

RESEARCH ARTICLE

ASTM: Developing the web service for anthrax related spatiotemporal characteristics and meteorology study

Shiwei Fan^{1,†}, Ming Xiao^{1,†}, Boyu Sun¹, Weizhong Zhou², Qingrong Chen², Weimin Lv³, Pengfei Zhang^{4,5}, Le Zhang^{1,*}

¹ College of Computer Science, Sichuan University, Chengdu 610065, China

² Gannan Center for Disease Control and Prevention, Hezuo 747000, China

³ Gannan Blood Station, Hezuo 747000, China

⁴ Department of Medical Oncology, Cancer Center, West China Hospital, Sichuan University, Chengdu 610065, China

⁵ Med-X Center for Informatics, Sichuan University, Chengdu 610065, China

* Correspondence: zhangle06@scu.edu.cn

Received August 4, 2021; Revised September 11, 2021; Accepted October 10, 2021

Background: Anthrax is a zoonotic infectious disease caused by *Bacillus anthracis*. Investigating the spatiotemporal characteristics of anthrax and the impact of meteorological factors on the incidence of anthrax is helpful for the prevention and control of anthrax.

Methods: At first, we applied the Granger causality test to explore the spatiotemporal characteristics of anthrax transmission between the counties and cities of Gannan Tibetan Autonomous Prefecture, Gansu Province of China. Then, we constructed three generalized linear models to analyze the impact of meteorological factors on the monthly number of anthrax cases in Gannan Tibetan Autonomous Prefecture. Finally, we developed an easy-to-use online web server that integrates the above functions.

Results: This study developed an online service website (ASTM Anthrax in Gannan, Zhang Lab) for the analysis and visualization of anthrax, which not only can investigate the correlation of anthrax among different regions in Gannan Tibetan Autonomous Prefecture, but also can analyze the correlation between meteorological factors and the number of anthrax cases.

Conclusions: Our study not only explored spatiotemporal characteristics of anthrax transmission, but also analyzed the impact of seven meteorological factors on the monthly number of anthrax cases. Meanwhile, the online service website which integrates the above functions is useful for the prevention and control of anthrax.

Keywords: time series analysis; correlation analysis; visualization; web service; data mining

Author summary: We tried to use the anthrax cases and meteorological data of Gannan Tibetan Autonomous Prefecture, Gansu Province, from 2005 to 2019 to investigate the spatiotemporal characteristics of anthrax incidence and its relationship with the meteorological factors. In addition, we developed a visualization platform for data sharing and analysis. Finally, our study can provide references for the prevention and control of anthrax for China and other countries.

INTRODUCTION

Anthrax is a zoonotic disease caused by *Bacillus anthracis*, and it has a very serious impact on the social

economy and public health safety of the population [1–3]. Although the current incidence of anthrax is gradually decreasing due to the reasonable disposal of contaminated materials and vaccination, anthrax is still

[†] These authors contributed equally this work.

spreading worldwide [4]. For the effective prevention and control of anthrax, the spatiotemporal transmission characteristics and the connection between meteorological factors and the infectious ability of anthrax must be understood.

Previous studies have already investigated the spatial or temporal characteristics of anthrax that broke out in many regions in China to understand the distribution characteristics of anthrax in different regions [5–7]. For example, Lyu *et al.* [5] applied a spatial autocorrelation analysis to study the temporal distribution characteristics of anthrax that occurred in Gannan Tibetan Autonomous Prefecture (Gannan), Gansu Province, China. And the results indicated a significant temporal aggregation of anthrax incidence from 2011 to 2017, mainly in 2015–2017. Additionally, Yu *et al.* [6] used a variety of molecular typing methods to analyse the genetic relationship of the *Bacillus anthracis* strains that broke out in Min County, Dingxi City, Gansu Province in August 2016 to explore whether the strains that broke out in Min County came from Gansu, Sichuan, and Qinghai. Moreover, Li *et al.* [7] found that the newly discovered areas of anthrax are generally connected with areas previously affected by anthrax. However, few studies have investigated the temporal and spatial characteristics of anthrax transmission among the counties and cities in Gannan based on time series of anthrax cases. Therefore, identifying a method of integrating the time series of anthrax cases that occurred in recent years to understand the temporal and spatial characteristics of anthrax transmission between the counties and cities of Gannan is our first scientific question.

Recent research indicates that meteorological factors such as temperature, precipitation, and humidity are closely associated with bacterial proliferation and spread, such as *Brucella* [8] and *Salmonella* [9]. Since anthrax is a kind of bacterial disease, there may be a certain correlation between meteorological factors and the incidence of anthrax. In recent years, some researchers have analysed the correlation between meteorological factors and the incidence of anthrax [10,11]. However, the types of meteorological factors covered by these studies have not been comprehensively studied. Therefore, identifying a method of exploring the precise quantitative correlation between multiple meteorological factors (such as temperature, wind speed, duration of sunshine, relative humidity, atmospheric pressure, precipitation, and heat index) and the number of anthrax cases to provide a reference for the prevention and control of anthrax has become our second scientific question.

To provide information on the transmission and incidence of anthrax, Chen *et al.* [10] used ArcGIS [12] and other statistical software for data analysis and

visualization. Additionally, Carlson *et al.* [13] used boosted regression trees to model and visualize anthrax that breaks out around the world. However, since these studies require high programming skills for researchers, they will complicate and impede the research progress. Therefore, identifying a method of developing an easy-to-use anthrax analysis and visualization platform is our third scientific question.

For these reasons, we developed an online web server called ASTM (Anthrax Related Spatiotemporal Characteristics and Meteorology analysis), which not only offers an easy-to-use online service to analyse the spatiotemporal correlation of anthrax among different regions but also provides an analysis of the impact of different meteorological factors on the incidence of anthrax. Moreover, ASTM visualizes anthrax cases in Gannan from multiple prospects.

RESULTS

Figure 1 shows the ASTM website (Anthrax in Gannan, Zhang Lab) with three function modules. And we are going to introduce their functions in detail.

Data overview

Figure 1 shows that the “Data Overview” link of the homepage has three modules. The first is the “Data Statistics” module, which describes the distribution of gender, age, occupation, location, and season of all the anthrax cases. The second is the “Spatial Distribution” module, which describes the spatial distribution of the anthrax cases in Gannan. The last is the “Time Distribution” module, which describes the time distribution for all anthrax cases from 2005 to 2019.

