Article

Prediction of Human Brucellosis in China Based on Temperature and NDVI

Yongqing Zhao [1,2,3], Rendong Li [1,*], Juan Qiu [1], Xiangdong Sun [4,*], Lu Gao [4] and Mingquan Wu [5]

1. Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan 430077, China;
   zhaoyongqing@asch.whigg.ac.cn (Y.Z.); qiujuan@asch.whigg.ac.cn (J.Q.)
2. University of Chinese Academy of Sciences, Beijing 100049, China
3. Department of geological mapping engineering, Shanxi Institute of Energy, Yuci 030600, China
4. China Animal Health and Epidemiology Center, Qingdao 266032, China; gaolu@cahec.cn
5. The State Key Laboratory of Remote Sensing Science Institute of Remote Sensing and Digital Earth,
   Chinese Academy of Sciences, Beijing 100101, China; wumq@aircas.ac.cn
* Correspondence: lrd@asch.whigg.ac.cn (R.L.); sunxiangdong@cahec.cn (X.S.)

check for updates

Abstract: Brucellosis occurs periodically and causes great economic and health burdens. Brucellosis prediction plays an important role in its prevention and treatment. This paper establishes relationships between human brucellosis (HB) and land surface temperature (LST), and the normalized difference vegetation index (NDVI). A seasonal autoregressive integrated moving average with exogenous variables (SARIMAX) model is constructed to predict trends in brucellosis rates. The fitted results (Akaike Information Criterion (AIC) = 807.58, Schwarz Bayes Criterion (SBC) = 819.28) showed obvious periodicity and a rate of increase of 138.68% from January 2011 to May 2016. We found a significant effect between HB and NDVI. At the same time, the prediction part showed that the highest monthly incidence per year has a decreasing trend after 2015. This may be because of the brucellosis prevention and control measures taken by the Chinese Government. The proposed model allows the early detection of brucellosis outbreaks, allowing more effective prevention and control.

1. Introduction

Brucellosis is a common zoonosis afflicting humans and animals and causes heavy economic and health burdens around the world [1–3]. There are four types of brucellosis [4,5] affecting sheep, cows, pigs, and dogs, with transmission to humans mainly occurring through close contact with infected animals and their wastes, or by eating unsterilized animal products such as fresh milk [4,6–8]. Aborted sheep fetuses are highly infectious. Farmers, butchers, and veterinarians who come into contact with animals are at high risk of infection. Brucellosis often causes a series of symptoms in infected people [9,10], and it leads to animal production and abortion [11–14]. Brucellosis is mainly distributed in Asia, the Middle East, sub-Saharan Africa and the Balkan Peninsula [1,15,16]. More than 500,000 cases occur globally every year [17]. Sheep brucella vaccine rev.1 is not suitable for human use [18,19]. The United States, New Zealand, and other countries have adopted quarantine or culling measures [20,21] but these measures cannot eradicate brucellosis completely.

As a country with a vast agricultural industry, China's brucellosis prevention and control situation is very serious [13]. The number of brucellosis cases is on the rise and the affected region is gradually spreading from the north to the south. The north is a high-incidence area [22], especially in Inner Mongolia. Meanwhile, the south has sporadic outbreaks [23], especially in recent years in Hainan Province, which was previously free of the disease. The first case of brucellosis occurred in China in

1905 and high rates of incidence occurred before 1980 that have since improved [13,24]. This is related to a series of measures taken by the Chinese Government: (1) A reporting system was established in 1950 [13]; (2) vaccination began in 1964 in high-risk areas; (3) a definition of the disease, test standards, control measures, and treatment programs were formulated in 1977; and (4) in 1980, a monitoring project of serum-positive rates at brucellosis sentinel sites was implemented. However, from the middle to late 1990s to the beginning of this century, the incidence of brucellosis increased sharply and has continued to increase in recent years, which presents great challenges to public health and social economic systems. Therefore, new theories and forecasting methods are urgently needed to meet these challenges.

Previous publications have focused on the epidemiological characteristics and spatial distribution of the disease [15,25–27], while time series-based predictive models have been used to provide early warnings of disease occurrence. Common models include residual autoregression models, the exponential smoothing method (ES) [28], grayscale models, negative binomial regression models, artificial neural network models [26], autoregressive integrated moving average models (ARIMAs) [29,30], and their syntheses [31]. ARIMA models are optimal when the time series contains long-term trend, periodicity, and disturbance terms [32–35]. Previous ARIMA models only considered the model itself without considering the influence of other variables. An improved ARIMA model, the autoregressive integrated moving average with exogenous variables (ARIMAX) model, overcomes this defect by incorporating other factors. Some epidemics involve factors such as hydatids [36] and *Oncomelania* snails. These factors have been modelled in other papers using factors such as temperature, precipitation, insolation time, relative humidity, average wind speed, vegetation cover, etc. [24]. Remote sensing has been applied in many fields as it can estimate many of these factors [16,37–43]. Although ARIMAX has been used to study air pollution [44], corn yields [45], and to predict various illnesses [46], no one has used it to study brucellosis in recent years. In addition, the use of ARIMAX models helps to explore the relationship between time series and other factors. Therefore, this study aims to fill this research gap. Brucellosis bacteria are particularly sensitive to temperature [26], especially surface temperature. In addition, vegetation cover can provide suitable hydrothermal conditions for bacterial survival. In this paper, on the basis of predecessors' ARIMA models that consider seasonal factors, we use remote sensing data on land surface temperature and vegetation cover to construct an ARIMAX model of the short-term onset of brucellosis.

The aims of this paper are to: (1) Establish an ARIMAX model using surface temperature and vegetation cover as input variables; and (2) use it to predict the incidence of brucellosis in mainland China.

2. Materials and Methods

2.1. Data Sources

2.1.1. Human Brucellosis Data

Data on human brucellosis (HB) cases were collected on a monthly basis by the Chinese Centre for Disease Control and Prevention (CDC) [47] from January 2011 to December 2016. Brucellosis is a class b infectious disease, as stipulated by the law of prevention and treatment of infectious diseases in China. Cases of brucellosis must be reported within 24 h and suspected cases are subject to clinical and serological testing or isolation as required by the World Health Organization. In this study, monthly means of HB cases across China from January 2011 to May 2016 were used in the model to fit the ARIMAX model, while the data from June 2016 to May 2017 were used for validation.

2.1.2. Land Surface Temperature Data

Land surface temperature (LST) has an important impact on surface-dwelling organisms and is an important variable for studying biological change. In the present study, LST data were obtained

for January 2011 to May 2016, and monthly means were used as input for the ARIMAX model of brucellosis. The data was provided by the International Scientific Data Mirror website [48] of the Computer Network Information Centre of the Chinese Academy of Sciences. We used the MODLT1T China 1 km surface temperature monthly composite product. Then, ArcGIS10.4 software (Environment Systems Research Institute, Redlands, CA, USA) was used to obtain monthly mean daytime and night-time surface temperatures across China (the monthly product includes daily and evening monthly average images), and a monthly mean national temperature was calculated.

2.1.3. Normalized Difference Vegetation Index Data

The normalized difference vegetation index (NDVI) is an important indicator of land cover. It is provided by the International Scientific Data Mirror website [48] of the Computer Network Information Centre of the Chinese Academy of Sciences. The MODND1M China 500 m NDVI monthly synthesis product from January 2011 to May 2016 was adopted, and ArcGIS software was used to calculate national mean monthly NDVI values.

