



## ORIGINAL ARTICLE

## The Incidence of Tuberculosis in Brazil from 2001 to 2018: Use of Polynomial Regression Combined with a Stochastic Volatility Model

Jorge Alberto Achcar<sup>1</sup>, Ricardo Puziol de Oliveira<sup>2</sup> and Emerson Barili<sup>2\*</sup>

<sup>1</sup>Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil

<sup>2</sup>State University of Maringá, Maringá, PR, Brazil

\*Corresponding author: Emerson Barili, State University of Maringá, Maringá, PR, Brazil



### Abstract

Monthly tuberculosis (TB) counts from January 2001 to December 2018 showed an upward trend (data from SINAN - Information System for Notifiable Diseases). The increase in the incidence of TB in general is associated with an increase in the rate of extreme poverty, an increase in AIDS cases and other factors. Combining the polynomial linear regression and stochastic volatility models, the purpose of this study was to analyze monthly count data as well as the AIDS/TB, extreme poverty/TB and urban/TB rates associated with the incidence of TB. The results were obtained using the Monte Carlo method in Markov chains under a Bayesian approach. In analyzing the data with the presence of great volatility, despite the growth of TB cases in Brazil, some rates such as AIDS/TB and extreme poverty/TB are decreasing in the same period, indicating that TB cases have increased in the last few years for all sectors of the population of Brazil. The proposed model can be of great use for other epidemiological data, in addition to modeling TB counts and rates associated with TB.

### Keywords

TB count, Polynomial regression models, Stochastic volatility models, Bayesian methods, MCMC simulation

### Introduction

Tuberculosis (TB) is an infectious disease, caused by *Mycobacterium tuberculosis*. The disease is usually presented as pulmonary TB, but it can also affect other organs (extrapulmonary TB). Tuberculosis usually progresses slowly from the latent phase (infection without active disease, LTBI-Latent Tuberculosis Infection)

to active TB, except in HIV-infected patients where the progression can be rapid and fatal. Only a small proportion (5% to 10%) of people infected with *M. tuberculosis* progress to active TB as reported by the [1]. Despite great scientific advances, the evolution of treatment and the discovery of a vaccine have contributed to the control of TB, it is still among the infectious diseases of high lethality in the world. Global trends show that the incidence of tuberculosis (TB), prevalence and mortality rates are gradually decreasing worldwide. TB counts in all countries and regions of the world were in great decline until the early 1980's when the first cases of patients infected with the HIV virus leading to AIDS emerged and consequently leading to a large increase in the incidence of TB in the World. After the discovery of new therapies against HIV, TB counts started to drop in practically all countries. Tuberculosis remains a poverty-related disease with a high burden of disease in low- and middle-income countries and in countries with a high incidence of HIV [2].

In Brazil, especially with the great economic crisis from the years 2014 with the increase of unemployed workers and significant increase of individuals living in extreme poverty, a new trend of increase in TB cases has been observed, regardless of the increase in AIDS cases. Many studies related to the count of individuals diagnosed with TB have been published in the literature in Brazil and in other countries.

In this direction [3], analyzed some estimates made



**Citation:** Achcar JA, Oliveira RP, Barili E (2021) The Incidence of Tuberculosis in Brazil from 2001 to 2018: Use of Polynomial Regression Combined with a Stochastic Volatility Model. Int J Clin Biostat Biom 7:035. doi.org/10.23937/2469-5831/1510035

**Accepted:** April 19, 2021; **Published:** April 21, 2021

**Copyright:** © 2021 Achcar JA, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

by the WHO and indicators of incidence, mortality, occurrence of multidrug resistance and association with HIV for tuberculosis in Brazil and worldwide. These authors present estimates (8.7 million new cases per year in the world and 116,000 new cases per year in Brazil) or actually verified (Brazil in 2000 notified 82,249 new cases) pointing out to a serious public health situation, mainly in developing countries, which it requires energetic and effective measures for its control. In a country of high prevalence like Brazil, the actions for the discovery of new cases, associated with the measures of biosafety, are of interest to all health professionals, mainly those who work in large hospitals or emergencies.

DC Costa [4] analyzed several aspects that influence the pattern of tuberculosis mortality and morbidity: The variation in living conditions, the effects of chemotherapy and the dynamics of disease transmission which characterizes the indicators used in the measurement of tuberculosis that are built from prevalence surveys and notification data. The main indicators are: risk of infection, mortality, incidence coefficient of tuberculous meningitis, incidence coefficients of tuberculosis cases, especially pulmonary bacilliferous among others. In that study, data on the incidence of bacilliferous pulmonary cases in the Brazilian provinces of Pará and Rio Grande do Sul and mortality from tuberculosis in the capitals of Pará, Rio Grande do Sul, Pernambuco and Rio de Janeiro were analyzed, which reveal in general, a downward trend at that time, still that at a different pace.

Many other studies are presented in the literature. LMW Acosta, et al. [5] relate some social determinants and the incidence of TB in Porto Alegre, RS, Brazil. Several recent studies consider different statistical models that capture temporal or spatial effects in the analysis of TB epidemiological series [6-12].

Some studies on the incidence of TB have used ARIMA [13] time series models (autoregressive integrated

moving average models) or ARCH [14] models (autoregressive conditional heteroscedasticity) in the analysis of epidemiological series. In this sense, RR Sarkar, et al. [15] consider the application of different time series models for epidemiological data with application to malaria. YL Zheng, et al. [16] use ARIMA models and a combination of ARIMA-ARCH models for predicting and monitoring monthly tuberculosis morbidity in Xinjiang, China. In this case, the ARCH model is suggested to provide tuberculosis surveillance, providing estimates on tuberculosis morbidity trends in Xinjiang, China. Other applications in epidemiology can also be seen in [17-25].

In the present study, a new formulation of statistical modeling (combination of polynomial linear regression models with stochastic volatility models) is proposed for the analysis of TB count data and associated rates using Bayesian methods of statistical inference. The article is formulated as follows: in section 2 the monthly data on TB in Brazil for the period 2001 to 2018 are presented; section 3 presents the material and methods; section 4 presents the obtained results; finally, section 5 presents conclusions and discussion of the obtained results.

### TB Data in Brazil for the Period 2001-2018

This study considers TB count data in Brazil between the years 2001 and 2018 and some existing relationships given by other time series in the same period such as rates of AIDS patients among patients diagnosed with TB, rates of patients living in urban centers among patients diagnosed with TB, rates of male patients among patients diagnosed with TB, rates of patients in extreme poverty among patients diagnosed with TB among several other factors. Figure 1(a) shows a graph of the yearly time series of TB patient counts in Brazil (data from the SINAN-Information System for Notifiable Diseases-<http://sinan.saude.gov.br/sinan/login/login.jsf>) for the period between January 2001 and December

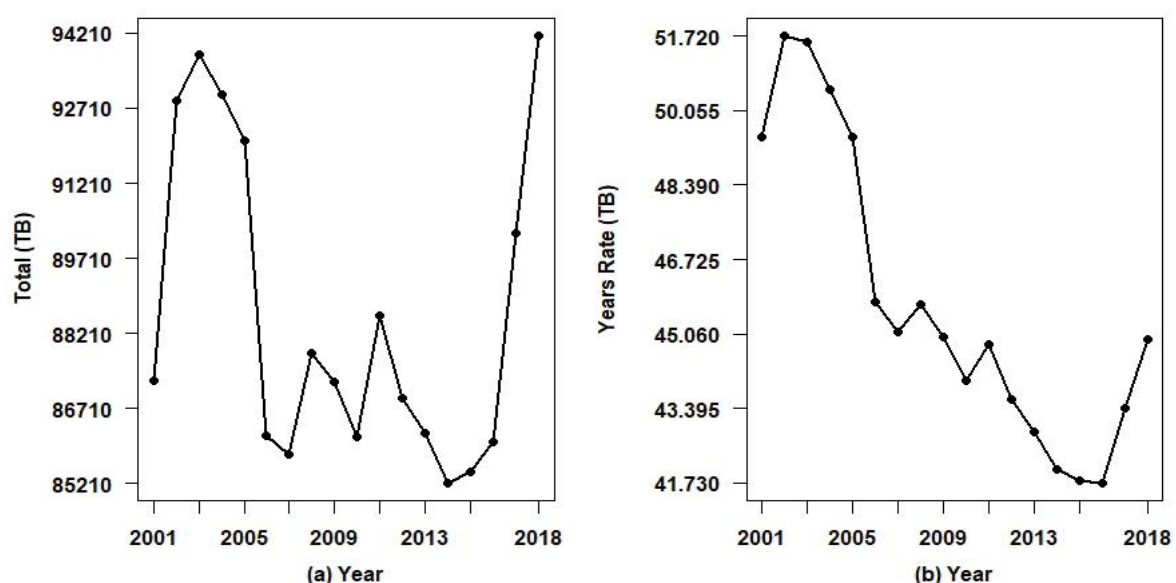


Figure 1: Yearly counts of TB patients in Brazil and yearly rates per 100,000.

