

RESEARCH ARTICLE

Open Access



Statistical methods for predicting tuberculosis incidence based on data from Guangxi, China

Yanling Zheng*, Liping Zhang, Lei Wang and Ramziya Rifhat

Abstract

Background: Tuberculosis (TB) remains a serious public health problem with substantial financial burden in China. The incidence of TB in Guangxi province is much higher than that in the national level, however, there is no predictive study of TB in recent years in Guangxi, therefore, it is urgent to construct a model to predict the incidence of TB, which could provide help for the prevention and control of TB.

Methods: Box-Jenkins model methods have been successfully applied to predict the incidence of infectious disease. In this study, based on the analysis of TB incidence in Guangxi from January 2012 to June 2019, we constructed TB prediction model by Box-Jenkins methods, and used root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) to test the performance and prediction accuracy of model.

Results: From January 2012 to June 2019, a total of 587,344 cases of TB were reported and 879 cases died in Guangxi. Based on TB incidence from January 2012 to December 2018, the SARIMA((2),0,(2))(0,1,0)₁₂ model was established, the AIC and SC of this model were 2.87 and 2.98, the fitting accuracy indexes, such as RMSE, MAE and MAPE were 0.98, 0.77 and 5.8 respectively; the prediction accuracy indexes, such as RMSE, MAE and MAPE were 0.62, 0.45 and 3.77, respectively. Based on the SARIMA((2),0,(2))(0,1,0)₁₂ model, we predicted the TB incidence in Guangxi from July 2019 to December 2020.

Conclusions: This study filled the gap in the prediction of TB incidence in Guangxi in recent years. The established SARIMA((2),0,(2))(0,1,0)₁₂ model has high prediction accuracy and good prediction performance. The results suggested the change trend of TB incidence predicted by SARIMA((2),0,(2))(0,1,0)₁₂ model from July 2019 to December 2020 was similar to that in the previous two years, and TB incidence will experience slight decrease, the predicted results can provide scientific reference for the prevention and control of TB in Guangxi, China.

Keywords: Tuberculosis, Prediction, SARIMA model

Background

Tuberculosis (TB) is a chronic respiratory infectious disease caused by the pathogen *Mycobacterium tuberculosis*. Infected people can spread TB germs from their mouth when they cough or sneeze. After suffering from TB, if the patients are not given timely, thorough

treatment, which can pose a serious threat to their health, even make them completely lost the ability to work, and the TB patients may also infect others [1]. At present, although great progress has been made around the world in the prevention and control of TB, many countries, especially in low-income and middle-income settings, are still afflicted with a chronic plague of TB with huge economies losses [2]. Moreover, TB remains one of the top 10 causes of death worldwide; it is

* Correspondence: zhengyl_math@sina.cn
College of Medical Engineering and Technology, Xinjiang Medical University, Urumqi 830011, People's Republic of China



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

estimated that globally there were 10.0 million new cases of TB in 2017, of which 1.3 million individuals' deaths were directly attributable to TB, and TB has killed more people than any other infectious disease in the past few decades [2, 3]. China is one of the countries with high burden of TB, number of TB patients ranked second in the world, accounting for a quarter of the world's patients, and about 250 thousand patients died of TB every year in China [2].

The Guangxi is a province of China, it is located in the south of China, the latitude $20^{\circ}54' \sim 26^{\circ}24' \text{ N}$, longitude $104^{\circ}26' \sim 112^{\circ}04' \text{ E}$, it covers a total area about 236,700 km^2 , with a population over 49.26 million in 2018, and is one of the Chinese provinces that is most affected by TB. From 2015 to 2017, the annual incidences (per 100,000 populations) of TB in China were 63.42, 61 and 60.53, respectively, while the annual incidences of TB in Guangxi province of China were 96.41, 86.27 and 87.86, respectively. These incidences of TB in Guangxi were much higher than that in the national level, so it is necessary to pay more attention to the prevention and control of TB in this area.

