



# Study on the relationship between the incidence of influenza and climate indicators and the prediction of influenza incidence

Yanling Zheng<sup>1</sup> · Kai Wang<sup>1</sup> · Liping Zhang<sup>1</sup> · Lei Wang<sup>1</sup>

Received: 21 May 2020 / Accepted: 13 August 2020 / Published online: 19 August 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

In recent 2 years, the incidence of influenza showed a slight upward trend in Guangxi; therefore, some joint actions should be done to help preventing and controlling this disease. The factors analysis of affecting influenza and early prediction of influenza incidence may help policy-making so as to take effective measures to prevent and control influenza. In this study, we used the cross correlation function (CCF) to analyze the effect of climate indicators on influenza incidence, ARIMA and ARIMAX (autoregressive integrated moving average model with exogenous input variables) model methods to do predictive analysis of influenza incidence. The results of CCF analysis showed that climate indicators (PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, CO, NO<sub>2</sub>, O<sub>3</sub>, average temperature, maximum temperature, minimum temperature, average relative humidity, and sunshine duration) had significant effects on the incidence of influenza. People need to take good precautions in the days of severe air pollution and keep warm in cold weather to prevent influenza. We found that the ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model has good predictive performance, which can be used to predict the influenza incidence in Guangxi, and the predicted incidence may be useful in developing early warning systems and providing important evidence for influenza control policy-making and public health intervention.

**Keywords** Influenza · Climate indicators · Prediction · ARIMA model · ARIMAX model

## Introduction

Influenza, an acute respiratory infectious disease caused by influenza virus, is the first global surveillance epidemic (Bekking et al. 2019). Globally, influenza surveillance is carried out in accordance with the definition of influenza-like cases recommended by the World Health Organization. Its main purpose is to master the characteristics and dynamics of influenza epidemics. The outbreak of influenza is often caught off guard, and it is especially severe in the winter and spring when the seasons change, influenza is often characterized by a wide range of epidemics and severe symptoms, vulnerable groups such as children and the elderly have poor resistance and are often the worst-hit areas for influenza outbreaks, and the prevalence of influenza has brought serious

harm to human health (Elhakim et al. 2019), so the preventive monitoring of influenza is of great significance (Priedhorsky et al. 2019).

Early prediction of infectious diseases can help early allocation of resources and take effective measures to prevent and control them (Chen et al. 2018). At present, there are many methods used in the prediction of infectious diseases, such as infectious disease dynamics model (Feng et al. 2013), neural network prediction method (Li et al. 2017), gray prediction method (Yang et al. 2018), Logistic Regression Model (Tuerlinckx et al. 2009), and ARIMA model method (Nikonahad et al. 2017; Gabriel et al. 2019; Wang et al. 2019; Yi-Yi et al. 2012; Wei et al. 2017; Rodrigues et al. 2018; Gharbi et al. 2011; Zhao et al. 2017; Tian et al. 2019). Different models have their own advantages and disadvantages. ARIMA model method can capture the periodicity, trend, and randomness of data, and it has high prediction accuracy and is widely used in the prediction of infectious diseases; for example, Mao et al. (Mao et al. 2018) forecasted well the incidence of tuberculosis in China by ARIMA method, and Liu et al. (Liu et al. 2016) predicted the incidence of hand, foot, and mouth disease accurately in Sichuan province, China, using the ARIMA model.

---

Responsible Editor: Lotfi Aleya

✉ Yanling Zheng  
zhengyl\_math@sina.cn

<sup>1</sup> College of Medical Engineering and Technology, Xinjiang Medical University, Urumqi 830011, People's Republic of China

Many studies have shown that climate indicators have a certain connection with disease outbreaks (Chong et al. 2019; Brattig et al. 2019; Cohen et al. 2017; Watts et al. 2018; Arikawa et al. 2019; Khan et al. 2019; Stewart-Ibarra et al. 2019); for this reason, incorporating climate indicators into the prediction model when making disease prediction might improve the prediction accuracy to some extent.

In recent 2 years, the incidence of influenza showed a slight upward trend in Guangxi, which increases the burden on the prevention and control of the disease. To analyze the factors of affecting influenza incidence and give the forecast and early warning may be able to provide reference for the prevention and control work.

This study first analyzed the effect of climate indicators on the incidence of influenza and further did the early warning prediction study based on ARIMA model method. In order to improve the prediction accuracy, climate indicator variables were incorporated into the prediction model; after statistical analysis, an ARIMAX model was used for prediction analysis of influenza incidence in Guangxi, China.

## Materials and methods

### Study area

Guangxi is located in the west of China's southern region, with an area of 236,700 km<sup>2</sup>; Nanning is the capital of the Guangxi. Guangxi plays an important role in the economic exchanges between China and Southeast Asia; it has 14 prefecture-level cities, of which the three in the heart of Guangxi are Nanning, Laibing, and Guigang. At the end of 2018, Guangxi had a population of about 49,260,000. The geographical location of Guangxi in China was shown in Fig. 1.

### Data sources

This study used the monthly data of air pollution variables from January 2015 to February 2020, and the six air pollution variables were PM<sub>2.5</sub>, PM<sub>10</sub>, CO, SO<sub>2</sub>, NO<sub>2</sub>, and O<sub>3</sub>; these data were from public web pages (<https://www.aqistudy.cn/historydata/>). The average values of six air pollution indicators in Nanning, Laibing, and Guigang of Guangxi were taken as the air pollution indexes data for Guangxi. The mean values of the monthly meteorological data (average temperature, maximum temperature, minimum temperature, average relative humidity, and sunshine duration) for Nanning and Laibing were approximated as the Guangxi meteorological data.

