

RESEARCH

Open Access



Spatio-temporal evolution and trend prediction of the incidence of Class B notifiable infectious diseases in China: a sample of statistical data from 2007 to 2020

Ruo-Nan Wang^{1†}, Yue-Chi Zhang², Bo-Tao Yu¹, Yan-Ting He¹, Bei Li^{1*†} and Yi-Li Zhang^{1*†}

Abstract

Background: With the accelerated global integration and the impact of climatic, ecological and social environmental changes, China will continue to face the challenge of the outbreak and spread of emerging infectious diseases and traditional ones. This study aims to explore the spatial and temporal evolutionary characteristics of the incidence of Class B notifiable infectious diseases in China from 2007 to 2020, and to forecast the trend of it as well. Hopefully, it will provide a reference for the formulation of infectious disease prevention and control strategies.

Methods: Data on the incidence rates of Class B notifiable infectious diseases in 31 provinces, municipalities and autonomous regions of China from 2007 to 2020 were collected for the prediction of the spatio-temporal evolution and spatial correlation as well as the incidence of Class B notifiable infectious diseases in China based on global spatial autocorrelation and Autoregressive Integrated Moving Average (ARIMA).

Results: From 2007 to 2020, the national incidence rate of Class B notifiable infectious diseases (from 272.37 per 100,000 in 2007 to 190.35 per 100,000 in 2020) decreases year by year, and the spatial distribution shows an “east-central-west” stepwise increase. From 2007 to 2020, the spatial clustering of the incidence of Class B notifiable infectious diseases is significant and increasing year by year (Moran’s I index values range from 0.189 to 0.332, $p < 0.05$). The forecasted incidence rates of Class B notifiable infectious diseases nationwide from 2021 to 2024 (205.26/100,000, 199.95/100,000, 194.74/100,000 and 189.62/100,000) as well as the forecasted values for most regions show a downward trend, with only some regions (Guangdong, Hunan, Hainan, Tibet, Guangxi and Guizhou) showing an increasing trend year by year.

Conclusions: The current study found that since there were significant regional disparities in the prevention and control of infectious diseases in China between 2007 and 2020, the reduction of the incidence of Class B notifiable infectious diseases requires the joint efforts of the surrounding provinces. Besides, special attention should be paid to provinces with an increasing trend in the incidence of Class B notifiable infectious diseases to prevent the

[†]Ruo-Nan Wang is the first author, with Yi-Li Zhang and Bei-Li as corresponding authors.

*Correspondence: 45510307@qq.com; zhangyi1223882953@163.com

¹School of Health Management, Southern Medical University, Guangzhou 510515, China

Full list of author information is available at the end of the article



© The Author(s) 2022. **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.

re-emergence of certain traditional infectious diseases in a particular province or even the whole country, as well as the outbreak and spread of emerging infectious diseases.

Keywords: Class B notifiable infectious diseases, Incidence, Spatial and temporal evolution, Trend prediction

Background

The occurrence and transmission patterns of infectious diseases are changing with the accelerated global integration and the impact of climatic, ecological and social environmental changes [1, 2]. Recent outbreaks of infectious diseases such as Ebola, Zika and Corona Virus Disease 2019 (COVID-19) demonstrate that emerging and traditional infectious diseases are posing and will continue to pose a threat to human life and health, which brings new challenges for the emergency response capacity of the whole world, especially developing countries with limited available resources [3–5]. Since the outbreak of the Severe acute respiratory syndrome (SARS) in 2003, China, as the world's largest developing country, has made many efforts to build a response system for public health emergencies, promulgating a number of emergency plans, overhauling the national information system for disease prevention and control, and launching National Major Scientific and Technological Special Project for Prevention and Control. Yet the latest epidemic profile of notifiable infectious diseases released by the Chinese Centre for Disease Control and Prevention show that the number of reported cases of notifiable infectious diseases in mainland China remained as many as 699,466, and that 1531 people died between 00:00 on 1 May 2021 and 24:00 on 31 May 2021 [6]. Thus China still remains under serious threat from infectious diseases.

The primary task for controlling infectious diseases is to understand the distribution characteristics of epidemiology. According to the existing research, about 80% of epidemiological data are spatial in nature [7]. Therefore, the exact analysis of the spatial distribution characteristics of infectious diseases is a must for the effective study of the causes and other influencing factors of diseases and also for the formulation of effective prevention and control strategies. In recent years, with a mature statistical method, spatio-temporal statistics, researchers can not only conduct a dynamic analysis of the temporal and spatial distribution characteristics of infectious diseases, but also summarize the spatio-temporal transmission patterns by considering the three-dimensional environment in which they occur and are prevalent [8, 9]. So far, extensive studies of the epidemiological characteristics of infectious diseases have been conducted by many scholars employing spatial statistical methods based on geographic information system (GIS) [10–14].

Modelling spatio-temporal trends in infectious diseases is of great importance, because real-time epidemiological forecasting can help to predict the geographic expansion of diseases as well as the number of cases. What's more, the exact prediction of the epidemic outbreak helps policy makers to prepare early so that they can better implement public health interventions. As the infectivity of pathogens and the availability of drugs and vaccines change over time, the application of updated shared data is much necessary for the evaluation and prediction of disease hazard. In China, the National Health Commission of the People's Republic of China regularly releases on its official website the annual incidence of Class A, Class B and Class C notifiable infectious diseases, providing a platform for further study of the epidemiological characteristics of notifiable infectious diseases in China. In the Law of the PRC on the Prevention and Treatment of Infectious Diseases enacted in 1989, infectious diseases are classified into three categories, Class A, Class B and Class C, among which the former two can both cause large-scale severe epidemics within a short period of time, while the third is less infectious and causes only minor outbreaks. The number of notifiable infectious diseases in the three categories is ever-changing with the outbreak of emerging ones. For example, on October 2, 2020, the National Health Commission issued a draft of the revised Law on Prevention and Control of Infectious Diseases for consultation, which clearly states that two new types of infectious diseases, namely human H7N9 avian influenza and novel coronavirus, have been added to Class B. Currently, the three categories contain altogether 41 notifiable infectious diseases, as shown in Additional file 1: Annex 1. Epidemic incidence data have been shown to be valuable epidemiological tools for real-time assessment and prediction of trends and transmission potential [15–18]. By means of the data, predictive models can be used to help provide timely forecasts of disease incidence and geographic spread of emerging epidemics. Auto Regressive Integrated Moving Average (ARIMA) is a time-domain tool for time series analysis that has been widely used for infectious disease prediction [19, 20]. For example, many scholars have recently used ARIMA models to predict the incidence of COVID-19 [21–23], as it was used to predict the incidence of other infectious diseases such as viral hepatitis [24], malaria [25, 26] and measles [27].

Despite the accumulated findings, there are still limitations in the existing research. In general, most of the current researches on infectious diseases are the study of one specific disease, and the hotspots of research have been focused on a single study of the current epidemiological status, prediction of incidence trends and spatial attributes of the disease, lacking systematic research on the epidemiological characteristics and trend prediction of multi-species combinations. Since there is growing evidence of the complexity of disease interactions and disease etiology, a multi-species analysis to capture the overall trends of infectious diseases may be a worthwhile approach. Therefore, by using the method of spatial econometric analysis and time series forecasting combined and taking 31 Chinese provinces (Hong Kong, Macao and Taiwan are excluded) as the study unit, this study aims to focus on the current spatial and temporal distribution of the incidence of highly infectious Class B notifiable infectious diseases, and to predict the future development trend of the incidence of Class B notifiable infectious diseases. Hopefully, it will help Chinese public policy makers to formulate better health policy intervention measures, build better response of China's public health epidemic prevention system to disease outbreaks, and provide for other countries valuable information to develop better intervention strategies to prevent and control the spread of emerging and re-emerging infectious diseases.

Data and methods

Data sources

Firstly, with Class B notifiable infectious diseases as the research object, only data on the incidence of Class B were collected, because Class C require only appropriate surveillance since they do not cause any serious consequences, and Class A, which include only two types of infectious diseases, i.e. plague and cholera, are relatively few in number and have been almost eradicated in China. Secondly, given the availability of data, the data on the incidence of Class B notifiable infectious diseases used in this study were that of 31 provinces in mainland China (except Hong Kong, Macao and Taiwan) and that of the whole country for each year during the period 2007–2020. The data used in this study were obtained from public sources: (1) Data on the incidence of Class B notifiable infectious diseases were obtained from the China Health Statistical Yearbook 2008–2012 and the China Health and Family Planning Statistical Yearbook 2013–2021 published by the National Health Commission of the People's Republic of China (<http://www.nhc.gov.cn>). (2) Geographic information based data were obtained from the standard map service website of the National

Bureau of Surveying, Mapping and Geographic Information (<http://bzdt.ch.mnr.gov.cn>).

