



# Time series analysis of foodborne diseases during 2012–2018 in Shenzhen, China

Siguo Li<sup>1</sup> · Zhao Peng<sup>2</sup> · Yan Zhou<sup>1</sup> · Jinzhou Zhang<sup>1</sup>

Received: 30 April 2021 / Revised: 14 September 2021 / Accepted: 23 September 2021 / Published online: 22 October 2021  
© The Author(s) 2021

## Abstract

The present study aimed to use the autoregressive integrated moving average (ARIMA) model to forecast foodborne disease incidence in Shenzhen city and help guide efforts to prevent foodborne disease. The data of foodborne diseases in Shenzhen comes from the infectious diarrhea surveillance network, community foodborne disease surveillance network, and student foodborne disease surveillance network. The incidence data from January 2012 to December 2017 was used for the model-constructing, while the data from January 2018 to December 2018 was used for the model-validating. The mean absolute percentage error (MAPE) was used to assess the performance of the model. The monthly foodborne disease incidence from January 2012 to December 2017 in Shenzhen was between 954 and 32,863 with an incidence rate between 4.77 and 164.32/100,000 inhabitants. The ARIMA (1,1,0) was an adequate model for the change in monthly foodborne disease incidence series, yielding a MAPE of 5.34%. The mathematical formula of the ARIMA (1,1,0) model was  $(1 - B) \times \log(\text{incidence}_t) = 0.04338 + \varepsilon_t / (1 + 0.51106B)$ . The predicted foodborne disease incidences in the next three years were 635,751, 1,069,993, 1,800,838, respectively. Monthly foodborne disease incidence in Shenzhen were shown to follow the ARIMA (1,1,0) model. This model can be considered adequate for predicting future foodborne disease incidence in Shenzhen and can aid in the decision-making processes.

**Keywords** Foodborne disease · ARIMA · Time series analysis · Forecasting foodborne diseases · Foodborne disease incidence

## 1 Introduction

Foodborne disease is one of the most important public health issues in both developed and developing countries (Saulat 2012). Generally, foodborne disease results from the consumption of food contaminated with pathogens such as bacteria, viruses, parasites or with poisonous chemicals or bio-toxins (Bintsis 2017). Although the disease is usually mild and self-limiting, due to the high number of individuals affected each year, foodborne disease exerts a substantial socioeconomic burden on the population and healthcare

system (Wu et al. 2018). The World Health Organization (WHO) estimated that 31 foodborne hazards caused 600 million foodborne disease cases and 420,000 deaths worldwide in 2010 (WHO 2017). China faces various and unprecedented challenges in all aspects of the food chain (Liu et al. 2018), and 94.117 m cases of bacterial foodborne diseases occur every year, of which 3.357 m are hospitalized and 8530 die with a mortality rate of 9.1/100,000 (Mao et al. 2011). Furthermore, a minority of patients with foodborne diseases seek formal medical care, and informative tests are reported for only a fraction (Alcorn and Ouyang 2012). The burden of foodborne disease in China may be even greater.

China is currently developing and implementing a foodborne disease surveillance system across the country. However, the foodborne disease surveillance system remains in the early stage of development in a stepwise fashion, and the system has various limitations (Liu et al. 2018). In Shenzhen, a city of Guangdong Province, a citywide surveillance network on infectious diarrhea patients was established in 2010. Samples were collected at intervals from both

✉ Siguo Li  
270452220@qq.com

<sup>1</sup> Shenzhen Center for Disease Control and Prevention, Longyuan Road 8, Shenzhen 518055, China

<sup>2</sup> Department of Nutrition and Food Hygiene, Hubei Key Laboratory of Food Nutrition and Safety, Tongji Medical College, Huazhong University of Science and Technology, Hangkong Road 13, Wuhan 430030, China

children and adults among diarrhea outpatients in hospitals and tested for viruses and bacteria in Shenzhen Center for Disease Control and Prevention (CDC). Acute gastroenteritis and diarrhea are clinical outcomes of foodborne diseases, with data showing that up to 70% of acute gastroenteritis results from foodborne diseases (Lakhan et al. 2013). Additionally, in order to monitor patients with foodborne diseases more comprehensively, Shenzhen city has also established two other foodborne disease surveillance networks, namely community foodborne disease surveillance network (2012) and student foodborne disease surveillance network (2012), whose main monitoring objects are community population and primary and secondary school students. The three major systems provide a good guarantee for the surveillance of foodborne diseases, and a good data basis for the prediction of foodborne diseases in Shenzhen. The accurate prediction of a foodborne disease epidemic is crucial for public health authorities.