2.2. Methods

2.2.1. General Concepts

The steps used to build the overall model are as follows. The input HB, LST and NDVI data are differentiated to stabilize them. Independent variables of LST and NDVI are fitted to the ARIMA model. Each independent variable model is used to filter the HB data and determine a fitted model, as shown in Figure 1. The detailed data processing procedure is shown in the following section.
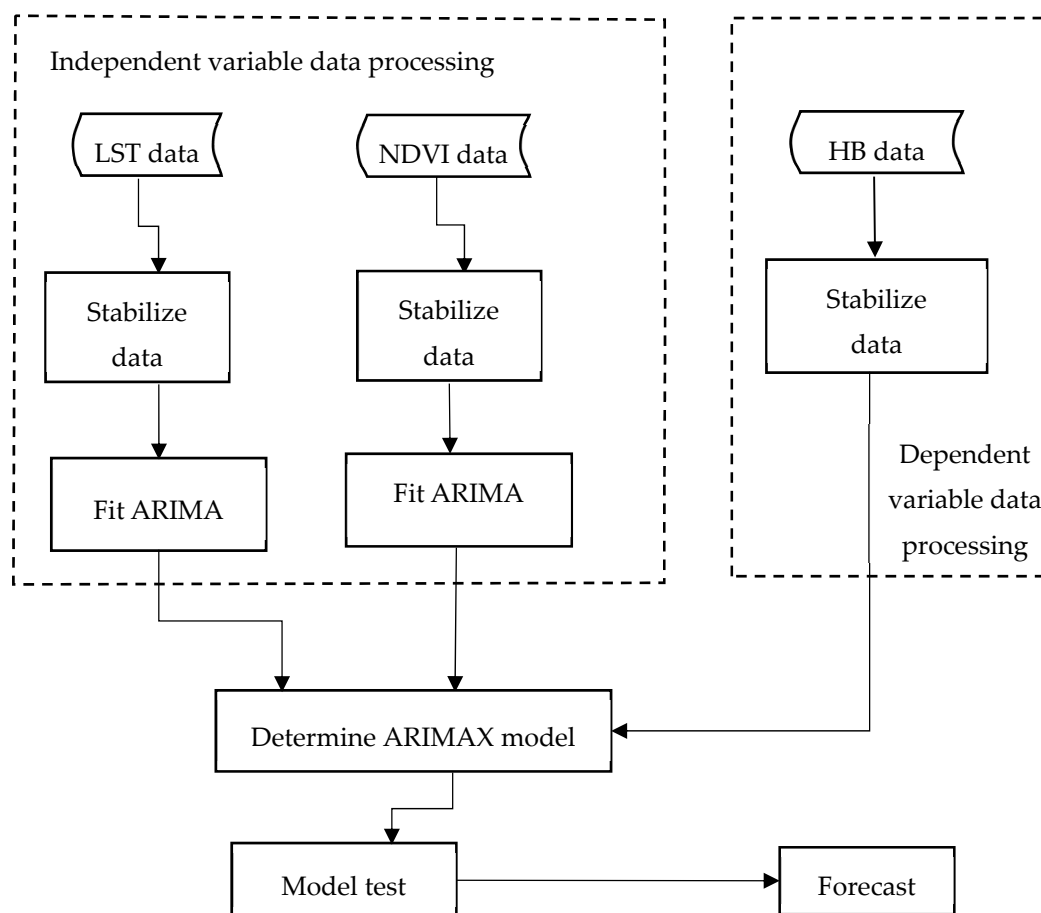

Figure 1. Data processing and modelling flowchart. LST represents land surface temperature, NDVI expresses normalized difference vegetation index, HB indicates human brucellosis.

2.2.2. Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX) Models

As a classical time series model [29], ARIMA (p, d, q) is a further development of the AR (p) and MA (q) models, where p is the number of AR autoregression terms, d is the difference order, and q is the number of MA sliding average terms. If the information of the sequence cannot be fully extracted when the time series is stable, an $SARIMA(p,d,q)(P,D,Q)_S$ model [49] needs to be considered, where P refers to the maximum lag order of the seasonal autoregression term, Q is the maximum lag order of the moving average operator, D is the seasonal difference order, and S is the seasonal difference cycle step. When we consider input variables such as LST, NDVI, we need to use an ARIMA model with input variables, namely, an ARIMAX model. A SARIMAX model can also be considered to be an ARIMA model with an interference sequence or dynamic regression model, which contains a regression section with input variables and an ARIMA section. Its general model is shown as formula (1):

$$
\begin{cases}
y_t = \mu + \sum\limits_{i=1}^{k} \beta_i B^{li} x_{it} + \varepsilon_t \\
\varepsilon_t = \dfrac{\theta_q(B)\Theta_Q\left(B^S\right)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\, a_t
\end{cases}
\tag{1}
$$

Here, μ represents a constant, β_i is regression coefficients of external variables, $\Phi_P(B)$ refers to the seasonal autoregressive coefficient polynomial with P-order, $\Theta_Q(B)$ denotes the polynomial of the seasonal moving average coefficient with Q-order, $\phi_p(B)$ stands for the autoregressive coefficient polynomial with p-order, $\theta_q(B)$ is the polynomial of the seasonal moving average coefficient with q-order, $(1 - B)^d$ indicates differencing operator of order d, $(1 - B^S)^D$ remarks seasonal differencing operator of order D, S expresses the seasonal length, ε_t shows the residuals, a_t is white noise time series, $\mu + \sum\limits_{i=1}^{k} \beta_i B^{li} x_{it}$ suggests the regression part, and ε_t indicates the disturbance part.

When setting up an ARIMAX model, the steps are as follows: (1) A stationarity test is carried out for the response sequence and each input variable. If the input data is not stationary, it should be stabilized by a differential method. No matter which difference [50] method is adopted, it cannot meet the requirement of stationarity and the seasonal variation is obvious. Therefore, a seasonal difference can be considered. Using a seasonal differential method, the sequence was determined to be in a steady state. Since the surface temperature data and NDVI (Figure 2) are themselves in the stationary condition, the next step can be taken directly. Similarly, HB data requires both first-order difference and seasonal difference processing to be a stationary series. The method used to check whether the data is stationary is the Augmented Dickey-Fuller (ADF) test [51], with a result of $p < 0.05$ being indicative of stability. (2) White noise testing. A stationary time series is of no value if it is simply white noise. On the condition that it is not white noise, further study is necessary, then further pattern recognition. All the above three input variables are non-white-noise sequences. (3) ARIMA and prewhitening. The ARIMA model is fitted to each input variable—LST, NDVI—to make the residual of the model a white noise sequence. Then, the output variable brucellosis sequence is filtered by this model to calculate the correlation coefficient between the input sequence and the output sequence. (4) The structure of the SARIMAX model is determined by using the co-correlativity coefficient, and then the model is fitted. (5) Checking the fitted model. If the requirements are not met, adjustment continues until the white noise test and the significance of the parameters meet the required conditions.