To master the regularity of infectious diseases, analyze and know the epidemic situation of infectious diseases by using the existing surveillance data, then predict the future, which can provide scientific reference for disease prevention and control. The Box-Jenkins method is a representative time series analysis and prediction method, which can take into account trend changes, periodic changes, and random disturbances in time series. It is very useful in modeling temporal dependence structure of a time series. At present, this method has been widely used in the prediction of infectious diseases, and has achieved successful prediction results, for instance, Tian C W et al. [4] forecasted monthly cases of hand-foot-mouth disease successfully in China; Wang T et al. [5] suggested that $\text{ARIMA}(3,1,1)(2,1,1)_{12}$ model was reliable with a high validity, which could be used to predict hemorrhagic fever with renal syndrome incidence in Zibo; Myriam Gharbil et al. [6] predicted the dengue incidence in Guadeloupe based on time series analysis; López-Montenegro LE [7] predicted dengue cases in Colombia from 2018 to 2022 based on Auto-Regressive Integrated Moving Average (ARIMA) model; Zheng Y-L et al. [8] and Liao Z [9] forecasted TB incidence successfully using SARIMA model, etc. [10–17].

The incidence of TB in Guangxi is very high, but there are few related prediction studies so far. In order to do a better job of prevention and control, in the study, the prediction research was carried out. Firstly, we briefly analyzed the change trend of the TB incidence in Guangxi over the years, and then, based on the data characteristics of the TB incidence in Guangxi, China, we established the best SARIMA model for prediction.

Finally, the TB incidence in the future was predicted, which can provide scientific reference for prevention and control of TB in Guangxi.

Methods

Data source

The data of the TB cases in Guangxi from January 2012 to June 2019 was obtained from the Guangxi center for Disease Control and Prevention, China; Population data was obtained from the official website of Guangxi Bureau of Statistics, based on the population data and the reported number of TB cases, we calculated the monthly incidence of TB (per 100,000 populations). The data used in this study is provided as Additional file 1.

SARIMA model descriptions

The Box-Jenkins method is a famous time series prediction method proposed by Box and Jenkins in the early 1970s, it includes the $\text{ARIMA}(p,d,q)$ model called. Auto-regressive Integrated Moving Average Model, AR is auto regression, p is the number of auto regression term, MA is moving average, q is the number of moving average terms [18, 19]. If the time series contains a seasonal cycle, it is often necessary to do a seasonal difference to establish a SARIMA model, the SARIMA model with s observations per period, denoted by $\text{SARIMA}(p, d, q)(P, D, Q)_s$. Generally, the standard statistical methodology to construct an $\text{SARIMA}(p, d, q)(P, D, Q)_s$ model includes four steps:

First step, data stationary test. Usually, data set needs to be divided into two subsets for model: one for training set, and the other one for testing set. The training set needs to be stationary time series. If the original training set data is not stationary, common differential or seasonal difference is required, d is the order of the ordinary difference, and D is the order of the seasonal difference. Augmented Dickey-Fuller (ADF) test can determine whether the time series was stationary, the significance level of the test is 0.05 (if the test Prob is less than 0.05, then, the data is stationary).

Second step, based on the data of stationary time series, to plot the graphs of the autocorrelation function (ACF) and partial autocorrelation function (PACF). According to the analysis of ACF and PACF, we can determine the possible values of p , q , P and Q , this process requires both skill and experience. Generally, more than one tentative model is chosen in this step.

Third step, to do parameter estimation and hypothesis test of all tentative SARIMA models by least square method. These model passed by the parameter test is feasible, furthermore, to do diagnostic checking of their residuals, if residuals are almost equivalent to white noises (significant level $\text{Prob} > 0.05$) by using the Box-Jenkins Q test, then SARIMA model has good

performance. Then, to select the best SARIMA model by the Akaike information criterion (AIC) and Schwarz criterion (SC). The preferred model is the one with the lowest AIC and SC values.

Fourth step, to predict the TB incidence based on the preferred SARIMA model, then, to calculate forecast accuracy indexes, such as root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). Good fitting and prediction performance of SARIMA model are demonstrated with RMSE, MAE and MAPE as small as possible.

Data processing and analysis

All analyses were performed using ArcGIS 10.4, Eviews7.2, R3.6.2 and Matlab 2012b.

Results

From January 2012 to June 2019, a total of 587,344 cases of TB and 879 deaths of TB were reported in Guangxi. It can be seen from Fig. 1 that the TB incidence was decreasing year by year, and there was certain seasonality. The TB incidence in the second and third quarters were higher than that in the first and fourth quarters.

We used R3.6.2 software to decompose TB incidence data, and found that TB incidence data have obvious seasonality, periodicity and randomness (see Fig. 2), so it is suitable to establish SARIMA model for prediction analysis.