Research methodology

Global spatial autocorrelation [28]

Global spatial autocorrelation analysis is performed at the national macro level by comparing the mean values of attributes aggregated over the overall region with the values of attributes on each spatial unit to derive the average degree of association between the incidence of Class B notifiable infectious diseases in each province and region at the national level, i.e. to determine whether there is any clustering of the incidence of Class B notifiable infectious diseases at the national level. Global spatial autocorrelation analysis usually uses the Moran's I index to determine whether the spatial distribution of regional variables is statistically clustered or dispersed, and is calculated using the formula:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (x_i - \bar{x})^2}$$

In the formula: n is the number of provincial administrative districts; x_i is the incidence of Class B notifiable infectious diseases in the i th provincial administrative district; W_{ij} is the spatial weight matrix, which is determined here by means of spatial geographical adjacency, i.e. if two places are adjacent, the corresponding element in the matrix takes 1, otherwise it takes 0; \bar{x} is the mean incidence rate of Class B notifiable infectious diseases aggregated across all n provincial administrative districts.

Moran's I index values generally range from $[-1, 1]$. At a given level of significance, $I > 0$ indicates a positive spatial correlation, with larger values indicating more pronounced spatial aggregation; $I < 0$ indicates a negative spatial correlation, with smaller values indicating more pronounced spatial dispersion; and $I = 0$ indicates that the observations are randomly distributed in space.

ARIMA model [29]

The ARIMA model is one of the most commonly used methods for infectious disease time series forecasting and has been shown to be highly accurate [30]. In ARIMA (p, d, q), AR(p) is an autoregressive model and MA(q) is a moving average model, with p, d and q being the number of autoregressive terms, the number of differences and the number of moving average terms respectively. The ARIMA (p, d, q) model is an autoregressive moving average process with smoothness obtained after d differences of the AR(p) process and the MA(q) process.

ARIMA modelling includes smooth constructive simulation fitting, parameter estimation and model diagnosis. The modelling process in this study consists of three main steps. Firstly, as the annual data used in this study did not show seasonal variation and were non-stationary, the series were differenced at the non-seasonal level, and the autocorrelation and partial autocorrelation plots were plotted to determine whether the differenced time series were stationary. Secondly, the best model was screened using the Bayesian Information Criterion Error (BIC) (the smaller the value, the better the model) based on the Box-Ljung test that the residual series were white noise. Finally, the accuracy of the model was evaluated in two ways. On the one hand, the fit of the ARIMA model between the actual and predicted values was determined by observing whether the actual values were within the 95% confidence interval of the predicted values. On the other hand, the accuracy of the ARIMA model was evaluated by calculating the relative error (RE) and the mean absolute percentage error (MAPE). The prediction accuracy of the model is generally considered to be high when the MAPE is <10%. RE and MAPE are calculated as $RE = \left| \frac{\hat{y}_t}{y_t} - 1 \right| \times 100\%$ and $MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{\hat{y}_t}{y_t} - 1 \right|$, where the equations y_t and \hat{y}_t denote the actual and predicted values at time point respectively.

As COVID-19 epidemic broke out in China in 2020, the 2020 data for Class B notifiable infectious diseases may affect the selection of the best model and in turn, the accuracy of the prediction results. Therefore, we first used 2007–2019 data for the above model construction for accurate forecast, followed by verifying the reliability of the best model selected in the above modelling process with 2020 data. If the relative error between the forecasted and actual incidence of Class B notifiable infectious diseases in 2020 is still less than 10%, it proves the reliability of the best model selected, which will continue to be used for short-term forecasting of the incidence of Class B notifiable infectious diseases.

Spatial correlation analysis in this study was performed using Stata 14.0 software; ARIMA model prediction using SPSS 26.0 software and mapping using ArcGIS 10.2 software. $\alpha = 0.05$ was used as the significance criterion.

Results

Analysis of the spatial and temporal evolution of the incidence of Class B notifiable infectious diseases in China

Temporal evolutionary characteristics

In order to study the trends in the incidence rates of Class B notifiable infectious diseases in different regions of China, this study divided the 31 provinces of China into three regions (east, central and west) in the light of

economic development level and geographical location, with the east the highest in the economic development level while the west the lowest [31]. Figure 1 shows the trend of the incidence rate of Class B notifiable infectious diseases in this study in order to analyse regional differences in the incidence rate of infectious diseases in China. From 2007 to 2020, China's national incidence of Class B notifiable infectious diseases shows a decreasing trend year by year (from 272.37/100,000 in 2007 to 190.35/100,000 in 2020), with the east dropping from 249.93/100,000 in 2007 to 165.45/100,000 in 2020, the central from 262.18/100,000 in 2007 to 191.63/100,000 in 2020 and the west from 360.07/100,000 in 2007 to 232.18/100,000 in 2020. The national incidence of Class B notifiable infectious diseases is on a gradual increase from the east to the central and then to the west, with the east the lowest incidence of Class B notifiable infectious diseases, while the west the highest. As for the decreasing range, the biggest is in the west (35.52%), followed by the east (33.80%) and the central (26.91%).

Spatial distribution characteristics

The incidence data of Class B notifiable infectious diseases in 2007 and 2020 in China were selected and visualized. Besides, in order to show the differences between regions more clearly, the same interval method in ArcGIS 10.2 software was used to classify the incidence into five classes: low (0/100,000~), slightly low (115/100,000~), medium (230/100,000~), slightly high (345/100,000~) and high (460/100,000~), see Fig. 2. In terms of spatial distribution, the incidence rate of Class B notifiable infectious diseases in each province has obvious spatial variation, with an overall increasing distribution of "east-central-west", showing a clear correlation with the level of economic development. In the economically developed eastern region, with the exception of a few provinces, the incidence rates of Class B notifiable infectious diseases are all at or below the lower end of the provincial incidence rankings. For example, the incidence rates of Class B notifiable infectious diseases in Jiangsu Province in 2007 and 2020 were 172.3 per 100,000 and 98.23 per 100,000 respectively, ranking 30th and 29th among all the provinces. Compared with the east and central regions, the incidence rates of Class B notifiable infectious diseases in the less developed western regions are relatively high. For example, the levels of Class B notifiable infectious diseases in Xinjiang (657.52/100,000) and Qinghai (463.63/100,000) in 2007 and Qinghai (376.80/100,000) and Xinjiang (324.99/100,000) in 2020 are all at the higher and above, ranking in the top four among provinces. Compared with 2007, the incidence rates of Class B notifiable infectious diseases decreased in all provinces in 2020, except Guangdong, Hainan, Hunan and Tibet,

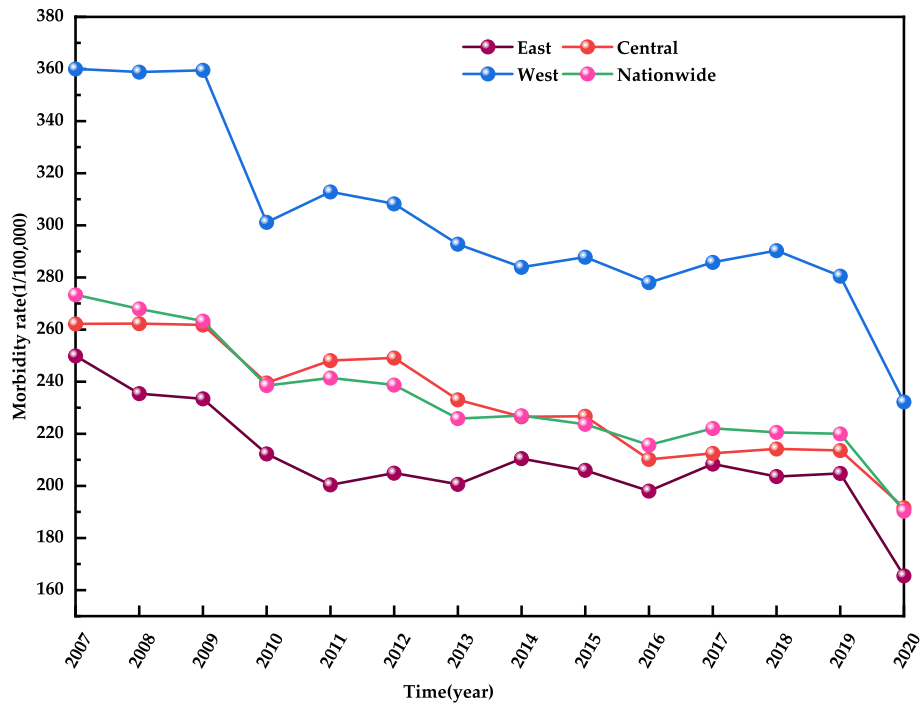


Fig. 1 Regional incidence rates of Class B notifiable infectious diseases, China, 2007–2020