Int. J. Environ. Res. Public Health **2019**, *16*, 4289

5 of 15
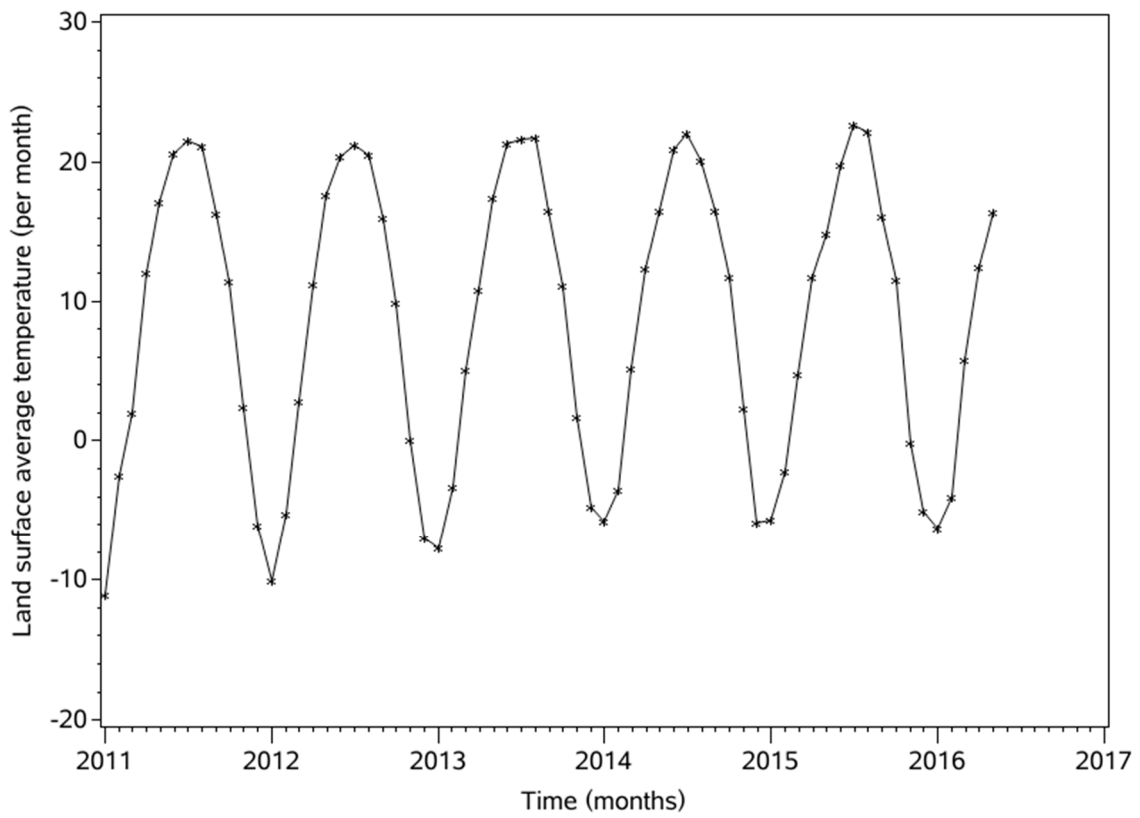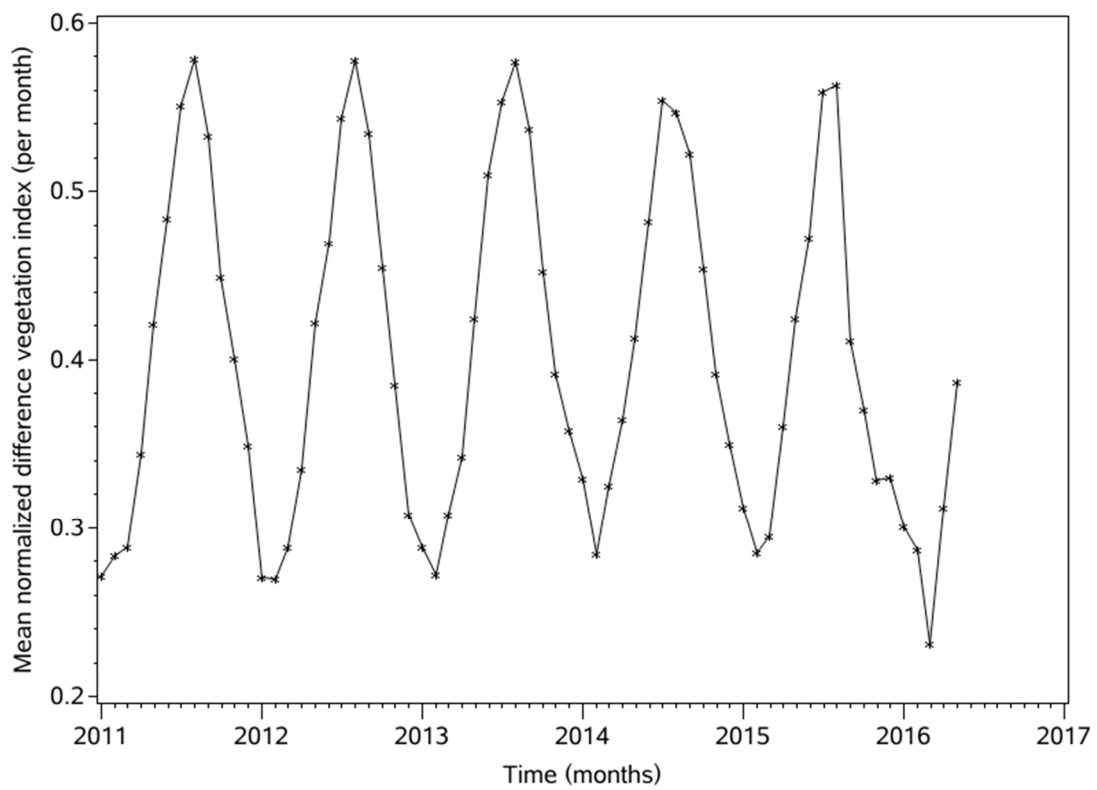
(**a**) LST trend

(**b**) NDVI trend

Figure 2. Trends of input variables, (**a**) LST trend, (**b**) NDVI trend.

2.2.3. Model Test

The root mean square error (RMSE), mean absolute percentage error (MAPE), and mean error rate (MER) are used to evaluate the fit and predictive effect of the established model. These are calculated according to Equations (2)–(4).

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \widehat{X_i})^2} \tag{2}$$

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{n}\frac{\left|X_i - \widehat{X_i}\right|}{X_i} \tag{3}$$

$$\text{MER} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|X_i - \widehat{X_i}\right|}{\overline{X_i}} \tag{4}$$

In the formulas, X_i denotes actual values, $\overline{X_i}$ is the mean of the actual values, $\widehat{X_i}$ is the fitted or predicted values, and N refers to the number of simulated or predicted values. The software used in this study was SAS 9.4 (SAS Institute Inc., Cary, NC, USA), ARCGIS 10.4 and EXCEL 2016 (Software, Redmond, CA, USA). The LST and NDVI data were read by ARCGIS 10.4, and the BDI data were imported into EXCEL 2016. Data fitting, model establishment, inspection, and mapping were all carried out in SAS 9.4.