2018 (data set in [Appendix 1](#) at the end of the paper). [Figure 1\(b\)](#) also shows the rates (number of TB cases in the year divided by the estimates of Brazil's population in the year, source IBGE). The data are presented in [Table 1](#), from where it is observed an increase in TB rates between the years 2001 and 2003; from 2003 onwards there was a decline in yearly TB rates until 2016, possibly due to the efficacy of drugs to control AIDS; after that year, an increase in TB rates is observed again, perhaps as a result of the great economic crisis affecting Brazil (increase in extreme poverty, increase in the

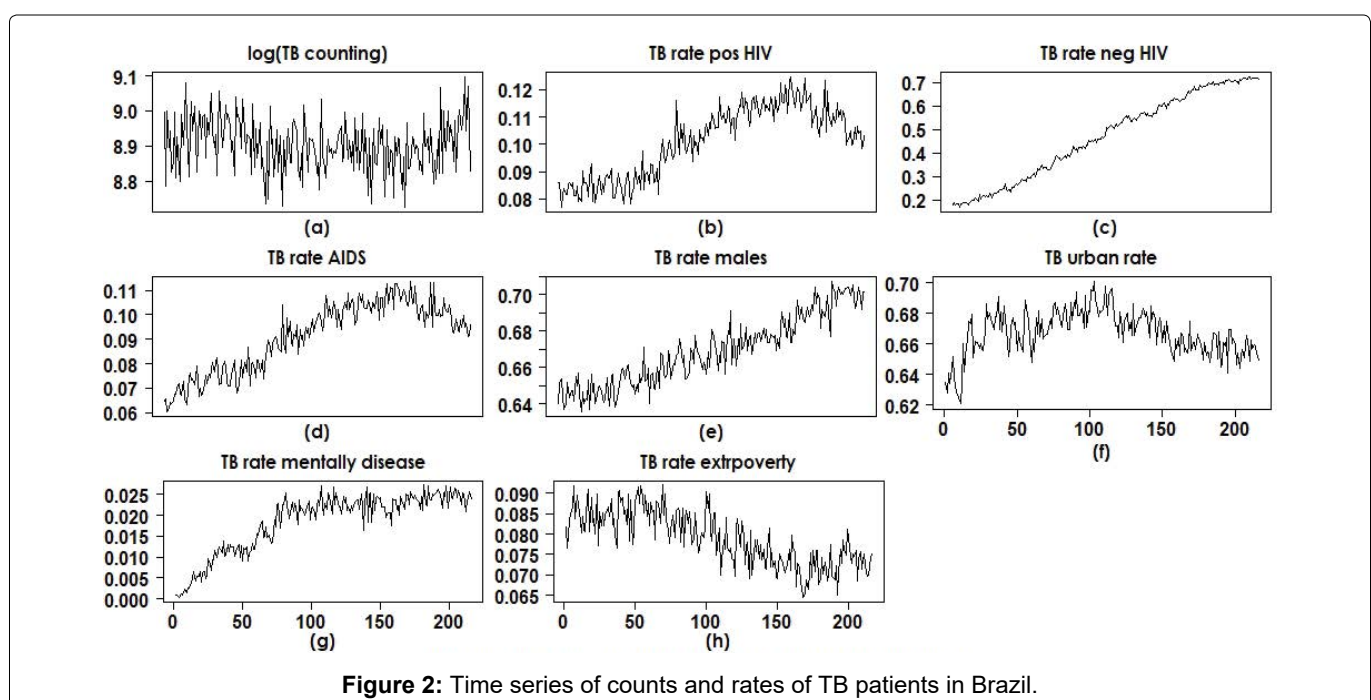
unemployed rates, lower investments in health, among many other factors). From [Figure 1](#), it is observed two change-points: One change-point around 2003 and a second change-point close to the year 2017.

[Figure 2](#) shows the monthly TB counts in Brazil (data from SINAN) from January 2001 to December 2018 (216 months). From the graphs presented in [Figure 2](#), it can be seen that:

**Table 1:** Yearly counts of TB patients in Brazil and yearly TB rates.

Year	Total TB	Estimated pop BR	Yearly TB rate/100,000
2001	87265	177196054	49.25
2002	92859	179537520	51.72
2003	93773	181809246	51.58
2004	92980	184006481	50.53
2005	92056	186127103	49.46
2006	86160	188167356	45.79
2007	85779	190130443	45.12
2008	87822	192030362	45.73
2009	87227	193886508	44.99
2010	86146	195713635	44.02
2011	88560	197514536	44.84
2012	86905	199287299	43.61
2013	86208	201035912	42.88
2014	85213	202763739	42.03
2015	85443	204471769	41.79
2016	86026	206163053	41.73
2017	90210	207833823	43.40
2018	94155	209469323	44.95

- The monthly counts of patients diagnosed with TB in the period between January 2001 and December 2018 (original scale and logarithmic scale) show seasonality for the seasons and there is clearly a yearly decrease in counts until the year 2014 where the beginning of an increase in TB cases is observed.
- It is also observed that until the mid-2014 there is a tendency to increase the rates of AIDS patients among TB patients; in that year there is a beginning of a decrease in the rates of AIDS patients among those infected with tuberculosis. This is also seen for rates of HIV-positive patients among TB patients. It is important to note that the incidence of TB was in great decline worldwide until the mid-1980s, the year of the beginning of the AIDS epidemic, a factor that led to the reappearance of TB in large numbers. From these graphs it is possible to conclude that there is a strong indication of the increase in TB in Brazil for all segments of the population regardless of the effect of HIV, as the rates of HIV-infected patients diagnosed with TB are declining, despite the increase in TB cases.
- It is also observed that the rates of urban patients among those infected with TB had an annual increase until mid-2012; as of that year, there has been a decline in the number of patients living in



**Figure 2:** Time series of counts and rates of TB patients in Brazil.

urban centers (people most susceptible to HIV infection) among TB patients. This also shows that the incidence of TB is gradually increasing for all segments of the population in Brazil, a result of great serious impact to the Brazilian health authorities. The same is observed for patients in extreme poverty (annual decrease in the rates of individuals in extreme poverty among TB patients).

Figure 2 shows, in most cases, for the data analysis and statistical data modeling, the need for polynomial linear regression models in the presence of the covariate (independent variable) month (temporal order) that can 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 covariate month squared or cube, and not the regression coefficient, the model remains a linear regression model [26,27]. An application of polynomial regression models with epidemiology data is given by [28]. The presence of a quadratic term in the model creates an inverted U- or U-shaped curve, as seen in the graphs in Figure 2(b), Figure 2(d), Figure 2(f), Figure 2(g) and Figure 2(h). A cubic term has two distinct parts: one facing up and one facing down, that is, the curve goes down, back up and back again (or vice versa). This is also seen in some graphs in Figure 2(a), Figure 2(c) and Figure 2(e).

The main goals of the study are:

- To model the monthly TB counts in Brazil with data on the logarithmic scale from 2001 to 2018, assuming stochastic volatility models to estimate the time series volatilities;
- Also to model some rates of factors associated with the incidence of TB using stochastic volatility models such as rates of patients diagnosed with TB and AIDS, rates of patients diagnosed with TB and urban dwellers and rates of patients diagnosed with TB and belonging to the socioeconomic groups of extreme poverty.
- To study the variability in the rates of patients diagnosed with TB and AIDS, rates of patients diagnosed with TB and residents of urban centers, rates of patients diagnosed with TB and belonging to extreme poverty socioeconomic groups that are decreasing despite the growth of TB cases seen in Brazil in recent years.
- To model yearly TB rates in Brazil from 2001 to 2018, assuming linear regression models in the presence of change-points.