Monthly influenza cases from January 2015 to October 2019 were from website of Guangxi Health Commission. Influenza case monitoring data are based on the definition of influenza-like cases recommended by WHO and the National

Influenza Centre. In China, influenza is a nationally notifiable disease, departments set up by medical institutions at all levels concerned with the diagnosis, and treatment of influenza established out-patient logbooks, in-patient registers, and registers of infectious diseases, and they must report every case of influenza very seriously to the local health authority within 24 h. Local health authorities later report monthly influenza case totals to higher the national level CDC (Center for Disease Control and Prevention) for surveillance purposes. A total of 93,484 cases of influenza surveillance were reported in Guangxi from January 2015 to October 2019. Using population data and monthly influenza cases, we calculated the incidence of influenza (per 100,000 populations). Descriptive statistics of climate indicators variables and influenza incidence data were shown in Table 1.

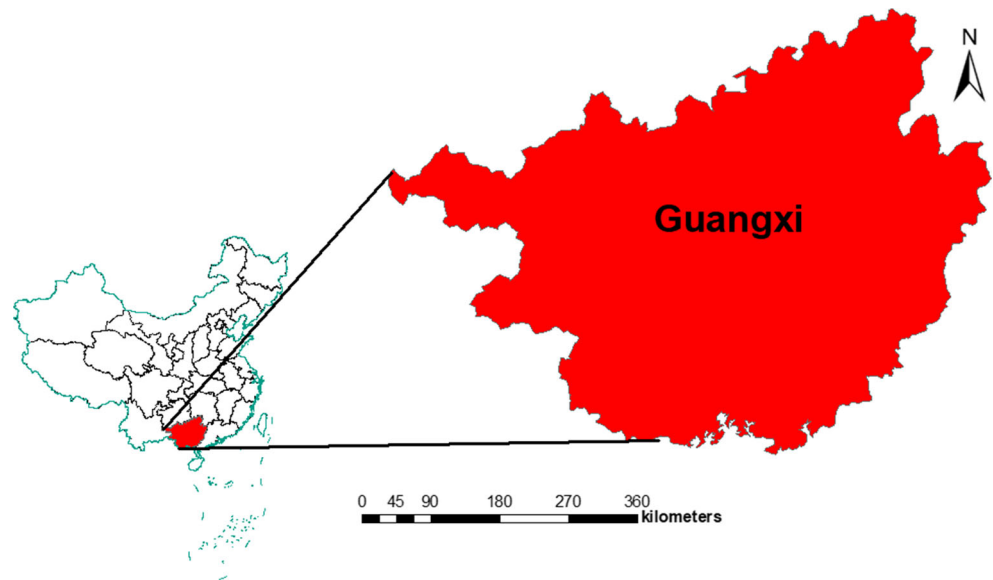
## Methods

First, the time-delay correlation between variables was analyzed by cross-correlation analysis function (CCF).

Second, to do the forecast analysis using auto regressive integrated moving average (ARIMA) model method, the model takes the form of the ARIMA( $p,d,q$ ), where  $d$  is the data difference times,  $p$  is the moving average order, and  $q$  is the sliding average order, when the data is periodic and the seasonal period is  $s$ ; the model is transformed into ARIMA( $p,d,q$ )( $P,D,Q$ ) $s$  form, where  $D$  is the seasonal difference number,  $P$  is the seasonal moving order, and  $Q$  is the seasonal sliding order. The method, proposed by American scholar Box and British statistician Jenkins in 1976 (Box and Jenkins 1976), is now widely used in the prediction and early warning analysis of infectious diseases. If exogenous variables are included in ARIMA model, it changes into ARIMAX model. There are three steps to modeling ARIMA model.

- Step 1: Model recognition. The stability and seasonality of the data are analyzed, and the difference method is used to stabilize the data and then to determine  $d$  and  $D$ . Augmented Dickey-Fuller (ADF) test is used to determine the stationary of the data (The null hypothesis is that the data is unstable), and then the alternative models of  $p$ ,  $q$ ,  $P$ , and  $Q$  are estimated according to autocorrelation function (ACF) and partial autocorrelation function graphs (PACF).
- Step 2: Parameter estimation and model test. The model that passes the parameter test is feasible, and then, to determine the best model according Akaike information criterion (AIC) and Schwarz-Bayes Criterion (SBC), the model with minimum value of AIC and SBC is the best model. The Box-Ljung statistic is used to do white noise test on the residual sequence;

**Fig. 1** Guangxi and its geographical location in China



if *p* value is more than 0.05, the model can pass the test and can be used for predictive analysis.

**Step 3:** Prediction applications. The selected models are used to fit the modeling data and to predict the values for a future period. The root mean square error (RSME) is used to measure the ability of fitting and forecasting. The smaller the RMSE, the better the model.