Data statistics

After clicking the “Data Statistics” link, we can visualize the gender, age, occupation, location, and season distribution of the anthrax cases in Gannan from 2005 to 2019 (Fig. 2). As described in Fig. 2, we can click the button above the pie chart to switch the classification method and hover the cursor over the pie chart to display the corresponding number of a certain category. For example, we found that summer and autumn were high-incidence periods of anthrax that accounted for more than 60% of the total cases after selecting “Season” as the classification label (Fig. 2B).

Spatial distribution

After clicking the “Spatial Distribution” link, we can

Anthrax analysis and visualization in Gannan Tibetan Autonomous Prefecture

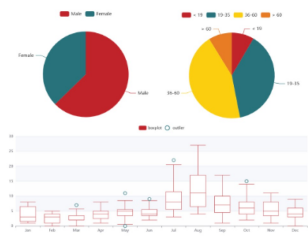
Introduction

Anthrax is a infectious disease in Gannan Tibetan Autonomous Prefecture, Gansu Province, China. This website is the analysis and visualization of anthrax cases in Gannan between 2005 and 2019. In addition, this website contains regression models to investigate the association between meteorological factors and anthrax incidence.

Data Overview

Simple visualization of anthrax cases in Gannan Prefecture.

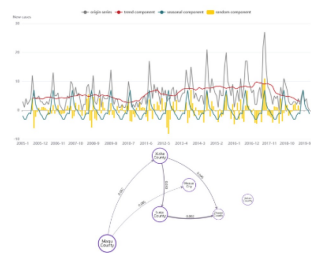
Data Statistics
Spatial Distribution
Time Distribution



Time Series Analysis

Time series analysis of monthly new cases.

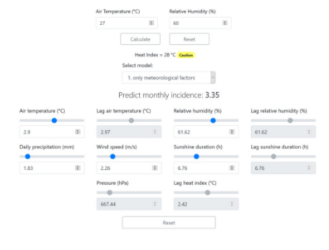
Time Series Decomposition
Holt-Winters Forecast
Granger Casualty Test



Regression Analysis

Regression analysis for correlation between anthrax cases and meteorological factors.

Heat Index Calculator
Regression Models



© 2021 College of Computer Science, Sichuan University

Figure 1. ASTM homepage.

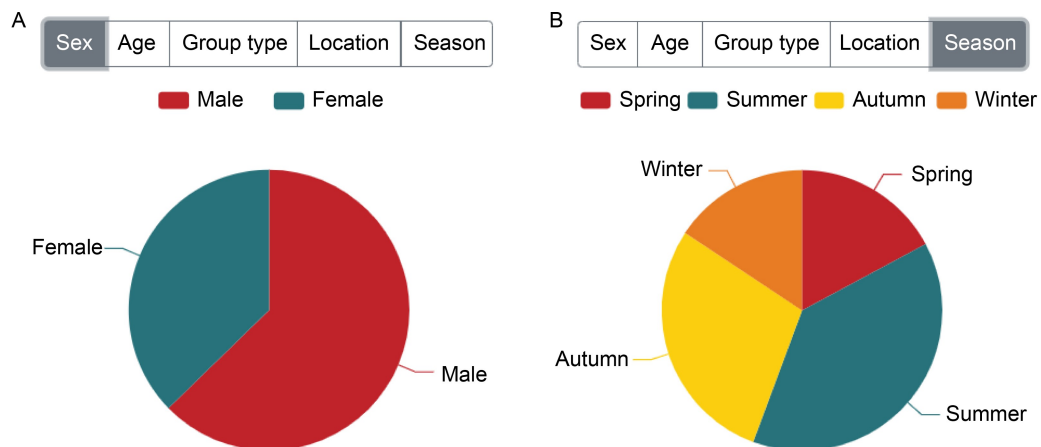


Figure 2. Summary of anthrax cases in Gannan Tibetan Autonomous Prefecture from 2005 to 2019. (A) Anthrax cases statistics by sex. (B) Anthrax cases statistics by season.

visualize the spatial distribution of anthrax cases in Gannan from 2005 to 2019. As shown in Fig. 3, we use bar chart to show the number of anthrax cases in a

certain time interval for different regions of the map. For example, we can select the time interval through the drop-down box and click “Submit” to obtain the results.

Moreover, we can hover the cursor over a county on the bar chart to display the number of anthrax cases in this county during a certain time interval. In addition, we can save the current result as a PNG image by clicking the button on the right of the bar chart.

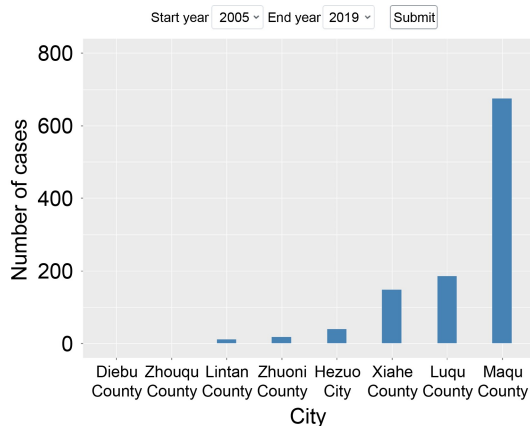


Figure 3. Spatial distribution of anthrax cases in Gannan Tibetan Autonomous Prefecture.