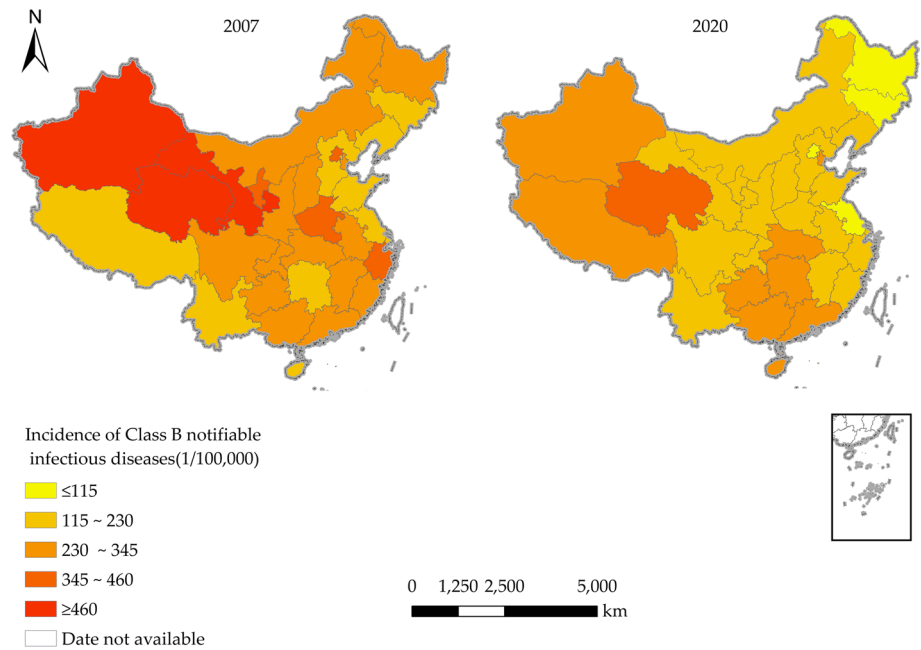


Fig. 2 Provincial distribution of the incidence of Class B notifiable infectious diseases, 2007 and 2020

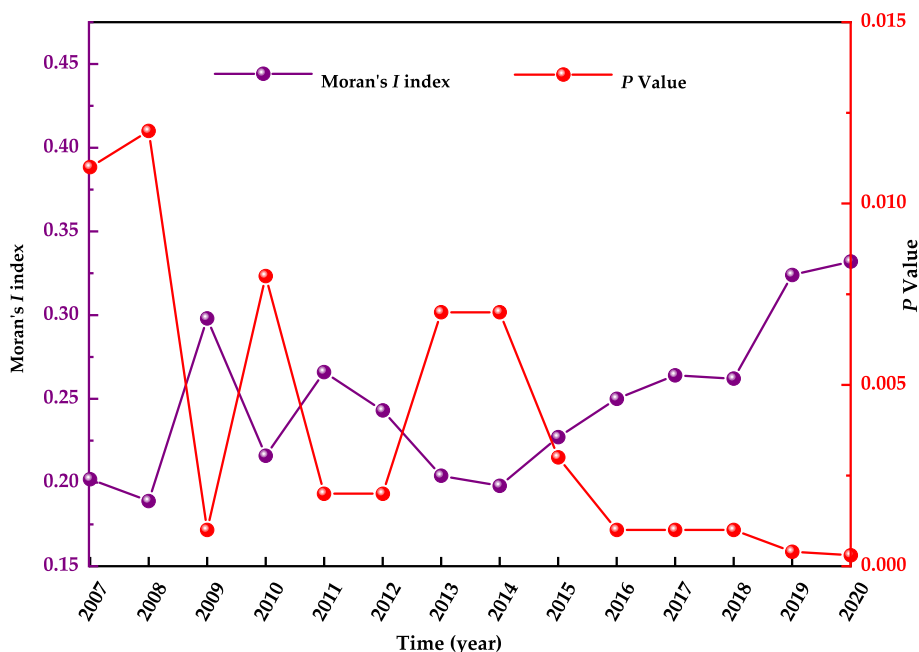


Fig. 3 Moran's *I* index chart of the incidence of Class B notifiable infectious diseases, 2007–2020

where the incidence rates of Class B notifiable infectious diseases increased.

Spatial autocorrelation analysis of the incidence of Class B notifiable infectious diseases in China

Figure 3 shows the trend of Moran's *I* index for the incidence of Class B notifiable infectious diseases in China during the period 2007–2020, derived from the global spatial autocorrelation analysis of this study. The Moran's *I* index values for the incidence of Class B notifiable infectious diseases in China during the period 2007–2020 ranged from 0.189 to 0.332, with an overall increasing trend and reaching significance levels in all years ($P < 0.05$) (see Fig. 3).

Predicting the incidence of Class B notifiable infectious diseases in China

The national incidence rate of Class B notifiable infectious diseases was differenced 1 time ($d=1$) to obtain an autocorrelation function (ACF) plot and a partial autocorrelation function (PACF) plot (see Fig. 4). The results show that the confidence intervals are within the range of $[-0.5, 0.5]$, so the series after the first-order difference tends to be smooth.

The ARIMA model was fitted with the national incidence rate of Class B notifiable infectious diseases as the dependent variable, and the model was screened by Box-Ljung test and BIC index. Considering that p and q values generally did not exceed 2, a trial subscription

from 0 to 2 was executed. Only five models passed the Box-Ljung test ($P > 0.05$): ARIMA (0, 1, 1), ARIMA (0, 1, 2), ARIMA (1, 1, 1), ARIMA (1, 1, 0), and ARIMA (2, 1, 1). Table 1 shows the statistics, BIC and parameter estimation results for the five models. The ARIMA (1, 1, 1) model was finally chosen as the optimal model with $BIC = 4.631$, standard $R^2 = 0.543$ and $MAPE = 2.169\%$.

Table 2 shows the results of the fit using the ARIMA (1,1,1) model, which shows that the RE values of the data for each period is less than 10%, indicating that the model fits well.

Figure 5 shows the backgeneration and prediction of the national incidence of Class B notifiable infectious diseases using the ARIMA (1,1,1) model. According to Fig. 5, the actual values are within the 95% confidence interval of the predicted values, a further indication that the model fits well. The reliability of the ARIMA (1,1,1) model was verified using the data for 2020, and the RE value between the predicted value (207.67/100,000) and the actual value (190.35/100,000) for 2020 was 8.34% ($RE < 10\%$), which indicates that the model has a good prediction effect. By comparing the predicted value with the actual value in 2020, we can find that the predicted value is larger than the actual value. The predicted incidence rates of Class B notifiable infectious diseases for 2021–2024 are 205.26/100,000, 199.95/100,000, 194.74/100,000 and 189.62/100,000 respectively, showing a decreasing trend year by year, as shown in Fig. 5.