The ARIMA model, which is also known as the Box–Jenkins model, was proposed by George Box and Gwilym Jenkins in the early 1970s. An ARIMA model is a statistical model for analyzing and forecasting time series data, and the model can explain a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that the equation can be used to forecast future values (Altekruse and Swerdlow 1996). The ARIMA model has become one of the most popular and convenient models in time series analysis and been widely used as classical method for various infectious disease prediction, especially in infectious diarrhea (Fang et al. 2020; Li et al. 2010). In China, so far, there have been few studies using the ARIMA model to predict foodborne disease, and there are some differences in the models used by different researchers. Based on the infectious diarrhea (excluding cholera, dysentery, typhoid and paratyphoid) morbidity data and meteorological data from 2012 to 2016 in Jiangsu province, a univariate ARIMA model  $(1,0,1) (1,0,0)_{52}$  ( $AIC = -575.92$ ,  $BIC = -558.14$ ) and a multivariable ARIMAX model  $(1,0,1) (1,0,0)_{52}$  with 0–1 week lag precipitation ( $AIC = -578.58$ ,  $BIC = -578.13$ ) were developed to predict the incidence of infectious diarrhea (Fang et al. 2020). It also suggested that the introduction of meteorological factors did not significantly optimize the prediction accuracy of the ARIMA model (Fang et al. 2020). This finding was consistent with other previous studies. Based on the monthly incidence rate of dysentery in Shanghai from 1990 to 2007, the model  $ARIMA(1,1,1) (0,1,2)_{12}$ , namely  $(1 - 0.443B)(1 - B)(1 - B^{12})Z_t = (1 - 0.806B)(1 - 0.543B^{12})(1 - 0.321B^{2 \times 12})\mu_t$ , had a good correlation with the changes of incidence rate of dysentery and could forecast the future incidence rate in Shanghai (Li et al. 2010). Another ARIMA  $(1,1,1) \times (1,1,2)_{12}$  model fitted exactly with the number of bacillary dysentery cases during January 2004 to December

2014 in Jiangsu and could predict bacillary dysentery incidence during the period January to August 2015 (Wang et al. 2016). However, one study conducted in Lanzhou of China showed that the incidence of other types of infectious diarrhea (other than cholera, dysentery, typhoid, and paratyphoid) had typical seasonal changes, and easily occurred in weather conditions with a higher temperature, humidity, and lower pressure. The number of people with other types of infectious diarrhea increased by 66.71%, 5.24%, 7.1% and 6.93% per day with an increase of the inter-quartile range determined by average temperature, wind speed, relative humidity, and daily sunshine hours respectively (Tao et al. 2015). Another study also concluded that the etiological and meteorological factors had age-specific effects on the prevalence of infectious diarrhea in Jiangsu (Fang et al. 2019). However, even if the ARIMA models are constructed based on foodborne diseases, there were still some differences in model parameters in different regions (Fang et al. 2020; Li et al. 2010; Tao et al. 2015; Wang et al. 2016).

In order to develop an early warning system for foodborne diseases to facilitate preventive strategies in Shenzhen, foodborne disease incidence data from 2012 to 2017 in Shenzhen was used to construct the prediction model, and the data from 2018 were used for validating model. The present study aimed to use ARIMA model to forecast foodborne disease incidence in Shenzhen city and help guide efforts to prevent foodborne disease.

## 2 Material and methods

### 2.1 Study area

Shenzhen is a major sub-provincial city on the east bank of the Pearl River estuary on the central coast of southern Guangdong province, People's Republic of China. Shenzhen's registered population was estimated to be 12,905,000 in 2017, but local police and authorities estimated the population to be about 20 m, due to large populations of short-term residents, unregistered floating migrants, part-time residents, commuters, visitors, as well as other temporary residents. With a total area of 1992 km<sup>2</sup>, Shenzhen has a population density of 6889 inhabitants per km<sup>2</sup>.

### 2.2 Data source

The data comes from the three surveillance systems from 2012 to 2018, namely infectious diarrhea surveillance network, community foodborne disease surveillance network, and student foodborne disease surveillance network. The surveillance system consisted of three levels: hospital, community and school for case finding, sampling and

information collection; district-level CDCs for sample testing; and the municipal CDC for management and quality control. The three levels could share information through a dedicated online system.

### 2.3 Case definition

All suspected cases in the surveillance system must be reviewed and confirmed by experts based on the criteria for foodborne diseases. Foodborne disease cases were defined according to the criteria from WHO: (1) patients with diarrhea as the main symptom, with three or more stools per day, and changes in stool characteristics; or with less than three stools per day, but with vomiting as the main symptom, and comprehensively judged as foodborne disease or suspected foodborne disease by asking the medical history; (2) patients diagnosed by a doctor as acute gastroenteritis; (3) patients without diarrhea as the main symptom, but with clinical manifestations and an epidemiological history that indicates foodborne disease or suspected foodborne disease.

### 2.4 Data check and process

All data of foodborne disease patients in three surveillance systems from 2012 to 2018 were reviewed by clinical doctors and downloaded from the online systems. The three databases were merged, and duplicate data was removed according to variables such as patient ID, gender, and diagnosis time. We used the incidence data from January 2012 to December 2017 as the model-constructing dataset and the incidence data from January 2018 to December 2018 as the model-validating dataset.