## Material and Methods

### Use of standard linear regression models

In a first data analysis, it is assumed a standard poly-

nomial linear regression model given by:

$$Y(t) = \beta_0 + \beta_1 \text{months}(t) + \beta_2 [\text{months}(t)]^2 + \beta_3 [\text{months}(t)]^3 + \epsilon(t) \quad (1)$$

for  $t = 1, 2, \dots, 216$  (months) where  $\epsilon(t)$  are noises considered independent and identically distributed random variables with a normal distribution  $N(0, \sigma_\epsilon^2)$ . Under a classic approach, the regression parameters are usually estimated using the least squares method (EMQ). An alternative of epidemiological interest would be to model the series not only to estimate monthly averages, but also to estimate monthly variances (volatilities) that are of interest to the public health area, possibly relating these volatilities to the occurrence of non-observed factors associated with the months (season, unemployment rate, extreme poverty rate, among others). For this purpose, time series models are used in the data analysis, which simultaneously estimate the monthly average and the monthly volatility.

### Use of polynomial regression models and stochastic volatility for TB data in Brazil

Stochastic volatility (SV) models have been widely used to analyze financial time series [29] as a powerful alternative to some existing regression models in the literature, such as the ARCH (conditional heteroscedastic autoregressive) models introduced by [30] and the generalized autoregressive conditional heteroscedastic models (GARCH) introduced by [31], but not widely used in the health field [32-34]. In one of these few health applications, [35] studied neural signals using a database of medial temporal lobe (MTL) recordings from 96 neurosurgical patients, using time series models with volatility described by a multivariate stochastic latent variable process and lagged interactions between signals in different brain regions providing new insights into the dynamics of brain function. In the financial area, these models are considered for modeling the logarithms of financial returns between current data and previous data (can be hours, days or months) without taking into account the modeling of the means depending on covariates.

The novelty of this study is the introduction of stochastic volatility models for epidemiological time series, in particular, for tuberculosis counting series in Brazil.

For the definition of the models,  $N \geq 1$  is a fixed integer that records the amount of observed data (in our case, it will represent the monthly TB counts on the logarithmic scale and also the rates of patients in certain groups among those infected with TB).

In the presence of heteroscedasticity, that is, variances depending on time  $t$ , assume that the time series  $Y(t)$ ,  $t = 1, 2, \dots, N$  assume the combined polynomial linear regression model with the stochastic volatility model for the analysis of TB count data on the logarithmic scale and some rates related to the TB epidemic:



$$Y(t) = \beta_0 + \beta_1 \text{months}(t) + \beta_2 [\text{months}(t)]^2 + \beta_3 [\text{months}(t)]^3 + \sigma(t)\epsilon(t) \quad (2)$$

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  $\sigma_j(t)$  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  $\sigma_\epsilon^2 e^{h(t)}$  where  $h(t)$  depends on an unobserved latent variable.

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

The mean of  $Y(t)$  is given by  $E[Y(t)] = \beta_0 + \beta_1 \text{months}(t) + \beta_2 [\text{months}(t)]^2 + \beta_3 [\text{months}(t)]^3$  since  $E[\sigma(t)\epsilon(t)] = 0$ .

The variance of  $Y(t)$  is given by  $\text{var}[Y(t)] = \text{var}[\sigma(t)\epsilon(t)] = \sigma^2(t)$  since we are assuming that  $\text{var}[\epsilon(t)] = \sigma_\epsilon^2 = 1$ .

To analyze the data set, it is introduced a latent variable (unobserved variable) defined by an auto-regressive model AR (2), for  $t = 1, 2, 3, \dots, N$  ( $N = 216$  months).

$$h(1) = \mu + \zeta(1), t = 1,$$

$$h(2) = \mu + \varphi_1[h(1) - \mu] + \zeta(2) \quad (3)$$

$$h(t) = \mu + \varphi_1[h(t-1) - \mu] + \varphi_2[h(t-1) - \mu] + \zeta(t), t = 3, 4, \dots, N,$$

where  $\zeta(t)$  is a noise considered to be independent and identically distributed random variables with a normal distribution  $N(0, \sigma_\zeta^2)$ , which is associated to a latent variable  $h(t)$ . The quantities  $\sigma_\zeta^2$ ,  $\mu$ ,  $\varphi_1$  and  $\varphi_2$  are unknown parameters that should be estimated ( $0 < \varphi_1 < 1$ ,  $0 < \varphi_2 < 1$ ).

Bayesian inference procedures, based on Markov Chain Monte Carlo simulation methods (MCMC) [34,36,37] 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 dimensionality and likelihood function without closed form among other factors.

For a Bayesian analysis of the model (1), we assume that the prior distributions for the parameters  $\mu$ ,  $\varphi_v$  and  $\zeta = 1/\sigma_\zeta^2$ ,  $v = 1, 2$  are respectively given by the normal distribution, 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  $d/e^2$ . The hyperparameters  $a, b, c, d, e$  are assumed to be known and specified previously. Also assume that the regression parameters  $\beta_j$  have independent normal distributions with known hyperparameters,  $j = 0, 1, 2, 3$ .

### Change-points for the yearly rate series of TB in Brazil

To complete the statistical data analysis of the TB

counting in Brazil for the period 2001-2018 we know consider the TB yearly counting in place of the month TB counting. Considering the yearly rates of TB in Brazil, we can see (Table 1 and Figure 1) the presence of two change-points around the years 2003 and 2016 (upward rate until 2003, downward rate between the year 2003 and year 2016, rising rate after year 2016. In this case we also can use a linear regression Bayesian model to be fitted by the data in the presence of two points of change. In this case it is assumed the linear regression model,

$$Y(t) = \beta_0[j] + \beta_1[j]\text{years}(t) + \epsilon(t) \quad (4)$$

for  $t = 1, 2, \dots, N$  and  $j = 1, 2, 3$  where  $\beta_0[1]$  and  $\beta_1[1]$  are regression parameters for the model (data set in logarithmic scale) corresponding to  $\text{time}[t] \leq \zeta_1$  ( $\zeta_1$  is the first change-point);  $\beta_0[2]$  and  $\beta_1[2]$  are regression parameters for the model (data set in logarithmic scale) corresponding to  $\zeta_1 < \text{time}[t] \leq \zeta_2$  ( $\zeta_2$  is the second change-point);  $\beta_0[3]$  and  $\beta_1[3]$  are regression parameters for the model (data set in logarithmic scale) corresponding to  $\text{time}[t] \geq \zeta_2$  and  $\epsilon_j(t)$  is the error assumed to be independently and identically distributed random variables with a normal distribution  $N(0, \sigma_\epsilon^2[j])$ ,  $j = 1, 2, 3$  where  $\sigma_\epsilon^2[1]$  for  $\text{time}[t] \leq \zeta_1$ ,  $\sigma_\epsilon^2[2]$  for  $\zeta_1 < \text{time}[t] \leq \zeta_2$  and  $\sigma_\epsilon^2[3]$  for  $\text{time}[t] \geq \zeta_2$ . In the Bayesian analysis, discrete uniform prior distributions are assumed for the change-points  $k_1$  and  $k_2$  concentrated around the years 2003 and 2016; Gamma(1,1) prior distributions for the parameters  $\zeta_j = 1/\sigma_\epsilon^2[j]$ ,  $j = 1, 2, 3$ ;  $\beta_0[1] \sim N(-7,6;1)$ ,  $\beta_0[2] \sim N(-7,7;1)$ ,  $\beta_0[3] \sim N(-7,75;1)$ ; and  $\beta_1[j] \sim N(0;0,01)$ . We further assume prior independence.

### Results

From a preliminary analysis of the data, it is observed that in addition to the count times series for the monthly number of TB in Brazil, other series are related for the data analysis of the monthly counting of TB in Brazil from January 2001 to December 2018. In this study we consider four series in the data analysis:

- Time series for the logarithm of the TB count versus time order,
- Time series for the rate AIDS/TB versus time order,
- Time series for the rate urban patients/TB versus time order,
- Time series for the rate of patients in extreme poverty/TB versus time order.