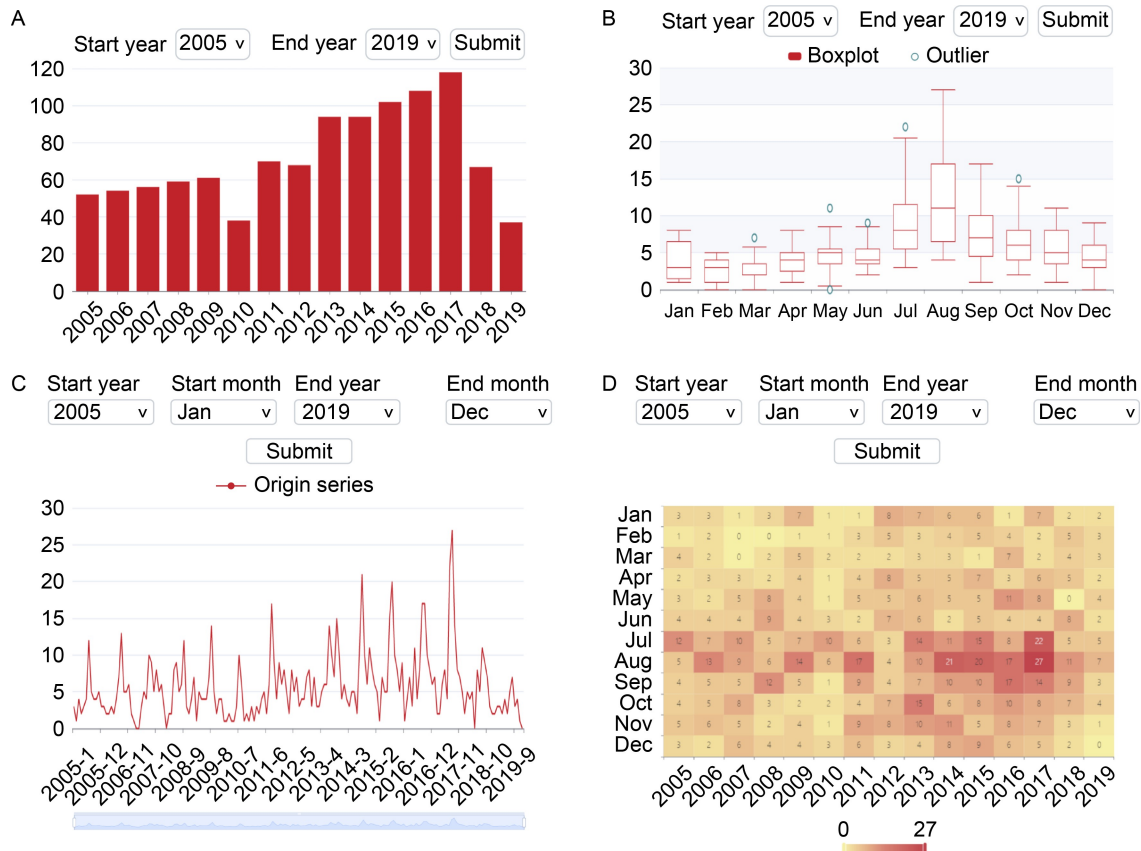


Figure 4. Time distribution of the anthrax cases in Gannan Tibetan Autonomous Prefecture from 2005 to 2019. (A) Bar chart of yearly anthrax cases. (B) Boxplot of monthly anthrax cases. (C) Line chart of monthly anthrax cases. (D) Heatmap of monthly anthrax cases.

Time distribution

After clicking the “Time Distribution” link, we can visualize the time distribution of the anthrax cases in Gannan from 2005 to 2019 (Fig. 4). Figure 4A shows the yearly number of anthrax cases. We can hover the cursor over the bar chart to display the corresponding number of anthrax cases for each year. At the same time, Fig. 4B–D show the number of monthly anthrax cases by boxplot, line chart, and heatmap, respectively. For example, after selecting “2005” and “2019” as the start and end years, respectively, we can click “Submit” to obtain results. Figure 4A demonstrates that the number of anthrax cases between 2014 and 2017 was the greatest. Additionally, Fig. 4B–D shows that the number of anthrax cases between July and August is the greatest for every year.

Time series analysis

The “Time Series Analysis” link of the homepage (Fig. 1) provides three functional modes. The first is the “Time

Series Decomposition” mode, which can decompose time series into different components. The second is the “Holt-Winters Forecast” mode, which can predict the new incidence for anthrax cases. The last is the “Granger Causality Test” mode, which can investigate the connections of anthrax cases between different regions.

Time series decomposition

After clicking the “Time Series Decomposition” link, we can decompose the Gannan anthrax case time series into trend, seasonal, and irregular components by the *decompose()* function of the R package *stats* [14]. We can select the start and end year by the drop-down box and obtain the analytical results after clicking the “Submit” button. Next, the user can click the button at the right of the chart to download the detailed result. For example, Fig. 5 shows the decomposition result of the anthrax time series from 2005 to 2019, where the original value, trend value, and seasonal value are presented by a line chart, and the irregular value is presented by a bar chart.

Holt-Winters forecast

After clicking the “Holt-Winters Forecast” link, we can predict the number of new anthrax cases in Gannan by the *HoltWinters()* function of the R package *stats* [14]. For example, we can select “2005” and “2019” as the start and end years, respectively. After clicking the “Submit” button, Fig. 6 shows the forecast result by the Holt-Winters model.

Granger causality test

After clicking the “Granger Causality Test” link, we can investigate the spatiotemporal correlation of anthrax for different regions. The correlation graph (Fig. 7) employs a purple bordered circle to represent the region (county or city in Gannan). In Fig. 7A, the radius of the circle represents the number of anthrax cases in the region; the larger the radius is, the greater the number of anthrax cases. A directed edge from “x” to “y” indicates that “x” is the Granger cause for “y”. The thickness of the edge represents the degree of significance; the thicker the edge, the stronger the significance. The value on the edge represents the p -value of the Granger causality test [15]. Figure 7B shows the p -value of the Granger causality test; the lower the value is, the darker the colour of the grid is.

For example, anthrax in Maqu County is the Granger cause of anthrax in Hezuo City and Xiahe County,

anthrax in Xiahe County is the Granger cause of anthrax in Zhuoni County and Luqu County, and anthrax in Luqu County is the Granger cause of anthrax in Zhuoni County, as shown in Fig. 7A.

Regression analysis

Finally, the “Regression Analysis” link of the homepage (Fig. 1) provides two functional modes. One is the “Heat Index Calculator” mode, which can calculate the corresponding heat index based on temperature and relative humidity. The other is the “Regression Models” mode, which can carry out predictive analysis for the number of anthrax cases based on meteorological factors.

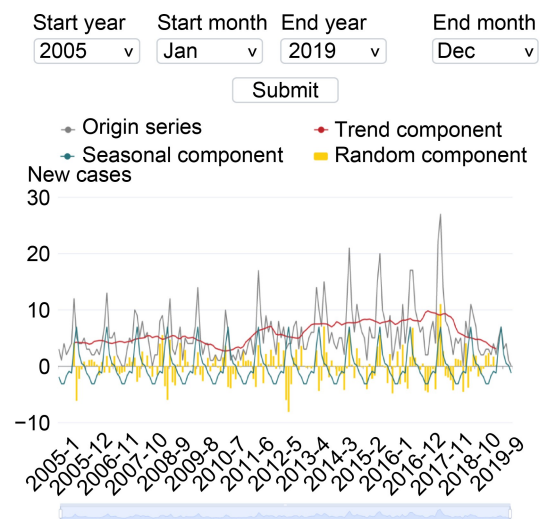


Figure 5. Time series decomposition of monthly anthrax cases.