The data from January 2012 to June 2019 was divided into two parts, the part from January 2012 to December 2018 was used to construct the SARIMA(p,d,q)(P,D,Q)_s model, and the other part from January 2019 to June 2019 was used to test the prediction performance of the SARIMA(p,d,q)(P,D,Q)_s model.

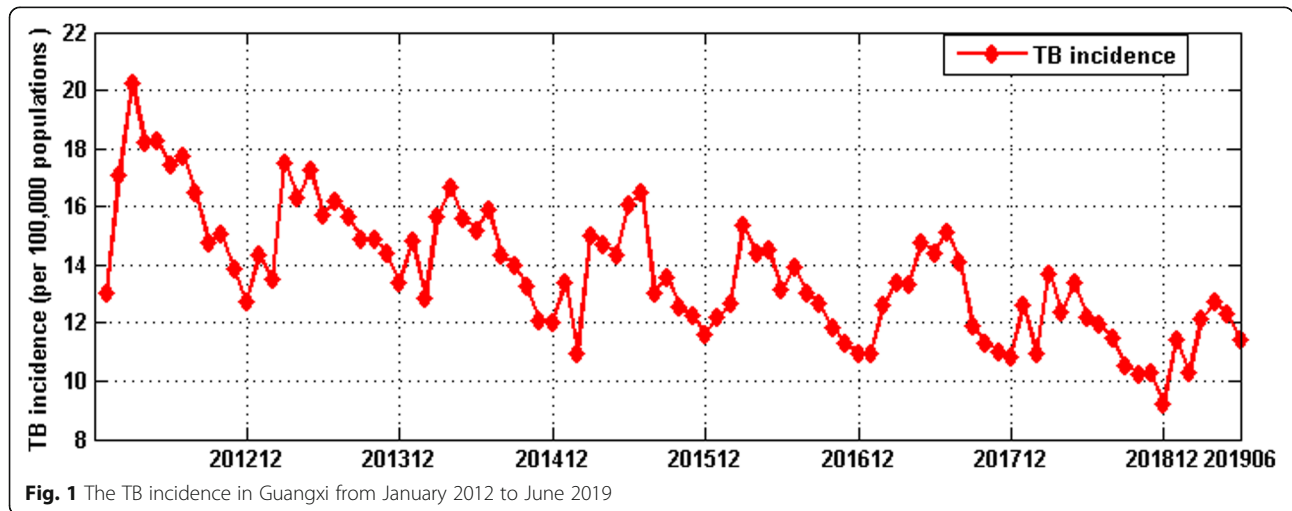
The SARIMA(p,d,q)(P,D,Q)_s model method requires data to be stationary, otherwise, neither of backcast or forecast of the series can be available. First, ADF was

used to test the stability of original series, and the tested Prob value was 0.94 greater than 0.05, which showed that the series was not stationary. Because there was obvious seasonality in the TB incidence series in Guangxi (see Fig. 2), we did the first-order seasonal difference with period 12 on original series, and then, did ADF test of the seasonal difference data again, and the tested Prob value was less than 0.01, therefore, after the first-order seasonal difference, the data was stationary, then, d = 0, D = 1 and s = 12. The test results were shown in Table 1.

Second, to draw ACF and PACF graphs of stationary data (see Fig. 3). According to the analysis of the ACF and PACF graphs, we established eight tentative models, SARIMA(1,0,1)(0,1,0)₁₂, SARIMA(1,0,(2))(0,1,0)₁₂, SARIMA((2),0,1)(0,1,0)₁₂, SARIMA((2),0,(2))(0,1,0)₁₂, SARIMA(2,0,(2))(0,1,0)₁₂, SARIMA(2,0,1)(0,1,0)₁₂, SARIMA(1,0,2)(0,1,0)₁₂, and SARIMA(2,0,2)(0,1,0)₁₂. Then, the least square method was used to test the parameters of the eight models, and the AIC and SC values of these models were calculated, the test results were shown in Table 2. It could be seen that only the SARIMA((2),0,(2))(0,1,0)₁₂ model with lowest AIC and SC passed the parameter test (all Prob values were less than 0.05).

Finally, we did the diagnostic checking of residuals of the SARIMA((2),0,(2))(0,1,0)₁₂ model by using the Box-Jenkins Q test, the test Prob was more than 0.05, therefore, according to these analyses, the SARIMA((2),0,(2))(0,1,0)₁₂ model was feasible for the prediction of TB incidence in Guangxi.