These four series are analyzed in this section. Initially, polynomial regression models are considered in the form (1) considering constant variances and classic least squares (LSE) estimation techniques. We consider four regression models denoted by,

- Model M1: The regression model (1) associated with the logarithms of TB counts versus time order,

- Model M2: Associated with rates of patients with AIDS among TB patients versus time order,
- Model M3: Associated with rates of urban patients among TB patients versus time order,
- Model M4: Associated with the rates of patients living in extreme poverty among TB patients versus temporal order.

### Use of a standard polynomial regression model

Table 2 shows the LSE estimators for the regression parameters of each proposed regression model obtained using Minitab® software version 17. The necessary assumptions (normality and constant variance of the residuals were verified from the residuals graphics).

### Use of polynomial regression models and stochastic volatility

Assuming the volatility model introduced in section 2 (equations (1) and (2)) for the four series above, also denoted here by model M1 (log (TB counts)), model M2 (AIDS / TB rate), model M3 (rate (urban patients)/TB) and model M4 (rate (extreme poverty patients)/TB), let us consider the following prior distributions for the parameters of each model:

- Model M1:  $\varphi_v \sim U(0,1)$ ,  $v = 1,2$ ;  $\zeta \sim \text{Gamma}(1,1)$ ;  $\mu \sim N(0,1)$ ;  $\beta_0 \sim N(9,100)$ ;  $\beta_j \sim N(0,1)$ .
- Model M2:  $\varphi_v \sim U(0,1)$ ,  $v = 1,2$ ;  $\zeta \sim \text{Gamma}(1,1)$ ;  $\mu \sim N(0,1)$ ;  $\beta_0 \sim N(0.07, 0.01)$ ;  $\beta_j \sim N(0, 0.01)$ .
- Model M3:  $\varphi_v \sim U(0,1)$ ,  $v = 1,2$ ;  $\zeta \sim \text{Gamma}(1,1)$ ;  $\mu \sim N(0,1)$ ;  $\beta_0 \sim N(0.63, 0.01)$ ;  $\beta_j \sim N(0, 0.01)$ .
- Model M4:  $\varphi_v \sim U(0,1)$ ,  $v = 1,2$ ;  $\zeta \sim \text{Gamma}(1,1)$ ;  $\mu \sim N(0,1)$ ;  $\beta_0 \sim N(0.082,0.01)$ ;  $\beta_j \sim N(0, 0.01)$ .

Further, let us assume prior independence among the parameters. The choice of the hyperparameters for the prior distributions of the regression parameters was made based on previous knowledge obtained from the classical regression models for the regression model (1) assuming normal errors without the presence of stochastic volatility (LSE given in Table 2). In this way, empirical Bayesian methods were used [35].

In the simulation of samples for the joint posterior distribution of interest, it was used the OpenBUGS software [36], which simplifies the computational work, since this software only requires the definition of the likelihood function and the prior distributions for the pa-

rameters of the model. Thus, the conditional posterior distributions for each parameter needed for the Gibbs and Metropolis-Hastings algorithms are not presented in this study. The convergence of the sample simulation algorithm of the joint posterior distribution of interest (Gibbs/Metropolis-Hastings algorithms) was verified from time series plots of the generated Gibbs samples. It was considered a “burn-in-sample” with 51,000 samples to eliminate the effect of the initial values in the iterative method to simulate Gibbs samples for the M1 model; after this, it was generated other 120,000 samples choosing each tenth sample which gives a final sample size of 1200 to be used to obtain the posterior summaries of interest. Similarly, for models M2, M3 and M4, it was considered a “burn-in-sample” with 111,000 samples; after that, it was generated other 100,000 samples choosing each tenth sample totaling a final sample size of 1,000 to be used to obtain the posterior summaries of interest. The posterior summaries of interest (posterior means, posterior standard deviations and 95% credibility intervals) for each parameter for the four models are shown in Table 3.

From the results presented in Table 3, it is observed that the covariate month has significant effects on the response in terms of linear effect of month (regression parameter  $\beta_1$ ), quadratic effect of month ( $\beta_2$ ) and cubic effect of month ( $\beta_3$ ) (intervals corresponding 95% credibility does not contain a zero value) for models M3 (rate (urban patients)/TB) and M4 (rate (extreme poverty patients)/TB); for the M1 model (log (TB rates)) there is a significant effect of the quadratic and cubic terms and for the M2 model (AIDS/TB rate), there is a significant effect of the linear and cubic terms.

Figure 3 shows the graphs of the observed and the fitted series assuming the models (Monte Carlo estimators of the responses mean) for the four series. A good fit is observed in all cases. Figure 4 shows graphs of the quadratic roots of the estimated volatilities.

It is important to point out that from the results obtained from Table 2 and Table 3 (LSE and Bayesian estimators), it can be seen that the LSE and Bayesian estimators have close point values where in most cases related to each regression parameter, we have concordance of p-value < 0.05 and credible intervals not containing the zero value. Only for the parameter  $\beta_1$  (linear effect) related to models M1 (log (TB counts) versus time), and M2 (AIDS/TB rate versus time) it is observed significance under the classic approach but no significance under the Bayesian approach and for the parameter  $\beta_2$  (quadratic effect) for model M4 (rate (extreme poverty patients)/TB versus time), we have no significance under the classic approach and significance under the focus Bayesian we have different inferential results in terms of significant effects.

**Table 2:** LSE for the four assumed models.

Models	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$
M1	8.9200	0.00107 <sup>(*)</sup>	-0.00002 <sup>(*)</sup>	< 0.00001 <sup>(*)</sup>
M2	0.06678	0.000052 <sup>(*)</sup>	0.000004 <sup>(*)</sup>	< 0.00001 <sup>(*)</sup>
M3	0.6345	0.001222 <sup>(*)</sup>	-0.00001 <sup>(*)</sup>	< 0.00001 <sup>(*)</sup>
M4	0.0817	0.08174 <sup>(*)</sup>	< 0.00001	< 0.00001 <sup>(*)</sup>

<sup>(\*)</sup>significative at 5% (p-value < 0.05).

**Table 3:** Posterior summaries for the assumed models.

Model		Mean	SD	Lower 95%	Upper 95%
M1	$\beta_0$	8.932	0.02269	8.8860	8.9740
	$\beta_1$	0.0008	0.00085	- 0.00073	0.002499
	$\beta_2$	- 0.000019	0.0000089	- 0.000037	- 0.000004
	$\beta_3$	0.00000008	0.000000027	0.00000003	0.00000013
	$\mu$	- 5.058	0.6944	- 5.5500	- 2.7900
	$\phi_1$	0.1775	0.1411	0.005846	0.5415
	$\phi_2$	0.3424	0.2518	0.01661	0.9020
	$\zeta$	4.251	1.575	1.8620	8.1220
M2	$\beta_0$	0.06851	0.001675	0.06574	0.07243
	$\beta_1$	-0.000002	0.00006	-0.000138	0.000103
	$\beta_2$	0.0000043	0.00000061	0.0000032	0.0000057
	$\beta_3$	- 0.00000002	0.000000002	- 0.00000001	- 0.00000001
	$\mu$	- 4.259	1.036	- 6.3700	- 2.2760
	$\phi_1$	0.6631	0.1781	0.3372	0.9741
	$\phi_2$	0.3359	0.1778	0.02611	0.6653
	$\zeta$	2.168	0.8817	1.001	4.358
M3	$\beta_0$	0.637	0.004801	0.6281	0.6477
	$\beta_1$	0.001195	0.000151	0.000857	0.000289
	$\beta_2$	-0.0000099	0.00000142	-0.000012	-0.0000068
	$\beta_3$	0.00000002	0.000000004	0.000000014	0.00000003
	$\mu$	-4.203	1.044	-6.115	-2.182
	$\phi_1$	0.6546	0.2077	0.252	0.9783
	$\phi_2$	0.3427	0.2079	0.0218	0.7514
	$\zeta$	2.324	0.8576	1.076	4.418
M4	$\beta_0$	0.08113	0.00176	0.07796	0.08485
	$\beta_1$	0.000243	0.0000622	0.000102	0.000357
	$\beta_2$	-0.0000036	0.0000006	-0.0000046	-0.0000022
	$\beta_3$	0.00000001	0.000000002	0.000000007	0.000000014
	$\mu$	-4.359	0.9985	-6.237	-2.372
	$\phi_1$	0.6784	0.1938	0.3097	0.9825
	$\phi_2$	0.3219	0.1943	0.01728	0.6966
	$\zeta$	2.552	0.8985	1.172	4.655

