332 cases were reported in Cambodia, Malaysia, the Philippines and Vietnam - the four countries in the Western Pacific region with the highest number of cases and deaths[1]. An interruption of dengue transmission in much of the Americas was the result of a campaign to eradicate the Aedes Aegypti mosquito, mainly in the 1960s and early 1970s. However, vector surveillance and control measures were not sustained and there were subsequent mosquito reinfestations, followed by outbreaks in the Caribbean, Central and South America[1]. Since then, dengue has spread with cyclical outbreaks that occur every 3 to 5 years. The biggest outbreak occurred in 2002 with more than 1 million reported cases. From 2001 to 2007, more than 30 countries in the Americas reported a total of 4,332,731 dengue cases. In this period of time, 64.6% (2,798,601) of all dengue cases in the Americas were reported in the sub-region including Argentina, Brazil, Chile, Paraguay and Uruguay. Although dengue exists in the WHO African region, surveillance data is poor.

* . Corresponding Author E-Mail: ebarili2@uem.br

Please cite this article in press as Achcar JA, De-Oliveira RP, Barili E. Use of Stochastic Volatility Models in Epidemiological Data: Application to a Dengue Time Series in São Paulo City, Brazil. *2020; 6(1):19-29*

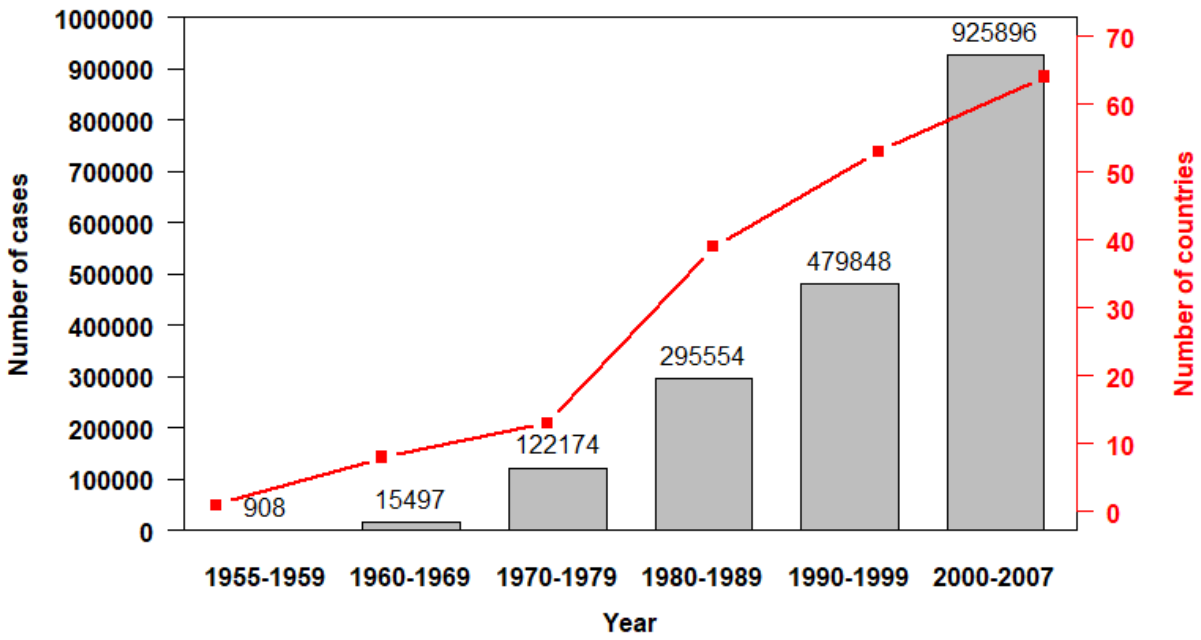

Figure 1. Average annual number of dengue and hemorrhagic dengue cases reported by WHO and countries reporting dengue, 1955-2007.

Different statistical models have been used in the literature to be fitted by dengue times series considering the count data in the original or transformed scale. In these studies, different statistical techniques are used in the analysis of the data, especially with data related to the time series of notified cases in a region in a given period of time, especially in the forecast of new cases. In this direction, [2]analyzed temporal patterns of dengue incidence for the period from 2001 to 2014 with forecasts for 2015 in two Brazilian cities: Goiania and Recife considering dengue surveillance data reported by SINAN (a Brazilian public health dataset office) using Moving Average or ARIMA (autoregressive integrated moving averag) times series models[3].Forecasting models (95% forecast range) were developed to predict the number of dengue cases in 2015 for these two cities. Several other authors have used ARIMA or SARIMA (seasonal autoregressive integrated moving average) models to model dengue incidence time series[4-7]. Other studies consider the use of space-time models in modeling dengue data[8].

On other hand, for the statistical analysis of the dengue count in the original scale, standard counting models based on regression Poisson models (generalized linear models, see for example,[9] also could be used, but the use of stochastic volatility models in the logarithm scale of the counting data gives more flexibility to simultaneously model the mean and variance (volatility) of the epidemiological time series. The use of stochastic volatility models is becoming very popular in the analysis of financial time series, as it can be verified in some studies as the paper published by[10] that carry out related empirical applications of financial risk incorporating stochastic volatility probability models in presence of random measures under a Bayesian approach. In other study, [11]proposed a spatial stock model in which latent log volatility measures follow an autoregressive process to estimate the return on residential property prices in the Chicago metropolitan area and [12]presented a Bayesian approach on the earning shocks and its volatility in financial crises and its subsequent recovering, but it is rarely used to analyze epidemiological time

series. The seasonality and volatility of dengue counting are studied using statistical volatility models that link the dependence of observed factors to the responses (dengue counting) and also non-observed factors linked to the volatilities.