**Analysis software:** R3.6.2, Eviews7.2 and Arcmap10.2.

## Results

To visually analyze the incidence of influenza, climate indicators variables over the years in Guangxi, the monthly

**Table 1** The descriptive statistics of data variables

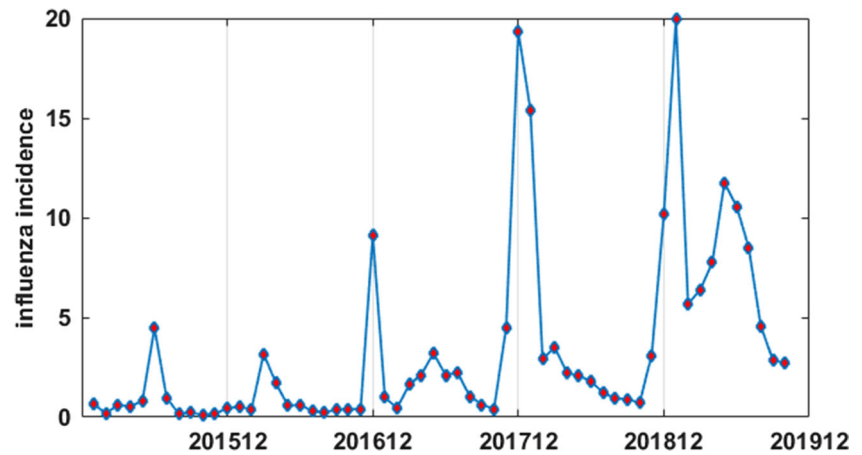
Variables	Range	Mean ± S.D.
Influenza	0.12–19.94	3.3 ± 4.6
PM2.5 (µg/m <sup>3</sup> )	18.3–79.0	37.8 ± 15.1
PM10 (µg/m <sup>3</sup> )	34.3–109.7	59.6 ± 18.5
SO <sub>2</sub> (ppb)	7.7–24.7	14.4 ± 4.3
CO (ppm)	0.6–1.3	0.9 ± 0.2
NO <sub>2</sub> (ppb)	16.0–40.3	24.9 ± 6.1
O <sub>3</sub> (ppm)	37.7–139.0	78.5 ± 20.0
Relative humidity(%RH)	66.1–87.23	79.03 ± 5.07
Average temperature(°C)	11.77–29.70	21.54 ± 5.91
Maximum temperature(°C)	14.55–34.15	26 ± 6.31
Minimum temperature(°C)	8.25–25.90	18.52 ± 5.71
Sunshine duration (h)	28.50–242.10	120.07 ± 55.86

incidence of influenza and PM2.5, PM10, CO, SO<sub>2</sub>, NO<sub>2</sub>, O<sub>3</sub>, average temperature, maximum temperature, minimum temperature, average relative humidity, and sunshine duration from January 2015 to October 2019 were plotted (see Figs. 2, 3 and 4). From the time series graph of influenza incidence, we can see that in recent years, the incidence of influenza was the highest around December; in general, the incidence of influenza showed a slight upward trend. From the change trend of the six air pollution variables in the past years, it could be seen that the air pollution was heavy almost every winter; in Guangxi, at this time the incidence of influenza was also high. The time series of meteorological variables appeared peaks around June every year and troughs in winter.

It is generally believed that climate indicators have latent effect on disease and there were lag effects. To this end, six index variables of air pollution and five meteorological variables and influenza incidence were analyzed with 0–12 order delay correlation. The calculated correlation coefficients and tests were shown in Table 2, there were the correlation coefficients between influenza incidence and eleven climate indicators variables, and there were the maximum delay correlation coefficient (Corr-Coeff) and the corresponding lag order. As can be seen from the data in Table 2, NO<sub>2</sub>, O<sub>3</sub>, average temperature, maximum temperature, minimum temperature, and sunshine duration were directly and significantly associated with influenza incidence; the CCF correlation analysis showed that the lag variables of the other four air pollution variables and average relative humidity were significantly correlated with influenza incidence.

Next, we attempted to build ARIMA model and ARIMAX model for predicting the incidence of influenza in Guangxi. Before modeling, we analyzed the stability of the modeling data to see if it needed differential processing. The data of influenza incidence and the eleven climate indicators variables

**Fig. 2** Time series of influenza incidence in Guangxi from January 2015 to October 2019 (Per 100,000 population)

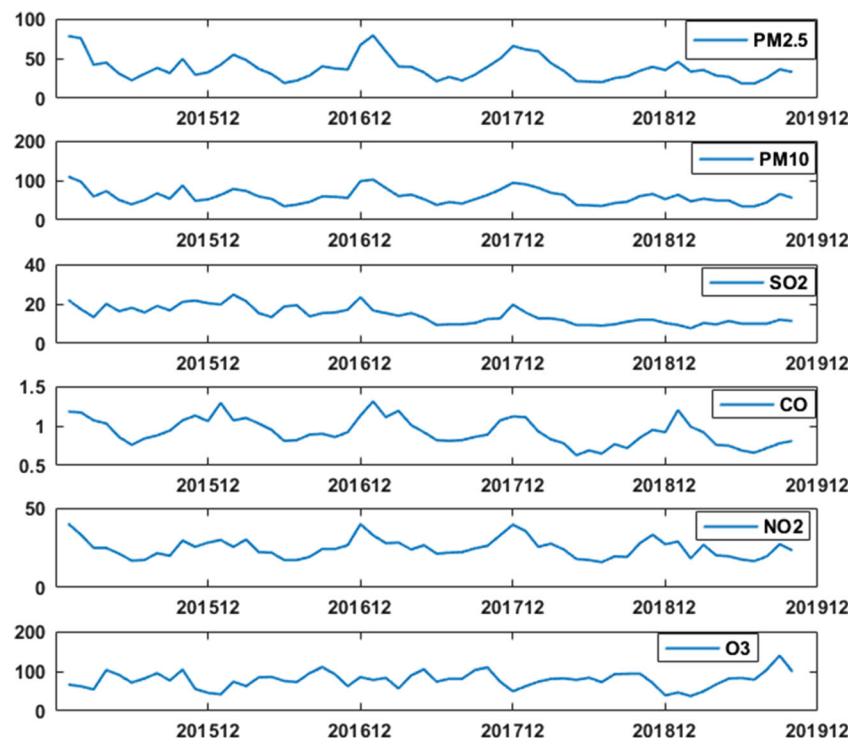


in Guangxi were tested by ADF test; the results showed that the twelve test  $p$  value values were less than 0.05, indicating that these data were stationary and did not need to be differentially processed. Therefore, both  $d$  and  $D$  in the ARIMA ( $p, d, q$ ) ( $P, D, Q$ )<sub>s</sub> model were 0.