### 2.5 Statistical analysis

The ARIMA model is generally referred to as an ARIMA ( $p, d, q$ ), where  $p$  is the number of autoregressive terms,  $d$  is the number of nonseasonal differences needed for stationarity, and  $q$  is the number of lagged forecast errors in the prediction equation (Li et al. 2019). The patterns of the plot of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) were used to determine the order of autoregressive (AR) and moving average (MA) included in the ARIMA model (Grahn 1995; Juang et al. 2017). The fitting of the ARIMA model involves the following three essential steps (Fang et al. 2020): first, a time series graph and an augmented Dickey–Fuller test is conducted to detect whether the original time series is stationary (statistical properties such as the mean and variance are all constant over time). If not, a logarithmic transformation or/and difference is adopted to achieve stability. Second, ARIMA models are established for a stationary time series, and the model with the minimum Akaike information criterion (AIC) and Schwartz Bayesian

information criterion (SBC) values is considered the optimal model. AIC and SBC are model selection criteria based on the log-likelihood, and AIC is defined as  $2\log(L) + 2(p + q)$ , SBC is defined as  $2\log(L) + \log(n)(p + q)$ . Here,  $L$  is the likelihood evaluated at the maximum likelihood estimate,  $n$  is the length of the time series. The model parameters are then estimated using the conditional least squares method. Third, to verify the adequacy of the ARIMA model, a Box-Ljung test, a type of statistical test of whether any of a group of autocorrelations of a time series are different from zero, is conducted to check whether the residual series is a white noise sequence. A white noise sequence is a purely random time series without an autocorrelation, and useful information has been extracted from the sequence for model fitting. If not, the model must be reestablished. Finally, a prospective prediction is conducted using the optimal model which was applied to predict the infectious diarrhea incidence and compare this prediction with the validating dataset. The mean value of MAPE is used to assess the level of forecast accuracy.  $MAPE = \frac{1}{n} \sum_{t=1}^n |(x_t - \hat{x}_t)/x_t|$ , where  $x_t$  and  $\hat{x}_t$  represents observed and predicted value at time  $t$ . The smaller the MAPE value, the higher the level of forecast accuracy. All computations are done by using SAS software (SAS® 9.3, SAS Institute Inc., Cary, NC, USA).

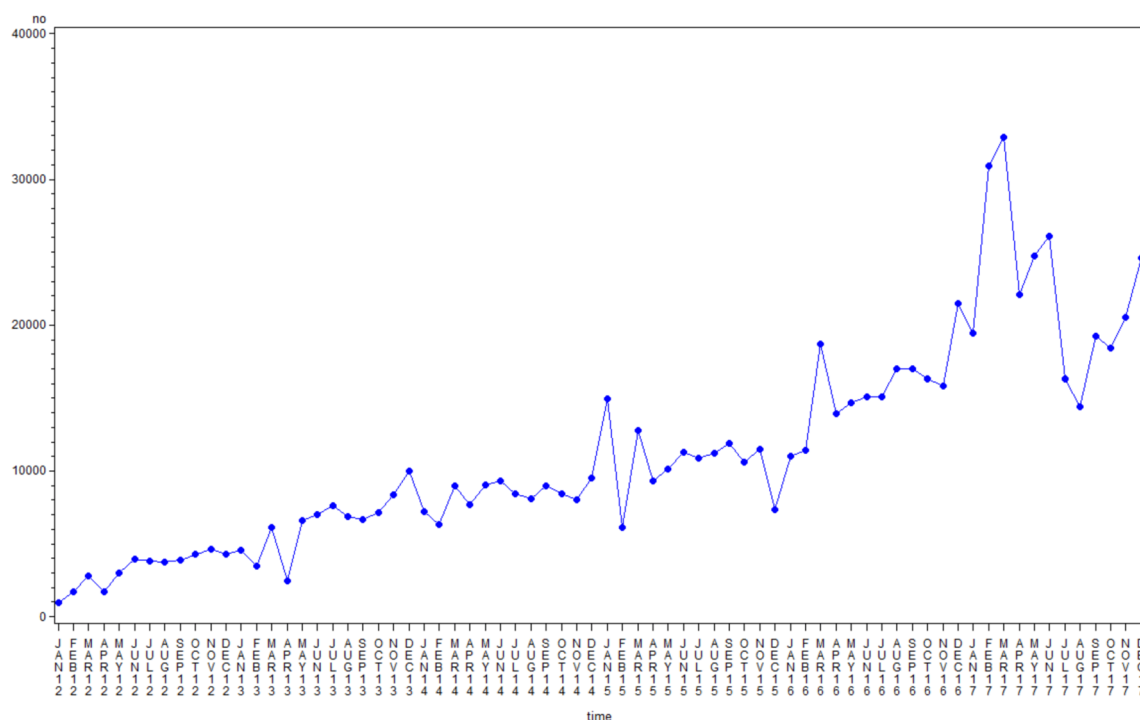
## 3 Results

### 3.1 Model identification

The monthly foodborne disease incidence from January 2012 to December 2017 in Shenzhen was between 954 and 32,863, with an incidence rate between 4.77 and 164.32/100,000 inhabitants. As shown in Fig. 1, the incidence exhibited a clear increasing long-term trend during these 6 years. In terms of the temporal distribution of the foodborne disease cases, Fig. 1 did not show that the monthly incidence exhibited obvious seasonal variation. After logarithmic conversion and first order difference, the temporal distribution was stationary and without seasonality (Fig. 2A, Table 1). So, a non-seasonal ARIMA model was used for model-constructing.