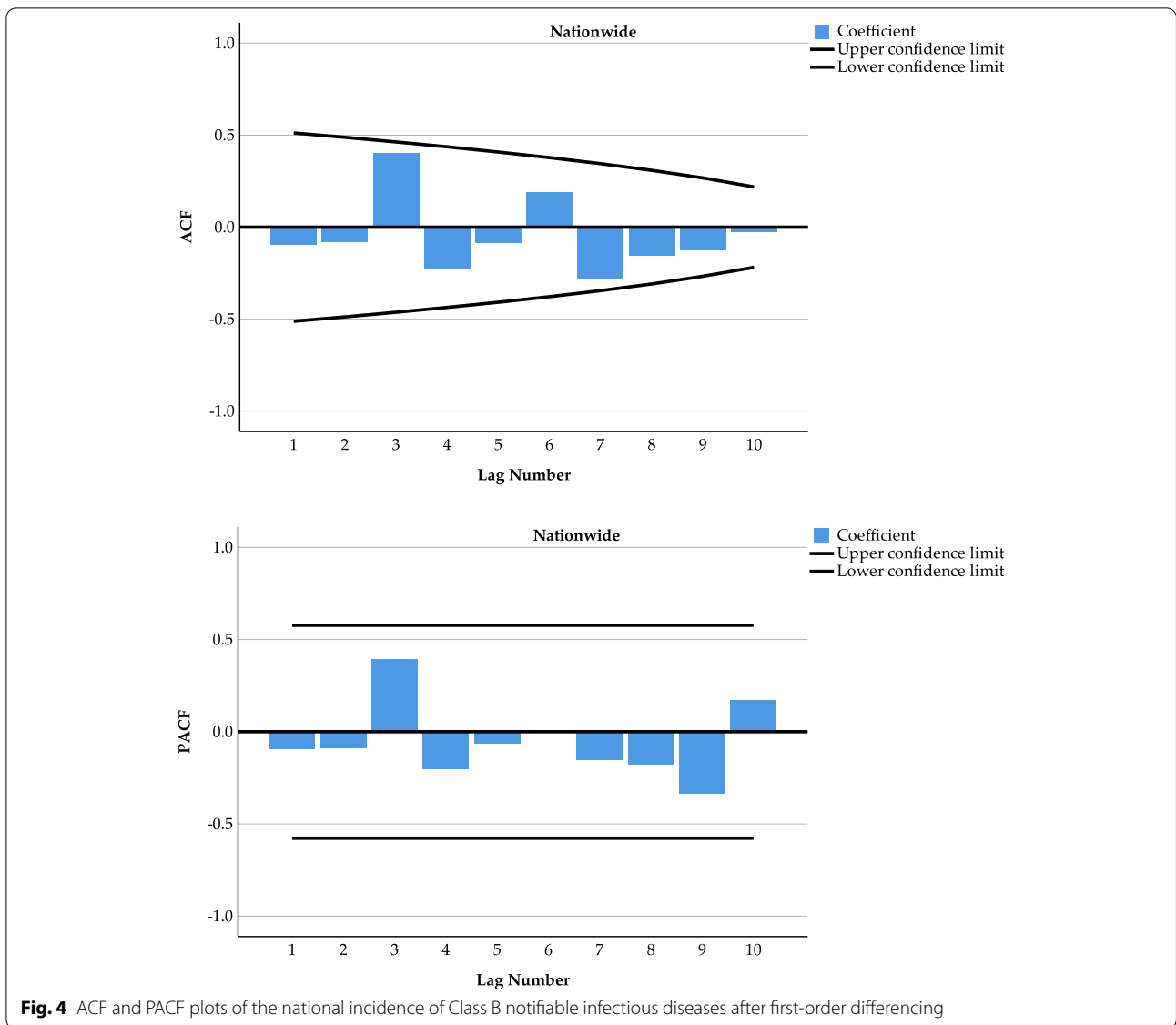


Fig. 4 ACF and PACF plots of the national incidence of Class B notifiable infectious diseases after first-order differencing

Table 1 Parameter estimation and model validation of the ARIMA model

Models	Fitted Model Statistics			Ljung-Box Q(18)	
	Stationary R ²	MAPE	BIC	Statistics	Sig.
ARIMA(0,1,1)	0.47	2.358	4.831	14.469	0.481
ARIMA(0,1,2)	0.219	3.427	5.983	21.53	0.198
ARIMA(1,1,1)	0.543	2.169	4.631	15.472	0.378
ARIMA(1,1,0)	0.453	2.523	4.934	17.571	0.329
ARIMA(2,1,1)	0.273	2.743	5.778	20.678	0.217

The ARIMA model was used to fit the incidence rates of Class B notifiable infectious diseases in 31 provinces, municipalities and autonomous regions of China from

2007 to 2019, and the MAPE values for all provinces were less than 10%. Meanwhile, the results of the back-generation validation using 2020 data showed that the RE values between the predicted and actual values in 2020 were also less than 10% for all the provinces except Hubei, indicating good fitting results for each province. See Table 3. It is worth mentioning that Hubei Province, as the initial outbreak place of the epidemic in China, may have been hit so hard by the epidemic that the best model selected did not pass the backgeneration validation in 2020, and if the selected model continues to be used for forecasting it will possibly result in a large bias, so we conducted prediction by using the best model selected for each province excluding Hubei, and the incidence data for 2021–2024 were obtained for each

Table 2 National ARIMA model fitting results for the incidence of Class B notifiable infectious diseases

Year	Actual value	Forecast	Relative error (%)
2007	272.37	272.47	0.04
2008	267.93	269.68	0.65
2009	263.29	264.69	0.53
2010	238.47	251.91	5.64
2011	241.44	237.75	1.53
2012	238.75	237.78	0.41
2013	225.8	232.02	2.75
2014	226.97	221.99	2.19
2015	223.6	223.24	0.16
2016	215.68	219.68	1.85
2017	222.06	208.09	6.29
2018	220.51	214.99	2.50
2019	219.98	211.92	3.66

province and city, as shown in Table 3. What’s more, the predicted incidence of Class B notifiable infectious diseases was visualised for each province and is shown in Fig. 6. The predicted incidence in most regions of China from 2021 to 2024 shows a decreasing trend year by year, while it shows the opposite in such regions as Guangdong, Hainan, Hunan, Guangxi, Guizhou and Tibet, whose predicted incidence rates of Class B notifiable infectious diseases in 2021–2024 range from 301.41/100,000 to 311.27/100,000, 375.67/100,000 to 436.99/100,000, 306.55/100,000 to 326.71/100,000, 303.16/100,000 to 307.43/100,000, 264.32/100,000 to 280.62/100,000, and 389.49/100,000 to 476.63/100,000

respectively. The overall incidence rate of Class B notifiable infectious diseases in China still shows a distribution pattern of high in the west and low in the east, but the difference between the east and the west is decreasing year by year.

Discussions

This study uses the incidence rates of Class B notifiable infectious diseases published on the official website of the National Health Commission of the People’s Republic of China from 2007 to 2020 as a data source to analyse the characteristics of the spatial and temporal distribution of the incidence rates of Class B notifiable infectious diseases in China, and to forecast their development trends. The results of the study show that, in terms of spatial distribution, the overall incidence of Class B notifiable infectious diseases in China during the period 2007–2020 showed an increasing distribution of “East-Central-West”, indicating that the incidence of Class B notifiable infectious diseases has a certain correlation with the level of economic development of the region. The incidence rate of Class B infectious diseases is correlated with the economic development level of the region. Previous studies have confirmed that higher levels of economic development can be beneficial to the reduction of likelihood of infectious disease outbreaks, as well as the disease hazards associated with infectious disease outbreaks [32]. For example, Wu et al. [33] explored the impact of climate change on human infectious diseases and found that developing countries face greater health risks from infectious disease outbreaks than developed countries because their public health systems lack the resources and capacity to respond effectively to the challenges. The present study confirms that the incidence of Class