The novelty of this study is the introduction of a new statistical modeling approach (combination of linear polynomial regression models with stochastic volatility models) to analyze dengue count data on the logarithmic scale considering a dengue times series data set of São Paulo city, Brazil from the period ranging from January 2007 to December 2016, using a Bayesian approach with Markov Chain Monte Carlo (MCMC) methods[13] to simulate samples of the joint posterior distribution for the parameters of the model. Stochastic volatility (SV) models have been widely used to analyze financial time series[14,15] as a powerful alternative to the self-regressive models in the literature, such as ARCH (conditional heteroscedastic autoregressive) models introduced by[16] and the generalized autoregressive conditional heteroscedastic models (GARCH) introduced by Bollerslev (1986), but not widely used in the health field[17-19]. In the financial area, these models are considered for modeling the logarithms of financial returns between current data and previous data (it can be hours, days or months) without taking into account the modeling of the means depending on covariates.

Usually there are great computational difficulties to get the posterior summaries of interest considering SV models. These difficulties can appear in the form of high dimensionality and likelihood function without closed form, among other factors. In this way, existing MCMC methods like the Gibbs and the Metropolis-Hastings algorithms have been used to get the inferences (Bayesian point estimators and credibility intervals) for the parameters of the proposed model. In the simulation of samples of the joint posterior distribution[13], $\pi(\theta/\text{data})$ where θ is the vector of all parameters, using Gibbs or Metropolis-Hastings algorithms, it is needed to sample each parameter from the posterior conditional distributions $\pi(\theta r/\theta(r), \text{data})$, where $\theta(r)$ denotes the vector of all parameters except θ_r and r is associated to each one of the parameters of the model. To simplify the computational work in the iterative procedure to get the Bayesian inferences, the literature presents different free softwares to simulate samples of the joint posterior distribution of interest. In this study, it is used the OpenBugs software[20] in the simulation of samples of the joint posterior distribution of interest which simplifies the computational work, since this software only requires the definition of the likelihood function for θ and the prior distribution $\pi(\theta)$.

The paper is organized as follows: in Section 2, it is presented monthly dengue data in São Paulo city, Brazil for the period 2007 to 2016 (period of 120 months or 10 years); in Section 3, a polynomial regression model and stochastic volatility is presented to analyze dengue data in the city of São Paulo; Section 4 presents the obtained results; finally, Section 5 presents some conclusions.

Goals of the Study and the Data Set

This study considers monthly dengue count data in the city of São Paulo, Brazil between the years 2007 and 2016 and some existing relationships with some covariates in the same period such as total monthly rainfall in the previous month and average minimum temperature in the previous month (lagged data). Figure (2) shows a graph of the count monthly time series of patients with dengue in the city of São Paulo, Brazil on the original scale and on the logarithmic scale (data obtained from the Brazilian public health data site SINAN – Information System for Notifiable Diseases - \url{http://sinan.saude.gov.br/sinan/login/login.jsf }) for the period between January 2007 and December 2016. Figure (2) also shows the total monthly rainfall data for the previous month and the average minimum temperature in the previous month (data obtained from the Brazilian meteorological data website INMET - Instituto Nacional de Meteorologia - \url{http://www.inmet.gov.br/portal/index.php?r= bdmep/bdmep}). Figure (3) shows the monthly dengue counts in the city of São Paulo from January 2007 to December 2016 from where we can conclude that:

