

Article

A Remote Sensing Data Based Artificial Neural Network Approach for Predicting Climate-Sensitive Infectious Disease Outbreaks: A Case Study of Human Brucellosis

Jiao Wang ^{1,2,†}, Peng Jia ^{3,*,†}, Diego F. Cuadros ^{4,5}, Min Xu ⁶, Xianliang Wang ², Weidong Guo ⁷, Boris A. Portnov ⁸, Yuhai Bao ⁹, Yushan Chang ⁹, Genxin Song ¹, Nan Chen ¹⁰ and Alfred Stein ³ 

¹ Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, Ministry of Education, Henan University, Kaifeng 475004, China; jwanghjs@163.com (J.W.); shengliking@126.com (G.S.)

² National Institute of Environmental Health, Chinese Center for Disease Control and Prevention, Beijing 100021, China; ahyxwzl@163.com

³ Department of Earth Observation Science, Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede 7500, The Netherlands; a.stein@utwente.nl

⁴ Department of Geography and Geographic Information Science, University of Cincinnati, Cincinnati, OH 45221, USA; cuadrodo@ucmail.uc.edu

⁵ Health Geography and Disease Modeling Laboratory, University of Cincinnati, Cincinnati, OH 45221, USA

⁶ State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, China; xumin@radi.ac.cn

⁷ Inner Mongolia Center for Disease Control and Prevention, Hohhot 010031, China; gwd.com@126.com

⁸ Department of Natural Resources and Environmental Management, Faculty of Management, University of Haifa, Haifa 3498838, Israel; portnov@research.haifa.ac.il

⁹ Inner Mongolian Key Laboratory of Remote Sensing and GIS, Inner Mongolia Normal University, Hohhot 010022, China; baoyuhai@imnu.edu.cn (Y.B.); yushangis@163.com (Y.C.)

¹⁰ Key Laboratory of Spatial Data Mining & Information Sharing of Ministry of Education, Fuzhou University, Fuzhou 350003, China; fjc99@163.com

* Correspondence: jiapengff@hotmail.com; Tel.: +31-5-3489-4038

† These two authors contributed equally to this work.

Received: 30 June 2017; Accepted: 22 September 2017; Published: 30 September 2017

Abstract: Remote sensing technologies can accurately capture environmental characteristics, and together with environmental modeling approaches, help to predict climate-sensitive infectious disease outbreaks. Brucellosis remains rampant worldwide in both domesticated animals and humans. This study used human brucellosis (HB) as a test case to identify important environmental determinants of the disease and predict its outbreaks. A novel artificial neural network (ANN) model was developed, using annual county-level numbers of HB cases and data on 37 environmental variables, potentially associated with HB in Inner Mongolia, China. Data from 2006 to 2008 were used to train, validate and test the model, while data for 2009–2010 were used to assess the model's performance. The Enhanced Vegetation Index was identified as the most important predictor of HB incidence, followed by land surface temperature and other temperature- and precipitation-related variables. The suitable ecological niche of HB was modeled based on these predictors. Model estimates were found to be in good agreement with reported numbers of HB cases in both the model development and assessment phases. The study suggests that HB outbreaks may be predicted, with a reasonable degree of accuracy, using the ANN model and environmental variables obtained from satellite data. The study deepened the understanding of environmental determinants of HB and advanced the methodology for prediction of climate-sensitive infectious disease outbreaks.

Keywords: infectious disease; climate-sensitive; artificial neural network (ANN); remote sensing; brucellosis; disease outbreak prediction; environmental health; Inner Mongolia

1. Introduction

Brucellosis, an infectious disease caused by bacteria of the genus *Brucella*, remains rampant worldwide in both domesticated animals and humans, especially in the Mediterranean region, Asia, the Middle East, Sub-Saharan Africa, Latin America and the Balkan Peninsula [1–3]. Over 500,000 new cases of human brucellosis (HB) are reported each year, with the incidence rate of more than 10 cases per 100,000 residents in some countries [4]. Despite relatively low mortality rates, brucellosis can lead to reduced milk and meat production and subsequent economic losses [5]. Humans become infected with brucellosis via two main pathways: (1) by contact with animals that carry the pathogenic bacteria of the genus *Brucella*, or with infectious materials during animal husbandry and meat processing; and (2) by consuming unpasteurized dairy products [6]. The infection attacks the immune system and results in symptoms, including weaknesses and disabilities, which can trigger a substantial increase in medical costs and a decrease in productivity [1]. When untreated or treated inappropriately, brucellosis can cause non-specific flu-like symptoms and more severe health complications in different organs [7]. The incurred costs and impacts on human health call for HB eradication and control measures [8].