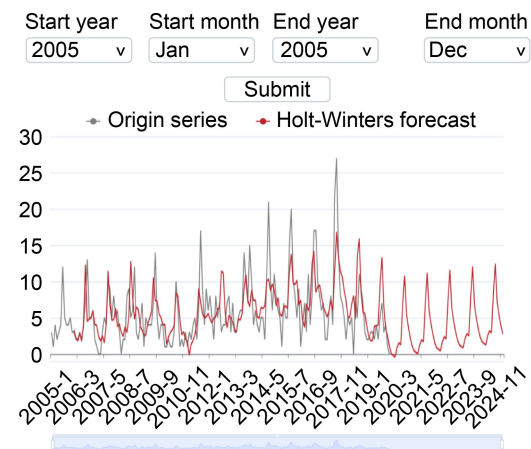


Figure 6. Holt-Winters Exponential Smoothing forecast result. Here, the grey line represents the original time series and the red line represents the forecast time series.

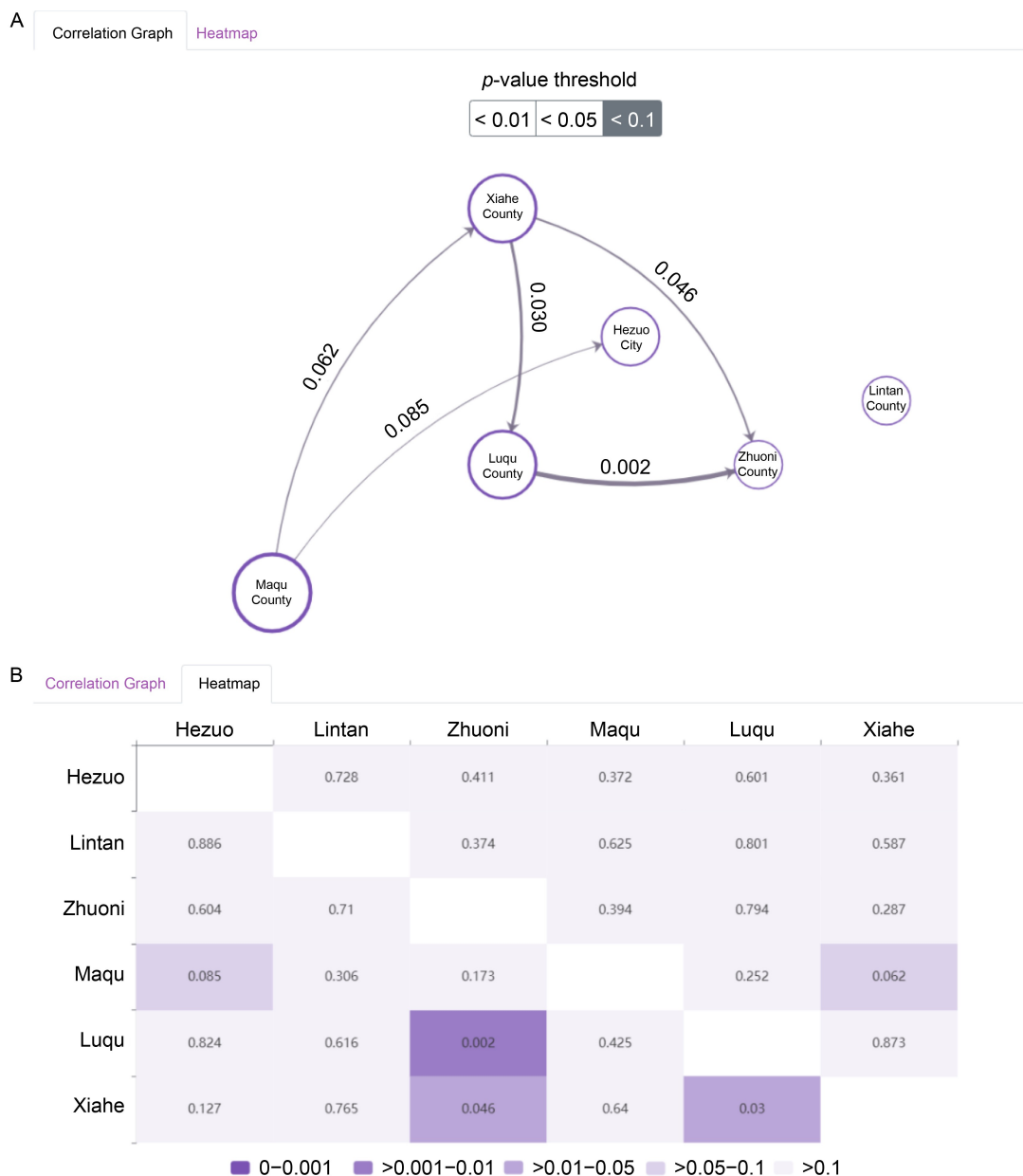


Figure 7. Correlation of anthrax between different regions. (A) Correlation graph of anthrax in different regions. (B) Heatmap of the p -value of the correlation of anthrax in different regions.

Heat index calculator

After clicking the “Heat Index Calculator” link of Fig. 8, we can calculate the heat index and corresponding category by inputting the air temperature ($^{\circ}\text{C}$) and relative humidity (%) in the input box. After clicking the “calculate” button, we can obtain the analytical results.

There are five categories given by the U.S. National Weather Service (NWS): safety, caution (fatigue is possible with prolonged exposure and physical activity), extreme caution (sunstroke and muscle cramps are possible), danger (heat cramps and heat exhaustion are

likely), and extreme danger (sunstroke is likely). For example, after inputting an air temperature ($^{\circ}\text{C}$) of “35” and relative humidity (%) of “35” and then clicking the “Calculate” button, the results show that the heat index is “36 $^{\circ}\text{C}$ ” and the corresponding category is “extreme caution” (Fig. 8).

Regression models

After clicking the “Regression Models” link, three generalized linear models can be selected to predict the number of anthrax cases in Gannan (Fig. 9). After

selecting a certain model described in “Materials and methods” through the drop-down box and setting the value of the independent variable through the input box or the slider below the input box, the number of anthrax cases will be computed. For example, after we select the first model and set the value of the independent variable of Fig. 9, the predictive number of anthrax cases is “3.35”. In addition, we can restore all input values to the default by clicking the “Reset” button. Table 1 and Fig. 10 show the coefficient of each meteorological factor in the three regression models, which turn out that the precipitation, temperature, heat index, and sunshine duration have a positive impact on the number of anthrax cases, whereas the relative humidity, wind speed, and air pressure have a negative impact on the number of anthrax cases. In addition, Fig. 11 demonstrates that the MAEs of the three models are all approximately 1.87, implying sufficient predictive ability. Moreover, since the second model has the smallest MAE, it has the best predictive ability.