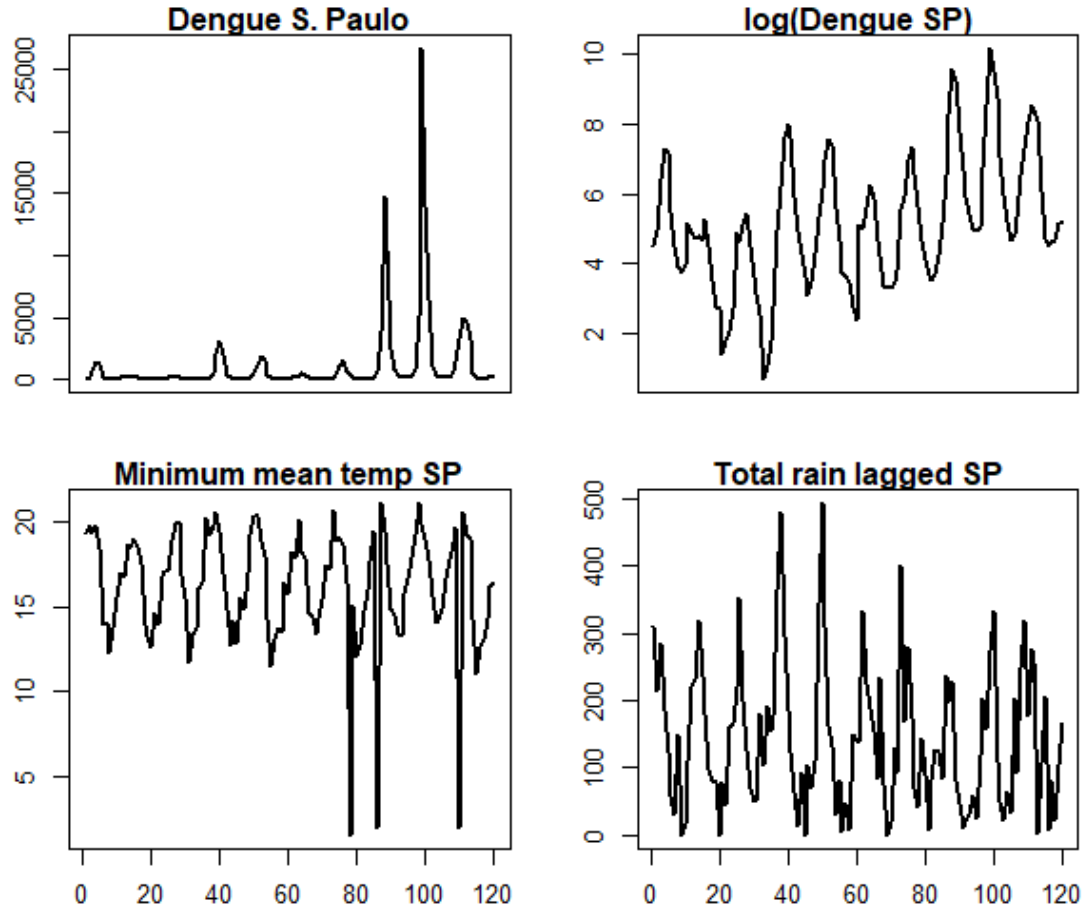

Figure 2: Counts of reported patients with dengue in São Paulo city in temporal order.

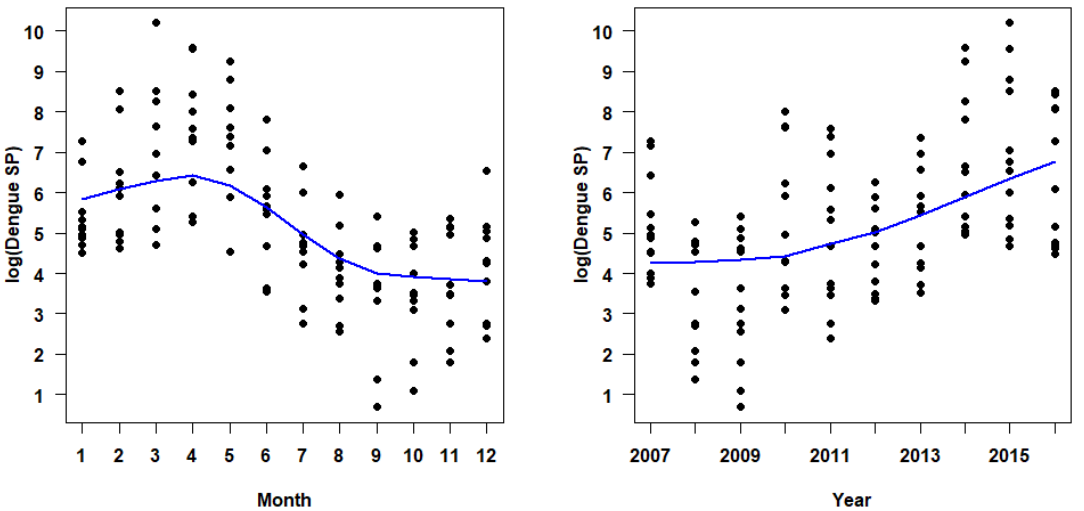

Figure 3: Monthly and anual counts of dengue in São Paulo city on the logarithmic scale.

The monthly counts of patients diagnosed with dengue in the period between January 2007 and December 2016 (logarithmic scale) show seasonality and there is clearly a significant increase in the months from March to May of each year; a decrease in the period between September and October; after this period there is an increase from November of each year.

The yearly counts of patients diagnosed with dengue in the period between January 2007 and December 2016 (logarithmic scale) show a decrease from 2007 to 2009; from 2009, it starts a consecutive increase from year to year, with a peak

in the years 2014 and 2015; thereafter, a decrease begins until the year 2016.

In Figure (3), it can be seen that for data analysis and statistical modeling of the data, it is necessary to consider polynomial linear regression models in the presence of the covariates (independent variables) month and year (temporal order) that are able to capture possible effects of linearity and that also have regression coefficients associated with quadratic and cubic effects. A quadratic term or cubic term transforms a linear regression model into a curve. Since the regression model has the covariates squared or cube month, and not the regression coefficient, the model remains a linear regression model[21,22].

The presence of a quadratic term in the model creates a U-shaped curve or an inverted U, as seen in the graphs of Figure (3). A cubic term has two distinct parts: one facing up and one facing down, that is, the curve go down, back up and back again.

In addition to the covariates month and year, we also included in the linear regression model the total precipitation and average minimum temperature covariates (lagged in previous months).

Also are considered seasonal effects included from an AR (2) model in the total dengue counting in the previous two months Y(t-1) and Y(t-2).

For the definition of the models, $N \geq 1$ is a fixed integer that registers the amount of data observed (in our case, it represents the monthly dengue counts on the logarithmic scale). Thus, initially let us assume the following linear regression model for the analysis of monthly data on the logarithmic scale:

$$Y(t) = \beta_o + \beta_1 month(t) + \beta_2 [month(t)]^2 + \beta_3 [month(t)]^3 + \beta_4 year(t) + \beta_5 [year\ (t)]^2 + \beta_6 lagged.average.minimum.temp[t] + \beta_7 lagged.total.precipitation[t] + \beta_8 Y(t-1) + \beta_9 Y(t-2) + \epsilon(t) \quad (1)$$

for t = 1,2,...,120 (months) where $\epsilon(t)$ are noises considered as independent and identically distributed random variables with a normal distribution $N(0, \sigma_\epsilon^2)$ and Y(t) are the monthly dengue counts on the logarithmic scale. This model is denoted by "model 1". Under a classical statistical approach, the regression parameters are usually estimated using the least squares method (LSE). In this study, we opted for the use of Bayesian methods.