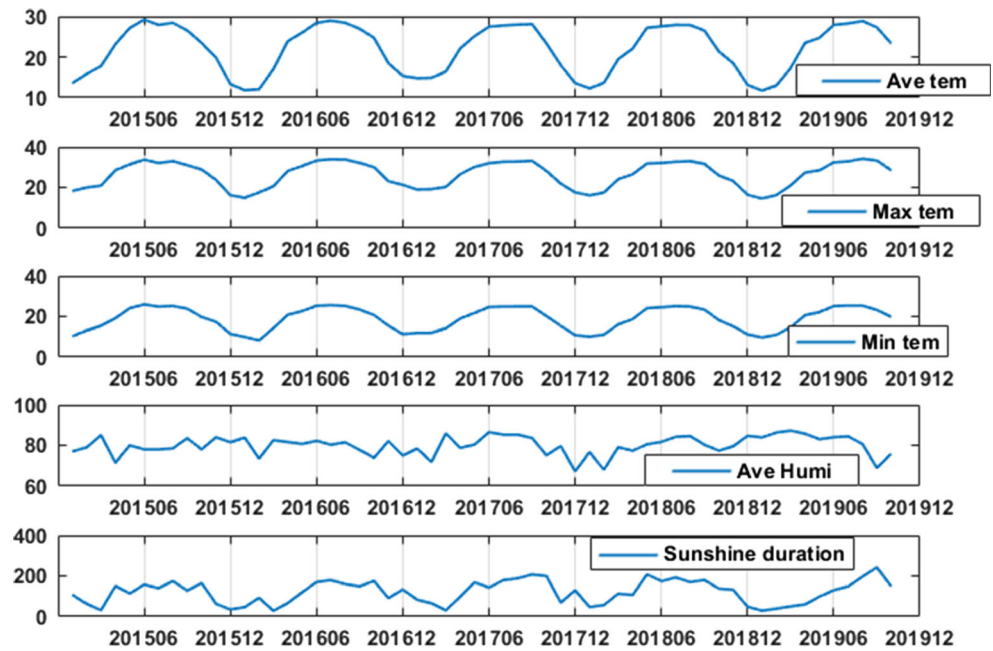
First, we used only influenza data from January 2015 to October 2019 to build the prediction model. In order to examine the periodicity of the data and determine the value of  $s$ , we decomposed the data into trend, season, and random items, as can be seen from Fig. 5; there was a significant seasonality in influenza incidence data with a seasonal cycle of 12, which indicated that  $s$  was 12, so the ARIMA ( $p, d, q$ ) ( $P, D, Q$ )<sub>s</sub> model became ARIMA ( $p, 0, q$ ) ( $P, 0, Q$ )<sub>12</sub>. Then, we divided the data into two parts: data from January 2015 to June 2019 for building model and data from July 2019 to October 2019

for testing the model's prediction performance. The ACF and PACF diagrams of the modeling data were drawn to determine the values of  $p, q, P,$  and  $Q$  in the model. See Fig. 6, autocorrelation coefficient first-order truncation, partial correlation number first order truncation, and the autocorrelation and partial correlation coefficients of the 12th order delay were large; therefore, it was preliminarily determined that  $p, q, P,$  and  $Q$  were 0 or 1. We tested various combinations of  $p, q, P,$  and  $Q$  with 0 or 1 and did the model hypothesis test analysis and finally determined that only the ARIMA (1,0,1)(0,0,1)<sub>12</sub> model parameters passed the test and the AIC and SBC of the model were the smallest; they were 5.14 and 5.25, respectively; the residual of this model was tested by Box-Ljung test, and almost all the  $p$  values of the residual lag were greater than 0.05, indicating that the residual

**Fig. 3** Time series of the six air pollution variables from January 2015 to October 2019



**Fig. 4** Time series of the meteorological variables (average temperature, maximum temperature, minimum temperature, average relative humidity, and sunshine duration) from January 2015 to October 2019



was white noise and this model passed the test (this indicated that the model has a good ability to do fit and prediction analysis). Using this model to do fitting analysis for the modeling data, the fitting RMSE was 2.98; using this model to predict influenza incidence from July 2019 to October 2019, the predicting RMSE was 7.61.

Second, considering that climate indicators variables might influence the incidence of influenza to some extent, they were considered for inclusion in the ARIMA (1,0,1)(0,0,1)<sub>12</sub> model, and the ARIMAX model analysis was done. Firstly, the maximum time lag correlation variables of each climate indicators variables were brought into ARIMA (1,0,1)(0,0,1)<sub>12</sub> model, respectively, to construct ARIMAX model. Secondly, due to the strong correlation between some climate indicators variables; in other words, these variables are not independent each other; in order to avoid the phenomenon of pseudo regression, we used the idea of stepwise regression to include these variables in the ARIMA (1,0,1)(0,0,1)<sub>12</sub>

model for ARIMAX model analysis; and after the model test analysis, we found that only four ARIMAX models passed all the parameters test; see Table 3 for details.