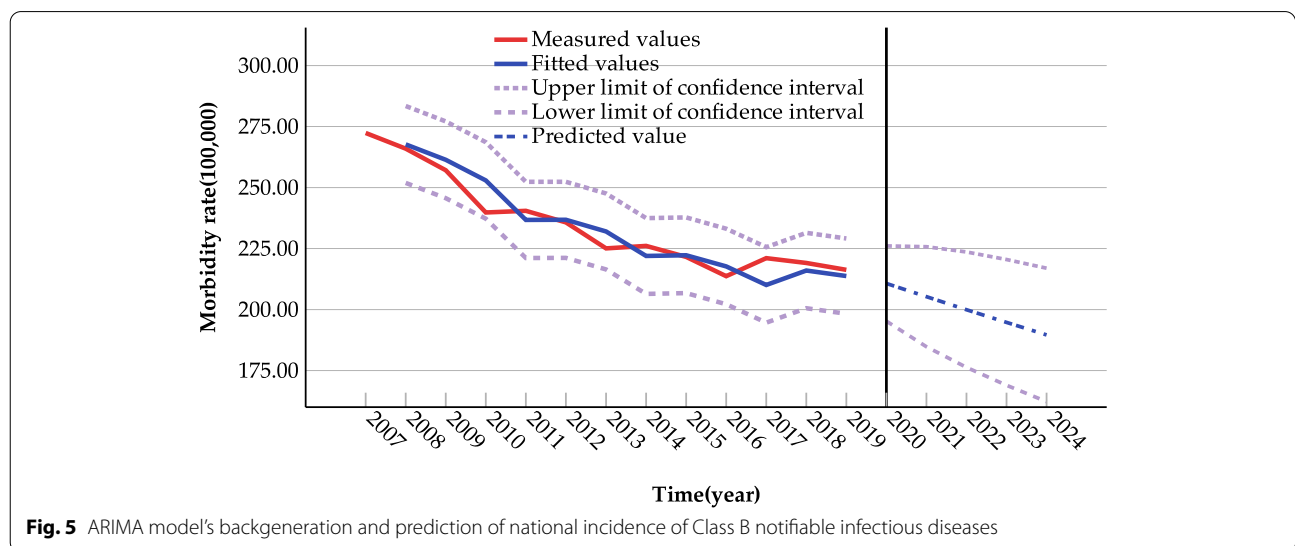


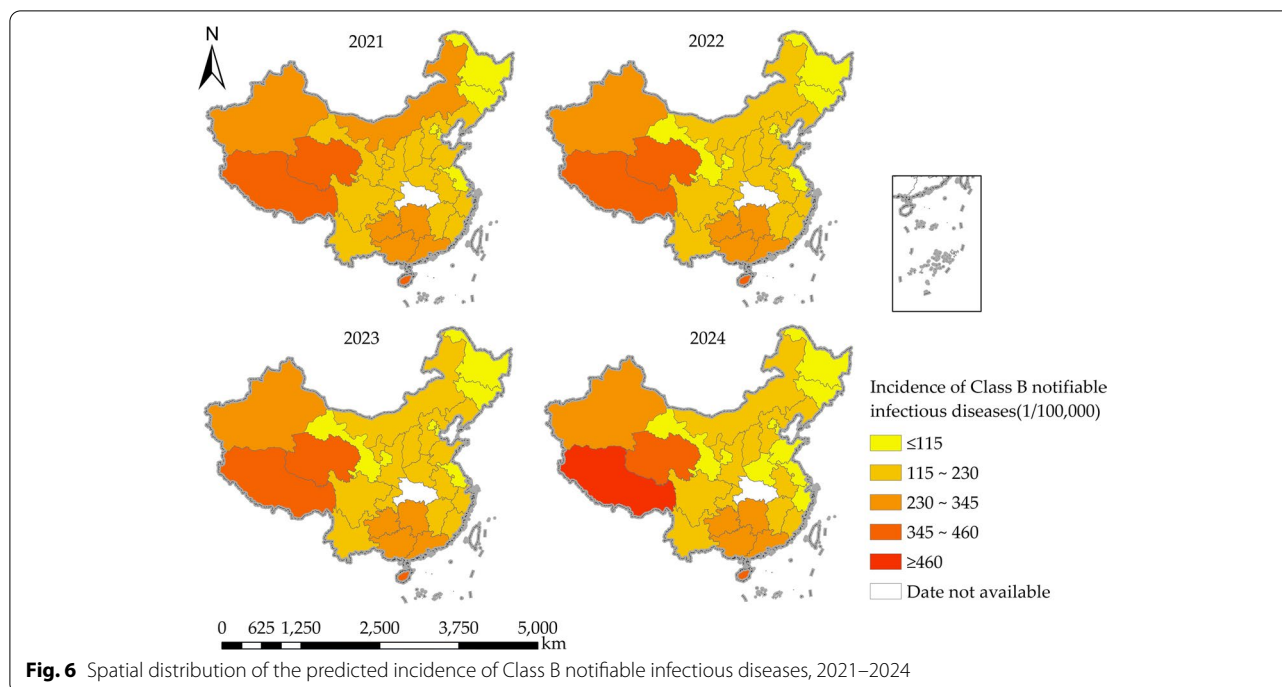
Fig. 5 ARIMA model’s backgeneration and prediction of national incidence of Class B notifiable infectious diseases

Table 3 Fitting and prediction results of ARIMA model for the incidence of Class B notifiable infectious diseases by district

Region	Best model	MAPE(%)	Actual value	RE(%)	Forecast value				
			2020	2020	2020	2021	2022	2023	2024
Anhui	(0, 1, 1)	2.34	225.08	1.18	227.73	225.74	223.55	221.35	219.74
Beijing	(1, 1, 1)	3.11	80.70	9.43	88.31	81.69	75.44	70.12	65.58
Fujian	(1, 1, 1)	2.18	223.99	7.87	241.61	226.61	211.42	198.46	185.73
Gansu	(2, 1, 1)	1.47	150.52	1.22	152.36	128.79	110.47	96.53	78.58
Guangdong	(0, 1, 1)	2.98	272.49	1.72	277.19	301.41	305.21	308.19	311.27
Guangxi	(0, 2, 1)	4.76	263.62	2.57	270.40	303.16	305.60	306.45	307.43
Guizhou	(1, 1, 1)	3.90	248.35	4.01	258.32	264.32	269.17	275.41	280.62
Hainan	(0, 1, 1)	3.54	339.12	4.29	353.68	375.67	397.33	415.97	436.99
Hebei	(1, 2, 1)	2.18	134.26	8.49	145.66	138.68	134.28	129.73	125.71
Henan	(1, 1, 1)	2.11	152.77	2.26	156.22	149.60	134.98	120.36	105.74
Heilongjiang	(0, 1, 1)	5.34	105.84	7.35	113.62	110.54	106.46	101.37	96.29
Hubei	(1, 2, 1)	2.18	290.81	19.55	233.95	–	–	–	–
Hunan	(1, 1, 1)	2.76	269.01	3.65	278.83	306.55	313.27	319.99	326.71
Jilin	(1, 1, 0)	4.72	87.30	8.68	94.88	93.99	90.61	88.79	87.55
Jiangsu	(0, 1, 1)	3.11	98.23	5.48	103.61	96.12	91.53	86.85	81.07
Jiangxi	(0, 1, 1)	4.15	199.30	7.12	213.49	204.15	195.76	187.49	181.49
Liaoning	(0, 2, 1)	3.13	155.22	6.94	165.99	152.99	139.21	131.87	123.72
Neimenggu	(1, 1, 1)	6.11	225.64	7.78	243.19	232.16	221.12	211.09	195.05
Ningxia	(0, 1, 1)	2.18	168.75	6.15	179.13	161.17	148.23	131.27	117.32
Qinghai	(2, 1, 1)	4.81	376.80	6.84	402.58	398.67	388.58	382.72	377.18
Shandong	(1, 1, 1)	2.31	131.73	8.77	143.28	132.21	125.17	119.81	112.05
Shanxi	(0, 1, 1)	1.74	202.95	8.33	219.86	209.24	201.35	192.98	184.65
Shaanxi	(0, 1, 1)	2.73	150.48	7.87	162.33	154.35	148.41	142.45	138.49
Shanghai	(1, 1, 1)	2.28	128.63	7.04	137.69	131.03	124.27	118.52	112.77
Sichuan	(1, 2, 0)	1.37	196.03	3.91	203.69	195.03	188.27	180.52	172.77
Tianjin	(1, 1, 1)	2.43	109.28	9.18	119.31	105.11	92.04	83.48	72.01
Xizang	(0, 1, 2)	5.69	337.10	8.79	366.74	389.49	419.56	441.91	476.63
Xinjiang	(1, 1, 2)	4.82	324.99	8.75	353.42	341.57	329.14	313.63	294.12
Yunnan	(0, 2, 1)	4.51	189.61	9.04	206.75	201.90	195.03	190.16	184.29
Zhejiang	(1, 1, 1)	1.32	146.47	6.26	155.64	147.54	132.59	123.39	111.53
Chongqing	(1, 2, 0)	3.18	202.07	5.47	213.13	210.60	209.49	207.61	204.85

B notifiable infectious diseases is lower in the economically developed eastern region of China than the central and western regions, which is easily explained by the fact that the eastern region has more medical and educational resources and better infrastructure than the central and western regions, and will suffer less from the burden of disease caused by infectious diseases, especially vaccine-accessible infectious diseases such as measles, whooping cough and tuberculosis [34]. However, the incidence of Class B notifiable infectious diseases decreased the most in the west, and the regional differences in incidence rates was on a decreasing trend. The analysis suggests that although the western region is relatively lagging behind in economic development, it has benefited from the Western Development Policy in recent years, making

it possible to achieve significant results in the prevention and treatment of infectious diseases in the western region. The identification of regional differences in the incidence of Class B notifiable infectious diseases is conducive to the adoption of differentiated infectious disease prevention and control intervention strategies for different regions. Since effective prevention and control of infectious diseases in the central and western regions is crucial to reducing the overall incidence of infectious diseases nationwide, it is much necessary to continue to increase the investment and talent introduction policies for the central and western regions. Through the “policy dividend” [35], the economic development of the central and western regions and the level of health prevention and control can be improved. Meanwhile, the eastern



region, as the driving force in the improvement of infectious disease prevention and control, should interact well with the central and western regions and play a leading role in order to reduce the health hazards of infectious diseases to the Chinese population.