We used the SARIMA((2),0,(2))(0,1,0)₁₂ model to fit the TB incidence data from March 2013 to December 2018, and the RMSE, MAE, and MAPE were 0.98, 0.77 and 5.8 respectively; We used the SARIMA((2),0,(2))(0,1,0)₁₂ model to predict the TB incidence from January 2019 to June 2019, and the RMSE, MAE, and MAPE were 0.62, 0.45 and 3.77, respectively. Both the fitting accuracy values and the prediction accuracy values were



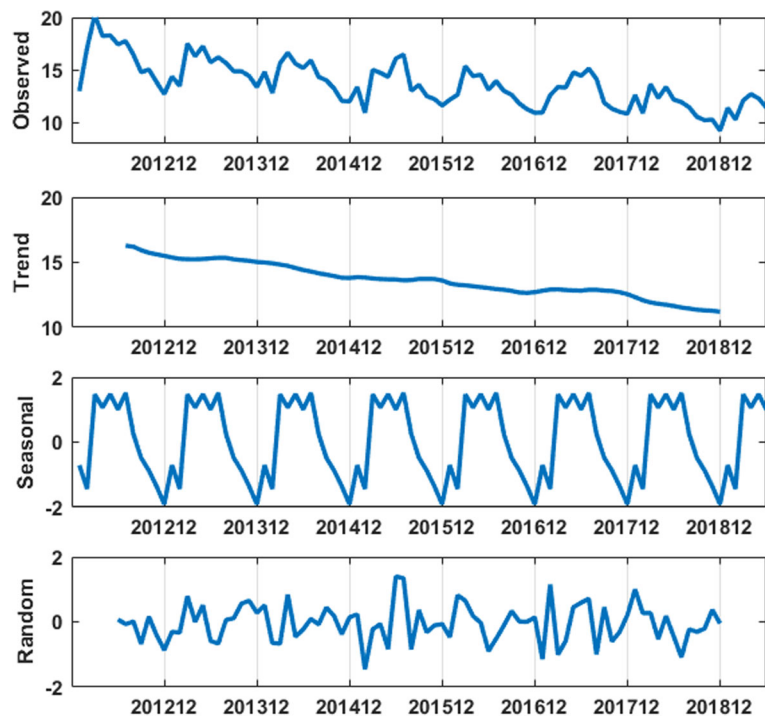


Fig. 2 Time series decomposition of TB incidence from January 2012 to June 2019

very small, which indicated that the SARIMA((2),0,(2))(0,1,0)₁₂ model was very good and its prediction accuracy was high. Based on the SARIMA((2),0,(2))(0,1,0)₁₂ model, we predicted the TB incidence in Guangxi from July 2019 to December 2020, these predicted values were shown in Table 3, and the fitted and predicted incidence were compared with the observed incidence in Fig. 4.