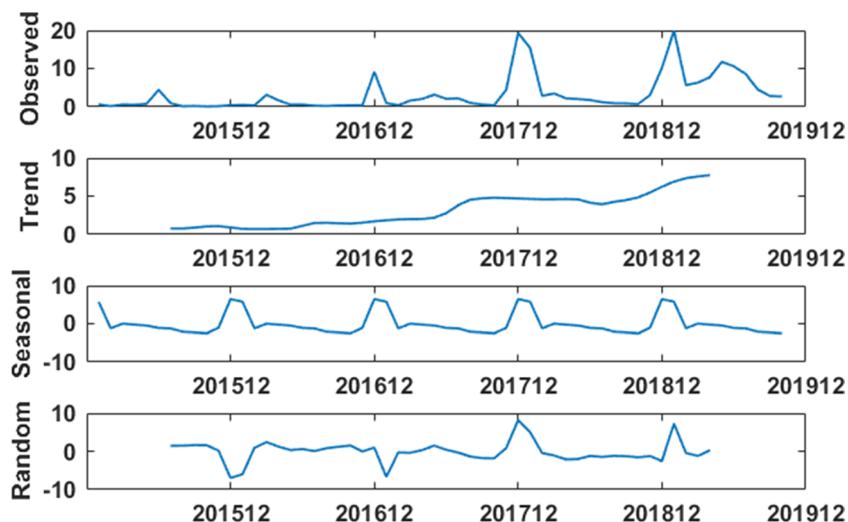
Comparing the model in Table 3, we found that the model of ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> had the smallest AIC and SBC values, indicating that the model had the highest fitting accuracy and the residual of the model was tested by Box-Ljung test and almost all the *p* values of the residual lag were greater than 0.05, which indicated that the model residual was white noise; therefore, the ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model had good prediction performance. The AIC and SBC of the ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model were less than ARIMA (1,0,1)(0,0,1)<sub>12</sub> model, which indicated that the ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model has higher prediction accuracy and it was more suitable for predicting influenza incidence in Guangxi. Using the model to fit the influenza incidences from January 2015 to June 2019, the fitting RMSE was 2.67; using the model to

**Table 2** The correlation coefficients and maximum lag correlation coefficients between influenza incidence and eleven variables

	PM2.5	PM10	SO2	CO	NO2	O3
Corr-Coef	0.14	0.13	-0.2	0.07	0.26*	-0.35**
Max lag Corr-Coef and	-0.36**	-0.4**	-0.49**	-0.38**	0.26*	-0.35**
Its lag order	6	6	6	7	0	0
	Ave-tem	Max-tem	Min-tem	Ave-Humi	Sunshine	
Corr-Coef	-0.368**	-0.388**	-0.351**	-0.046	-0.307*	
Max lag Corr-Coef and	-0.368**	-0.388**	-0.351**	0.351**	-0.315*	
Its lag order	0	0	0	6	1	

\* means significant at 0.05 level; \*\* means significant at 0.01 level

**Fig. 5** The data of influenza incidence in Guangxi were decomposed into trend part, seasonal part and random part



predict the influenza incidences from July 2019 to October 2019 and the prediction RMSE was 2.57 and to predict influenza incidence from November 2019 to February 2020, the predicted values were 2.0, 1.92, 15, and 4.73, respectively. Chart of fitting and predicting influenza incidence based on ARIMAX was shown in Fig. 7.

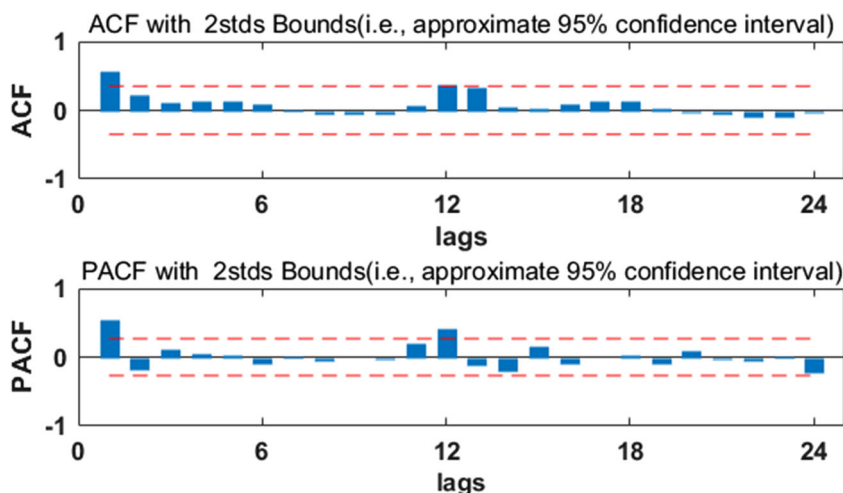
## Discussion

Influenza is an acute febrile respiratory infectious disease caused by influenza virus, characterized by short incubation period, rapid transmission, wide spread, variable antigen, general susceptibility, and difficulty in control, prone to outbreak or epidemic. In recent 2 years, the incidence of influenza in Guangxi showed a slight upward trend, so it is necessary to analyze the factors affecting the incidence of influenza and do a good job of early warning and prevention.