An alternative of epidemiological interest would be to model the series not only to estimate the monthly averages in the period considered (January 2007 to December 2016), but also to estimate the monthly variances (volatilities) that are of interest to the public health researchers, possibly relating these volatilities to the occurrence of factors associated with the months (total monthly rainfall or minimum average temperatures in the previous month, that is, considering lagged effects relative to the previous month). For this purpose, a time series model is considered that simultaneously estimate the monthly average and the monthly volatility.

A Polynomial Regression Combined with a Stochastic Volatility Model for Dengue Data in São Paulo City

In the presence of heteroscedasticity, that is, variances depending on time t, assume that the time series Y(t), t = 1, 2 ,. . . , N assume a combination of a polynomial linear regression model for months and years with a stochastic volatility model, and the inclusion of lagged effects of counts (an auto-regressive model) and some factors that may be related to the incidence of dengue (total precipitation of monthly rainfall and minimum monthly average temperatures, considered as lagged effects relative to the previous month) for the analysis of dengue count data in São Paulo city on the logarithmic scale:

$$Y(t) = \beta_o + \beta_1 month(t) + \beta_2 [month(t)]^2 + \beta_3 [month(t)]^3 + \beta_4 year(t) + \beta_5 [year\ (t)]^2 + \beta_6 lagged.average.minimum.temp[t] + \beta_7 lagged.total.precipitation[t] + \beta_8 Y(t-1) + \beta_9 Y(t-2) + \sigma(t)\epsilon(t) \quad (2)$$

where it is assumed that $\epsilon(t)$ are noises considered independent and identically distributed random variables with a normal distribution $N(0, \sigma_\epsilon^2)$ and $\sigma(t)$ is the square root of the variance of (1) (for simplicity, we can assume $\sigma_\epsilon^2 = 1$). The variance of Y(t) is modeled by $\sigma_\epsilon^2 eh(t)$ where h(t) depends on an unobserved latent variable. This model is denoted by "model 2". It is important to point out

that the inclusion of quadratic and cubic effects of months and years were based on the behavior of the plots presented in Figure (3). Thus it is included cubic effects only for the covariate month.

Remark: From model (2), it is observed that:

The mean of Y(t) is given by
$E[Y(t)] = \beta o + \beta 1 month(t) + \beta 2[month(t)]2 + \beta 3[month(t)]3 + \beta 4 year(t) + \beta 5[year(t)]2 + \beta 6 lagged.average.minimum.temp[t] + \beta 7 lagged.total.precipitation[t] + \beta 8 Y(t-1) + \beta 9 Y(t-2)$, since $E[\sigma(t)\epsilon(t)] = 0$.
The variance of Y(t) is given by,
$var[Y(t)] = var[\sigma(t)\epsilon(t)] = \sigma 2(t)$ since we are assuming $var[\epsilon(t)] = \sigma_\epsilon^2 = 1$.

To analyze the data set, a latent variable (unobserved variable) defined by an auto-regressive model AR(2) is also introduced, for t = 1, 2, 3, ..., N (N = 120 months) .

$$\begin{aligned} h(1) &= \mu + \zeta(1), \quad t = 1 , \\ h(2) &= \mu + \phi 1[h(1) - \mu] + \zeta(2) \\ h(t) &= \mu + \phi 1[h(t-1) - \mu] + \phi 2[h(t-2) - \mu] + \zeta(t), \quad t = 3, 4, \ldots, N, \end{aligned} \tag{3}$$

where $\zeta(t)$ is a noise with a normal distribution $N(0, \sigma_\zeta^2)$, that is associated with the latent variable h(t). The quantities σ_ζ^2, μ, $\phi 1$ and $\phi 2$ are unknown parameters that must be estimated ($0 < \phi 1 < 1$, $0 < \phi 2 < 1$).

Bayesian inference procedures, based on Markov Chain Monte Carlo (MCMC) methods, 13 have been widely used to analyze stochastic volatility models. The main reason for using Bayesian methods is that, in general, we may have great difficulties in obtaining inferences (point and interval estimation) for the parameters of interest of the stochastic volatility model when using a standard classical inference approach. These difficulties can appear in the form of high dimensional and likelihood function without closed form, among other factors. For a Bayesian analysis of the model defined by (2), it is assumed prior distributions for the parameters μ, ϕv and $\zeta = 1/\sigma_\zeta^2$, v = 1, 2 given respectively by a normal N(0,a2) distribution, a Beta(b, c) distribution and a Gamma(d, e) distribution, where Beta(b, c) denotes a Beta distribution with mean b/(b+c) with variance bc/[(b+c)2(b+c+1)] and Gamma(d,e) denotes a gamma distribution with mean d/e and variance d/e2. The hyperparameters a, b, c, d, and e are assumed to be known and previously specified. Also assume that the regression parameters βj, j = 0, 1, …, 9 have independent normal distributions with known hyperparameters. Further, it is assumed prior independence between the parameters.

Results

For data analysis, initially consider the regression model ("model 1") given by (1). For a Bayesian analysis of the model, consider the following prior distributions for the model parameters: $\beta 0 \sim N(3, 1)$; $\beta j \sim N(0,0.1)$, j = 1,2,...,9 and $\zeta = 1/\sigma_\epsilon^2 \sim$ Gamma(1,1). That is, we are assuming approximately non-informative priors. The choice of the hyperparameters of the prior distributions were base on some preliminary data analysis of the data, as considered for the regression parameters βj (small values that could be negative or positive) and for the parameter ζ. The OpenBugs software20 was used in the simulation of samples of the joint posterior distribution of interest. Thus, the posterior conditional distributions13, $\pi(\theta r/ \theta(r)$, data), needed for the Gibbs and Metropolis-Hastings algorithms where $\theta = (\beta 0, \beta 1, ..., \beta 9, \zeta)$ is the vector of all parameters, are not presented in this paper.