Discussion

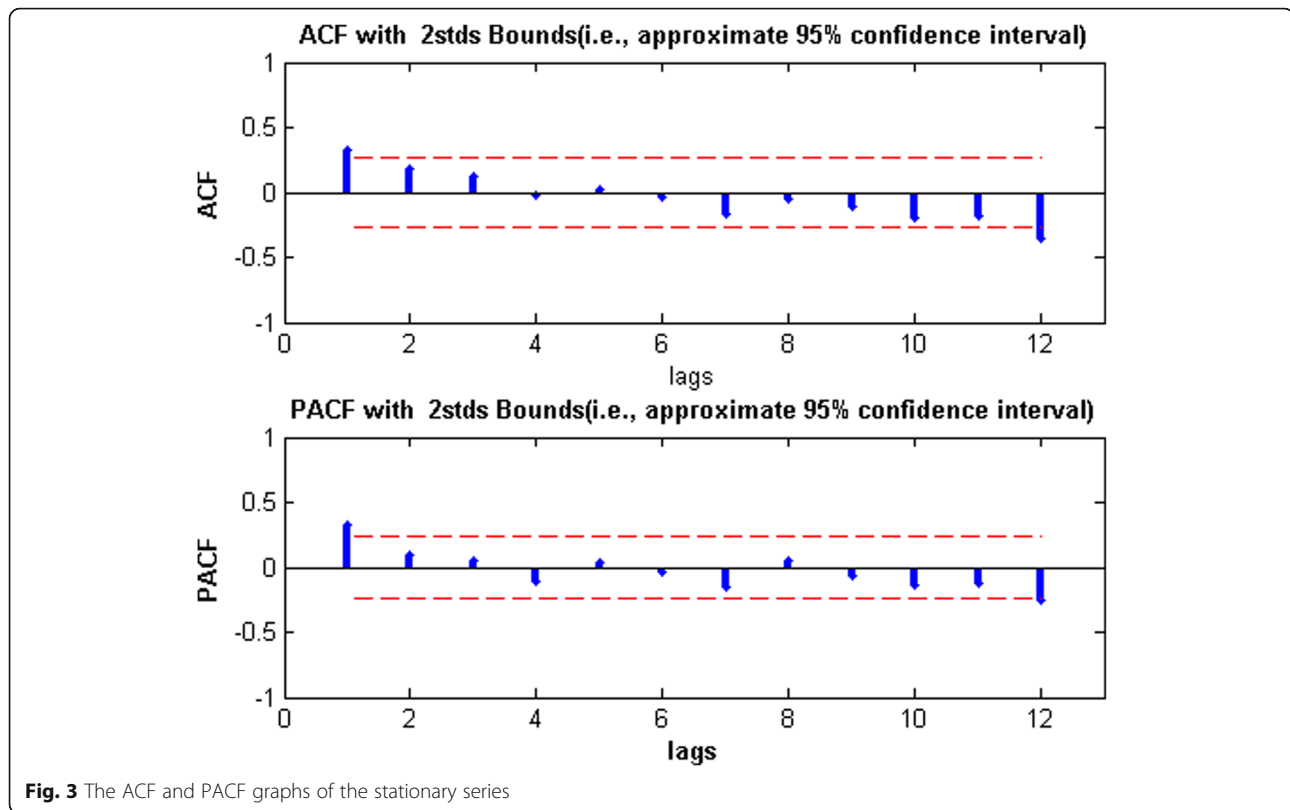
Currently, the annual TB incidence in Guangxi is much higher than that in the national level, although it has been slightly decreasing annually; the potential achievement is diminished by an increasing large-scale transient population, the emergence of MDR-TB, along with the co morbid conditions of AIDS and non-communicable diseases, which have led to a resurgence of TB in recent years [20–22]. Additionally, WHO initiated the End TB Strategy with the target of a 90% reduction in new TB

cases by 2035 compared with 2015, and a milestone of reducing the TB incidence by 50% by 2025 relative to 2015 [2], in order to accelerate progress towards such a daunting task, corresponding measures and actions are expected at both the national and international levels. At the national level, every province should make efforts, especially in provinces with high incidence, such as Guangxi. Appropriate plans may fail to be becomingly formulated without getting a clear perspective of the past, current and future temporal levels of this disease, therefore, advanced detection and early response systems for epidemics have formed an integral part of the effective precautions against TB and the reasonable allocation of available health resources.

In this study, the historical trend of TB incidence in Guangxi was carefully analyzed, then, the prediction model of TB incidence in Guangxi was established by using Box Jenkins model method, and this method is

Table 1 The ADF tests of Modeling data

Before first-order seasonal difference			After first-order seasonal difference		
t-Statistic		Prob	t-Statistic		Prob
Augmented Dickey-Fuller test statistic	-0.13	0.94	Augmented Dickey-Fuller test statistic	-6.09	< 0.01
Test critical values:	1% level	-3.52	Test critical values:	1% level	-3.52
	5% level	-2.9		5% level	-2.9
	10% level	-2.59		10% level	-2.59



one of the most widely used time series forecasting techniques because of its structured modeling basis and acceptable forecasting performance. Through the analysis of the change trend and decomposition graph of original TB incidence, we found that the data had obvious seasonality, trend and randomness, so it is suitable to establish SARIMA model for prediction analysis. For SARIMA model construction, monthly TB incidence from January 2012 to December 2018 was used; for testing the predictive ability of this model, TB incidence from January 2019 to June 2019 was used.

SARIMA model requires data to be stationary, Table 1 showed that the Prob value of ADF test was 0.94 more than 0.05, indicating that the original data was not stable. Considering the seasonal variation of TB incidence data, we did the first-order seasonal difference with a period of 12, after that, we used ADF to test the stationarity of the seasonal-difference data, the Prob value of the test was less than 0.01(see Table 1), which indicated that the difference data was stable and could be used to build SARIMA model. Then, in order to determine the p , q , P and Q in SARIMA($p,0,q$)($P,1,Q$)₁₂ model, the ACF and PACF graphs were drawn, then, eight tentative models were established by the analysis of ACF and PACF graphs.