It is generally believed that climate indicators affect influenza epidemics by influencing the survival of viruses, crowd

and individual behavior. This study analyzed the correlation between the air pollution variables (PM<sub>2.5</sub>, PM<sub>10</sub>, CO, SO<sub>2</sub>, NO<sub>2</sub>, and O<sub>3</sub>) and the incidence of influenza, and the correlation between meteorological variables (average temperature, maximum temperature, minimum temperature, average relative humidity, and sunshine duration) and the incidence of influenza in Guangxi. Considering that there might be the lagging effect of air pollution and meteorological factors on disease, the cross-correlation function was used to analyze the lag correlation of 0–12 order between them; the results showed that the correlations of PM<sub>2.5</sub>, PM<sub>10</sub> and SO<sub>2</sub> were most significant at lag 6, their correlation coefficients were  $-0.36$ ,  $-0.4$ ,  $-0.49$  and  $-0.38$  respectively; the correlation of CO was most significant at lag 7, and the correlation coefficients was  $-0.38$ , while the correlation coefficients of NO<sub>2</sub> and O<sub>3</sub> were highest at lag 0 and their correlation coefficients were 0.26 and  $-0.35$ . The correlations of average temperature, maximum temperature, and minimum temperature were most significant at lag 0; their correlation coefficients were  $-0.368$ ,  $-0.388$ , and  $-0.351$  respectively; the correlation of sunshine

**Fig. 6** The ACF and PACF charts for influenza incidence data from January 2015 to June 2019



**Table 3** Four ARIMAX models with their parameters test results

	ARIMA(1,0,1) (0,0,1) <sub>12</sub>	ARIMAX(1,0,1) (0,0,1) <sub>12</sub> with PM2.5(lag6)	ARIMAX(1,0,1) (0,0,1) <sub>12</sub> with PM10(lag6)	ARIMAX(1,0,1) (0,0,1) <sub>12</sub> with SO <sub>2</sub> (lag6)	ARIMAX(1,0,1) (0,0,1) <sub>12</sub> with NO <sub>2</sub>
Variable		-0.11*	0.01*	0.08*	0.25**
AR(1)	1.05**	1.03**	1.02**	1.05**	0.87**
MA(1)	-1.3**	-0.98**	-1.3**	-1.32**	-1.25**
SMA(1)	0.46**	0.5**	0.48**	0.47**	0.62**
AIC	5.14	5.53	5.30	5.29	4.96
SBC	5.24	5.69	5.46	5.44	5.11

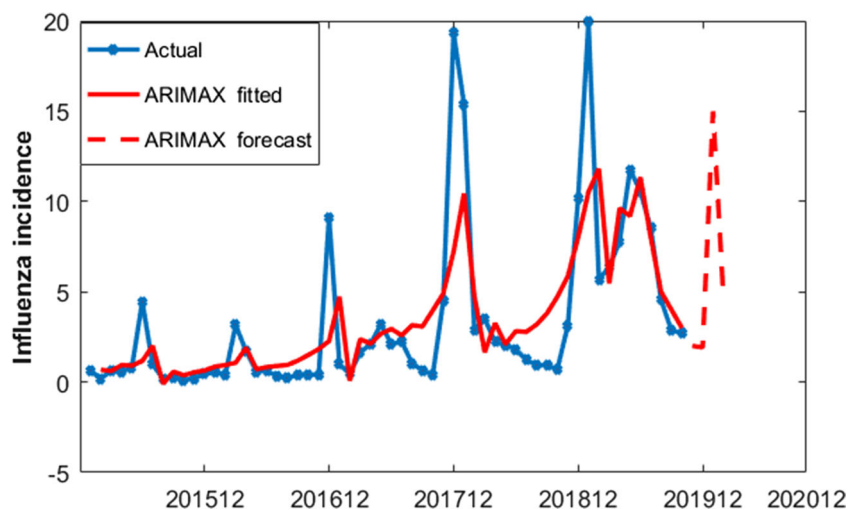
\* means significant at 0.05 level; \*\* means significant at 0.01 level

duration was most significant at lag 1, and the correlation coefficients was -0.315; the correlation of average relative humidity was most significant at lag 6, and the correlation coefficients was 0.351. Among them, most air pollution variables were negatively associated with influenza incidence; it was because that people in Guangxi often decide whether to take part in activities based on the weather forecast, and when air pollution was severe, they might reduce their outdoor activities; this reduced close contact behavior, which reduced the chance of being infected with the influenza. There was a positive correlation between NO<sub>2</sub> and the incidence of influenza, because NO<sub>2</sub> mainly damaged the respiratory tract and there were only slight irritation symptoms of eyes and upper respiratory tract, such as throat discomfort, dry cough, etc. Delayed pulmonary edema often develops after an incubation period of several to more than 10 h or more and then develops into pneumonia; at this point, on the one hand, the symptoms of the disease are similar to influenza; on the other hand, it reduces the body’s immunity and thus increases the risk of infection influenza. The three temperature variables and sunshine duration variable were negatively correlated with influenza incidence, indicating that the higher the temperature, the

lower the influenza incidence, and the average relative humidity was not significantly correlated with influenza incidence; however, there was a significant positive correlation between average relative humidity at lag 6 and the incidence of influenza, indicating that the relative humidity also had a long-term lag positive correlation effect on influenza incidence.