DISCUSSION

In general, this study provides an informative and interactive platform for the visualization and analysis of

Air temperature (°C)
35

Relative humidity (%)
35

Calculate Reset

Heat index = 36°C Extreme caution

Figure 8. Heat index calculator.

anthrax cases in Gannan Tibetan Autonomous Prefecture, Gansu Province, China. Additionally, it answers our three scientific questions, which are discussed in the introduction section as follows: (1) identify a method of integrating temporal and spatial information to understand the transmission characteristics of anthrax; (2) determine a method of exploring the impact of meteorological factors on the incidence of anthrax; and (3) develop an easy-to-use anthrax analysis and visualization platform.

For the first question, Figs. 5–7 show the spatiotemporal correlation of anthrax cases in different counties by the Granger causality test. Figure 6 indicates that the number of anthrax cases will decrease in the next 4 years due to the stronger livestock quarantine protocols enacted since 2017 [16]. Meanwhile, Fig. 7 implies that anthrax may spread from Maqu County to Hezuo City and Xiahe County, from Xiahe County to Zhuoni County and Luqu County, and from Luqu County to Zhuoni County.

For the second question, Figs. 8–10 and Table 1 describe the impact of various meteorological factors on the number of anthrax cases by a regression analysis. Specifically, as shown in Table 1 and Fig. 10, the temperature and heat index have a strong positive impact on the number of anthrax cases, whereas the relative humidity, wind speed, and air pressure have a strong negative impact on the number of anthrax cases. Additionally, meteorological factors such as temperature, duration of sunshine, and relative humidity have a delayed impact on the number of anthrax cases.

For the third question, Fig. 1 shows an easy-to-use online platform for anthrax that can not only visualize anthrax cases in Gannan from 2005 to 2019 but also

Correlation between anthrax cases and meteorological factors

Selected model
1. Only meteorological factors

Predict monthly cases: 3.35

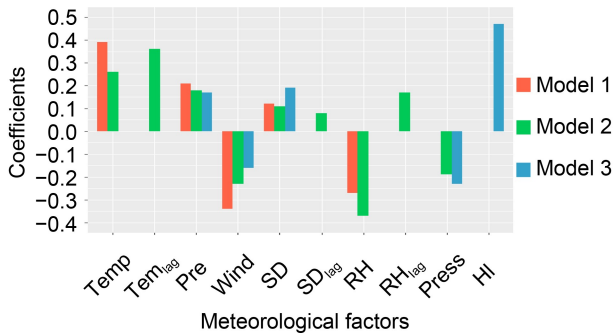
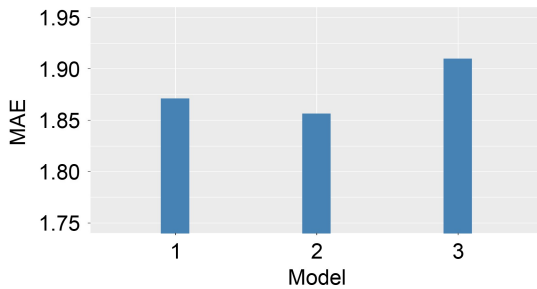
Air temperature (°C)	Lag air temperature (°C)	Relative humidity (%)	Lag relative humidity (%)
2.9	2.97	61.62	61.62
Daily precipitation (mm)	Wind speed (m/s)	Sunshine duration (h)	Lag sunshine duration (h)
1.83	2.26	6.76	6.76
Pressure (hPa)	Lag heat index (°C)		
667.44	2.42		

Reset

Figure 9. Three predictive models for anthrax cases.

Table 1 Coefficients of the three regression models

Model	Temp	Temp _{lag}	Pre	Wind	SD	SD _{lag}	RH	RH _{lag}	Press	HI
1	0.39	–	0.21	–0.34	0.12	–	–0.27	–	–	–
2	0.26	0.36	0.18	–0.23	0.11	0.08	–0.37	0.17	–0.19	–
3	–	–	0.17	–0.16	0.19	–	–	–	–0.23	0.47

**Figure 10.** Visualization of the coefficients for three regression models.**Figure 11.** MAE of the three-anthrax case predictive models.

show the analytic results for our first and second questions.

However, this study still has some limitations. First, since we assume a linear correlation between different meteorological factors and the number of anthrax cases, the regression analysis results are not sufficiently accurate due to the potential nonlinear correlation. In detail, when the value of a meteorological factor is less than a certain threshold, its increase may have a positive impact on the number of anthrax cases, and when it is greater than the threshold, its increase may have a negative impact, just like the curve of $y = -x^2$. For example, the temperature and relative humidity have the nonlinear effect on dengue incidence in Guangzhou, China [17]. Second, because the number of anthrax cases is not large enough, an underfitting problem will be observed for the model training.

Therefore, to overcome these shortcomings, we will investigate the nonlinear correlation between different meteorological factors and the number of anthrax cases in further studies. Finally, since anthrax still exists

around the world, we will develop more in-depth methods to further analyse the pathogenic factors and transmission characteristics of anthrax in the future.

MATERIALS AND METHODS

Data collection and computing methods

Meteorological factor data collection

We obtained the daily average data of 6 meteorological factors [18] (temperature, wind speed, duration of sunshine, relative humidity, atmospheric pressure, and precipitation) from 2005 to 2019, monitored by 2 meteorological monitoring stations located in Gannan Tibetan Autonomous Prefecture, Gansu Province, China. Table 2 lists the six meteorological factors and their corresponding symbols in the subsequent equations.

Stationarity test for time series data of anthrax cases

We obtained clinical data [19] of a total of 1078 anthrax cases from 2005 to 2019 from medical institutions in Gannan, Gansu Province, China. Before investigating the spatial-temporal correlation of anthrax in different regions by the Granger causality test [15], we employ both the augmented Dickey-Fuller test (ADF) [20,21] and the Kwiatkowski–Phillips–Schmidt–Shin test (KPSS) [22] to determine whether the time series of monthly anthrax cases is stationary, as shown in Fig. 12.