The parameters of these tentative models were tested and these models performance were compared by AIC and SC, Table 2 showed SARIMA((2),0,(2))(0,1,0)₁₂ model had smallest AIC and SC, as well as, all the Prob values of its parameter test were less than 0.01, and the Prob value of the Box-Jenkins Q test was more than 0.05, which indicated that the SARIMA((2),0,(2))(0,1,0)₁₂ model was feasible to predict the TB incidence in Guangxi. Using SARIMA((2),0,(2))(0,1,0)₁₂ model to fit original TB incidence from January 2012 to December 2018, the RMSE(0.98), MAE(0.77), and MAPE(5.80) were very small; Using SARIMA((2),0,(2))(0,1,0)₁₂ model to predict TB incidence from January 2019 to June 2019, the RMSE(0.62),MAE(0.45),and MAPE(3.77) were very small too, which indicated that the SARIMA((2),0,(2))(0,1,0)₁₂ model was very good and its prediction accuracy was very high. We predicted the TB incidence in Guangxi based on the SARIMA((2),0,(2))(0,1,0)₁₂ model from July 2019 to December 2020(see Table 3 and Fig. 4), the results suggested the change trend of predicted TB incidence was similar to change trend in the previous two years, and TB incidence will experience slight decrease, the predicted results can provide scientific reference for the prevention and control of TB in Guangxi, China.

Table 2 The Parameter estimates of the tentative models with their AIC and SC

Model	Variable	Coefficient	Std. Error	t-Statistic	prob	AIC	SC
SARIMA(1,0,1)(0,1,0) ₁₂	C	-0.81	0.16	-4.92	< 0.01		
	AR(1)	-0.16	0.22	-0.73	0.47	3.00	3.10
	MA(1)	0.51	0.20	2.52	0.01		
SARIMA(1,0,(2))(0,1,0) ₁₂	C	-0.84	0.21	-4.00	< 0.01		
	AR(1)	0.23	0.12	1.93	0.06	3.03	3.12
	MA(2)	0.27	0.12	2.23	0.03		
SARIMA((2),0,1)(0,1,0) ₁₂	C	-0.78	0.20	-3.96	< 0.01		
	AR(2)	0.14	0.12	1.21	0.23	2.94	3.04
	MA(1)	0.37	0.12	3.15	< 0.01		
SARIMA((2),0,(2))(0,1,0) ₁₂	C	-0.77	0.15	-5.22	< 0.01		
	AR(2)	-0.59	0.10	-6.07	< 0.01	2.87	2.98
	MA(2)	0.96	0.02	54.39	< 0.01		
SARIMA(2,0,(2))(0,1,0) ₁₂	C	-0.78	0.16	-4.83	< 0.01		
	AR(1)	0.17	0.10	1.77	0.08	2.87	3.00
	AR(2)	-0.61	0.10	-6.26	< 0.01		
	MA(2)	0.97	0.02	46.44	< 0.01		
SARIMA(2,0,1)(0,1,0) ₁₂	C	-0.71	0.22	-3.21	0.00		
	AR(1)	0.64	0.25	2.57	0.01	2.91	3.03
	AR(2)	-0.02	0.14	-0.17	0.86		
	MA(1)	-0.31	0.27	-1.15	0.26		
SARIMA(1,0,2)(0,1,0) ₁₂	C	-0.81	0.18	-4.46	< 0.01		
	AR(1)	-0.16	0.29	-0.54	0.59	3.01	3.13
	MA(1)	0.50	0.30	1.65	0.10		
	MA(2)	0.18	0.14	1.29	0.20		
SARIMA(2,0,2)(0,1,0) ₁₂	C	-0.72	0.22	-3.31	< 0.01		
	AR(1)	0.65	0.26	2.52	0.01		
	AR(2)	-0.09	0.21	-0.44	0.66	2.93	3.09
	MA(1)	-0.33	0.28	-1.17	0.25		
	MA(2)	0.12	0.23	0.52	0.60		

Table 3 The observed TB incidence and predicted TB incidence by SARIMA((2),0,(2))(0,1,0)₁₂ model from January 2019 to December 2020