The convergence of the simulation algorithm (Gibbs / Metropolis-Hastings) was verified from time series graphs of the generated Gibbs samples not presented for space saving. It was considered a "burn-in-sample" of size 111,000 discarded to eliminate the effect of the initial values in the iteractive method; thereafter, it was generated more 400,000 Gibbs samples from where it was taken each 100th sample (a final sample size of 4,000) to obtain the posterior summaries of interest. The posterior summaries of interest (posterior means, posterior standard deviations and 95% high

posterior density intervals) for each parameter of the model are presented in Table (1).

Table 1. Posterior summaries for "model 1" assuming dengue data in São Paulo city.

			95% HPD Int.	
	Mean	Std. Dev.	Lower	Upper
β0	2.1275	0.4840	1.2040	3.0941
β1	-0.0023	0.2150	-0.4198	0.4248
β2	-0.0747	0.0390	-0.1541	-0.0009
β3	0.0056	0.0020	0.0018	0.0097
β4	-0.0816	0.0861	-0.2414	0.0970
β5	0.0153	0.0079	0.0002	0.0305
β6	0.0392	0.0200	0.0004	0.0769
β7	-0.0002	0.0008	-0.0020	0.0014
β8	0.7554	0.0852	0.5883	0.9194
β9	-0.0634	0.0813	-0.2193	0.0958
ζ	2.5763	0.3482	1.8857	3.2495

From the results in Table (1), it can be seen that the covariates month2 (quadratic effect of month), month3 (cubic effect of month), year2 (quadratic effect of year), lagged average minimum temperature and Y(t-1) (lagged total of the previous month) have significant effects on the log response (dengue count) as the 95% high posterior density intervals for the corresponding regression parameters do not contain the zero value. Figure (4) shows the graphs of the observed series and fitted by the model (Monte Carlo estimators of the mean responses) from where it is observed a good fit. The needed assumptions of the fitted "model 1" (normality and non correlated residuals are verified from normal probability plots and ACF (autocorrelation function) of the residuals presented in Figure (5) where it is observed that the needed assumptions are well verified for "model 1".

Assuming the volatility model introduced in Section 3 equations (2) and (3}) that is, "model 2", for the dengue series in São Paulo city on the logarithmic scale, consider the following prior distributions for the parameters of the model: $\phi v \sim U(0,1)$, $v = 1,2$; $\zeta \sim$ Gama(1,1); $\mu \sim N(0,1)$; $\beta 0 \sim N(3, 1)$; $\beta j \sim N(0,0.1)$, $j = 1,2,...,9$. Observe that the choice of the hyperparameters was based on a preliminary data analysis (especially for the regression parameters) and non-informative priors for the other parameters (ϕv , ζ and μ) where the prior where defined for the possible values of each parameter ($\zeta > 0$, $-\infty < \mu < \infty$, $0 < \phi 1 < 1$, $0 < \phi 2 < 1$). Some sensitivity analysis was made considering other hyperparameter values for the prior distributions, but the obtained inference results were very similar. Also, it is assumed prior independence between the parameters. Using the OpenBugs, the convergence of the simulation algorithm for the joint posterior distribution (Gibbs / Metropolis-Hastings) was verified from time series plots of the generated Gibbs samples.

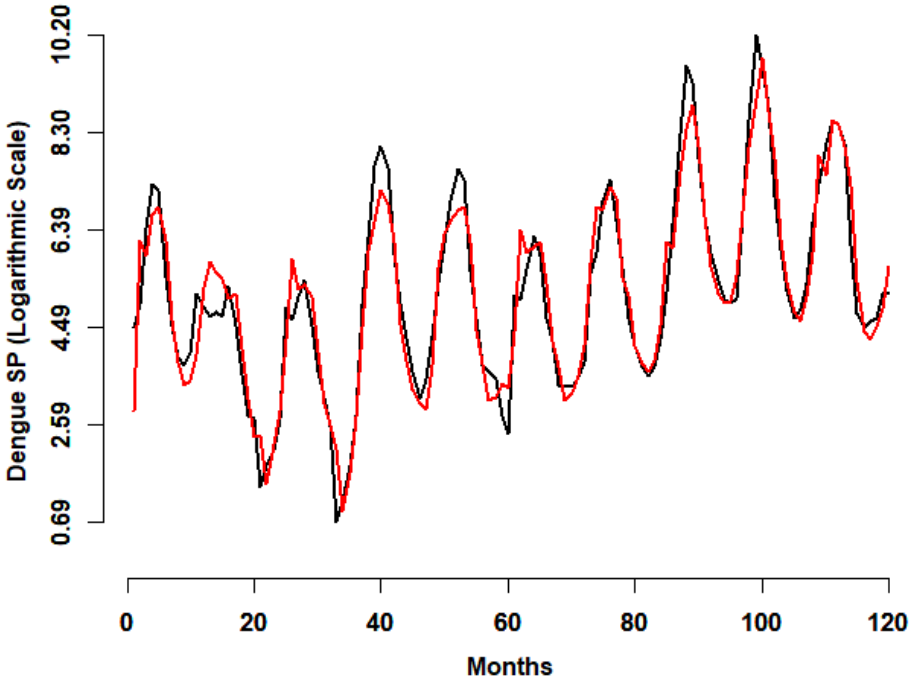

Figure 4: Observed counts and fitted model (logarithmic scale in "model 1").