**Table 4:** Posterior summaries for a model with two change-points.

	Mean	SD	Lower 95%	Upper 95%
$\beta_0[1]$	-7.59	0.7826	-9.196	-5.954
$\beta_0[2]$	-7.585	0.235	-8.075	-7.125
$\beta_0[3]$	-7.718	0.8682	-9.433	-6.032
$\beta_1[1]$	-0.001627	0.09955	-0.194	0.197
$\beta_1[2]$	-0.01256	0.02407	-0.06036	0.03539
$\beta_1[3]$	-0.00085	0.05843	-0.1162	0.1135
k1	2.104	0.9501	1.0	4.0
k2	16.71	0.976	15.0	18.0
$\zeta_1 = 1/\sigma_\epsilon^2[1]$	1.327	1.161	0.06158	4.355
$\zeta_2 = 1/\sigma_\epsilon^2[2]$	7.317	2.771	2.911	13.58
$\zeta_3 = 1/\sigma_\epsilon^2[3]$	1.744	1.375	0.1363	5.238

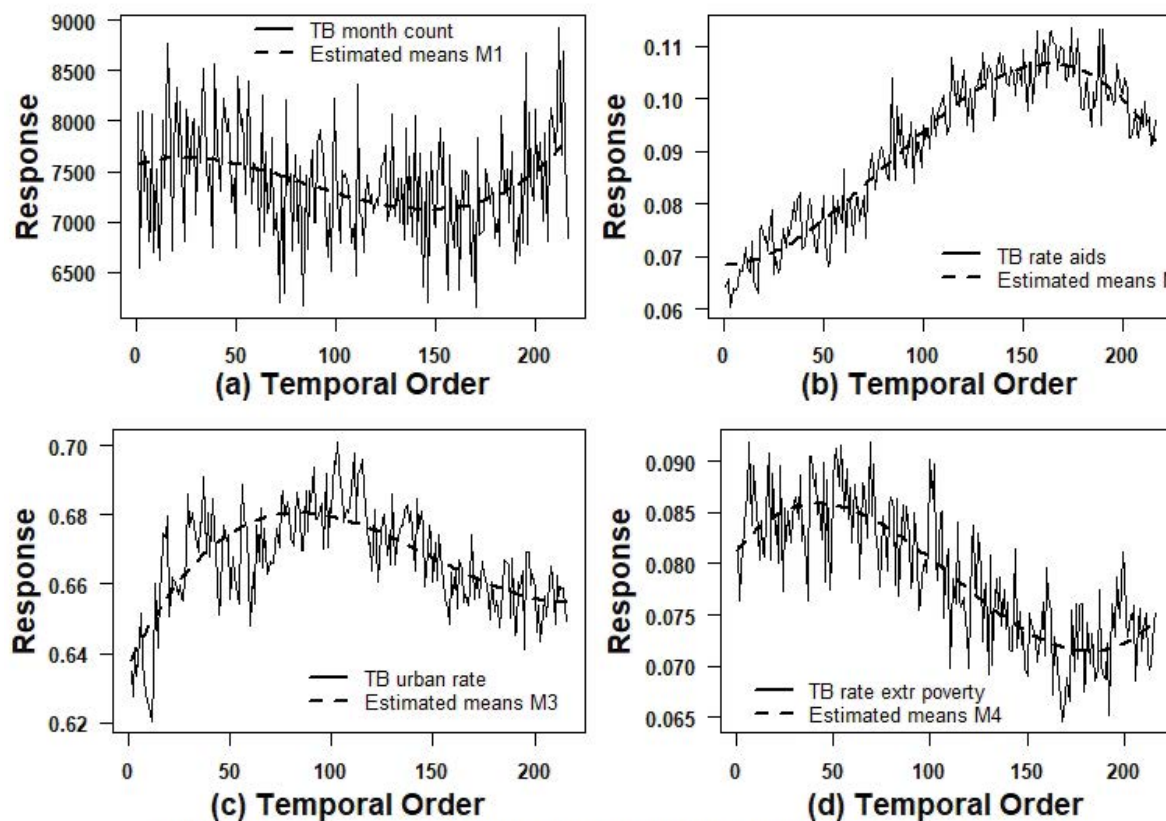


Figure 3: Graphs of the observed series and estimated means by the models.

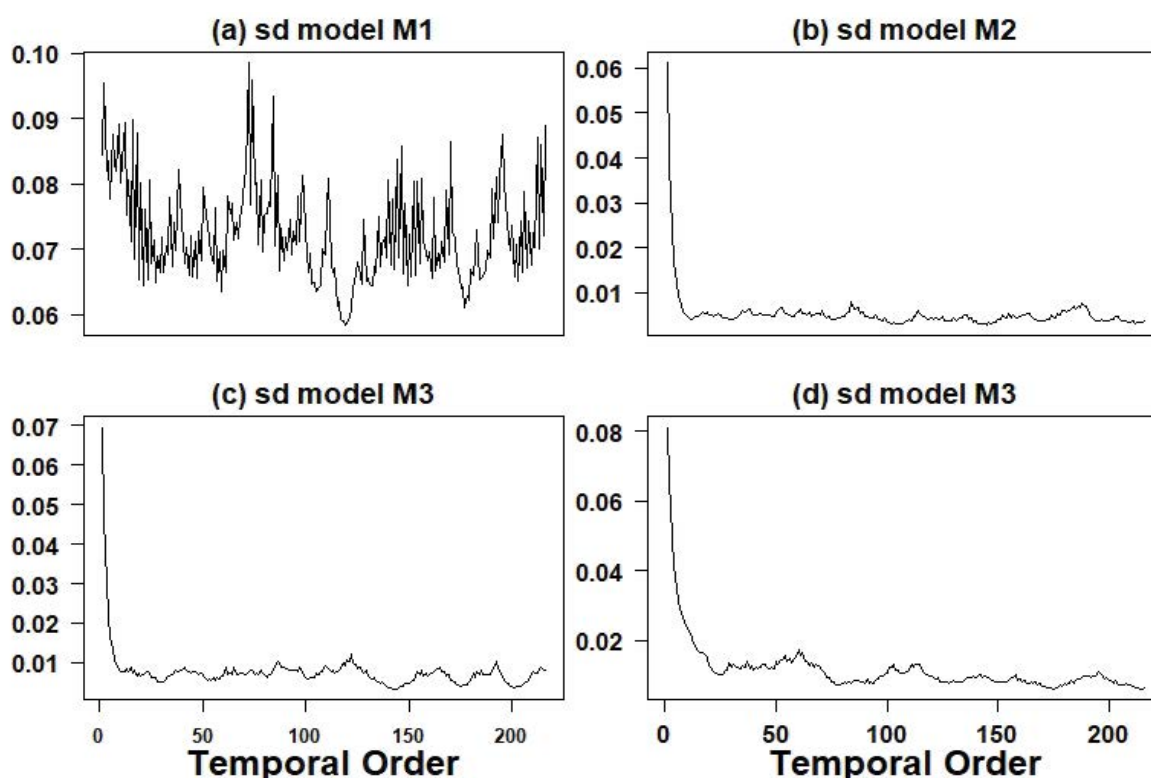


Figure 4: Graphs of the quadratic roots of the estimated volatilities.

### Use of a change-point model for the yearly rate series of TB in Brazil

Table 4 presents the posterior summaries obtained using the OpenBUGS software (burn-in sample = 1,000;

10,000 additional Gibbs samples). Figure 5 shows the fitted model (presence of two change-points, one change-point around the year 2003 which corresponds to the  $k_1$  estimated by 2.104 and a second change-point in the year 2016 which corresponds to the  $k_2$  estimated





Figure 5: Fitted model (presence of two change-points).