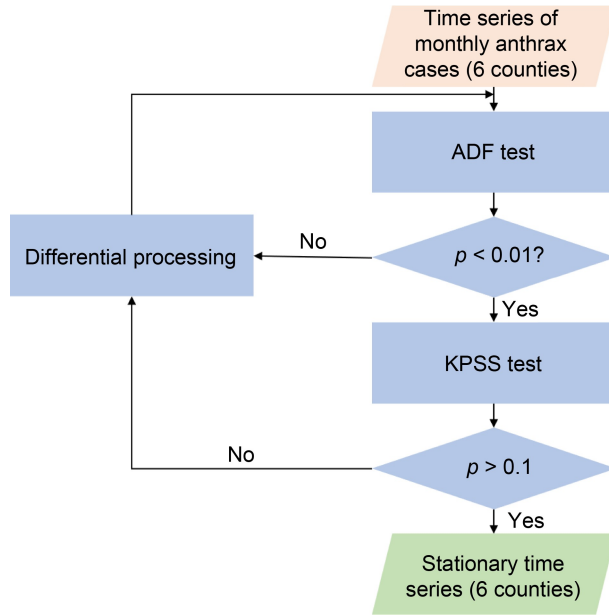
Equation (1) shows the ADF test model, where α_0 is the offset term, $\alpha_1 t$ is the linear time trend term, Δy_t is the first-order difference of time series y_t , ϵ_t is the error term and p is the order of lag difference. The null hypothesis is that the time series has a unit root ($\mu=0$, Eq. (1)). When the p -value is less than 0.01, we consider that the null hypothesis can be rejected, thus indicating that the time series is stationary.

$$\Delta y_t = \alpha_0 + \alpha_1 t + \mu y_{t-1} + \sum_{i=1}^p \theta_i \Delta y_{t-i} + \epsilon_t. \quad (1)$$

For the KPSS test, we develop the regression equation as Eq. (2) in the beginning. Then, we use ordinary least squares to obtain the residual series by Eq. (3). Finally, we calculate the LM statistics [22] by Eq. (4) to determine the p -value [23–27]. The null hypothesis is

Table 2 Six meteorological factors

Meteorological factor	Unit	Symbol in equation
Temperature	°C	Tem
Precipitation	mm	Pre
Wind speed	m/s	Wind
Relative humidity	h	SD
Sunshine duration	%	RH
Atmospheric pressure	hPa	Press

**Figure 12.** Time series stationarity test workflow.

that the time series is stationary. If the p -value is greater than 0.1, then we consider that the null hypothesis cannot be rejected, indicating that the time series is stationary.

$$y_t = \delta x_t + \mu_t \text{ where } t = 1, 2, 3, \dots, T, \quad (2)$$

$$\bar{\mu}_t = y_t - \bar{\delta}x_t, \quad (3)$$

$$LM = \sum_{t=1}^T S(t)^2 / (T^2 f_0), \text{ where } S(t) = \sum_{i=1}^t \bar{\mu}_i, \quad (4)$$

where x_t , y_t , $\bar{\mu}_t$ and f_0 are the predicted time series, original time series, residual series, and frequency zero spectrum estimation, respectively. If any stationarity test results show that the time series is not stationary, then the time series are differentiated. The above process is repeated until the time series is stationary.

Spatiotemporal correlation analysis of the anthrax

For stationary time series x_t and y_t , we use ordinary least squares [28–30] for regression analysis by Eq. (5).

$$y_t = \alpha_0 + \sum_{i=1}^k \alpha_i y_{t-i} + \sum_{i=1}^k \beta_i x_{t-i} + \varepsilon_t, \quad (5)$$

where k represents the lag order and ε_t is the error term. We use the F-test [23,31–39] to verify the null hypothesis (Eq. (6)). If the null hypothesis is rejected, then the alternative hypothesis (Eq. (7)) is accepted, indicating that x_t is the Granger reason for y_t . If the null hypothesis is accepted, x_t is not a Granger reason for y_t . We choose $k = 2$ as the optimal parameter with respect to the Akaike information criterion described by Luciano *et al.* [40].

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0, \quad (6)$$

$$H_a : \exists 1 \leq i \leq k \text{ such that } \beta_i \neq 0. \quad (7)$$

Impact of meteorological factors on anthrax

To analyse the impact of meteorological factors on the monthly number of anthrax cases in Gannan, we construct three generalized linear models (GLMs) with a Poisson distribution. Table 1 shows each symbol in the subsequent equations.

Eq. (8) shows the first model, in which the monthly number of anthrax cases (I_t) is the dependent variable and the monthly averages of the six meteorological factors are independent variables.

$$\log(I_t) = \alpha_0 + \alpha_1 \text{Tem}_t + \alpha_2 \text{Pre}_t + \alpha_3 \text{Wind}_t + \alpha_4 \text{SD}_t + \alpha_5 \text{RH}_t + \alpha_6 \text{Press}_t, \quad (8)$$

where i is the index of the month and $\alpha_j (j = 1, 2, \dots, 6)$ represents the coefficient.

Based on the first model, we developed our second model (Eq. (9)) by adding the monthly average lag value of the six meteorological factors.

$$\begin{aligned} \log(I_t) = & \alpha_0 + \alpha_{1,1} \text{Tem}_t + \alpha_{1,2} \text{Tem}_{t-1} + \alpha_{2,1} \text{Pre}_t + \alpha_{2,2} \text{Pre}_{t-1} \\ & + \alpha_{3,1} \text{Wind}_t + \alpha_{3,2} \text{Wind}_{t-1} + \alpha_{4,1} \text{SD}_t + \alpha_{4,2} \text{SD}_{t-1} \\ & + \alpha_{5,1} \text{RH}_t + \alpha_{5,2} \text{RH}_{t-1} + \alpha_{6,1} \text{Press}_t + \alpha_{6,2} \text{Press}_{t-1}, \end{aligned} \quad (9)$$

where α_0 and $\alpha_{j,k} (j = 1, 2, \dots, 6; k = 1, 2)$ represent regression coefficients.

To explore the impact of heat exposure on the incidence of anthrax, we incorporate the heat index (HI) [41] into the third model (Eq. (10)) to describe the impact of heat on humans [42,43] by considering both the dry bulb temperature and relative humidity.

$$\begin{aligned} \log(I_t) = & \alpha_0 + \alpha_{1,1} \text{HI}_t + \alpha_{1,2} \text{HI}_{t-1} + \alpha_{2,1} \text{Pre}_t + \alpha_{2,2} \text{Pre}_{t-1} \\ & + \alpha_{3,1} \text{Wind}_t + \alpha_{3,2} \text{Wind}_{t-1} + \alpha_{4,1} \text{SD}_t + \alpha_{4,2} \text{SD}_{t-1} \\ & + \alpha_{5,1} \text{Press}_t + \alpha_{5,2} \text{Press}_{t-1}. \end{aligned} \quad (10)$$

We use the MAE (Mean Absolute Error, Eq. (11)) to evaluate the predictive ability of the regression model. A

less MAE indicates a better predictive ability of the model [44].