3. Results

3.1. General Characteristics

A total of 71 observed values of brucellosis data were obtained from January 2011 to November 2016, among which the first 65 were used as training data for the establishment of an ARIMAX model, and the last six values were used for comparison to test the predictive effect of the model. Looking at the data from previous years, the monthly data ranged from a minimum of 1123 cases per month (January 2012) to a maximum of 8102 cases per month (July 2014). The total number of cases in more than five years was 274,800, the monthly average was 4227 cases per month (HB/mo), and the standard error was 1875 HB/mo. Between 2011 and 2015, HB cases increased from 43,827 to 60,782, an increase of 138.68%. According to the data, the incidence of brucellosis has obvious seasonality and periodicity, with a peak period of May to August, as shown in Figure 3.
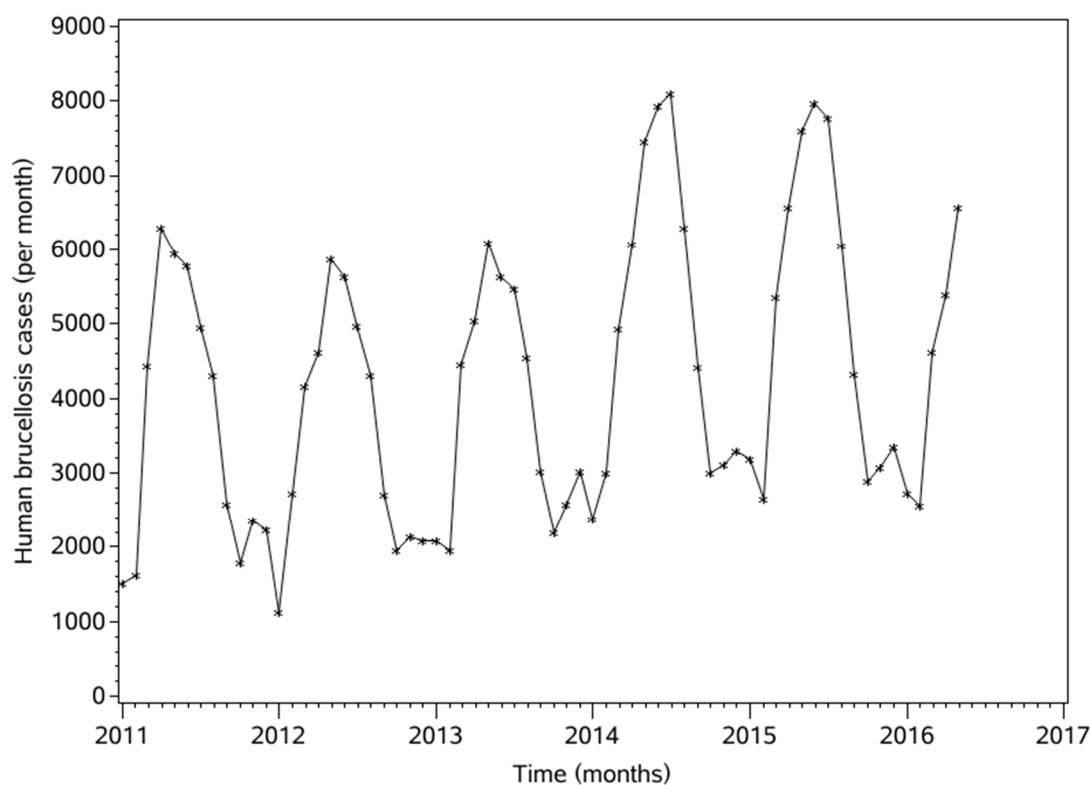
Int. J. Environ. Res. Public Health **2019**, *16*, 4289

7 of 15

Figure 3. Trends in monthly human brucellosis cases.

3.2. HB Data Stationarity

As a response sequence, HB data is required to be stationary for modelling purposes. It can be seen from the scatter diagram that it is non-stationary and has obvious seasonality. After the difference method is adopted, it is found that the data is first order twelve-step differential stability. The test results are shown in Table 1.

Table 1. Augmented Dickey-Fuller Unit Root Tests.

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	−69.6921	<0.0001	−11.03	<0.0001		
	1	−110.947	0.0001	−7.51	<0.0001		
	2	−90.4384	<0.0001	−5.33	<0.0001		
	3	−76.5233	<0.0001	−3.75	0.0004		
Single Mean	0	−69.7764	0.0004	−10.96	0.0001	60.15	0.001
	1	−111.796	0.0001	−7.46	0.0001	27.81	0.001
	2	−90.5227	0.0004	−5.26	0.0001	13.91	0.001
	3	−77.2761	0.0004	−3.71	0.0067	6.89	0.0042
Trend	0	−70.1527	<0.0001	−10.87	<0.0001	59.35	0.001
	1	−116.393	0.0001	−7.56	<0.0001	28.57	0.001
	2	−107.518	0.0001	−5.74	0.0001	17.26	0.001
	3	−112.044	0.0001	−4.11	0.0113	8.59	0.0126

3.3. LST and NDVI Models

A scatter diagram of surface temperature and NDVI data from January 2011 to May 2016 was made and it showed that the data were approximately stable. The ADF test result of $p < 0.05$ showed that the data were stable, so the difference method was not needed. Meanwhile, the white noise test showed a non-white noise sequence. According to the ACF diagram, PACF diagram, and AIC

Int. J. Environ. Res. Public Health **2019**, *16*, 4289

8 of 15

criterion, the optimal temperature model is ARMA (2, 1), whose corresponding AIC and SBC values are 340.43 and 349.13, respectively. All parameters of this model are significant, and the residual error is white noise. The NDVI ADF test shows that the NDVI model is stable and the residuals are white noise. Also, with the aid of the ACF and PACF diagrams, the orders of models were determined. Finally, the selected models are shown in Table 2; the optimization model is ARMA (2, 1). The results for LST and NDVI are shown in Figure 2.

Table 2. LST and NDVI models.

Variable	Model	AIC Value	SBC Value	Optimization Model
LST	ARMA (2, 1)	340.43	349.13	ARMA (2, 1)
NDVI	ARMA (2, 0)	237.51	230.99	ARMA (2,1)
	ARMA (2, 1)	253.63	244.94	
	ARMA (3, 2)	249.83	236.79	

Note: AIC= Akaike Information Criterion.

3.4. SARIMAX Model

The optimal brucellosis model was filtered by the optimal models of LST and NDVI, then a regression equation was determined, with the model fitted according to the correlation number. If the parameters of the fitted model are not significant or the residuals do not pass the white noise test, the parameters need to be adjusted constantly. Finally, the model with significant parameters (Table 3) that passed the white noise test (Table 4) after multiple adjustments is shown as Formula (5), and its corresponding AIC and SBC values are 807.58 and 819.28, respectively. This model is used for prediction, with the predicted results shown in Table 5 and in Figure 4. After establishing the final model, RMSE, MAPE and MER are adopted to evaluate the ARIMAX model established above.

$$(1 - B)(1 - B^{12})y = (-36.80113 + 25.29822B)x_1 + (5000.2 - 4680.1B)x_2 + (1 - 0.53724B)(1 - 0.57166B^{12})\varepsilon_0 \quad (5)$$

Table 3. Parameter significance test for HB by conditional least squares estimation.

Parameter	Estimate	Standard Error	*t*-Value	Approx Pr > |t|	Lag	Variable	Shift
MA1,1	0.53724	0.12945	4.15	0.0001	1	HB	0
MA2,1	0.57166	0.14307	4.00	0.0002	12	HB	0
NUM1	−36.80113	16.58224	−2.22	0.0314	0	LST	0
NUM1,1	−25.29822	12.04109	−2.10	0.0412	1	LST	0
NUM2	5000.2	1720.1	2.91	0.0056	0	NDVI	0
NUM1,1	4680.1	1571.9	2.98	0.0046	1	NDVI	0

Table 4. Autocorrelations with lags.

Lag	χ^2	*df*	Pr > χ^2	Autocorrelations					
6 m	1.12	4	0.8908	0.052	−0.118	0.022	0.047	−0.009	−0.020
12 m	3.26	10	0.9746	−0.036	0.063	−0.099	−0.058	0.117	−0.007
18 m	7.37	16	0.9654	−0.081	−0.015	0.038	0.153	0.144	0.019
24 m	10.26	22	0.9837	−0.079	0.002	−0.051	−0.097	−0.092	0.066

Note: *df* = degree of freedom.

Table 5. Brucellosis prediction results.