Disease epidemic dynamic estimation and prediction are important links of infectious disease prevention and control, and they are the main basis for health management to formulate prevention and control countermeasures and allocate resources. However, because the influenza epidemic is affected by many uncertain factors, it is difficult to identify early, which leads to the lag of corresponding preventive and control measures. How to timely and effectively warn the influenza epidemic has been the focus and difficulty of disease prevention and control. The effect of climate indicators on the incidence of influenza also suggests that significant explanatory variables should be taken into account in influenza prediction rather than only confined to historical data on influenza incidence. ARIMA model is the most commonly used time series prediction method, which has been widely used in various fields of preventive medicine; this study first considered

**Fig. 7** Chart of fitting and predicting influenza incidence based on ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub>



establishing a predictive model containing only influenza historical data. After careful analysis, the ARIMA (1,0,1)(0,0,1)<sub>12</sub> model was established. The AIC and SBC of this model were 5.14 and 5.25. Using this model to do fitting analysis for the modeling data, the fitting RMSE was 2.98; using this model to predict influenza incidence from July 2019 to October 2019, the prediction RMSE was 7.61. Secondly, considering the influence of climate indicators on the incidence of influenza, the ARIMAX model was established by including some lag variables of air pollution and meteorological factors in ARIMA (1,0,1)(0,0,1)<sub>12</sub> models; after careful analysis by model test and comparing the AIC and SBC of models, it was found that ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> had the smallest AIC and SBC values; they were 4.96 and 5.11, respectively; the fitting RMSE of this model was 2.67, and the prediction RMSE of this model was 2.57. Comparing the ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model with the ARIMA(1,0,1)(0,0,1)<sub>12</sub> model, it was found that the AIC, SBC values, and the fitting and prediction RMSE of the ARIMAX model were all less than these values of the ARIMA (1,0,1)(0,0,1)<sub>12</sub> model, indicating that the ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model was superior to the ARIMA (1,0,1)(0,0,1)<sub>12</sub> model.

## Conclusions

The incidence of influenza in Guangxi showed a slight upward trend in recent two years. It is necessary to analyze the factors affecting the incidence of influenza and do a good job of early warning. Considering the influence of climate indicator variables on the incidence of influenza and considering the lag effect of climate indicators on the incidence of influenza, we did the CCF correlation analysis between PM<sub>2.5</sub>, PM<sub>10</sub>, CO, SO<sub>2</sub>, NO<sub>2</sub>, O<sub>3</sub>, average temperature, maximum temperature, minimum temperature, average relative humidity, sunshine duration, and the incidence of influenza. The results found that when air pollution was severe, people might reduce outdoor activities; this reduced close contact behavior, which reduced the chance of being infected with the influenza. The more serious the NO<sub>2</sub> pollution, the higher the incidence of influenza; therefore, it is necessary to take good precautions to reduce the incidence of influenza when the air pollution is serious. The higher the temperature, the lower the incidence of influenza; therefore, people should keep warm in cold weather to prevent influenza. In this study, the ARIMA (1,0,1)(0,1)<sub>12</sub> model based on the historical incidence data of influenza in Guangxi and the ARIMAX (1,0,1)(0,1)<sub>12</sub> with NO<sub>2</sub> model were established; after comparison, it was found that ARIMAX(1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub> model was better than ARIMA(1,0,1)(0,0,1)<sub>12</sub> model, predicting influenza incidence in Guangxi by ARIMAX (1,0,1)(0,0,1)<sub>12</sub> with NO<sub>2</sub>

model may provide an evidence for influenza control policy-making and public health intervention.

**Funding information** This work was supported by the Natural science funding of Xinjiang Uygur Autonomous Region (Grant No.2019D01C206), China.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Arikawa G, Fujii Y, Abe M, Mai NT, Mitoma S, Notsu K, Nguyen HT, Elhanafy E, Daous HE, Kabali E, Norimine J, Sekiguchi S (2019) Meteorological factors affecting the risk of transmission of HPAI in Miyazaki, Japan. *Vet Rec Open* 6(1):e000341
- Bekking C, Yip L, Groulx N, Doggett N, Finn M, Mubareka S (2019) Evaluation of bioaerosol samplers for the detection and quantification of influenza virus from artificial aerosols and influenza virus-infected ferrets. *Influenza Other Respir Viruses* 13(6088):564–573
- Box GE, Jenkins GM (1976) *Time series analysis: forecasting and control* rev. ed. Oakland, California. Holden-Day 31(4):238–242
- Brattig NW, Tanner M, Bergquist R, Utzinger J (2019) Impact of environmental changes on infectious diseases: Key findings from an international conference in Trieste, Italy in May 2017. *Acta Trop*: 105165
- Chong KC, Lee TC, Bialasiewicz S et al (2019) Association between meteorological variations and activities of influenza A and B across different climate zones: a multi-region modeling analysis across the globe. *J Inf Secur* 30(19):S0163–S44533
- Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K, Brunekreef B, Dandona L, Dandona R, Feigin V, Freedman G, Hubbell B, Jobling A, Kan H, Knibbs L, Liu Y, Martin R, Morawska L, Pope CA III, Shin H, Straif K, Shaddick G, Thomas M, van Dingenen R, van Donkelaar A, Vos T, Murray CJL, Forouzanfar MH (2017) Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the global burden of diseases study 2015. *Lancet* 389(10082):1907–1918
- Chen EQ, Zeng F, Zhou LY et al (2018) Early warning and clinical outcome prediction of acute-on chronic hepatitis B liver failure. *World J Gastroenterol* 42:92–101
- Elhakim M, Hafiz R M, Fahim M, et al. (2019) Epidemiology of severe cases of influenza and other acute respiratory infections in the eastern Mediterranean region, July 2016 to June 2018. *J Infect Publ Health*
- Feng Z, Velasco-Hernandez J, Tapia-Santos B (2013) A mathematical model for coupling within-host and between-host dynamics in an environmentally-driven infectious disease. *Math Biosci* 241(1):49–55
- Gabriel A et al (2019) Dengue outbreaks: unpredictable incidence time series. *Epidemiol Infect* 147:E116,1–E116,7
- Gharbi M, Quenel P, Gustave J et al (2011) Time series analysis of dengue incidence in Guadeloupe, French West Indies: forecasting models using climate variables as predictors. *BMC Infect Dis* 11: 166
- Khan MD, Thi Vu HH, Lai QT, Ahn JW (2019) Ahn JW. Aggravation of human diseases and climate change nexus. *Int J Environ Res Public Health* 16(15):2799