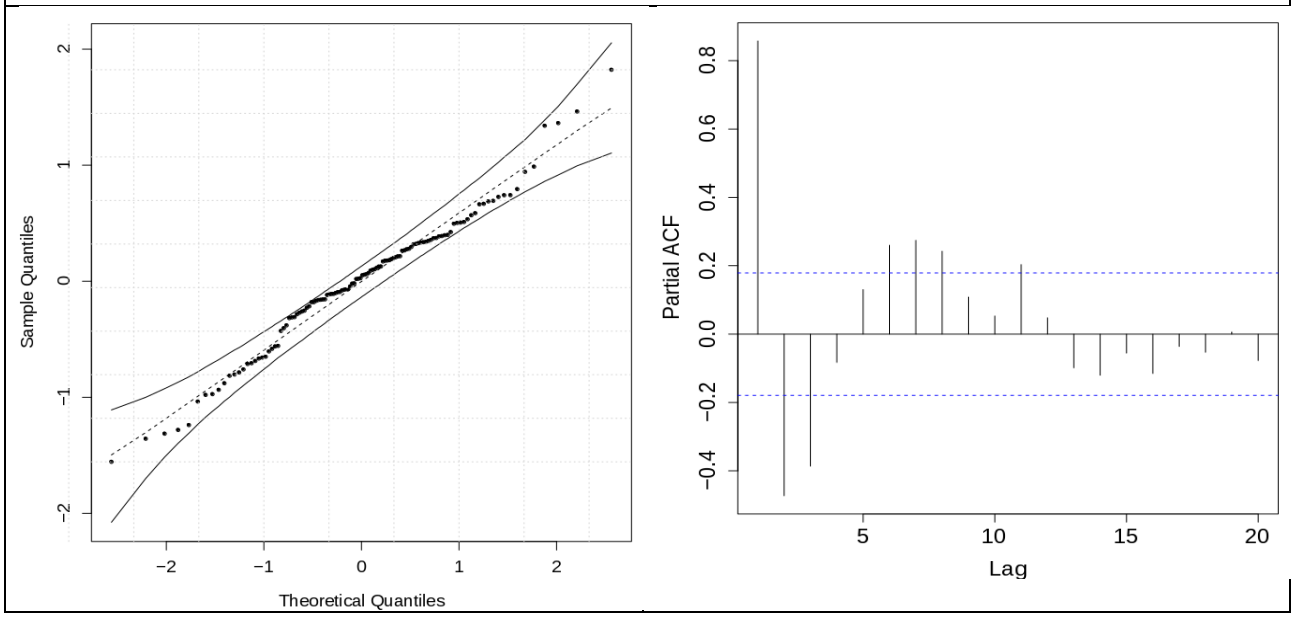

Figure 5: Normal probability plots and PACF (Autocorrelation Function) for the residuals (logarithmic scale in "model 1")

It was considered a "burn-in-sample" of size 111,000 samples discarded to eliminate the effect of the initial values in the iterative method; thereafter, we generated other 400,000 samples choosing each 100th sample

(a final sample of size 4000) to obtain the posterior summaries of interest. The posterior summaries of interest for each parameter of the model are shown in Table (2).

Table 2. Posterior summaries for "model 2" assuming dengue data in São Paulo city.

			95% HPD Int.	
	Mean	Std. Dev.	Lower	Upper
β0	2.2235	0.4908	1.2667	3.1813
β1	-0.1197	0.2157	-0.5270	0.3178
β2	-0.0490	0.0378	-0.1229	0.0236
β3	0.0042	0.0019	0.0006	0.0080
β4	-0.0370	0.0858	-0.2019	0.1342
β5	0.0109	0.0074	-0.0027	0.0260
β6	0.0289	0.0188	0.0043	0.0686
β7	-0.0002	0.0008	-0.0017	0.0013
β8	0.8205	0.0949	0.6367	1.0090
β9	-0.1149	0.0898	-0.2852	0.0651
μ	0.3014	0.3014	-1.7493	-0.6465
φ1	0.1820	0.1820	0.0015	0.6362
φ2	0.1527	0.1527	0.0001	0.4829
ζ	0.9603	0.9603	0.3873	3.5969

From the results of Table (2), it is observed that the covariates month3 (cubic effect of month), lagged average minimum temperature and Y(t-1) (previous count month) have significant effects on the log(response) as the 95% credibility intervals for the corresponding regression parameters do not contain the zero value. Figure (6) shows the graphs of the observed series and fitted by the model (Monte Carlo estimators of the mean responses) from where also it is observed a good fit of the model. Figure (6) also shows the graphs of the quadratic roots of the estimated volatilities. Despite some covariates did not show significant effects on the response log(dengue count), the inclusion of all covariates in the model is important to get better fit and better forecasts.