### 3.2 ARIMA model-constructing

The ACF plot in Fig. 2B shows a significantly positive spike at lag 0, revealing a non-seasonal parameter MA of one order ( $q = 0$ ). The PACF plot shown in Fig. 2C reveals a significantly positive spike at lag 1, suggesting a non-seasonal parameter AR of one order ( $p = 1$ ). In addition, we also conducted ARIMA (0,1,1) and ARIMA (1,1,1) model. Among the three models, ARIMA (1,1,0) model has the smallest value of AIC (19.99) and SBC (24.51).



**Fig. 1** Monthly reported foodborne disease incidence from January 2012 to December 2017 in Shenzhen, China

The next step was to verify the adequacy of this model. Figure 3 shows the residual correlation and white noise test plots, which indicated that the residuals were uncorrelated, and the normality plots showed no departure from the normal distribution. The model passed autocorrelation check of residuals ( $p > 0.05$ ). So, the ARIMA (1,1,0) was adequate for the monthly change in foodborne disease incidence series.

The parameter estimation was conducted next, and the results were reported in Table 2. There were two parameters in the model. The mean term was labelled as MU (the constant term, namely the average difference in the time series); its estimated value was 0.04338, and the  $p$  value for MU indicated that this term marginal significantly contributed to the model ( $p = 0.0491$ ). The AR parameter was labelled as AR1,1; this was the coefficient of the lagged value of the change in foodborne disease incidence, and its estimate was  $-0.51106$  ( $p < 0.0001$ ). So, the mathematical formula of the fitted ARIMA (1,1,0) model was given by the following:  $(1 - B) \times \log(\text{incidence}_t) = 0.04338 + \varepsilon_t / (1 + 0.51106B)$ , where  $B$  is the back operator.

Based on this equation, the values and plot of the forecast from January 2012 to December 2017 were produced. As shown in Fig. 4, all the actual monthly foodborne disease incidences almost fell in the 95%-confidence interval of the forecasted value, which further indicated the accuracy of the constructed ARIMA (1,1,0) model.

### 3.3 Prediction performance comparison

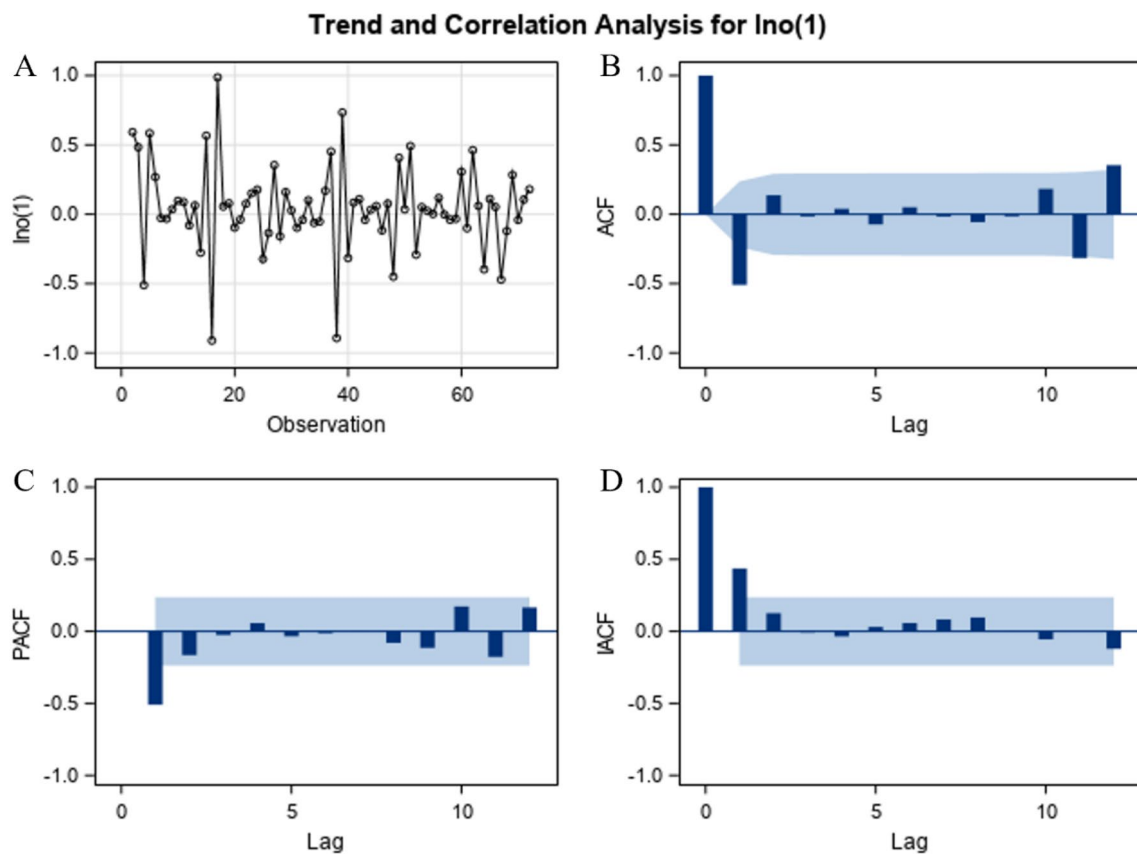
As shown in Table 3, the monthly foodborne disease data in 2018 were predicted using the ARIMA (1,1,0) model, the results of which suggested that the predicted values fitted well with the actual values. Notably, the actual values in 2018 fell in the 95% confidence intervals of the ARIMA (1,1,0) model, and the MAPE was 5.34%.