A lot of studies have shown significant spatial clustering in the incidence of Class B notifiable infectious diseases such as typhoid and paratyphoid fever, dengue fever and novel coronavirus [36–39]. According to the current study, the Moran's I index for the incidence of Class B notifiable infectious diseases in China from 2007 to 2020 was significantly positive, indicating that the incidence of Class B notifiable infectious diseases showed a strong clustering in spatial distribution, which is consistent with the results of previous studies. Besides, the Moran's I index values show an increasing trend year by year, indicating a gradual increase in the degree of clustering in the spatial distribution of the incidence of Class B notifiable infectious diseases. In other words, there is a spatial correlation between the incidence of Class B notifiable infectious diseases in each province of China and that in neighbouring provinces, and the degree of correlation is gradually increasing. This result confirms the fact that no provincial unit can control infectious diseases without cooperating with other provinces in the fight against them, and no region can protect itself from an infectious disease crisis alone. Therefore, in order to reduce the incidence of Class B notifiable infectious diseases in a particular province, the level of economic

development and the level of sanitary and epidemiological protection in the province as well as in the surrounding provinces need to be taken into account as important factors affecting the effectiveness of infectious disease prevention and control [40–42]. What's more, in order to reduce the overall incidence of Class B notifiable infectious diseases in China, it is recommended that a regional community for improvement in infectious disease prevention and control be formed by breaking down geographical and administrative barriers and delineating a multi-level regional framework, and that inter-provincial mechanisms for joint prevention and control of infectious diseases be established so as to give full play to the positive spatial spillover effects of the positive prevention and control of infectious diseases across the region.

According to the incidence rates of Class B notifiable infectious diseases in China from 2007 to 2020, the national overall rates as well as the rates in each province are decreasing except a certain rise in four provinces (Guangdong, Hainan, Hunan, and Tibet). This indicates that we have achieved notable results in the prevention and control of infectious diseases in China, which is consistent with the results of previous studies [43]. According to the research analysis, the fact that the four provinces have not seen a decline in the incidence of Class B notifiable infectious diseases may be mainly related to their special geographical location and climatic conditions [44–49]. In spite of a certain economic boost in Tibet, a border region in the southwest, the increasing trade

flows results in the bigger size of the mobile population, which leads to the problem of new and traditional infectious diseases and their increasingly prominent cross-border transmission. Guangdong, Hunan and Hainan all have tropical or subtropical monsoon climates, which is temperate and conducive to the prevalence of climate-sensitive mosquito-borne viral diseases such as dengue, malaria and encephalitis B.

After verification and adoption of the selected best model using the data of 2020, the short-term forecast for 5 years from 2020 to 2024 was carried out using the best model. The results show that, firstly, compared with the predicted 2020 incidence assuming no COVID-19 epidemic broke out, the actual 2020 incidence of Class B infectious diseases are smaller for both China as a whole and all the provinces (except Hubei Province), suggesting that the prevention and control measures taken in response to the COVID-19 epidemic are conducive to controlling the occurrence and development of other Class B notifiable infectious diseases, which is consistent with the results of existing studies [50–52]. The study suggests that it is the government's mandatory prevention and control strategies as well as the public's increasing awareness of personal health that works. On the one hand, studies have shown that non-pharmaceutical interventions implemented by the government during the COVID-19 epidemic (e.g. school closures, movement restrictions and social distancing) contributed to a decline in the incidence of infectious diseases such as pertussis, scarlet fever and hand, foot and mouth disease (HFMD) [53]. On the other hand, both the initial period of rigorous epidemic prevention and control and the subsequent period of normal practice of them saw an increase in public awareness of personal health. Such intervention measures as wearing masks, social distancing, hand washing and ventilation also effectively prevent the spread of other infectious diseases transmitted through respiratory tract, intestinal tract or intimate contact, for example, whooping cough, scarlet fever, tuberculosis, and brucellosis, etc. Secondly, the forecast of the incidence of Class B notifiable infectious diseases in China from 2021 to 2024 shows that the incidence of infectious diseases is still on a downward trend nationwide and in most provinces, but the predicted incidence of Class B notifiable infectious diseases in such provinces as Guangdong, Guizhou, Hunan, Hainan, Tibet and Guangxi is on an upward trend, which suggests that the government should focus more attention upon provinces where the incidence of infectious diseases is on an upward trend, and tailor specific measures to the actual situation of each province to avoid the re-emergence of certain infectious diseases in a province or even nationwide and the outbreak of new infectious diseases due to

relaxed vigilance. The ARIMA model is highly accurate (within 10%) in predicting the incidence of Class B notifiable infectious diseases, and can effectively compensate for the current lack of capacity to develop, evaluate, manufacture, distribute and manage effective medical countermeasures (e.g. vaccines, diagnostics, etc.), effectively address the unmet disease burden associated with outbreaks or prevalence of traditional and emerging infectious diseases, and effectively guide policy decisions such as rational allocation of health resources and pre-deployment of emergency supplies [54].

There are, of course, certain limitations in this study. The first limitation is about the prediction accuracy of the ARIMA model. On the one hand, the prediction accuracy of the ARIMA model is easily affected by the sample size. The bigger the sample size, the higher the prediction accuracy of the model, with the number of variables unchanged. In this study, based on the principle of indicator data availability, 13 years of data (2007–2019) were finally selected for fitting the ARIMA model, using 2020 data for model validation and forecasting the incidence of Class B notifiable infectious diseases in China as well as in each province from 2021 to 2024. Although the relative errors of the selected models were less than 10% for China as a whole and for each province (except Hubei Province), the prediction accuracy could be further improved in the future by increasing the sample size. On the other hand, it is worth mentioning that while this study recognises that notifiable infectious diseases in Class B are more infectious and may lead to outbreaks or epidemics with relatively high sensitivity and specificity compared with notifiable infectious diseases in category C. The use of incidence data for Class B notifiable infectious diseases can effectively reduce the reduction in predictive accuracy of ARIMA models due to missing reports. However, it is worth mentioning that the prediction results may deviate from the actual values over time due to the changing external environment of the host, e.g. policy instability and vaccine availability [55]. Therefore, it is much necessary to update the forecast of the incidence of Class B notifiable infectious diseases in accordance with the availability of data. Secondly, this study uses the annual incidence of Class B notifiable infectious diseases, a multi-disease joint data, for the prediction study. While this is good for getting a full picture of infectious diseases, it also has shortcomings because even if problems can be identified, it still takes time to determine the specific situation. Future applications will need to be tailored to the specific characteristics of each disease. Finally, the study on the incidence of Class B notifiable infectious diseases is mainly an exploratory spatial data analysis. Future studies can be made on the

basis of the existing studies by combining other variables such as economic factors, transport factors and population mobility factors to conduct empirical spatial data analysis through the establishment of spatio-temporal regression models and to explore the causes and driving mechanisms of the formation and changes in the spatio-temporal patterns of the incidence of Class B notifiable infectious diseases.

Conclusion

This study introduced GIS spatial statistics into the study of infectious diseases, and revealed the spatial association patterns and spatio-temporal evolution characteristics of Class B notifiable infectious diseases in China at a macro level. During the period 2007–2020, infectious disease prevention and control in China has been effective in a certain degree, but there is a clear “east-central-west” decreasing geographical difference in prevention and control effectiveness. The spatial distribution of the incidence rate of Class B notifiable infectious diseases in China is clustered and the degree of clustering is increasing year by year. This study also demonstrates that the ARIMA model can be used to predict the incidence of Class B notifiable infectious diseases, a multi-disease joint data. The results of this study will not only help to facilitate the application of geographic methods to health care, but assist in resource allocation and preparedness planning by accurately predicting the potential geographic extent of disease incidence and transmission of infectious diseases, providing effective information for evidence-based scientific prevention and control of infectious diseases.