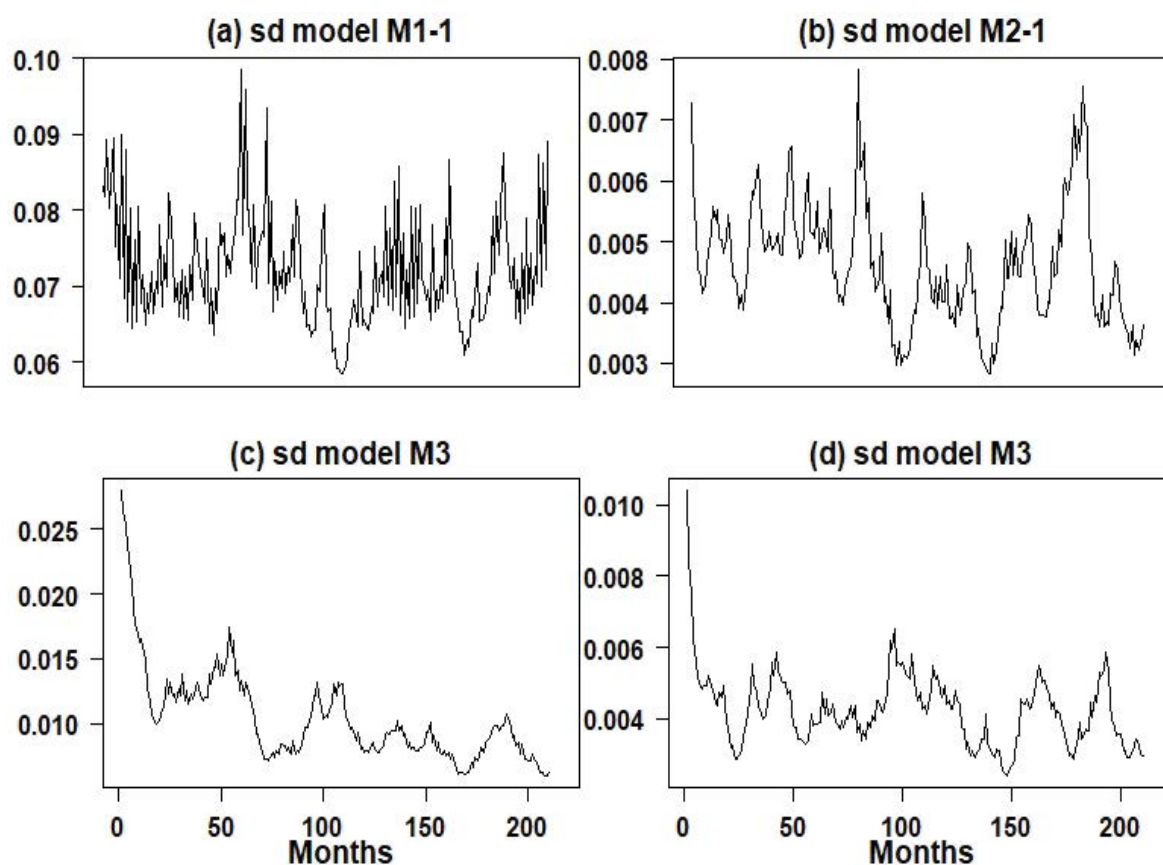


Figure 6: Graphs of the quadratic roots of the estimated volatilities (July 2001 to December 2018) for the four models.

by 16.71). The Openbugs program code is presented in [Appendix 2](#) at the end of the article.

### Discussion of the Obtained Results and Concluding Remarks

From the obtained results using model (2), it is observed a good fit of the models to the four data sets denoted here by models M1, M2, M3 and M4. Regarding to the estimated variability (volatility) of the data

presented in [Figure 6](#) for the period from July 2001 to December 2018, we have:

- For the TB count series considered on the logarithmic scale (model M1), there is great variability around the year 2007; from that year on, there is a decline in variability to be increased again especially at the end of the period 2001/2018.
- For the AIDS/TB rate series (model M2), there is

great variability around the year 2007/2008 and 2016/2017, an indication that the AIDS/TB rate was in transition, indicating that the number TB cases were increasing regardless of whether a person have AIDS or not. It is a worrying result, as it indicates that TB is increasing in Brazil regardless of the increase in AIDS cases.

- For the urban patient/TB rate series (model M3), it is observed that the variability is declining; with this result and the lower average of urban patient/TB rates, there is an indication that the TB epidemic is increasing throughout Brazil regardless of being a resident of large cities.
- For the extreme poverty/TB rate series (model M4), it is observed that the variability is approximately stable regardless of the great seasonality with a slight decline, also indicating that the number of TB cases is increasing in Brazil, regardless the increasing of poverty; these results show that the increase in TB cases in Brazil is in a increasing phase regardless of the increase in extreme poverty that has been observed in Brazil, especially since 2019.

In conclusion, we can emphasize that the proposed model can be of great use for other epidemiological data in addition to modeling TB counts and rates associated with TB. Another important point of an epidemiological character obtained from the data analysis: As the number of TB cases is increasing in Brazil and some rates such as AIDS/TB and extreme poverty/TB are falling with the presence of great volatility, there is an indication that TB is increasing indiscriminately in Brazil, as AIDS and extreme poverty are two factors widely discussed in the literature that are associated with TB. This is a problem of public health concern.

It is important to point out, that in the present study we considered univariate epidemiological time series (model M1 (log (TB counts)), model M2 (AIDS/TB rate), model M3 (rate (urban patients)/TB) and model M4 (rate (extreme poverty patients)/TB). Another possibility is to assume multivariate SV models and to study the structure of possible dependence among the series. This approach will be the goal of a future study.

## Acknowledgments

The authors would like to thank the referee's comments that led to a great improvement in the manuscript.

## References

1. World Health Organization (2012) Global tuberculosis report.
2. A San Pedro, RM Oliveira (2013) Tuberculose e indicadores socioeconômicos: Revisão sistemática da literatura. *Rev Panam Salud Publica* 33: 294-301.
3. MA Hajar, MJPR Oliveira, GM Teixeira (2001) A tuberculose no Brasil e no mundo. *Bol. Pneumol Sanit, Rio de Janeiro*, 9: 9-16.
4. DC Costa (1985) Considerações sobre a tendência da tuberculose no Brasil. *Cad Saúde Pública, Rio de Janeiro*, 1: 313-326.
5. LMW Acosta, SL Bassanesi (2014) The Porto Alegre paradox: Social determinants and tuberculosis incidence. *Rev Bras Epidemiol, São Paulo*, 17: 88-101.
6. MG Almeida, DRM Barbosa, DFS Almeida (2013) Epidemiologia e distribuição espacial de casos notificados de tuberculose multirresistente (TBMR) no Brasil, 2008-2012. *Rev Epidemiol Control Infect* 3: 117-122.
7. P Hino (2011) Padrões espaciais da tuberculose e sua associação à condição de vida no município de Ribeirão Preto. *Ciênc Saúde Coletiva* 16: 4795-4802.
8. DL Roza, MCGG Caccia-Bava, EZ Martinez (2012) Spatio-temporal patterns of tuberculosis incidence in Ribeirão Preto, State of São Paulo, southeast Brazil, and their relationship with social vulnerability: A Bayesian analysis. *Rev Soc Bras Med Trop* 45: 607-615.
9. SHF Vendramini, NSGM dos Santos, MLSG Santos, F Chiaravalloti-Neto, MAZ Ponce (2010) Análise espacial da co-infecção tuberculose/HIV: relação com níveis socioeconômicos em município do sudeste do Brasil. *Rev Soc Bras Med Trop* 43: 536-541.
10. R Abellana, C Ascaso, J Aponte, F Saute, D Nhalungo, et al. (2008) Spatio-seasonal modeling of the incidence rate of malaria in Mozambique. *Malar J* 7: 228.
11. IS Chang, SJ Fu, CH Chen, TH Wang, CA Hsiung (2009) Estimating temporal transmission parameters from infectious disease household data, with application to Taiwan SARS data. *Stat Biosci* 1: 80-94.
12. J Kleinschmidt, M Bagayoko, GP Clarke, M Craig, D Le Sueur (2000) A spatial statistical approach to malaria mapping. *Int J Epidemiol* 29: 355-361.
13. G Box, G Jenkins (1970) Time series analysis: Forecasting and control. (1<sup>st</sup> edn), Holden-Day, San Francisco.
14. RS Tsay (2002) Analysis of Financial Time Series. (1<sup>st</sup> edn), John Wiley and Sons, Canada.
15. RR Sarkar, C Chatterjee (2017) Application of different time series models on epidemiological data - Comparison and predictions for malaria prevalence. *SM J Biometrics Biostat* 2: 1022.
16. YL Zheng, LP Zhang, XL Zhang, K Wang, YJ Zheng (2015) Forecast model analysis for the morbidity of tuberculosis in Xinjiang, China. *PLoS One* 10: e0116832.
17. A Earnest, MI Chen, D Ng, LY Sin (2005) Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore. *BMC Health Serv Res* 11: 5-36.
18. Q Li, NN Guo, ZY Han, YB Zhang, SX Qi (2012) Application of an autoregressive integrated moving average model for predicting the incidence of hemorrhagic fever with renal syndrome. *Am J Trop Med Hyg* 87: 364-370.
19. M Gharbi, P Quenel, J Gustave, S Cassadou, G La Ruche, et al. (2011) Time series analysis of dengue incidence in Guadeloupe, French West Indies: Forecasting models using climate variables as predictors. *BMC Infect Dis* 11: 166.
20. Q Liu, X Liu, B Jiang, W Yang (2011) Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect Dis* 11: 218.