Time	Forecast	95% Confidence Limits	
June 2016	7270.13	6170.46	8372.49
July 2016	7512.87	6285.99	8820.31
August 2016	6194.6	4859.57	7645.41
September 2016	4070.9	2598.44	5576.09
October 2016	2621.38	1086.38	4242.86
November 2016	2712.06	1017.16	4374.35
December 2016	2948.77	1096.85	4673.4
January 2017	2285.18	382.38	4166.83
February 2017	1988.04	33.07	3989.59
March 2017	3884.78	1925.2	6020.5
April 2017	5013.8	2947.99	7172.93
May 2017	6347.48	4193	8560.69
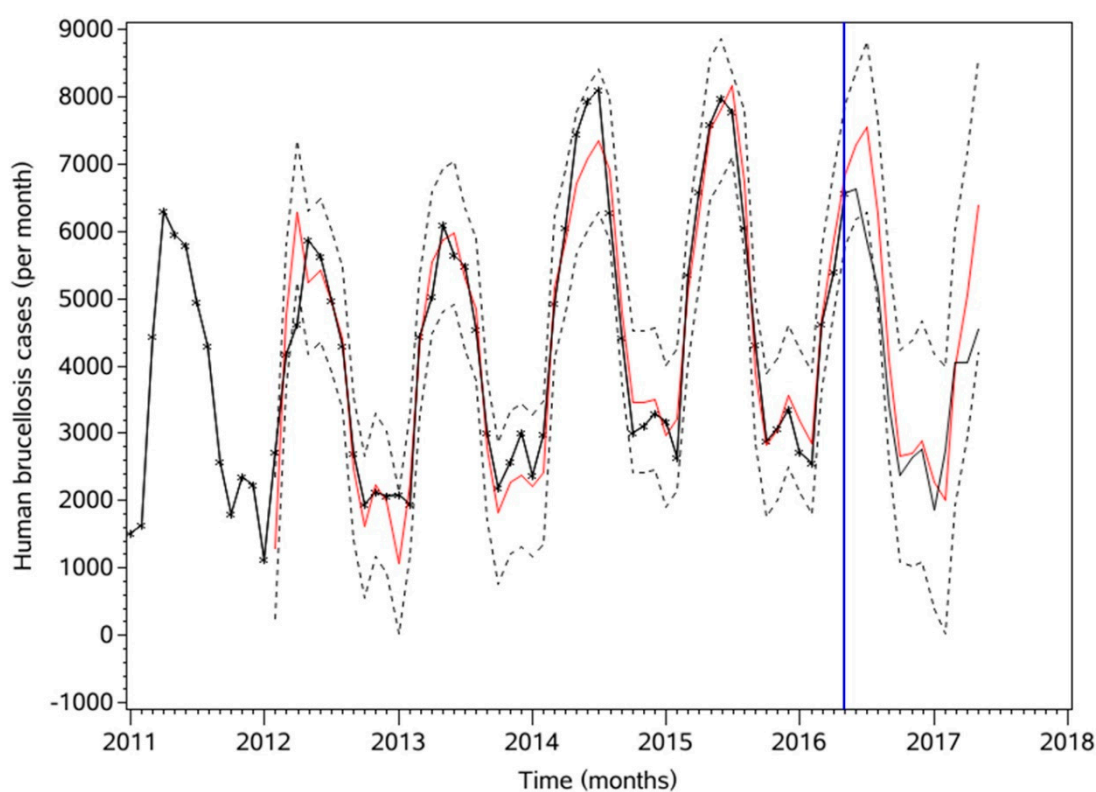
Figure 4. Brucellosis fitting and prediction (January 2011–May 2017). The black line represents actual observations. The red line to the left of the blue vertical line represents fitted values, while that to the right of the blue line represents predicted values. The dotted black lines represent 95% confidence intervals.

In the formula, y represents the incidence of HB, x_1, x_2 represent the input variables LST, NDVI, respectively, B is a lag factor, and ε_0 is the white noise residual sequence.

3.5. Assessing Model Performance

For the final evaluation result, RMSE, MAPE, MER [29], and other common evaluation methods are adopted to evaluate the difference between fitted and predicted values. For the fitted values, the RMSE = 506.896, MAPE = 0.104, and MER = 0.089. For the predicted values, RMSE = 921.215, MAPE = 0.179, and MER = 0.188. The predicted results are shown in Table 5.

4. Discussion

4.1. Trends and Seasonal Periodicity

As can be seen from Figure 1, the annual incidence of brucellosis has been increasing since 2011, with the annual peaks increasing. One important cause is the increasing scale of livestock husbandry caused by improvements in China's economy and demands for livestock products in recent years [52]. At the same time, it may be related to the free-range breeding methods used by farmers, which have much higher animal disease risks than conventional methods. In this case, susceptible animals can come into contact with infected animals or brucellosis bacteria. Sheep are the main source of infection in China, followed by cattle and pigs. The peak timing of brucellosis rates is also closely related to the lamb production season. The LST, NDVI, and HB data have similar trends, which indicates that there is a close relationship between them. See Figure 3 for more details.

Brucellosis shows obvious seasonal periodicity. The incidence of brucellosis is relatively high from April to August every year, with the peak occurring from May to July. This may be related to many factors [36], such as precipitation, lamb production, animal movement, wind speed, and sunshine levels. A more plausible explanation is that as temperatures rise in the spring, snow and ice melt, amounts of organic matter increase, and precipitation increases due to weather changes, all of which are conducive to the activation and incubation of bacteria. Spring is also the breeding season for many animals. With more animals moving around, the area affected by infected animals and their excretions becomes larger and the risk of infection of sensitive animals is higher. The net result may be a growing number of brucellosis cases that are distributed more widely across the country.

It can be seen from Figure 3 that from January 2011 to May 2016, parts of the fitted results are almost identical to the actual values. Only in July 2014 do the fitted and actual values have certain differences; the month where the abnormal values reached 8102, and which attained a maximum from 2011 to 2016. This is because the results of the ARIMAX simulation are continuous and smooth, so extreme values or outliers will make it relatively flat. From January 2011 to May 2016, the annual maximum values showed an increasing trend, however, the ARIMAX model could accurately predict from this increasing trend that the maximum monthly incidence decreased in 2016. It was less than in 2014 and 2015. This is consistent with actual observations, indicating that this model predicts trends well. In addition, since 2015, the rate of brucellosis cases has shown a downward trend. One important reason is that the measures taken by the Chinese Government to control brucellosis more strictly have been effective.

4.2. The Interrelations of Variables and Their Significance

Since LST and NDVI are two environmental factors that have important influence on HB, we chose these two factors mainly to explore their influence on HB. Although collinearity exists between LST and NDVI, collinearity does not affect the prediction accuracy of ARIMAX model. So, we do not deal with it. As shown in Figure 4, the fitted curve is basically consistent with that of NDVI. The coefficient of NDVI is the largest, indicating that its change does indeed indicate the basic trend of brucellosis onset and has the greatest influence. This is because vegetation coverage increases from January to July each year. This provides live *Brucella* bacteria with favorable hydrothermal equilibrium conditions. With increasing areas of favorable conditions, the number of *Brucella* hosts, such as sheep that feed on infected vegetation, increases. The highest incidence of brucellosis occurs in Inner Mongolia [23] in northern China, which is closely related to its widespread grasslands. This is the best example of the influence of vegetation on *Brucella*. Because the time at which spring ewes lamb is appropriate for *Brucella*, contact with amniotic fluid that carries *Brucella* bacteria can cause infection in humans. However, with increases in temperature, water evaporation increases and large numbers of bacteria multiply rapidly. The moisture content of the soil decreases because of water evaporation. At the same time, brucella has very high nutritional demands, while a large number of bacteria are competing for nutrients from brucellosis bacteria, so *Brucella* bacteria receive less nutrition. These conditions do

not facilitate the growth of bacteria, causing brucellosis cases to decrease. Hence, there is a negative correlation between LST and brucellosis cases. In summary, HB incidence is primarily determined by NDVI, while LST influence the amplitude of its change. Previous studies have considered temperature, moisture, and other environments in which bacteria infect hosts; however, it is better to also consider the effect of vegetation on temperature and moisture.