As one of the largest agricultural countries in the world, China has reported cases of HB in 25 out of its 34 provinces (or autonomous regions) over the past years, with an average annual economic loss of more than CNY 10 billion (approximately \$145 million US dollars) caused directly by brucellosis [9]. Among these provinces, the Inner Mongolia Autonomous Region in northern China has been severely affected by brucellosis in both humans and livestock since 1999, reporting almost 50% of HB cases across China during 2005–2010 [10]. The annual incidence of brucellosis in this region ranks in the top three among all Chinese provinces [11]. A study carried over seven years, has identified Inner Mongolia, a region characterized by wide pastures and highly developed animal husbandry, as a hotspot with an abnormally high risk of HB [12].

Despite the growing public health challenge of HB in China, environmental determinants of HB remain under-investigated. A recent study [1] used the MaxEnt model, an ecological niche modeling (ENM) approach, to predict the probability of the presence of HB in northern China. The researchers found that the density of small ruminants (especially sheep), altitude, seasonality of precipitation, and temperature annual range, were key factors associated with the presence of HB. Although the MaxEnt model overcomes several methodological obstacles inherent to other algorithms, such as Genetic Algorithm for Rule Set Production (GARP), it also has disadvantages. For example, MaxEnt uses a number of functional transformations of each predictor, known as features, in conjunction with regularization to penalize for model complexity and select the best set of features from a large candidate pool, which was recently criticized for overfitting [13–15]. Moreover, according to Merow and Silander (2014), the MaxEnt model poorly explains the probability of event occurrence because the assumption of random sampling of individuals is not the same as the assumption of random sampling of space associated with the occurrence probability [16].

Artificial neural network (ANN) is an alternative method that may help to overcome the limitations of conventional approaches and can create ENM models for complex problems. An ANN model is an information processing system that models the relationships between variables and performs predictions, by simulating the human brain function and learning from given examples. The concept of artificial neuron was first introduced by McCulloch and Pitts in 1943 [17]. Soon after, the first ANN model perceptron was developed to model how the human brain recognizes new objects from a prior supervised learning [18], which was widely used for a variety of purposes beyond studying human cognition in following studies [19–21]. Because of its predictability and a broad

application range, especially in solving complex and nonlinear problems [22,23], the ANN method has recently been proposed as an efficient tool for modeling and forecasting disease outbreaks [24].

Using the ANN approach, we developed a quantitative model for prediction of HB outbreaks based on the environmental variables primarily captured by remote sensing technologies, which have been applied to a variety of disciplines such as agriculture [25], archaeology [26–28], ecology [1,29], health [30,31] and meteorology [32]. The goal of the study was twofold: (1) to identify the important predictor variables for HB; and (2) to construct a model based on those most relevant explanatory variables to predict the annual numbers of HB cases in different counties of Inner Mongolia of China. Our study advanced the current state of knowledge about the environmental determinants of HB and improved its prediction tools. More importantly, the research approach developed in this study can be applied to the prediction of outbreaks of a broad range of climate-sensitive infectious diseases.

2. Materials and Methods

2.1. Ethics Statement

No work on human subjects was undertaken in this study. Human brucellosis case data were collected from the annual reports published by the Inner Mongolia Center for Disease Control and Prevention (CDC), where the reported cases were aggregated by county and year. All data were anonymized.