21. PM Luz, BVM Mendes, CT Codeco, CJ Struchiner, AP Galvani (2008) Time series analysis of dengue incidence in Rio de Janeiro, Brazil. *Am J Trop Med Hyg* 79: 933-939.
22. W Wu, P Guan, JQ Guo, BS Zhou (2008) Comparison of GM(1,1) gray model and ARIMA model in forecasting the incidence of hemorrhagic fever with renal syndrome. *Journal of China Medical University* 37: 52-55.
23. L Yu, L Zhou, L Tan, H Jiang, Y Wang, S Wei (2014) Application of a new hybrid model with seasonal auto-regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (NARNN) in forecasting incidence cases of HFMD in Shenzhen, China. *PLoS One* 9: e98241.
24. NR Draper, HS Smith, HS (1981) *Applied regression analysis*. Wiley Online Publisher, Canada.
25. GAF Seber, AJ Lee (2003) *Linear regression analysis*. (2<sup>nd</sup> edn), Wiley Series in Probability and Mathematical Statistics.
26. C Chatterjee, RR Sarkar (2009) Multi-step polynomial regression method to model and forecast malaria incidence. *PLoS One* 4: e4726.
27. J Danielsson (1994) Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. *Journal of Econometrics* 61: 375-400.
28. R Engle (1982) Autoregressive conditional heterosdasticity with estimates of the variance of UK inflation. *Econometrica* 50: 987-1007.
29. T Bollerslev (1986) Generalized autoregressive conditional heterosdasticity. *Journal of Econometrics* 31: 307-327.
30. E Ghysels (1996) Stochastic volatility, in 'Statistical methods on finance', North-Holland.
31. S Kim, N Shephard (1998) Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65: 361-393.
32. R Meyer, T Yu (2000) Bugs for a Bayesian analysis of stochastic volatility models. *Econometrics Journal* 3: 198-215.
33. AE Gelfand, AFM Smith (1990) Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 85: 398-409.
34. AFM Smith, GO Roberts (1993) Bayesian computation the Gibbs sampling and related Markov Chain Monte Carlo methods. *Journal of the Royal Statistical Society* 55: 3-23.
35. TD Phan, JA Wachter, EA Solomon, MJ Kahana (2019) Multivariate stochastic volatility modeling of neural data. *eLife* 8: e42950.
36. B Carlin, TA Louis (2000) *Bayes and Empirical Bayes Methods for Data Analysis*. (2<sup>nd</sup> edn), Chapman & Hall, London.
37. DJ Spiegelhalter, A Thomas, N Best, D Lunn (2003) *WinBUGS user manual*. MRC Biostatistics Unit, Cambridge, United Kingdom.

**Appendix 1: TB series in Brazil (January 2001 to December 2018).**

	TB	AIDS/TB	URB/TB	POVERTY/TB		TB	AIDS/TB	URB/TB	POVERTY/TB
1	8088	0.064416	0.634891	0.0817260	109	6987	0.099900	0.682553	0.0814370
2	6542	0.065729	0.627484	0.0764292	110	6471	0.100757	0.687838	0.0698501
3	8095	0.060531	0.640889	0.0795553	111	8373	0.097098	0.697719	0.0738087
4	7305	0.061602	0.633676	0.0835044	112	7113	0.093350	0.678335	0.0785885
5	7656	0.064002	0.643939	0.0858150	113	7039	0.095184	0.692002	0.0761472
6	6805	0.063630	0.651433	0.0865540	114	6701	0.107745	0.693180	0.0840173
7	6987	0.064692	0.637613	0.0918849	115	7178	0.105322	0.696155	0.0764837
8	8065	0.067452	0.628518	0.0838190	116	7463	0.099022	0.687391	0.0758408
9	6697	0.067344	0.626699	0.0895924	117	7216	0.103520	0.676968	0.0766353
10	7524	0.069777	0.624668	0.0867889	118	7273	0.098721	0.671662	0.0741097
11	6874	0.071720	0.620454	0.0853942	119	7096	0.100761	0.671787	0.0735626
12	6627	0.068206	0.633016	0.0804285	120	7236	0.105445	0.663903	0.0697899
13	8013	0.067141	0.660302	0.0842381	121	7173	0.101771	0.681305	0.0818347
14	7346	0.072965	0.641710	0.0811326	122	7335	0.095024	0.671302	0.0837082
15	7961	0.065444	0.659465	0.0806431	123	7548	0.101881	0.660837	0.0784314
16	8771	0.064531	0.661612	0.0884734	124	7619	0.098701	0.674629	0.0767817
17	7784	0.063207	0.674717	0.0908273	125	7765	0.093883	0.678558	0.0830650
18	6713	0.074631	0.673171	0.0804409	126	7008	0.102312	0.679937	0.0794806
19	7746	0.076039	0.679706	0.0888200	127	7138	0.102690	0.668535	0.0724293
20	8330	0.073229	0.650660	0.0828331	128	8062	0.104937	0.674274	0.0769040
21	7654	0.073948	0.662007	0.0837471	129	7427	0.102195	0.665814	0.0780934
22	8200	0.071220	0.661098	0.0800000	130	7052	0.108622	0.686188	0.0692002
23	7532	0.074615	0.660781	0.0896176	131	7439	0.101895	0.665815	0.0807904
24	6809	0.078866	0.656925	0.0772507	132	6994	0.105376	0.671433	0.0700601
25	8115	0.066790	0.659766	0.0853974	133	7435	0.106523	0.673571	0.0746469
26	7985	0.069881	0.655980	0.0821540	134	6830	0.104539	0.676867	0.0784773
27	7474	0.066899	0.655205	0.0850950	135	7936	0.096018	0.676789	0.0756048
28	7975	0.067586	0.667712	0.0820063	136	6981	0.097550	0.681135	0.0786420
29	8024	0.071286	0.685942	0.0844965	137	7649	0.105896	0.682965	0.0738659
30	7082	0.075402	0.676363	0.0858515	138	6861	0.108731	0.677744	0.0773940
31	7949	0.071833	0.680463	0.0865518	139	7317	0.105098	0.681017	0.0736641
32	7585	0.075676	0.676994	0.0843771	140	8060	0.108313	0.675434	0.0763027
33	8136	0.074115	0.673673	0.0886185	141	6785	0.103021	0.663670	0.0716286
34	8521	0.077808	0.669405	0.0822673	142	7679	0.102748	0.684725	0.0714937
35	7624	0.080273	0.677991	0.0836831	143	7010	0.102425	0.676034	0.0766049
36	7303	0.076818	0.678078	0.0798302	144	6362	0.105942	0.680918	0.0814209
37	7574	0.080803	0.691048	0.0764457	145	7502	0.104639	0.676220	0.0718475
38	6743	0.082308	0.669435	0.0904642	146	6209	0.100821	0.659204	0.0753745
39	8574	0.071495	0.680546	0.0905062	147	7102	0.105182	0.669107	0.0720924
40	8089	0.071826	0.668068	0.0876499	148	7696	0.104730	0.676975	0.0732848
41	7763	0.072008	0.684916	0.0888832	149	6996	0.102630	0.674099	0.0697541
42	7314	0.078616	0.675280	0.0849057	150	6953	0.109305	0.660866	0.0690349
43	7825	0.080895	0.669776	0.0857508	151	7379	0.104757	0.672042	0.0752134
44	8232	0.078839	0.651361	0.0823615	152	7931	0.097844	0.661077	0.0742655
45	7901	0.072902	0.655993	0.0898620	153	7377	0.107090	0.674393	0.0737427
46	8016	0.070734	0.670659	0.0782186	154	7789	0.107074	0.664270	0.0729234
47	7753	0.070940	0.677157	0.0882239	155	6939	0.099150	0.661623	0.0704712
48	7196	0.075875	0.672457	0.0774041	156	6335	0.105762	0.655564	0.0716654