High vegetation coverage is good for *Brucella* survival and spread, which is a kind of average result. Some high vegetation coverage and wide distribution is very good for *Brucella* bacteria survival, while some other coverages are not very good. The level of NDVI that causes the highest HB incidence requires further study. The Inner Mongolia prairie [23] is very favorable for *Brucella* survival and the highest HB incidence in China is found there. In some southern areas, the vegetation coverage is higher but is not necessarily suitable for *Brucella* survival. From the local area to the national level, no vegetation types were significantly adverse to brucellosis. Of course, some vegetation may play a greater role in *Brucella* survival, while others may play a smaller role.

The results of this study improve on previous research results [29]. NDVI, LST were found to be of certain significance in the prediction, prevention and control of brucellosis. The consistency of the NDVI and brucellosis incidence trends show that NDVI is of important significance in formulating accurate prevention and control strategies, as it can be used to predict the incidence of brucellosis in a country or a region according to its specific vegetation. The lagged effects of LST show that we need to allocate medical resources, personnel, and animal control and isolation measures according to changes in LST. According to these conditions, brucellosis prevention and control measures can be planned in advance, medical resources can be allocated appropriately, and public awareness can be carried out. The two variables above may have greater significance for certain regional infections due to their relatively easy availability, extensive coverage and close correlations with various surface-dwelling bacteria.

This model is established throughout the whole country and has certain applicability to high-risk areas. For high-risk areas, the idea and method of building the model are the same, but the coefficients of the model may be different. As long as we get LST and NDVI for these regions, we can build a similar model. In areas with high vegetation coverage (such as grasslands) and dense flocks of sheep, the incidence is high, such as in the grasslands of Inner Mongolia (as can be seen from Equation (5)), which is a typical high-risk area. A higher NDVI coefficient may be obtained from the model if these regions are selected, because the temperature in the north is lower and the influence of LST is greater. That is, the absolute value of the coefficient is smaller (because it is negative) and lower temperatures are more conducive to the survival of brucellosis. However, in some areas in the south, the number of sheep is relatively small and the area of vegetation cover (such as grassland) is smaller, so the incidence in these areas is relatively low. The NDVI coefficient of the model presented here is relatively small. Meanwhile, due to the high temperature in the south, LST will increase while its influence on the equation will decrease. In other words, its coefficient will decrease (the absolute value will increase).

4.3. Other Factors

In the analysis above, only significant factors were selected for research. Choosing the main influencing factors can help to grasp the essence of the problem and highlight its regularity so as to solve it. In addition to the above factors, there may be other factors that affect brucellosis incidence, such as relevant policies and animal numbers. In general, strict control policies may decrease the incidence of brucellosis while lax ones may increase it. From the 1970s, the incidence of brucellosis declined steadily; however, after the 1990s, the number of brucellosis cases increased continuously. After 2015, brucellosis began to decline again. These changes were related to the policies of the Chinese Government. Before the 1990s, China's brucellosis control was strict and brucellosis incidence steadily declined. After the 1990s, due to the relaxation of awareness, vigilance and policies related to brucellosis, the number of brucellosis cases increased, and the Government attached greater importance to it. After 2013, the Government gradually formulated and adopted stronger policies that had an effect, as suggested by the decline in brucellosis since 2015.

From the perspective of animal numbers, after the 1990s, the number of livestock increased significantly. This was coupled with a restructuring of the livestock industry; the proportion of pork production decreased, while that of beef and mutton increased. According to official annual data released by the National Bureau of Statistics, the numbers of sheep raised from 2011 to 2015 were 28,664.15 (×10,000 head), 28,512.74, 28,935.22, 30,391.28, 31,174.27, and 29,930.54, respectively; an overall increasing trend. In the same period, the numbers of cattle were 9384.00 (×10,000 head), 9137.25, 8985.76, 9007.28, 9055.79, and 8834.49, respectively; an overall decrease, and contrary to the trend in brucellosis incidence. The incidence of brucellosis in sheep greatly increased along with the large increase in sheep numbers, while the incidence of brucellosis in humans remained unchanged. At the same time, the incidence of brucellosis in sheep greatly increased due to its wide distribution range and strong pathogenicity, so the chance of contracting brucellosis in humans would also have been greatly increased. This indicates that sheep and brucellosis are closely related. In 2016, the number of sheep decreased and brucellosis incidence declined. At the same time, the timing of human brucellosis incidence coincided with the timing of sheep reproduction. From 2011 to 2016, cattle numbers showed a downward trend, which was opposite to the increase in human brucellosis, indicating that brucellosis in cattle did not affect trends in human brucellosis on the whole, indicating that the influence of cattle brucellosis on human brucellosis is relatively weak. The reason may be that cattle *Brucella* are less pathogenic and only occur locally. This suggests that the effect of sheep brucellosis on human brucellosis is much greater than that of cattle brucellosis.

On the one hand, *Brucella* is dependent on the spread of environmental factors, etc. High vegetation cover provides a suitable living environment for *Brucella*. People and animals are easily infected with *Brucella* after coming into contact with these infected substances, thus greatly increasing the risk of human infection. On the other hand, exposure is also an important factor. Farmers who keep sheep on their farms may also be at increased risk of contracting brucellosis due to frequent contact with sheep. Similarly, butchers who often come into contact with live animals such as sheep during slaughter are also susceptible to infection. Therefore, the incidence of the disease is higher in the northern areas where sheep are raised intensively. Because farm households are more exposed to animals such as sheep during lambing season, this leads to a higher incidence of brucellosis, which shows a distinct seasonality.

4.4. Limitations

There are some shortcomings to the present study. First, the effect of precipitation on brucellosis was not considered. We know that precipitation and temperature are two important factors affecting brucellosis disease. However, since we mainly studied the problem based on remote sensing data, we did not consider precipitation, which should be done in future. In addition, since the LST and NDVI data were only available from January 2011 to May 2016, further data accumulation may be able to further improve the models.

5. Conclusions

In recent years, brucellosis rates have stayed at high levels, and the disease is involved in almost all provinces of China. It is seriously dangerous to human health and causes huge economic losses to the country; therefore, its control is of great significance. This study established relationships between HB incidence and several influencing factors. Using remote sensing data, we established the specific relationships between HB and LST and NDVI using an ARIMAX model. NDVI has a positive correlation with HB and follows its basic trends, while LST are negatively correlated with HB and only affect its magnitude. Although brucellosis has been on the decline since 2015 because of the prevention and control efforts of the Chinese Government, the disease is still considered serious due to its continued high incidence. Accordingly, we propose the following suggestions: (1) In the process of risk control and prevention, vegetation coverage should be regarded as an important reference index, especially in grasslands with high vegetation coverage. (2) Due to the higher human incidence in the

lambing season, management of livestock practices should be strengthened at this time to prohibit the discharge of amniotic fluid and associated pollutants. (3) Since the low temperature of LST is conducive to the survival of brucellosis, during low spring temperatures, pathogens are spread in water due to snow and ice melts, resulting in greatly increased pathogen spreading and risk. Therefore, attention should be paid to strengthening the prevention and control of brucellosis at this time. (4) Since brucellosis has appeared repeatedly in China, prevention and treatment measures should be constantly reviewed to prevent decreases in the awareness of the problem. Our study has established a specific relationship between HB and several of its influences, which is of certain significance for the prevention and control of brucellosis in China. It also provides a reference for similar research in other countries.