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - h(X_i)|, \quad (11)$$

where X_i represents the six meteorological factors, y_i is the true value, $h(X_i)$ is the predicted value and n is the number of samples.

To eliminate the impact of regional differences, we grouped the data by counties and cities before performing the regression. At the same time, a stepwise regression was performed to select variables according to the Akaike information criterion (AIC) [45] to reduce the instability of regression coefficients caused by collinearity among the variables.

Website development

As our previous studies [46–49], we developed ASTM on the cloud Linux server [50] (CentOS 7.5.1804) by employing Nginx [51] (version 1.18) as the web server. The front end of the website is based on the Bootstrap framework [52] (version 4.5.3) and React framework [53] (version 17.0.1).

ACKNOWLEDGEMENTS

This work was supported by grants from the National Natural Science and Technology Major Project (No. 2018ZX10201002), China Postdoctoral Science Foundation (No. 2020M673221), and the Fundamental Research Funds for the Central Universities (No. 2020SCU12056).

COMPLIANCE WITH ETHICS GUIDELINES

The authors Shiwei Fan, Ming Xiao, Boyu Sun, Weizhong Zhou, Qingrong Chen, Weimin Lv, Pengfei Zhang and Le Zhang declare that they have no conflict of interest or financial conflicts to disclose.

All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

OPEN ACCESS

This article is licensed by the CC BY 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

1. Kamal, S. M., Rashid, A. K., Bakar, M. A. and Ahad, M. A. (2011) Anthrax: an update. *Asian Pac. J. Trop. Biomed.*, 1, 496–501
2. Goel, A. K. (2015) Anthrax: A disease of biowarfare and public health importance. *World J. Clin. Cases*, 3, 20–33
3. Hueffer, K., Drown, D., Romanovsky, V. and Hennessy, T. (2020) Factors contributing to anthrax outbreaks in the circumpolar north. *EcoHealth*, 17, 174–180
4. Nderitu, L. M., Gachohi, J., Otieno, F., Mogoa, E. G., Muturi, M., Mwatondo, A., Osoro, E. M., Ngere, I., Munyua, P. M., Oyas, H., *et al.* (2021) Spatial clustering of livestock Anthrax events associated with agro-ecological zones in Kenya, 1957–2017. *BMC Infect. Dis.*, 21, 191
5. Lyu, W., Wei, K., Liu, D., Zhou, W., Wang, W. and Chen, Q. (2019) Spatiotemporal clustering of anthrax in Gannan Tibetan Autonomous Prefecture of Gansu, 2011–2017. *Dis. Surveill.*, 34, 32–36
6. Yu, D., He, J., Zhang, E., Wang, P., Liu, D., Hou, Y., Zhang, H., Wei, K., Gou, F., Zhang, H., *et al.* (2018) Investigation and source-tracing of an anthrax outbreak in Gansu Province, China. *PLoS One*, 13, e0203267
7. Li, Y., Yin, W., Hugh-Jones, M., Wang, L., Mu, D., Ren, X., Zeng, L., Chen, Q., Li, W., Wei, J., *et al.* (2017) Epidemiology of human anthrax in China, 1955–2014. *Emerg. Infect. Dis.*, 23, 14–21
8. Cao, L. T., Liu, H. H., Li, J., Yin, X. D., Duan, Y. and Wang, J. (2020) Relationship of meteorological factors and human brucellosis in Hebei Province, China. *Sci. Total Environ.*, 703, 135491–135498
9. Nili, S., Khanjani, N., Bakhtiari, B., Jahani, Y. and Dalaei, H. (2021) The effect of meteorological variables on salmonellosis incidence in Kermanshah, West of Iran: a generalized linear model with negative binomial approach. *J. Environ. Health Sci. Eng.*, 19, 1171–1177
10. Chen, W.-J., Lai, S.-J., Yang, Y., Liu, K., Li, X.-L., Yao, H.-W., Li, Y., Zhou, H., Wang, L.-P., Mu, D., *et al.* (2016) Mapping the distribution of anthrax in mainland China, 2005–2013. *PLoS Negl. Trop. Dis.*, 10, e0004708
11. Brownlie, T., Bishop, T., Parry, M., Salmon, S. E. and Hunnam, J. C. (2020) Predicting the periodic risk of anthrax in livestock in Victoria, Australia, using meteorological data. *Int. J. Biometeorol.*, 64, 601–610
12. Scott, L. M. and Janikas, M. V. (2010) Spatial statistics in arcgis. In: *Handbook of Applied Spatial Analysis*, pp. 27–41. Springer
13. Carlson, C. J., Kracalik, I. T., Ross, N., Alexander, K. A., Hugh-Jones, M. E., Fegan, M., Elkin, B. T., Epp, T., Shury, T. K., Zhang, W., *et al.* (2019) The global distribution of *Bacillus anthracis* and associated anthrax risk to humans, livestock and wildlife. *Nat. Microbiol.*, 4, 1337–1343
14. R Core Team. (2021) R: A language and environment for