### 3.4 Forecasting

The ARIMA (1,1,0) model was then used to forecast the foodborne disease incidence in the next 5 years. The predicted incidences were shown in Fig. 4 with a significant upward trend in the next 5 years. The predicted incidences from 2019 to 2021 in Shenzhen were 635,751, 1,069,993 and 1,800,838, respectively.

## 4 Discussion

A variety of bacteria, viruses, parasites, and chemical hazards can be transmitted to humans via contaminated food and cause illness and death. The sources of these hazards and the routes of exposure are diverse, ranging from the environment and primary production at the beginning of the food chain through to domestic handling and food consumption (WHO



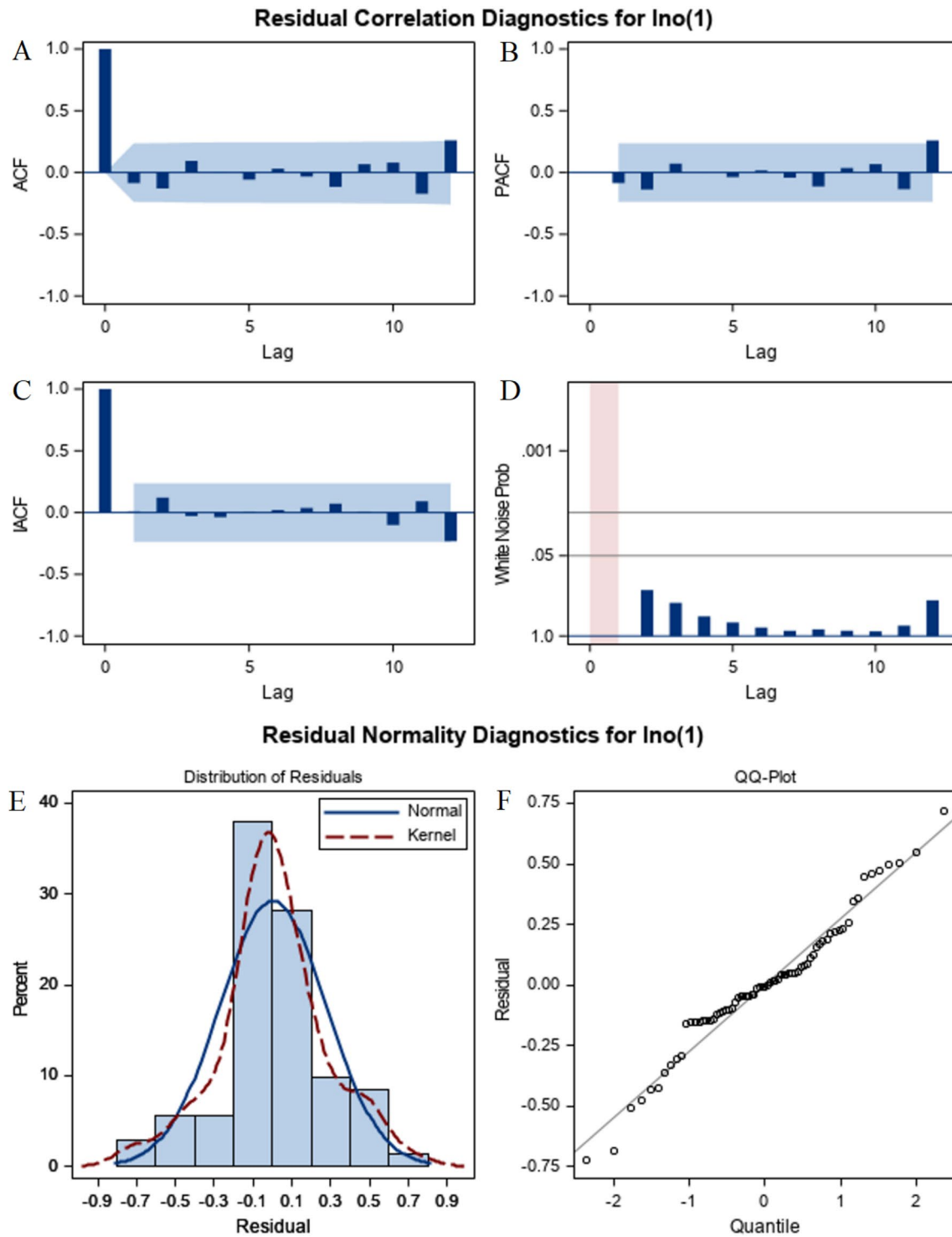
**Fig. 2** Autocorrelation function (ACF), partial autocorrelation function (PACF) and inverse autocorrelation function (IACF) for foodborne disease incidence time series based on log transformed data. The former **A** shows a significantly positive spike at lag 0, revealing a non-seasonal moving average component of one order. The latter