- Li H, Luo et al. (2017) An artificial neural network prediction model of congenital heart disease based on risk factors: a hospital-based case-control study 96(6):e6090
- Liu L, Luan RS, Yin F et al (2016) Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. *Epidemiol Infect* 144(01):144–151
- Mao Q, Zhang K, Yan W et al (2018) Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *J Infect Publ Health* 2018: S1876034118300455
- Nikonahad A, Khorshidi A, Ghaffari HR, Aval HE, Miri M, Amarloei A, Nourmoradi H, Mohammadi A (2017) A time series analysis of environmental and metrological factors impact on cutaneous leishmaniasis incidence in an endemic area of Dehloran, Iran. *Environ Sci Pollut Res* 24:14117–14123
- Priedhorsky R, Daughton AR, Barnard M, O’Connell F, Osthus D (2019) Estimating influenza incidence using search query deceptiveness and generalized ridge regression. *PLoS Comput Biol* 15(10): e1007165
- Rodrigues E, Machado A, Silva S, Nunes B (2018) Excess pneumonia and influenza hospitalizations associated with influenza epidemics in Portugal from season 1998/1999 to 2014/2015. *Influenza Other Respir Viruses* 12(1):153–160
- Stewart-Ibarra AM, Romero M, Hinds AQJ, Lowe R, Mahon R, van Meerbeeck CJ, Rollock L, Gittens-St. Hilaire M, St. Ville S, Ryan SJ, Trotman AR, Borbor-Cordova MJ (2019) Co-developing climate services for public health: stakeholder needs and perceptions for the prevention and control of Aedes-transmitted diseases in the Caribbean. *PLoS Negl Trop Dis* 13(10):e0007772
- Tian C et al (2019) Time-series modeling and forecasting of hand, foot and mouth disease cases in China from 2008 to 2018. *Epidemiol Infect* 147:E82,1–E82,3
- Tuerlinckx D, Bodart E, Jamart J, Glupczynski Y (2009) Prediction of Lyme meningitis based on a logistic regression model using clinical and cerebrospinal fluid analysis. *Pediatr Infect Dis J* 28(5):394–397
- Wang Y-w, Shen Z-z, Jiang Y (2019) Comparison of autoregressive integrated moving average model and generalized regression neural network model for prediction of hemorrhagic fever with renal syndrome in China: a time-series study. *BMJ Open* 9(6):e025773
- Watts N, Amann M, Arnell N, Ayeb-Karlsson S, Belesova K, Berry H, Bouley T, Boykoff M, Byass P, Cai W, Campbell-Lendrum D, Chambers J, Daly M, Dasandi N, Davies M, Depoux A, Dominguez-Salas P, Drummond P, Ebi KL, Ekins P, Montoya LF, Fischer H, Georgeson L, Grace D, Graham H, Hamilton I, Hartinger S, Hess J, Kelman I, Kiesewetter G, Kjellstrom T, Kniveton D, Lemke B, Liang L, Lott M, Lowe R, Sewe MO, Martinez-Urtaza J, Maslin M, McAllister L, Mikhaylov SJ, Milner J, Moradi-Lakeh M, Morrissey K, Murray K, Nilsson M, Neville T, Oreszczyn T, Owfi F, Pearman O, Pencheon D, Pye S, Rabbaniha M, Robinson E, Rocklöv J, Saxer O, Schütte S, Semenza JC, Shumake-Guillemot J, Steinbach R, Tabatabaei M, Tomei J, Trinanes J, Wheeler N, Wilkinson P, Gong P, Montgomery H, Costello A (2018) The 2018 report of the lancet countdown on health and climate change: shaping the health of nations for centuries to come. *Lancet*. 392(10163):2479–2514
- Wei W, Jiang J, Gao L et al (2017) A new hybrid model using an autoregressive integrated moving average and a generalized regression neural network for the incidence of tuberculosis in Heng County, China. *Am J Tropic Med Hygiene* 97(3):799–805
- Yang X, Zou J, Kong D, Jiang G (2018) The analysis of GM (1, 1) grey model to predict the incidence trend of typhoid and paratyphoid fevers in Wuhan City, China. *Medicine* 97(34):e11787
- Yi-Yi Z, Wei F, Qi Z et al (2012) Application of multiple seasonal ARIMA model in forecasting the incidence of hepatitis A in Shanghai. *Fudan Univ J Med Sci* 39(5):460–464
- Zhao D, Wang L, Cheng J et al (2017) Impact of weather factors on hand, foot and mouth disease, and its role in short-term incidence trend forecast in Huainan City, Anhui Province. *Int J Biometeorol* 61: 453–461

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