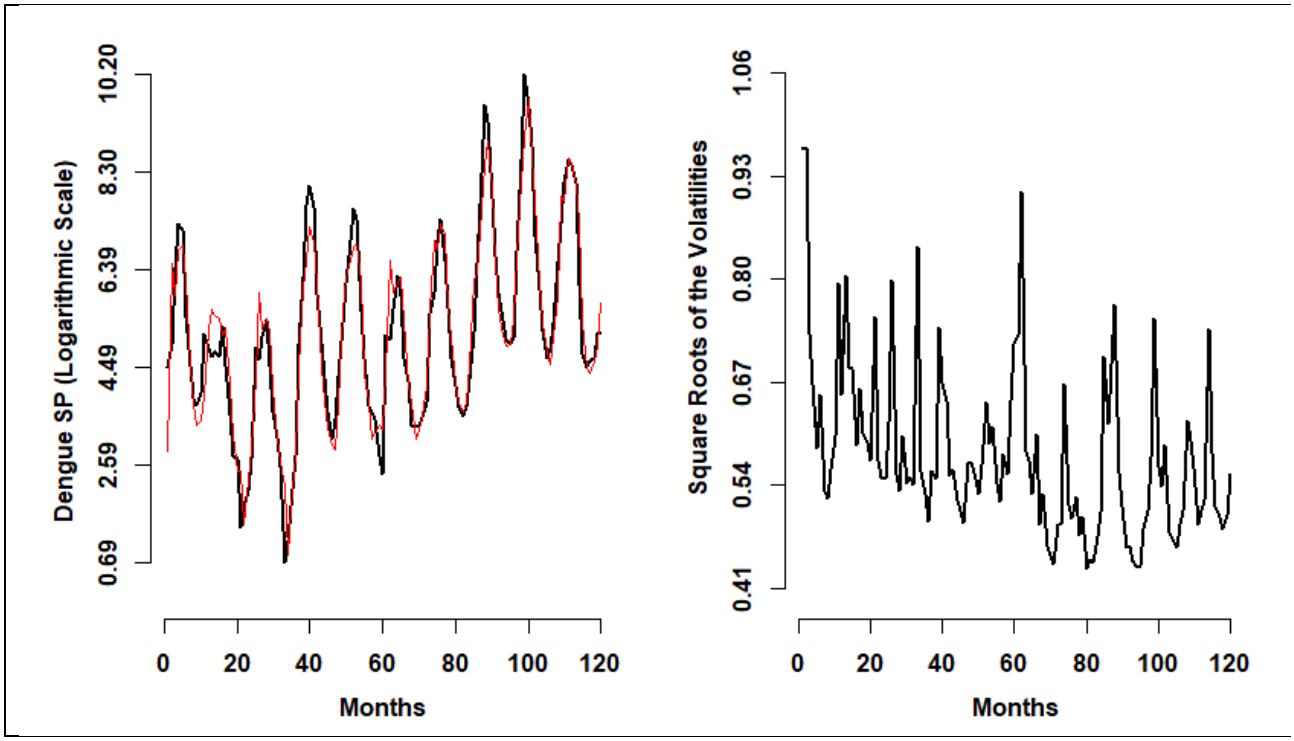

Figure 2. Observed countings, fitted model (logarithmic scale in "model 2") and square roots of the volatilities

The needed assumptions of the fitted "model 2" (normality and independence of the residuals) also were verified from normal probability plots and ACF (autocorrelation function) of the residuals (Figure (7)). From Figure (7), it is observed that the

needed assumptions are well verified for "model 2".

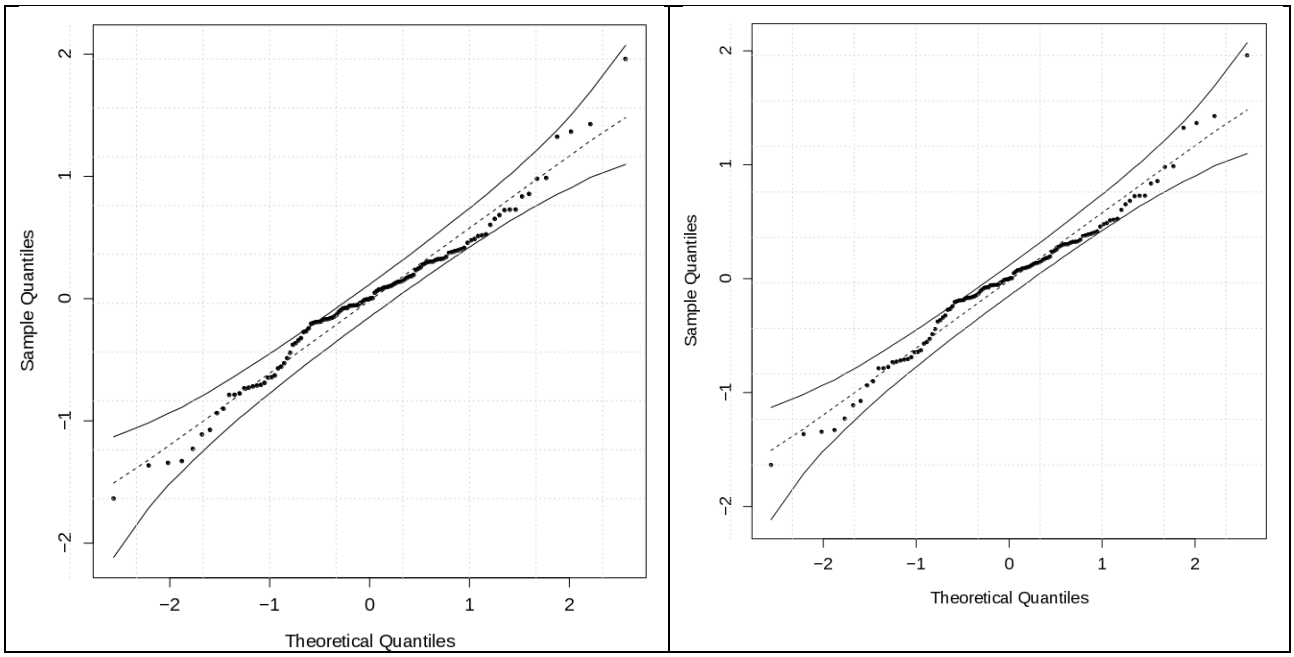

Figure 7: Normal probability plots and PACF (Autocorrelation Function) for the residuals (logarithmic scale in "model 2")

It is important to point out that other models in presence of volatilities could be considered in the data analysis. As a special case, it was also assumed ``model 2" in presence of cubic effect of year, but this model presented DIC (Deviance Information Criterion23) value equal to 347.4 that is larger to DIC value (equal to 262.0) of "model 2" not considering the presence of the cubic effect, an indication of better fit for the data of ``model 2" presented in (2). The deviance information criterion (DIC) is a hierarchical modeling generalization of the Akaike information criterion (AIC) (lower DIC values indicate better models).