**B–D** shown reveals a significantly positive spike at lag 1, suggesting a non-seasonal autoregressive component of one order ( $p=1$ ). In addition, no seasonal lag in the foodborne disease incidence series data was found within the period examined

**Table 1** Augmented Dickey-Fuller unit root tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero mean	0	-103.961	0.0001	-14.54	<0.0001		
	1	-148.236	0.0001	-9.28	<0.0001		
	2	-122.914	0.0001	-5.70	<0.0001		
	3	-97.6595	<0.0001	-4.49	<0.0001		
Single mean	0	-105.773	0.0001	-14.91	0.0001	111.20	0.0010
	1	-166.931	0.0001	-9.74	0.0001	47.44	0.0010
	2	-180.840	0.0001	-6.21	0.0001	19.26	0.0010
	3	-191.988	0.0001	-4.96	0.0002	12.32	0.0010
Trend	0	-106.636	0.0001	-14.97	<0.0001	112.17	0.0010
	1	-174.817	0.0001	-9.75	<0.0001	47.90	0.0010
	2	-217.293	0.0001	-6.32	<0.0001	20.04	0.0010
	3	-282.860	0.0001	-5.01	0.0006	12.73	0.0010

There are three kinds of tests under the augmented Dickey-Fuller tests: Rho-test, Tau-test, and F-test. The Rho-test is the regression coefficient test, which is also called the normalized bias test. The Tau-test is the studentized test. The F-test is a joint test for unit root



**Fig.3** Graphical check of the residuals for the ARIMA (1,1,0) model. We accepted the null hypothesis that the residuals were uncorrelated. The normality plots also showed no departure from the normal distribution. **A** Autocorrelation of the time series, **B** partial autocorrelation of the time series, **C** inverse autocorrelation of the time series,

**D** autocorrelation check for white noise, **E** histogram of the residuals, **F** normal quantile plot of the residuals. *ACF* autocorrelation function, *IACF* interaural cross-correlation function, *PACF* partial autocorrelation function



**Table 2** Parameter estimates for the ARIMA (1,1,0) model using the algorithm of maximum likelihood

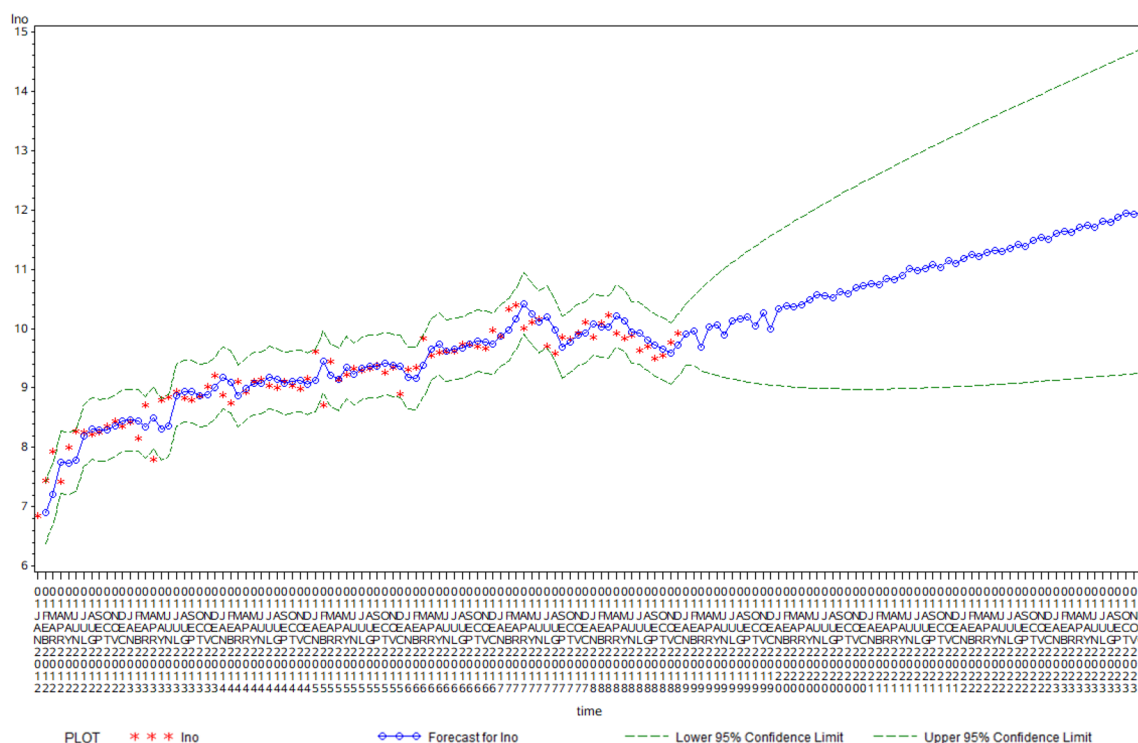
Parameter	Estimate	SE	t value	p value	Lag
MU	0.04338	0.02166	2.00	0.0491	0
AR1,1	−0.51106	0.10366	−4.93	<0.0001	1

ARIMA autoregressive integrated moving average, SE standard error, MU mean term of parameter in the model, AR autoregressive parameter

2017). Globalization, international trade and travel, and climate change also increase the spread of foodborne disease (Pozio 2020; WHO 2017). The results of this study show that the incidence of foodborne disease in Shenzhen exhibits a long-term gradual growth trend. Mathematical prediction models are urgently required to reinforce integrated management to monitor, control and prevent foodborne disease. The ARIMA (1,1,0) model was constructed and delivered a good accuracy in predicting the incidence of foodborne disease with a MAPE of approximately 5.34%. The model can relatively estimate the monthly foodborne disease incidence well in Shenzhen and can provide new evidence for policy-making and assist in the development of foodborne disease prevention and control strategies, such as funding and staffing for foodborne disease control. The ARIMA models can

also be used to evaluate the effectiveness of preventive and control measures.