References

1. Pappas, G.; Papadimitriou, P.; Akritidis, N.; Christou, L.; Tsianos, E.V. The new global map of human brucellosis. *Lancet Infect. Dis.* **2006**, *6*, 91–99. [CrossRef]

2. Zhong, Z.; Yu, S.; Wang, X.; Dong, S.; Xu, J.; Wang, Y.; Chen, Z.; Ren, Z.; Peng, G. Human brucellosis in the People's Republic of China during 2005–2010. *Int. J. Infect. Dis.* **2013**, *17*, e289–e292. [CrossRef] [PubMed]

3. Asiimwe, B.B.; Kansiime, C.; Rwego, I.B. Risk factors for human brucellosis in agro-pastoralist communities of south western Uganda: A case-control study. *BMC Res. Notes* **2015**, *8*, 405. [CrossRef] [PubMed]

4. Abdullayev, R.; Kracalik, I.; Ismayilova, R.; Ustun, N.; Talibzade, A.; Blackburn, J.K. Analyzing the spatial and temporal distribution of human brucellosis in Azerbaijan (1995–2009) using spatial and spatio-temporal statistics. *BMC Infect. Dis.* **2012**, *12*, 185. [CrossRef] [PubMed]

5. Zinsstag, J.; Roth, F.; Orkhon, D.; Chimed-Ochir, G.; Nansalmaa, M.; Kolář, J.; Vounatsou, P. A model of animal–human brucellosis transmission in Mongolia. *Prev. Vet. Med.* **2005**, *69*, 77–95. [CrossRef] [PubMed]

6. De Massis, F.; Di Girolamo, A.; Petrini, A.; Pizzigallo, E.; Giovannini, A. Correlation between animal and human brucellosis in Italy during the period 1997–2002. *Clin. Microbiol. Infect.* **2005**, *11*, 632–636. [CrossRef]

7. Marianelli, C.; Graziani, C.; Santangelo, C.; Xibilia, M.T.; Imbriani, A.; Amato, R.; Neri, D.; Cuccia, M.; Rinnone, S.; Di Marco, V.; et al. Molecular Epidemiological and Antibiotic Susceptibility Characterization of Brucella Isolates from Humans in Sicily, Italy. *J. Clin. Microbiol.* **2007**, *45*, 2923–2928. [CrossRef]

8. Memish, Z.A.; Balkhy, H.H. Brucellosis and International Travel. *J. Travel Med.* **2004**, *11*, 49–55. [CrossRef]

9. Punda-Polić, V.; Cvetnić, Ž. Human brucellosis in Croatia. *Lancet Infect. Dis.* **2006**, *9*, 540–541. [CrossRef]

10. Bossi, P.; Tegnell, A.; Baka, A.; Van Loock, F.; Hendriks, J.; Werner, A.; Maidhof, H.; Gouvras, G. Bichat guidelines for the clinical management of brucellosis and bioterrorism-related brucellosis. *Eurosurveillance* **2004**, *9*, 33–34. [CrossRef]

11. Corbel, M.J. Brucellosis: An Overview. *Emerg. Infect. Dis.* **1997**, *3*, 213–221. [CrossRef] [PubMed]

12. Earhart, K.; Vafakolov, S.; Yarmohamedova, N.; Michael, A.; Tjaden, J.; Soliman, A. Risk factors for brucellosis in Samarqand Oblast, Uzbekistan. *Int. J. Infect. Dis.* **2009**, *13*, 749–753. [CrossRef] [PubMed]

13. Shang, D.Q.; Xiao, D.L.; Yin, J.M. Epidemiology and control of brucellosis in China. *Vet. Microbiol.* **2002**, *90*, 165–182.

14. Sohn, A.H.; Probert, W.S.; Glaser, C.A.; Gupta, N.; Bollen, A.W.; Wong, J.D.; Grace, E.M.; McDonald, W.C. Human Neurobrucellosis with Intracerebral Granuloma Caused by a Marine Mammal Brucella spp. *Emerg. Infect. Dis.* **2003**, *9*, 485–488. [CrossRef]

15. Jia, P.; Joyner, A. Human brucellosis occurrences in inner mongolia, China: A spatio-temporal distribution and ecological niche modeling approach. *BMC Infect. Dis.* **2015**, *15*, 36. [CrossRef]

16. Mufinda, F.C.; Boinas, F.; Nunes, C. Prevalence and factors associated with human brucellosis in livestock professionals. *Rev. Saúde Pública* **2017**, *51*, 57. [CrossRef]

17. Sanjuan-Jimenez, R.; Colmenero, J.D.; Morata, P. Lessons learned with molecular methods targeting the BCSP-31 membrane protein for diagnosis of human brucellosis. *Clin. Chim. Acta* **2017**, *469*, 1–9. [CrossRef]

18. Jain-Gupta, N.; Contreras-Rodriguez, A.; Vemulapalli, R.; Witonsky, S.G.; Boyle, S.M.; Sriranganathan, N. Pluronic P85 enhances the efficacy of outer membrane vesicles as a subunit vaccine against Brucella melitensis challenge in mice. *FEMS Immunol. Med. Microbiol.* **2012**, *66*, 436–444. [CrossRef]

19. Banai, M. Control of small ruminant brucellosis by use of Brucella melitensis Rev.1 vaccine: Laboratory aspects and field observations. *Vet. Microbiol.* **2002**, *90*, 497–519. [CrossRef]

20. Godfroid, J.; Käsbohrer, A. Brucellosis in the European Union and Norway at the turn of the twenty-first century. *Vet. Microbiol.* **2002**, *90*, 135–145. [CrossRef]

21. Carneiro, P.A.; Kaneene, J.B. Bovine tuberculosis control and eradication in Brazil: Lessons to learn from the US and Australia. *Food Control* **2018**, *93*, 61–69. [CrossRef]

22. Chen, J.-D.; Ke, C.-W.; Deng, X.; Jiang, S.; Liang, W.; Ke, B.-X.; Li, B.; Tan, H.; Liu, M. Brucellosis in Guangdong Province, People's Republic of China, 2005–2010. *Emerg. Infect. Dis.* **2013**, *19*, 817–818. [CrossRef] [PubMed]

23. Wang, Y.; Zhang, W.; Ke, Y.; Zhen, Q.; Yuan, X.; Zou, W.; Li, S.; Sun, Y.; Cui, B.; Song, H.; et al. Human Brucellosis, a Heterogeneously Distributed, Delayed, and Misdiagnosed Disease in China. *Clin. Infect. Dis.* **2013**, *56*, 750–751. [CrossRef] [PubMed]

24. Lai, S.J.; Zhou, H.; Xiong, W.Y.; Gilbert, M.; Huang, Z.J.; Yu, J.X.; Yin, W.W.; Wang, L.P.; Chen, Q.L.; Li, Y.; et al. Changing Epidemiology of Human Brucellosis, China, 1955–2014. *Emerg. Infect. Dis.* **2017**, *2*, 184–194. [CrossRef]

25. Zhang, J.; Yin, F.; Zhang, T.; Yang, C.; Zhang, X.; Feng, Z.; Li, X. Spatial analysis on human brucellosis incidence in mainland China: 2004–2010. *BMJ Open* **2014**, *4*, 004470. [CrossRef]