Abbreviations

COVID-19: Coronavirus disease 2019; SARS: Severe acute respiratory syndrome; GIS: Geographic information system; ARIMA: Auto regressive integrated moving average; BIC: Bayesian information criterion error; RE: Relative error; MAPE: Mean absolute percentage error; ACF: Autocorrelation function.; PACF: Partial autocorrelation function..

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12889-022-13566-2>.

Additional file 1. The classification of reported infectious diseases in China.

Acknowledgements

We would like to thank all the participants in this study.

Authors' contributions

YLZ and RNW conceptualized the paper. RNW did the statistical analysis and drafted the manuscript. YCZ, YTH and BTY conducted the research and data collection. BL and YLZ reviewed and edited the writing. All authors made

significant intellectual contributions to multiple revisions of the draft. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by Guangdong Province Philosophy and Social Science Planning 2020 Project, grant number GD20CGL23.

Availability of data and materials

The original data can be found on the National Health Commission of the People's Republic of China (<http://www.nhc.gov.cn>) and the standard map service website of the National Bureau of Surveying, Mapping and Geo-graphic Information (<http://bzdt.ch.mnr.gov.cn>).

Declarations

Ethics approval and consent to participate

The ethical approval is not warranted for our work as geographic information data and data on the annual incidence rate of Class B notifiable infectious diseases in China are publicly available. A link to the raw data can be found in the method section. All methods in this study were carried out in accordance with relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interest.

Author details

¹School of Health Management, Southern Medical University, Guangzhou 510515, China. ²Business School, University of Aberdeen, Aberdeen, UK.

Received: 7 February 2022 Accepted: 23 May 2022

Published online: 17 June 2022

References

- Jones BA, Betson M, Pfeiffer DU. Eco-social processes influencing infectious disease emergence and spread. *Parasitology*. 2017;144(1):26–36. <https://doi.org/10.1017/S0031182016001414>.
- Rodó X, Pascual M, Doblas-Reyes FJ, Gershunov A, Stone DA, Giorgi F, et al. Climate change and infectious diseases: can we meet the needs for better prediction? *Clim Chang*. 2013;118(3):625–40. <https://doi.org/10.1007/s10584-013-0744-1>.
- Bogoch II, Brady OJ, Kraemer MU, German M, Creatore MI, Kulkarni MA, et al. Anticipating the international spread of Zika virus from Brazil. *Lancet*. 2016;387(10016):335–6. [https://doi.org/10.1016/S0140-6736\(16\)00080-5](https://doi.org/10.1016/S0140-6736(16)00080-5).
- Cunningham AC, Goh HP, Koh D. Treatment of COVID-19: old tricks for new challenges. 2020;24:91–2. <https://doi.org/10.1186/s13054-020-2818-6>.
- Nuzzo JB, Inglesby TV. Ramping up the response to Ebola. *N Engl J Med*. 2018;379(26):2490–1. <https://doi.org/10.1056/nejmp1814296>.
- Prevention TCCfDCa. National epidemiological profile of statutory infectious diseases in may 2021 22 June 2021 [7 July 2021]. Available from: <http://www.nhc.gov.cn/jkj/s3578/202106/60b4cadb563b4f01ab5f095ba02d54bd.shtml>.
- Castillo-Salgado C. Use of geographic information systems in epidemiology (GIS-epi). *Epidemiol Bull*. 1996;17(1):1–6.
- Mollalo A, Mao L, Rashidi P, Glass GE. A GIS-based artificial neural network model for spatial distribution of tuberculosis across the continental United States. *Int J Environ Res Public Health*. 2019;16(1):157–74. <https://doi.org/10.3390/ijerph16010157>.
- Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci Total Environ*. 2020;728(2):138884–92. <https://doi.org/10.1016/j.scitotenv.2020.138884>.
- Zambrano LI, Rodriguez E, Espinoza-Salvado IA, Fuentes-Barahona IC, de Oliveira TL, da Veiga GL, et al. Spatial distribution of dengue in Honduras during 2016–2019 using a geographic information systems