As described earlier, the ARIMA models for predicting foodborne diseases constructed by different studies were different (Fang et al. 2020; Li et al. 2010; Tao et al. 2015; Wang et al. 2016), which may be related to the different prevalence patterns and trends of foodborne diseases in different regions, and the latter was related to influencing factors. For example, economic level, new foodborne pathogens, the types of food that people eat, the sources of those foods, and the possible decline in public awareness of safe food preparation practices were the factors altering foodborne disease patterns (Altekruse et al. 1996). Additionally, climatic factors (temperature, relative humidity, rainfall, insolation, and cloudiness) can affect the incidence of foodborne diseases, and the interrelationships or higher-order interrelationships among these climatic factors played an important role in the incidence of foodborne diseases (Park et al. 2018). Food-borne diseases still showed spatio-temporal high-risk clusters in some regions (Yang et al. 2019). All those factors would increase the difficulty of modeling, and it is not difficult to see that although the ARIMA model was used, all those factors would also lead to differences in ARIMA models constructed by different studies. Therefore, when constructing the model, we need to consider which factors



**Fig. 4** Fitting status situation between actual and predicted foodborne disease incidence from 2012 to 2018, and with the predicted incidences from 2019 to 2023 in Shenzhen, China. The red line indicates the actual incidence values, and the blue dotted line indicates

the predicted incidence, while the two green lines indicate the upper and lower scope of 95% confidence intervals of predicted incidence of foodborne disease (Color figure online)

**Table 3** Comparison between actual and predicted monthly infectious diarrhea in 2018

Month	Actual value		Predicted value		Error absolute value <sup>a</sup>	Absolute percentage error <sup>b</sup>
	Number	Logarithm	Logarithm	95% CI		
Jan	19,027	9.85	10.08	9.55–10.62	0.23	0.02
Feb	24,132	10.09	10.16	9.56–10.76	0.07	0.01
Mar	27,490	10.22	10.19	9.47–10.91	−0.03	0.00
Apr	20,535	9.93	10.24	9.45–11.04	0.31	0.03
May	18,786	9.84	10.28	9.40–11.16	0.44	0.04
Jun	19,757	9.89	10.33	9.38–11.27	0.44	0.04
July	15,175	9.63	10.37	9.36–11.38	0.74	0.08
Aug	16,442	9.71	10.41	9.34–11.48	0.7	0.07
Sep	13,312	9.50	10.45	9.33–11.58	0.95	0.10
Oct	14,064	9.55	10.50	9.31–11.68	0.95	0.10
Nov	17,613	9.78	10.54	9.31–11.78	0.76	0.08
Dec	20,362	9.92	10.58	9.30–11.87	0.66	0.07

*MAPE* mean of absolute percentage error, and the value is 5.34%, *95% CI* 95% confidence interval

<sup>a</sup>Value between actual value and predicted value

<sup>b</sup>Ratio of error absolute value and actual value

should be included in the ARIMA, such as meteorological factors, which depended on the specific circumstances of each region. So, it is very meaningful to build ARIMA model based on data from different regions.

In the present study, seasonal fluctuation of foodborne disease was not found in Shenzhen, while other studies in China showed a distinct seasonality, i.e., two incidence peaks were observed during each year: namely a higher winter peak from December to February and a lower summer peak from July to September in Jiangsu (Fang et al. 2020). Although Shenzhen is situated about a degree south of the Tropic of Cancer, due to the Siberian anticyclone it has a warm, monsoon-influenced, humid subtropical climate, which also leads to little change in temperature throughout the year, so there may not be obvious seasonality. Unlike Shenzhen, Jiangsu Province, located along the eastern coast of China, has a typical temperate subtropical monsoon climate with mild temperature, moderate rainfall and a distinct four-season pattern (Fang et al. 2020). Additionally, the performance of the ARIMAX model was comparable to that of the ARIMA model with a MAPE reaching approximately 30% (Fang et al. 2020). However, the MAPE in the present study using the ARIMA (1,1,0) model was 5.34%, which indicated this model has a good prediction performance. Furthermore, even for the same disease in China, the prediction of foodborne disease incidence in different regions should consider different ARIMA models.