Conclusions

Epidemics time series usually assumes standard times series models as MA (moving average) or ARIMA models to be fitted by the data and to get forecasts. Alternatively, some studies introduced in the literature, consider polynomial regression Bayesian models in presence of lagged effects, presence of other covariates and normal errors with a constant variance using MCMC simulation methods as assumed in "model 1". Despite the good fit using "model 1" (see Figure 4) for the dengue data set of São Paulo city, in some situations the public health researchers could be interested of the jointly modeling of the mean and the variance of the epidemiological time series. In this way, "model 2" considering a polynomial regression model in presence of lagged effects and some covariates combined with existing volatility models under a Bayesian approach showed that it is possible to get very accurate fit for the count dengue data in São Paulo city (and forecasts) as compared with existing time series models like ARIMA or moving average models. Another great advantage of the proposed model: the estimation of the volatilities in each time of the disease counts which could be of great interest for public health researchers. It is important to point out that the ARIMA or moving average time series models are exploratory statistical approaches where in each application it is needed to reformulate the model. The estimated proposed Bayesian model has great advantages since it is not needed to update the model in each time to be used to get forecasts of the dengue counts in future time and also permits the inclusion of important factors (covariates) that could affect the epidemic counts. Better inferences (point estimators and more accurate interval inferences for the forecasts of future dengue counts) are obtained assuming the proposed model. Other advantage of the proposed model: possibility to incorporate prior opinion of health experts in form of more informative priors which usually

implies in better inferences. As a final remark, it is important to point out that the proposed modeling approach could be used to any other epidemiology time series, like COVID-19, tuberculosis, flu or influenza times series among many others.

References

1.Organization, WH, for Research, SP, in Tropical Diseases, T., of Control of Neglected Tropical Diseases WHO, Epidemic WHO, Alert P. Dengue: guidelines for diagnosis, treatment, prevention and control. World Health Organization, 2009.

2. Cortes F, Martelli CM, de Alencar Ximenes RA, Montarroyos UR, Junior JBS, Cruz OG, Alexander N, de Souza WV. Time series analysis of dengue surveillance data in two Brazilian cities. Acta Tropica. 2018;182:190–197.

3. Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons, 2015.

4. Luz PM, Mende BV, Code¸co CT, Struchiner, CJ, Galvani AP. Time series analysis of dengue incidence in Rio de Janeiro, Brazil. The American Journal of Tropical Medicine and Hygiene. 2008;79(6):933–939.

5. Martinez EZ, Silva EASD. Predicting the number of cases of dengue infection in Ribeir˜ao Preto, S˜ao Paulo state, Brazil, using a SARIMA model. Cadernos de Saude Publica. 2011; 27:1809–1818.

6. Nobre FF, Monteiro ABS, Telles PR, Williamson GD. Dynamic linear model and SARIMA: A comparison of their forecasting performance in epidemiology. Statistics in Medicine. 2001; 20(20):3051-3069.

7. SilawanT,Singhasivanon P, Kaewkungwal J, Nimmanitya S, Suwonkerd W, et al. Temporal patterns and forecast of dengue infection in Northeastern Thailand. Southeast Asian Journal of Tropical Medicine and Public Health. 2008; 39(1):90.

8. Morato DG, Barreto FR, Braga JU, Natividade MS, Costa MDCN, Morato V, Teixeira MDGLC. The spatiotemporal trajectory of a dengue epidemic in a medium-sized city. Memorias do Instituto Oswaldo Cruz. 2015; 110(4):528–533.

9. Nelder JA, Wedderburn RW.Generalized linear models. Journal of the Royal Statistical Society:Series A (General). 1972; 135(3):370–384.

10. Dimitrakopoulos S, Kolossiatis M. Bayesian analysis of moving average stochastic volatility models: modeling in-mean effects and leverage for financial time series. Econometric Reviews.2020; 39(4):319–343.

11. Taspinar S, Dogan O, Chae J, Bera AK. Bayesian inference in spatial stochastic volatility models: An application to house price returns in Chicago. 2019. Available at SSRN 3104611.

12. Guo X, McAleer M, Wong WK, Zhu L. A Bayesian approach to excess volatility, short-term underreaction and long-term overreaction during financial crises. The North American Journal of Economics and Finance. 2017; 42:346–358.

13. Gelfand AE, Smith AF. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association. 1990; 85(410):398–409.

14. Danielsson J. Stochastic volatility in asset prices estimation with simulated maximum likelihood. Journal of Econometrics.1994; 64(1-2):375–400.

15. Yu J. Forecasting volatility in the New Zealand stock market. Applied Financial Economics. 2002; 12(3):193–202.

16. Engle RF. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society. 1982; 987–1007.

17. Ghysels E, Harvey A, Renault E. Stochastic Volatility.1996.

18. Kim S, Shephard N, Chib S. Stochastic volatility: Likelihood inference and compar- ison with ARCH models. The review of economic studies. 1998; 65(3):361–393.

19. Meyer R, Yu J. BUGS for a Bayesian analysis of stochastic volatility models. The Econometrics Journal. 2000; 3(2):198–215.

20. Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. WinBUGS user manual, version 1.4 MRC Biostatistics Unit. Cambridge, UK. 2003.

21. Draper NR, Smith H. Applied regression analysis, volume 326. John Wiley & Sons.1998.

22.Seber GAF, Wild CJ. Nonlinear regression. New Jersey, John Wiley & Sons.2003.

23. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. J R Stat Soc Ser B Stat Methodol. 2002; 64(4):583-639