- (GIS)-dengue epidemic implications for public health and travel medicine. *Travel Med Infect Dis.* 2019;32(2):101517–27. <https://doi.org/10.1016/j.tmaid.2019.101517>.
11. Guo C, Du Y, Shen S, Lao X, Qian J, Ou C. Spatiotemporal analysis of tuberculosis incidence and its associated factors in mainland China. *Epidemiol Infect.* 2017;145(12):2510–9. <https://doi.org/10.1017/S0950268817001133>.
 12. F-s Y, Liu L, L-h L, Y-l Z, L-l Z, He F, et al. Epidemiological and spatiotemporal analyses of HIV/AIDS prevalence among older adults in Sichuan, China between 2008 and 2019: a population-based study. *Int J Infect Dis.* 2021;105(2021):769–75. <https://doi.org/10.1016/j.ijid.2021.02.077>.
 13. Li W-T, Feng R-H, Li T, Du Y-B, Zhou N, Hong X-Q, et al. Spatial-temporal analysis and visualization of scarlet fever in mainland China from 2004 to 2017. *Geospat Health.* 2020;15(1):831–45. <https://doi.org/10.4081/gh.2020.831>.
 14. Chen H, Chen Y, Sun B, Wen L, An X. Epidemiological study of scarlet fever in Shenyang, China. *BMC Infect Dis.* 2019;19(1):1–7. <https://doi.org/10.1186/s12879-019-4705-9>.
 15. Viboud C, Sun K, Gaffey R, Ajelli M, Fumanelli L, Merler S, et al. The RAPIDD ebola forecasting challenge: synthesis and lessons learnt. *Epidemics.* 2018;22(3):13–21. <https://doi.org/10.1016/j.epidem.2017.08.002>.
 16. Chowell G, Viboud C, Simonsen L, Merler S, Vespignani A. Perspectives on model forecasts of the 2014–2015 Ebola epidemic in West Africa: lessons and the way forward. *BMC Med.* 2017;15(1):1–8. <https://doi.org/10.1186/s12916-017-0811-y>.
 17. Chretien J-P, Riley S, George DB. Mathematical modeling of the West Africa Ebola epidemic. *Elife.* 2015;4(3):e09186. <https://doi.org/10.7554/eLife.09186>.
 18. Akhmetzhanov AR, Lee H, Jung S-m, Kayano T, Yuan B, Nishiura H. Analyzing and forecasting the Ebola incidence in north Kivu, the Democratic Republic of the Congo from 2018–19 in real time. *Epidemics.* 2019;27(2019):123–31. <https://doi.org/10.1016/j.epidem.2019.05.002>.
 19. Liu Q, Liu X, Jiang B, Yang W. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect Dis.* 2011;11(1):1–7. <https://doi.org/10.1186/1471-2334-11-218>.
 20. Zeng Q, Li D, Huang G, Xia J, Wang X, Zhang Y, et al. Time series analysis of temporal trends in the pertussis incidence in mainland China from 2005 to 2016. *Sci Rep.* 2016;6(1):1–8. <https://doi.org/10.1038/srep32367>.
 21. Awan TM, Aslam F. Prediction of daily COVID-19 cases in European countries using automatic ARIMA model. *Journal of public health. Research.* 2020;9(3):227–33. <https://doi.org/10.4081/jphr.2020.1765>.
 22. Roy S, Bhunia GS, Shit PK. Spatial prediction of COVID-19 epidemic using ARIMA techniques in India. *Model Earth Syst Environ.* 2021;7(2):1385–91. <https://doi.org/10.1007/s40808-020-00890-y>.
 23. Tandon H, Ranjan P, Chakraborty T, Suhag V. Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future. *arXiv preprint.* 2020;2004(07859):1–11. <https://doi.org/10.48550/arXiv.2004.07859>.
 24. Wang Y, Wang S, Pan J, Wang W. Trend analysis and prediction of viral hepatitis incidence in China, 2009–2018. *Zhonghua liu Xing Bing xue za zhi= Zhonghua Liuxingbingxue Zazhi.* 2020;41(9):1460–4. <https://doi.org/10.3760/cma.j.cn112338-20191024-00761>.
 25. Anokye R, Acheampong E, Owusu I, Isaac OE. Time series analysis of malaria in Kumasi: using ARIMA models to forecast future incidence. *Cogent Soc Sci.* 2018;4(23):1461544–57. <https://doi.org/10.1080/2331886.2018.1461544>.
 26. Anwar MY, Lewnard JA, Parikh S, Pitzer VE. Time series analysis of malaria in Afghanistan: using ARIMA models to predict future trends in incidence. *Malar J.* 2016;15(1):1–10. <https://doi.org/10.1186/s12936-016-1602-1>.
 27. Talirongan H, Orong MY, Talirongan FJB. Alleviating vulnerabilities of the possible outbreaks of measles: a data trend analysis and prediction of possible cases. *Mediterranean J Basic Appl Sci (MJBAS).* 2021;4(4):129–35. <https://doi.org/10.48550/arXiv.2101.01387>.
 28. Getis A. Spatial autocorrelation. *Handbook Appl Spatial Anal.* 2010;11(3):255–78.
 29. Box GE, Pierce DA. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *J Am Stat Assoc.* 1970;65(332):1509–26. <https://doi.org/10.1080/01621459.1970.10481180>.
 30. Sato RC. Disease management with ARIMA model in time series. *Einstein (Sao Paulo).* 2013;11(1):128–31. <https://doi.org/10.1590/S1679-4508013000100024>.
 31. Yao S, Zhang Z. On regional inequality and diverging clubs: a case study of contemporary China. *J Comp Econ.* 2001;29(3):466–84. <https://doi.org/10.1006/jceec.2001.1726>.
 32. Nii-Trebi NI. Emerging and neglected infectious diseases: insights, advances, and challenges. *Biomed Res Int.* 2017;2017(2):1–15. <https://doi.org/10.1155/2017/5245021>.
 33. Wu X, Lu Y, Zhou S, Chen L, Xu B. Impact of climate change on human infectious diseases: empirical evidence and human adaptation. *Environ Int.* 2016;86(2):14–23. <https://doi.org/10.1016/j.envint.2015.09.007>.
 34. Dong Y, Wang L, Burgner DP, Miller JE, Song Y, Ren X, et al. Infectious diseases in children and adolescents in China: analysis of national surveillance data from 2008 to 2017. *BMJ.* 2020;369(2):m1043–58. <https://doi.org/10.1136/bmj.m1043>.
 35. Wang L, Wang Y, Jin S, Wu Z, Chin DP, Koplan JP, et al. Emergence and control of infectious diseases in China. *Lancet.* 2008;372(9649):1598–605. [https://doi.org/10.1016/S0140-6736\(08\)61365-3](https://doi.org/10.1016/S0140-6736(08)61365-3).
 36. Wang T, Wang X, Tie P, Bai Y, Zheng Y, Yan C, et al. Spatio-temporal cluster and distribution of human brucellosis in Shanxi Province of China between 2011 and 2016. *Sci Rep.* 2018;8(1):1–10. <https://doi.org/10.1038/s41598-018-34975-7>.
 37. Mao Y, Zhang N, Zhu B, Liu J, He R. A descriptive analysis of the Spatio-temporal distribution of intestinal infectious diseases in China. *BMC Infect Dis.* 2019;19(1):1–14. <https://doi.org/10.1186/s12879-019-4400-x>.
 38. Wang Q, Dong W, Yang K, Ren Z, Huang D, Zhang P, et al. Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors. *Int J Infect Dis.* 2021;105(4):675–85. <https://doi.org/10.1016/j.ijid.2021.03.014>.
 39. Zhu G, Xiao J, Liu T, Zhang B, Hao Y, Ma W. Spatiotemporal analysis of the dengue outbreak in Guangdong Province. *China BMC infectious diseases.* 2019;19(1):1–11. <https://doi.org/10.1186/s12879-019-4015-2>.
 40. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature.* 2008;451(7181):990–3. <https://doi.org/10.1038/nature06536>.
 41. Kilpatrick AM, Randolph SE. Drivers, Dynamics, and control of emerging vector-borne zoonotic diseases. *Lancet.* 2012;380(9857):1946–55. [https://doi.org/10.1016/S0140-6736\(12\)61151-9](https://doi.org/10.1016/S0140-6736(12)61151-9).
 42. Lashley FR. Emerging infectious diseases: vulnerabilities, contributing factors and approaches. *Expert Rev Anti Infect Ther.* 2004;2(2):299–316. <https://doi.org/10.1586/14787210.2.2.299>.
 43. Yang S, Wu J, Ding C, Cui Y, Zhou Y, Li Y, et al. Epidemiological features of and changes in incidence of infectious diseases in China in the first decade after the SARS outbreak: an observational trend study. *Lancet Infect Dis.* 2017;17(7):716–25. [https://doi.org/10.1016/S1473-3099\(17\)30227-X](https://doi.org/10.1016/S1473-3099(17)30227-X).
 44. De Wu JW, Zhang Q, Zhong H, Ke C, Deng X, Guan D, et al. Chikungunya outbreak in Guangdong province, China, 2010. *Emerg Infect Dis.* 2012;18(3):493–5. <https://doi.org/10.3201/eid1803.110034>.
 45. Guan J, He Z, Qin M, Deng X, Chen J, Duan S, et al. Molecular characterization of the viral structural protein genes in the first outbreak of dengue virus type 2 in Hunan Province, inland China in 2018. *BMC Infect Dis.* 2021;21(1):1–11. <https://doi.org/10.1186/s12879-021-05823-3>.
 46. Li Y, Zhou G, Zhong S, Wang X, Zhong D, Hemming-Schroeder E, et al. Spatial heterogeneity and temporal dynamics of mosquito population density and community structure in Hainan Island, China. *Parasit Vectors.* 2020;13(1):1–11. <https://doi.org/10.1186/s13071-020-04326-5>.
 47. Li Z, Yin W, Clements A, Williams G, Lai S, Zhou H, et al. Spatiotemporal analysis of indigenous and imported dengue fever cases in Guangdong province. *China BMC Infectious Diseases.* 2012;12(1):1–9. <https://doi.org/10.1186/1471-2334-12-132>.
 48. Liu L, Wu T, Liu B, Nelly RMJ, Fu Y, Kang X, et al. The origin and molecular epidemiology of dengue fever in Hainan Province, China, 2019. *Front Microbiol.* 2021;12(3):623–31. <https://doi.org/10.3389/fmicb.2021.657966>.
 49. Xiao J-P, He J-F, Deng A-P, Lin H-L, Song T, Peng Z-Q, et al. Characterizing a large outbreak of dengue fever in Guangdong Province, China. *Infect Dis Poverty.* 2016;5(1):1–8. <https://doi.org/10.1186/s40249-016-0131-z>.
 50. Zeng P, Yang K, Chen S. Using Big Data to Monitor the Impact of the COVID-19 Epidemic on Notifiable Diseases Reported in China. 2020 6th International Conference on Big Data and Information Analytics (BigDIA). 2020;10(2):138–41. <https://doi.org/10.1109/BigDIA51454.2020.00030>.

51. Li H, Ling F, Zhang S, Liu Y, Wang C, Lin H, et al. Comparison of 19 major infectious diseases during COVID-19 epidemic and previous years in Zhejiang, implications for prevention measures. *BMC Infect Dis.* 2022;22(1):1–14. <https://doi.org/10.1186/s12879-022-07301-w>.
52. Bai B-k, Jiang Q-y, Hou J. The COVID-19 epidemic and other notifiable infectious diseases in China. *Microbes Infect.* 2022;24(1):104881. <https://doi.org/10.1016/j.micinf.2021.104881>.
53. Geng Y, Zhang L. Impact of non-pharmaceutical interventions during COVID-19 pandemic on pertussis, scarlet fever and hand-foot-mouth disease in China. *J Infect.* 2022;84(2):e13–e5. <https://doi.org/10.1016/j.jinf.2021.12.023>.
54. Lutz CS, Huynh MP, Schroeder M, Anyatonwu S, Dahlgren FS, Danyluk G, et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. *BMC Public Health.* 2019;19(1):1–12. <https://doi.org/10.1186/s12889-019-7966-8>.
55. Desai AN, Kraemer MU, Bhatia S, Cori A, Nouvellet P, Herringer M, et al. Real-time epidemic forecasting: challenges and opportunities. *Health Security.* 2019;17(4):268–75. <https://doi.org/10.1089/hs.2019.0022>.

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