49	7432	0.077234	0.674247	0.0805974	157	7670	0.112907	0.653064	0.0724902
50	6748	0.081654	0.655750	0.0881743	158	7040	0.104545	0.648438	0.0762784
51	8441	0.072859	0.668049	0.0900367	159	6754	0.106752	0.663459	0.0710690
52	7958	0.068233	0.661850	0.0913546	160	7260	0.111019	0.657025	0.0796143
53	8019	0.068837	0.659683	0.0875421	161	7177	0.110631	0.663648	0.0769124
54	7781	0.075954	0.654543	0.0916335	162	6322	0.102341	0.652800	0.0752926
55	7282	0.081571	0.677012	0.0891239	163	7516	0.112693	0.667110	0.0671900
56	8391	0.074365	0.688833	0.0852103	164	7124	0.112858	0.664374	0.0728523
57	7690	0.080624	0.678674	0.0893368	165	7505	0.110859	0.654897	0.0719520
58	7186	0.079043	0.676037	0.0843306	166	7470	0.110308	0.656894	0.0681392
59	7587	0.076447	0.662581	0.0873863	167	6905	0.105431	0.653874	0.0670529
60	7541	0.086593	0.647792	0.0820846	168	6470	0.108192	0.658733	0.0646059
61	7466	0.070855	0.663274	0.0877310	169	7114	0.110065	0.674445	0.0653641
62	6767	0.080095	0.654352	0.0854145	170	6162	0.109380	0.655794	0.0688088
63	8256	0.080305	0.676841	0.0856347	171	7836	0.104135	0.661052	0.0675089
64	6901	0.077815	0.666570	0.0785393	172	6877	0.103534	0.661189	0.0665988
65	7699	0.074295	0.681777	0.0879335	173	6885	0.106609	0.656064	0.0755265
66	7094	0.081336	0.664787	0.0873978	174	6915	0.113666	0.666377	0.0694143
67	7220	0.081440	0.676177	0.0853186	175	7529	0.104396	0.662903	0.0761057
68	7835	0.076707	0.662923	0.0825782	176	7466	0.104875	0.662202	0.0707206
69	6873	0.078714	0.664484	0.0919540	177	7191	0.111528	0.656098	0.0759282
70	7089	0.081676	0.666667	0.0846382	178	7344	0.106618	0.649646	0.0761166
71	6751	0.073619	0.666124	0.0897645	179	7284	0.098298	0.660626	0.0675453
72	6209	0.083427	0.670800	0.0834273	180	6840	0.097953	0.652047	0.0687135
73	7436	0.085933	0.676304	0.0809575	181	7025	0.102491	0.655089	0.0743060
74	6294	0.089609	0.670480	0.0776930	182	6756	0.104056	0.653641	0.0703079
75	8214	0.086681	0.684076	0.0808376	183	8050	0.099130	0.647578	0.0731677
76	7373	0.082734	0.686830	0.0846331	184	7335	0.094888	0.657805	0.0711656
77	7481	0.087555	0.679722	0.0820746	185	7315	0.101162	0.666986	0.0686261
78	6714	0.089812	0.683944	0.0792374	186	7520	0.098537	0.666755	0.0696809
79	7283	0.090897	0.681038	0.0856790	187	6993	0.100529	0.665523	0.0775061
80	7669	0.089842	0.673491	0.0864519	188	7702	0.113217	0.650091	0.0701117
81	6801	0.086458	0.670931	0.0854286	189	7035	0.095096	0.667662	0.0693674
82	7550	0.085298	0.678278	0.0778808	190	6590	0.113354	0.645068	0.0687405
83	6786	0.084586	0.686413	0.0863543	191	7036	0.101904	0.646674	0.0722001
84	6178	0.104079	0.679994	0.0768857	192	6669	0.099715	0.660819	0.0652272
85	7459	0.088886	0.669527	0.0833892	193	7475	0.099933	0.658729	0.0707692
86	6729	0.084559	0.669639	0.0857483	194	6790	0.101915	0.663623	0.0761414
87	7359	0.098519	0.686914	0.0826199	195	8676	0.099009	0.640964	0.0728446
88	7717	0.091227	0.675392	0.0855255	196	6774	0.099941	0.669472	0.0786832
89	6992	0.097254	0.686928	0.0826659	197	8093	0.106759	0.669220	0.0763623
90	7001	0.089273	0.683617	0.0775603	198	7386	0.102897	0.659762	0.0747360
91	7777	0.087052	0.693969	0.0786936	199	7221	0.101509	0.662928	0.0811522
92	7912	0.092139	0.678842	0.0850607	200	8112	0.101085	0.659024	0.0792653
93	7720	0.090933	0.678756	0.0813472	201	7434	0.104116	0.646220	0.0745225
94	7643	0.095774	0.681539	0.0754939	202	7793	0.101886	0.653664	0.0728859
95	6863	0.083928	0.683229	0.0760600	203	7447	0.092789	0.643615	0.0750638
96	6650	0.094135	0.670376	0.0787970	204	7009	0.092595	0.651733	0.0744757
97	7180	0.090111	0.692061	0.0803621	205	7878	0.094821	0.656512	0.0759076
98	6507	0.091747	0.670816	0.0791455	206	6803	0.098045	0.650595	0.0684992

99	8231	0.086988	0.681813	0.0816426	207	7893	0.093754	0.659698	0.0741163
100	7584	0.094937	0.680643	0.0901899	208	8126	0.092789	0.665272	0.0756830
101	7245	0.092478	0.695376	0.0864044	209	7878	0.097994	0.662224	0.0714648
102	6799	0.094426	0.694661	0.0897191	210	7688	0.094693	0.648413	0.0736212
103	7513	0.090643	0.701185	0.0805271	211	8085	0.099567	0.662585	0.0752010
104	7450	0.095705	0.688188	0.0783893	212	8932	0.094940	0.656740	0.0704210
105	7478	0.098288	0.687350	0.0850495	213	7663	0.096437	0.659533	0.0695550
106	7359	0.092947	0.677674	0.0852018	214	8691	0.091129	0.658842	0.0699574
107	7088	0.096924	0.681716	0.0747743	215	7682	0.091382	0.653606	0.0741994
108	6793	0.098042	0.680701	0.0791992	216	6836	0.095816	0.649210	0.0751902

**Appendix 2:** OpenBUGS code (change-points).

```

model
{
for (i in 1:N) {
  J[i] <- 1+step(i-k1)+step(i-k2)
  y[i] <- log(yearly.rate.TB[i])
  y[i] ~ dnorm(mean[i],tau[J[i]])
}
for (i in 1:N) {
mean[i] <- beta0[J[i]]+ beta1[J[i]]*years[i]
fitted.mean[i]<-exp(mean[i])
}
k1 ~ dcat(p1[])
k2 ~ dcat(p2[])
tau[1] ~ dgamma(1,1)
tau[2] ~ dgamma(1,1)
tau[3] ~ dgamma(1,1)
beta0[1] ~ dnorm(-7.60,1)
beta0[2] ~ dnorm(-7.70,1)
beta0[3] ~ dnorm(-7.75,1)
beta1[1] ~ dnorm(0,100)
beta1[2] ~ dnorm(0,100)
beta1[3] ~ dnorm(0,100)
}

```