- statistical computing. Available from the website of r-project
15. Granger, C. W. (1969) Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37, 424–438
16. Gannan Tibetan Autonomous Prefecture Statistics Bureau, and the Gannan Investigation Team of National Bureau of Statistics. (2021) Gannan statistical yearbook. Available from the website of statistics bureau in Gannan
17. Wu, X., Lang, L., Ma, W., Song, T., Kang, M., He, J., Zhang, Y., Lu, L., Lin, H. and Ling, L. (2018) Non-linear effects of mean temperature and relative humidity on dengue incidence in Guangzhou, China. *Sci. Total Environ.*, 628–629, 766–771
18. Qi, F., Ding, X. and Lu, S. (2021) National greenhouse data system.
19. World Health Organization. (2021) International statistical classification of diseases and related health problems. Available from the website of World Health Organization
20. Dickey, D. A. and Fuller, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.*, 74, 427–431
21. Dickey, D. A. and Fuller, W. A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49, 1057–1072
22. Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econom.*, 54, 159–178
23. Zhang, L., Fu, C., Li, J., Zhao, Z., Hou, Y., Zhou, W. and Fu, A. (2019) Discovery of a ruthenium complex for the theranosis of glioma through targeting the mitochondrial DNA with bioinformatic methods. *Int. J. Mol. Sci.*, 20, 4643–4658
24. Zhang, L., Zhao, J., Bi, H., Yang, X., Zhang, Z., Su, Y., Li, Z., Zhang, L., Sanderson, B. J., Liu, J., *et al.* (2021) Bioinformatic analysis of chromatin organization and biased expression of duplicated genes between two poplars with a common whole-genome duplication. *Hortic. Res.*, 8, 62
25. Xia, Y., Yang, C., Hu, N., Yang, Z., He, X., Li, T. and Zhang, L. (2017) Exploring the key genes and signaling transduction pathways related to the survival time of glioblastoma multiforme patients by a novel survival analysis model. *BMC Genomics*, 18, 950
26. Li, T., Cheng, Z. and Zhang, L. (2017) Developing a novel parameter estimation method for agent-based model in immune system simulation under the framework of history matching: A case study on influenza a virus infection. *Int. J. Mol. Sci.*, 18, 2592–2603
27. Gao, H., Yin, Z., Cao, Z. and Zhang, L. (2017) Developing an agent-based drug model to investigate the synergistic effects of drug combinations. *Molecules*, 22, 2209–2221
28. Zhang, L., Liu, G., Kong, M., Li, T., Wu, D., Zhou, X., Yang, C., Xia, L., Yang, Z. and Chen, L. (2021) Revealing dynamic regulations and the related key proteins of myeloma-initiating cells by integrating experimental data into a systems biological model. *Bioinformatics*, 37, 1554–1561
29. Zhang, L. and Zhang, S. (2017) Using game theory to investigate the epigenetic control mechanisms of embryo development: Comment on: “Epigenetic game theory: How to compute the epigenetic control of maternal-to-zygotic transition” by Qian Wang *et al.* *Phys. Life Rev.*, 20, 140–142
30. Zhang, L., Qiao, M., Gao, H., Hu, B., Tan, H., Zhou, X. and Li, C. M. (2016) Investigation of mechanism of bone regeneration in a porous biodegradable calcium phosphate (CaP) scaffold by a combination of a multi-scale agent-based model and experimental optimization/validation. *Nanoscale*, 8, 14877–14887
31. Gao, J., Liu, P., Liu, G.-D. and Zhang, L. (2021) Robust needle localization and enhancement algorithm for ultrasound by deep learning and beam steering methods. *J. Comput. Sci. Technol.*, 36, 334–346
32. Zhang, L., Lv, J., Xiao, M., Yang, L. and Zhang, L. (2021) Exploring the underlying mechanism of action of a traditional chinese medicine formula, youdujing ointment, for cervical cancer treatment. *Quant. Biol.*, 9, 292–303
33. Wu, W., Song, L., Yang, Y., Wang, J., Liu, H. and Zhang, L. (2020) Exploring the dynamics and interplay of human papillomavirus and cervical tumorigenesis by integrating biological data into a mathematical model. *BMC Bioinformatics*, 21, 152
34. Lei, W., Zeng, H., Feng, H., Ru, X., Li, Q., Xiao, M., Zheng, H., Chen, Y. and Zhang, L. (2020) Development of an early prediction model for subarachnoid hemorrhage with genetic and signaling pathway analysis. *Front. Genet.*, 11, 391–400
35. Zhang, L., Zheng, C., Li, T., Xing, L., Zeng, H., Li, T., Yang, H., Cao, J., Chen, B. and Zhou, Z. (2017) Building up a robust risk mathematical platform to predict colorectal cancer. *Complexity*, 2017, 8917258
36. Zhang, L., Xiao, M., Zhou, J. and Yu, J. (2018) Lineage-associated underrepresented permutations (LAUPs) of mammalian genomic sequences based on a Jellyfish-based LAUPs analysis application (JBLA). *Bioinformatics*, 34, 3624–3630
37. Zhang, L., Liu, Y., Wang, M., Wu, Z., Li, N., Zhang, J. and Yang, C. (2017) EZH2-, CHD4-, and IDH-linked epigenetic perturbation and its association with survival in glioma patients. *J. Mol. Cell Biol.*, 9, 477–488
38. You, Y., Ru, X., Lei, W., Li, T., Xiao, M., Zheng, H., Chen, Y. and Zhang, L. (2020) Developing the novel bioinformatics algorithms to systematically investigate the connections among survival time, key genes and proteins for Glioblastoma multiforme. *BMC Bioinformatics*, 21, 383
39. Zhang, L., Bai, W., Yuan, N. and Du, Z. (2019) Comprehensively benchmarking applications for detecting copy number variation. *PLOS Comput. Biol.*, 15, e1007069
40. Lopez, L. and Weber, S. (2017) Testing for granger causality in panel data. *Stata J.*, 17, 972–984
41. Anderson, G. B., Bell, M. L. and Peng, R. D. (2013) Methods to calculate the heat index as an exposure metric in environmental health research. *Environ. Health Perspect.*, 121, 1111–1119
42. Yin, Q. and Wang, J. (2018) A better indicator to measure the

- effects of meteorological factors on cardiovascular mortality: heat index. *Environ. Sci. Pollut. Res. Int.*, 25, 22842–22849
43. National Weather SERVICE. (2021) Heat index. Available from the website of National Weather Service
 44. Willmott, C. J. and Matsuura, K. (2005) Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Clim. Res.*, 30, 79–82
 45. Cavanaugh, J. E. and Neath, A. A. (2019) The akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdiscip. Rev. Comput. Stat.*, 11, e1460
 46. Zhang, L., Zhang, L., Guo, Y., Xiao, M., Feng, L., Yang, C., Wang, G. and Ouyang, L. (2021) MCDB: A comprehensive curated mitotic catastrophe database for retrieval, protein sequence alignment, and target prediction. *Acta Pharm. Sin. B*, 11, 3092–3104
 47. Xiao, M., Liu, G., Xie, J., Dai, Z., Wei, Z., Ren, Z., Yu, J. and Zhang, L. (2021) 2019ncovas: Developing the web service for epidemic transmission prediction, genome analysis, and psychological stress assessment for 2019-ncov. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 18, 1250–1261
 48. Xiao, M., Yang, X., Yu, J. and Zhang, L. (2020) CGIDLA: Developing the web server for CpG island related density and laups (lineage-associated underrepresented permutations) study. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 17, 2148–2154
 49. Zhang, L., Dai, Z., Yu, J. and Xiao, M. (2021) CpG-island-based annotation and analysis of human housekeeping genes. *Brief. Bioinform.*, 22, 515–525
 50. Red Hat Enterprise Linux. (2018) Centos linux. Available from the website of The CentOS Project
 51. F5 incorporation. (2021) Nginx 1.18. Available from the website of NGINX
 52. Bootstrap Team. (2021) Bootstrap 4.4.1.
 53. Facebook Inc. (2021) React 17.0.1.