To our best knowledge, only few studies have applied the ARIMA model for studying the annual variability of foodborne disease incidence based on the monthly data in China (Fang et al. 2020; Li et al. 2010; Wang et al. 2016), and the numbers of foodborne disease cases were all downloaded

from the National Notifiable Disease Surveillance System (NNDSS). However, some mild cases may use home therapies, and cases with atypical symptoms may be misdiagnosed, therefore, the reported data may underestimate the level of morbidity (Fang et al. 2020). In present study, the data from the three surveillance systems could provide all possible foodborne disease patients in the region, thereby improving the accuracy of the basic data. The present study suffered from a few limitations. First, we employed the univariate ARIMA model to forecast foodborne disease incidence. Other factor, i.e., meteorological factors, may influence the quantity of infectious diarrhea incidence (Tao et al. 2015). Second, the ARIMA model itself also has certain limitations. For instance, it proposes the linear or non-linear relation within time-series data by taking the time factor into consideration only, rather than pathogen, host, social economy, and natural environment factors (Lee et al. 2013).

In summary, monthly foodborne disease incidence in Shenzhen were shown to follow the ARIMA (1,1,0) model. This model can be considered adequate for predicting future foodborne disease incidence in Shenzhen and can be employed to aid in decision-making processes.

**Funding** Funding was provided by Shenzhen Science and Technology Innovation Commission (Grant no. JCYJ20180307102036312).

**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/>.

## References

- Alcorn T, Ouyang Y (2012) China's invisible burden of foodborne illness. *Lancet* (London, England) 379:789–790
- Altekruse SF, Swerdlow DL (1996) The changing epidemiology of foodborne diseases. *Am J Med Sci* 311:23–29
- Bintsis T (2017) Foodborne pathogens. *AIMS Microbiol* 3:529–563
- Fang X, Ai J, Liu W, Ji H, Zhang X, Peng Z, Wu Y, Shi Y, Shen W, Bao C (2019) Epidemiology of infectious diarrhoea and the relationship with etiological and meteorological factors in Jiangsu Province, China. *Sci Rep* 9:19571
- Fang X, Liu W, Ai J, He M, Wu Y, Shi Y, Shen W, Bao C (2020) Forecasting incidence of infectious diarrhea using random forest in Jiangsu Province, China. *BMC Infect Dis* 20:222
- Grahn T (1995) A conditional least squares approach to bilinear time series estimation. *J Time Ser Anal* 16:509–529
- Juang WC, Huang SJ, Huang FD, Cheng PW, Wann SR (2017) Application of time series analysis in modelling and forecasting emergency department visits in a medical centre in Southern Taiwan. *BMJ Open* 7:e018628
- Lakhan C, Badrie N, Ramsabhag A, Sundaraneedi K, Indar L (2013) Burden and impact of acute gastroenteritis and foodborne pathogens in Trinidad and Tobago. *J Health Popul Nutr* 31:30–42
- Lee HS, Her M, Levine M, Moore GE (2013) Time series analysis of human and bovine brucellosis in South Korea from 2005 to 2010. *Prev Vet Med* 110:190–197
- Li J, Wu HY, Li YT, Jin HM, Gu BK, Yuan ZA (2010) Establishing and applying of autoregressive integrated moving average model to predict the incidence rate of dysentery in Shanghai. *Zhonghua Yu Fang Yi Xue Za Zhi [chin J Prev Vet Med]* 44:48–53
- Li Z, Wang Z, Song H, Liu Q, He B, Shi P, Ji Y, Xu D, Wang J (2019) Application of a hybrid model in predicting the incidence of tuberculosis in a Chinese population. *Infect Drug Resist* 12:1011–1020
- Liu J, Bai L, Li W, Han H, Fu P, Ma X, Bi Z, Yang X, Zhang X, Zhen S et al (2018) Trends of foodborne diseases in China: lessons from laboratory-based surveillance since 2011. *Front Med* 12:48–57
- Mao X, Hu J, Liu X (2011) Epidemiological burden of bacterial foodborne diseases in China—preliminary study. *Chin J Food Hyg* 23:132–136
- Park MS, Park KH, Bahk GJ (2018) Interrelationships between Multiple Climatic Factors and Incidence of Foodborne Diseases. *Int J Environ Res Public Health* 15(11):2482
- Pozio E (2020) How globalization and climate change could affect foodborne parasites. *Exp Parasitol* 208:107807
- Saulat J (2012) Epidemiology of foodborne illness. Scientific, health and social aspects of the food industry. *INTECH* 336:790–793
- Tao Y, Liu ZM, Mi S, Song J, Qiang L (2015) Effects of meteorological factors on other types of infectious diarrhea. *J Lanzhou Univ Nat Sci* 51:646–651
- Wang K, Song W, Li J, Lu W, Yu J, Han X (2016) The use of an autoregressive integrated moving average model for prediction of the incidence of Dysentery in Jiangsu, China. *Asia Pac J Public Health* 28:336–346
- WHO (2017) WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group. World Health Organization, Geneva, pp 2007–2015
- Wu YN, Liu XM, Chen Q, Liu H, Dai Y, Zhou YJ, Wen J, Tang ZZ, Chen Y (2018) Surveillance for foodborne disease outbreaks in China, 2003 to 2008. *Food Control* 84:382–388
- Yang L, Sun YB, Zhong Q, Duan S, Liu SQ, Zhang Y (2019) Epidemiological characteristics and spatio-temporal patterns of foodborne diseases in Jinan, Northern China. *BES* 32:309–313

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.