Time	201,901	201,902	201,903	201,904	201,905	201,906
Observed	11.4	10.3	12.12	12.72	12.32	11.45
Predicted	11.34	9.94	13.19	11.71	12.45	11.35
Time	201,907	201,908	201,909	201,910	201,911	201,912
Observed						
Predicted	11.29	10.74	9.73	9.43	9.56	8.48
Time	202,001	202,002	202,003	202,004	202,005	202,006
Observed						
Predicted	9.28	9.05	11.94	11.32	11.68	11.07
Time	202,007	202,008	202,009	202,010	202,011	202,012
Observed						
Predicted	10.61	10.24	9.24	8.95	8.40	7.79

Conclusions

The incidence of tuberculosis in Guangxi is high, but there is little prediction study of the disease in recent years, advanced detection and early response systems have formed an integral part of the effective precautions against TB and the reasonable allocation of available health resources. In view of this, we used Box-Jenkins method to establish the SARIMA((2),0,(2))(0,1,0)₁₂ model for predicting the TB incidence in Guangxi. The RMSE, MAE and MAPE of the SARIMA((2),0,(2))(0,1,0)₁₂ were very small, which indicated that the model was successful, its prediction accuracy was high, and its prediction performance was good. Based on SARIMA((2),0,(2))(0,1,0)₁₂ model, we predicted the TB incidence of Guangxi from July

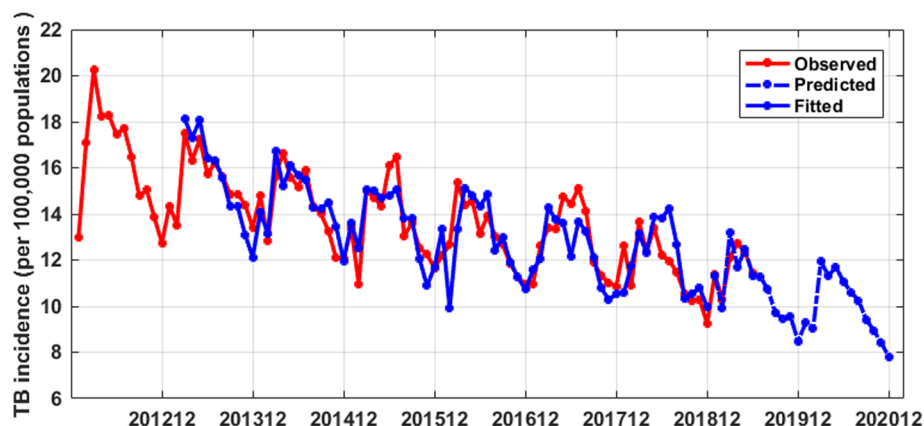


Fig. 4 The fitted and predicted TB incidence by the SARIMA((2),0,(2))(0,1,0)₁₂ model

2019 to December 2020, the results suggested the TB incidence will experience slight decrease, and its changing trend will be similar to before. The prediction results can provide help for reallocating resources so as to get better in control and prevention of TB in Guangxi, China.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12879-020-05033-3>.

Additional file 1: Table S1. Monthly TB incidence data from January 2012 to June 2019, Guangxi, China.

Abbreviations

TB: Tuberculosis; ACF: Autocorrelation function;; PACF: Partial autocorrelation function; AIC: Akaike information criterion; SC: Schwarz criterion; SARIMA: Seasonal autoregressive integrated moving average; ADF: Augmented Dickey-Fuller (ADF) test; RMSE: Root mean square error; MAE: Mean absolute error; MAPE: Mean absolute percentage error.

Acknowledgements

We would like to express our gratitude to peer reviewers for carefully revising our manuscript and useful comments.

Authors' contributions

YLZ conceived and designed the experiments. YLZ collected the data. YLZ analyzed the data. YLZ, LPZ, LW and RR wrote the manuscript. All authors read and approved the final version of the manuscript

Funding

This work was supported by the Natural science funding of Xinjiang Uygur Autonomous Region (2017D01C189), China.

Availability of data and materials

The data used in this study are available from the corresponding author on reasonable request and with permission of the Guangxi center for Disease Control and Prevention, China. The relevant data is provided as Additional file 1.