26. Wang, J.; Jia, P.; Cuadros, D.F.; Xu, M.; Wang, X.; Guo, W.; Portnov, B.A.; Bao, Y.; Chang, Y.; Song, G.; et al. A Remote Sensing Data Based Artificial Neural Network Approach for Predicting Climate-Sensitive Infectious Disease Outbreaks: A Case Study of Human Brucellosis. *Remote Sens.* **2017**, *9*, 1018. [CrossRef]

27. Li, Y.-J.; Li, X.-L.; Liang, S.; Fang, L.-Q.; Cao, W.-C. Epidemiological features and risk factors associated with the spatial and temporal distribution of human brucellosis in China. *BMC Infect. Dis.* **2013**, *13*, 547. [CrossRef]

28. Guan, P.; Wu, W.; Huang, D. Trends of reported human brucellosis cases in mainland China from 2007 to 2017: An exponential smoothing time series analysis. *Environ. Heal. Prev. Med.* **2018**, *1*, 23. [CrossRef]

29. Yang, L.; Bi, Z.W.; Kou, Z.Q.; Li, X.J.; Zhang, M.; Wang, M.; Zhang, L.Y.; Zhao, Z.T. Time-series analysis on human brucellosis during 2004-2013 in Shandong Province, China. *Zoonoses Public Health* **2015**, *3*, 228–235. [CrossRef]

30. Lee, H.S.; Her, M.; Levine, M.; Moore, G.E. Time series analysis of human and bovine brucellosis in South Korea from 2005 to 2010. *Prev. Vet. Med.* **2013**, *2*, 190–197. [CrossRef]

31. Wang, Y.; Xu, C.; Zhang, S.; Wang, Z.; Zhu, Y.; Yuan, J. Temporal trends analysis of human brucellosis incidence in mainland China from 2004 to 2018. *Sci. Rep.* **2018**, *8*, 15901. [CrossRef] [PubMed]

32. Akhtar, S.; Rozi, S. An autoregressive integrated moving average model for short-term prediction of hepatitis C virus seropositivity among male volunteer blood donors in Karachi, Pakistan. *World J. Gastroenterol.* **2009**, *15*, 1607–1612. [CrossRef] [PubMed]

33. Wangdi, K.; Singhasivanon, P.; Silawan, T.; Lawpoolsri, S.; White, N.J.; Kaewkungwal, J. Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: A case study in endemic districts of Bhutan. *Malar. J.* **2010**, *9*, 251. [CrossRef] [PubMed]

34. Liu, Q.Y.; Liu, X.D.; Jiang, B.F.; Yang, W.Z. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infect. Dis.* **2011**, *11*, 218. [CrossRef] [PubMed]

35. Martinez, E.Z.; Da Silva, E.A.S.; Fabbro, A.L.D. A SARIMA forecasting model to predict the number of cases of dengue in Campinas, State of São Paulo, Brazil. *Rev. Soc. Bras. Med. Trop.* **2011**, *44*, 436–440. [CrossRef] [PubMed]

36. Huang, D.; Li, R.; Qiu, J.; Sun, X.; Yuan, R.; Shi, Y.; Qu, Y.; Niu, Y. Geographical Environment Factors and Risk Mapping of Human Cystic Echinococcosis in Western China. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1729. [CrossRef] [PubMed]

37. Bastiaanssen, W.G.; Molden, D.J.; Makin, I.W. Remote sensing for irrigated agriculture: Examples from research and possible applications. *Agric. Water Manag.* **2000**, *46*, 137–155. [CrossRef]

38. Jia, P.; Nie, Y.; Song, G. Detection of underground remains by remote sensing and geophysics. In Proceedings of the 2010 18th International Conference on Geoinformatics, Beijing, China, 18–20 June 2010.

39. Jia, P.; Nie, Y.P.; Yang, L. Recognition and extraction of the ancient sites covered by thick vegetation in Hainan Province of China. In Proceedings of the 30th IEEE International Geoscience and Remote Sensing Symposium (IGARSS) on Remote Sensing—Global Vision for Local Action, Honolulu, HI, USA, 25–30 June 2010; pp. 3898–3901.

40. Zhu, J.F.; Peng, J.; Nie, Y.P. Analysis of the ancient river system in Loulan period in Lop Nur region. In Proceedings of the 17th China Conference on Remote Sensing—Remote Sensing of the Environment, Hangzhou, China, 27–31 August 2010.

41. Xu, M.; Cao, C.; Li, Q.; Jia, P.; Zhao, J. Ecological Niche Modeling of Risk Factors for H7N9 Human Infection in China. *Int. J. Environ. Res. Public Health* **2016**, *13*, 600. [CrossRef]

42. Jia, P.; Stein, A. Using remote sensing technology to measure environmental determinants of non-communicable diseases. *Int. J. Epidemiol.* **2017**, *46*, 1343–1344. [CrossRef]

43. Jia, P.; Joyner, A.; Sun, Y.L. Short-term Associations between Accumulated Rainfall and Atmospheric Moisture during Landfall of Three Atlantic Hurricanes. *Geogr. Bull.* **2014**, *55*, 49–62.

44. Ahani, I.K.; Salari, M.; Shadman, A. Statistical models for multi-step-ahead forecasting of fine particulate matter in urban areas. *Atmos. Pollut. Res.* **2019**, *10*, 689–700. [CrossRef]

45. Kim, M.; Chemere, B.; Sung, K. Effect of Heavy Rainfall Events on the Dry Matter Yield Trend of Whole Crop Maize (*Zea mays* L.). *Agriculture* **2019**, *4*, 75. [CrossRef]

46. Aekphachaisawat, N.; Sawanyawisuth, K.; Khamsai, S.; Chattakul, P.; Takahashi, K.; Chotmongkol, V.; Tiamkao, S.; Limpawattana, P.; Senthong, V.; Chindaprasirt, J.; et al. An ecological study of eosinophilic meningitis caused by the nematode, *Angiostrongylus cantonensis* (Chen, 1935) (Nematoda: Metastrongylidae). *Parasitol. Int.* **2019**, *72*, 101944. [CrossRef] [PubMed]

47. Chinese Center for Disease Control and Prevention (CDC). Available online: http://www.nhc.gov.cn/jkj/s2907/new_list.shtml (accessed on 20 August 2019).

48. International Scientific & Technical Data Mirror Site, Computer Network Information Center, Chinese Academy of Sciences. Available online: http://www.gscloud.cn (accessed on 20 August 2019).

49. Briet, O.J.; Vounatsou, P.; Gunawardena, D.M.; Galappaththy, G.N.; Amerasinghe, P.H. Models for short term malaria prediction in Sri Lanka. *Malar. J.* **2008**, *7*, 76. [CrossRef] [PubMed]

50. Moghram, l.; Rahrnan, S. Analysis and evaluation of five short-term load forecasting techniques. *IEEE Trans. Power Syst.* **1989**, *4*, 1484–1491. [CrossRef]

51. Zhang, X.; Liu, Y.; Yang, M.; Zhang, T.; Young, A.A.; Li, X. Comparative Study of Four Time Series Methods in Forecasting Typhoid Fever Incidence in China. *PLoS ONE* **2013**, *8*, e63116. [CrossRef]

52. National Development and Reform Commission of China. National Beef and Mutton Production Development Plan (2013–2020). Available online: http://www.ndrc.gov.cn/zcfb/zcfbghwb/201402/P020140221362074574291.pdf (accessed on 20 August 2019).