Ethics approval and consent to participate

The research did not involve any direct participation by human subjects. The TB data were extracted from monthly reports maintained by Guangxi center for Disease Control and Prevention.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 4 September 2019 Accepted: 15 April 2020

Published online: 22 April 2020

References

- Zhao Y, Li M, Yuan S. Analysis of transmission and control of tuberculosis in mainland China, 2005-2016, based on the age structure mathematical model. *Int J Environ Res Public Health*. 2017;14:1192.
- WHO. Global tuberculosis report 2018. http://www.who.int/tb/publications/global_report/en/. (Accessed on 4 Dec 2018).
- Moosazadeh M, Khanjani N, Nasehi M, et al. Predicting the incidence of smear positive tuberculosis cases in Iran using time series analysis. *Iran J Public Health*. 2015;44:1526–34.
- Tian CW, Wang H, Luo XM. Time-series modelling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. *Epidemiol Infect*. 2019;147:e82.
- Wang T, Zhou Y, Wang L, et al. Using an autoregressive integrated moving average model to predict the incidence of hemorrhagic fever with renal syndrome in Zibo, China, 2004-2014. *Jpn J Infect Dis*. 2016;69(4):279–84.
- Gharbi M, Quenel P, Gustave J, et al. Time series analysis of dengue incidence in Guadeloupe, French West Indies: Forecasting models using climate variables as predictors. *BMC Infect Dis*. 2011;11:166.
- López-Montenegro LE, Pulecio-Montoya AM, Marcillo-Hernández GA. Dengue cases in Colombia: mathematical forecasts for 2018-2022. *MEDICC Rev*. 2019;21(2–3):38–45.
- Zheng Y-L, Zhang L-P, Zhang X-L, et al. Forecast model analysis for the morbidity of tuberculosis in Xinjiang, China. *PLoS ONE*. 2015;10(3):e0116832.
- Liao Z, Zhang X, Zhang Y, et al. Seasonality and Trend Forecasting of Tuberculosis Incidence in Chongqing, China. *Interdiscip Sci*. 2019;11(1):77–85.
- Carvajal Thaddeus M, Viacrusis Katherine M, Hernandez Lara Fides T, et al. Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila, Philippines. *BMC Infect Dis*. 2018;18:183.
- Mao Q, Zhang K, Yan W, et al. Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *J Infect Public Health*. 2018;11(5):707–12.
- Anokye R, Acheampong E, Owusu I, et al. Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. *Cogent Soc Sci*. 2018;4(1):1461544.
- Withanage GP, Viswakula SD, Yi SGN, et al. A forecasting model for dengue incidence in the district of Gampaha, Sri Lanka. *Parasit Vectors*. 2018;11(1):262.

14. Siregar FA, Makmur T, Saprin S. Forecasting dengue hemorrhagic fever cases using ARIMA model: a case study in Asahan district. In: IOP Conference Series Materials Science and Engineering; 2018. p. 300.
15. Tohidinik HR, Mohebal M, Mansournia MA, et al. Forecasting zoonotic cutaneous leishmaniasis using meteorological factors in eastern Fars province, Iran: a SARIMA analysis. *Tropical Med Int Health*. 2018;23(8):860–9.
16. Xu Q, Li R, Liu Y, et al. Forecasting the incidence of mumps in Zibo City based on a SARIMA model. *Int J Environ Res Public Health*. 2017;14(8):925.
17. Wang H, Tian CW, Wang WM, et al. Time-series analysis of tuberculosis from 2005 to 2017 in China. *Epidemiol Infect*. 2018;146(8):935–9.
18. Box GEP, Jenkins GM, Reinsel GC, et al. Time series analysis: forecasting and control, 5th edition. *J Oper Res Soc*. 2015;22(2):199–201.
19. Box, George E.P, Jenkins, Gwilym M, Reinsel, Gregory C. Time series analysis. Forecasting and control. 3rd ed. *Journal of Time*. 2010;31(4):303.
20. Moon MS, Kim SS, Moon H. (i) Tuberculosis of the spine: Current views in diagnosis, management, and setting a global standard. *Orthopaedics Trauma*. 2013;27(4):185–94.
21. Maitra A, Bates S, Shaik M, et al. Repurposing drugs for treatment of tuberculosis: a role for non-steroidal anti-inflammatory drugs. *Br Med Bull*. 2016;118(1):138–48.
22. Berlin L. Tuberculosis: resurgent disease, renewed liability. *AJR Am J Roentgenol*. 2008;190(6):1438–44.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