2.2. Study Area

Inner Mongolia is an autonomous region located along the northern border of China that stretches for about 2400 km from west to east and for 1700 km from north to south. Several large minority ethnic groups populate the region, including Mongolian, Daur, Ewenki, Hui and others. The region shares a ~3000 km international border with Mongolia and a ~1000 km border with Russia (Figure 1). Inner Mongolia is the third largest provincial unit in China (after Xinjiang and Tibet), covering approximately 12.28% (~1.183 million km²) of China's total land area. Due to its elongated shape, extending from about 40° to 50°N and from about 107° to 125°E, a wide variety of regional climates exist in the region: arid and semiarid climates predominate in the north and west; humid continental climate is predominant in the northeast; and subarctic climate is found in the far north, in the part of the region that borders on Russia and Heilongjiang Province, China [33]. Inner Mongolia consists of 12 municipal units (nine cities and three leagues) and 89 sub-municipal units (i.e., counties).

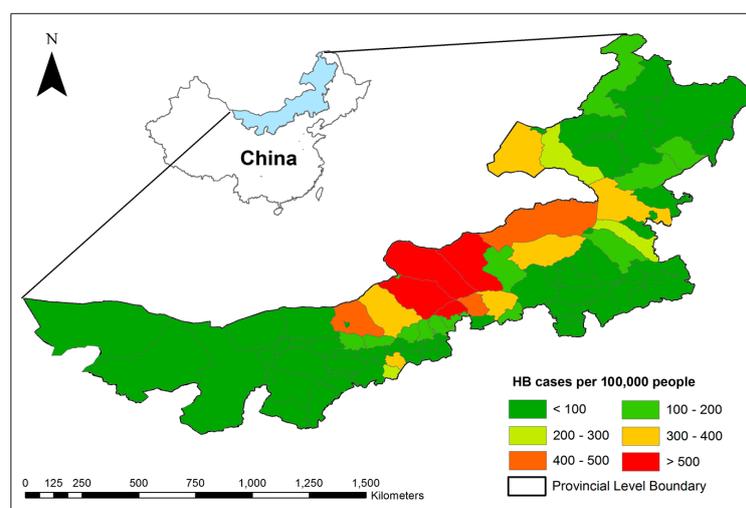


Figure 1. Eighty-nine counties of Inner Mongolia, China, with the numbers of human brucellosis (HB) cases per 100,000 people in 2010.

2.3. Data Sources

2.3.1. Epidemiological Data

The annual numbers of HB cases at the county level in Inner Mongolia during 2006–2010 were collected from the annual reports published by the Inner Mongolia CDC. Each reported case was confirmed by both clinical signs and serological tests, or by isolation in accordance with the HB case definition by the World Health Organization (WHO) [34].

2.3.2. Human and Animal Population Data

A population grid, representing the total number of people in each 1×1 km cell or population density, was obtained from the Global Rural-Urban Mapping Project, Version 1 (GRUMPv1) [35]. The access to major settlements can be a strong determinant of social, health and economic differences between regions [36]. Therefore, a gridded map of accessibility in 2000, at a resolution of 1×1 km, represented the length of travel time (minutes) from each pixel to the nearest major settlement of population size 50,000 or more [37] was used as a separate input variable.

Densities of cattle, sheep and goats in 2010 were obtained from the Food and Agriculture Organization (FAO) of the United Nations with a resolution of 5×5 km and resampled to 1×1 km grids. The worldwide Land Use Systems (Version 1.1), developed by the Land Tenure and Management Unit of the FAO, was used to extract livestock-related areas.

2.3.3. Environmental Variables

Minimum, maximum and mean middle infrared radiation (MIR), land surface temperature (LST) and enhanced vegetation index (EVI) were produced by the Temporal Fourier Analysis based on 1×1 km gridded data obtained from MODerate-resolution Imaging Spectroradiometer (MODIS) sensors on-board NASA's Terra and Aqua satellites during 2001–2008 [38]. These parameters were used to describe the physical environment, including water content (MIR), surface temperature (LST and MIR) and vegetation canopy greenness (EVI and MIR).

Monthly precipitation data between 1960 and 1990 were downloaded from WorldClim [39] and transformed into annual minimum, maximum and mean precipitation variables by implementing raster math calculations in the ArcGIS (Version 10.4.1, Esri, Redlands, CA, USA) to describe the precipitation trends [40,41]. In addition, altitude and 19 bioclimatic variables, with a resolution of 1×1 km, were also derived from WorldClim to represent the general climatic conditions between 1950 and 2000. The importance of these variables as factors affecting human health was identified by previous studies [36]. All environmental variables used in the study are listed in Table S1.

2.4. Statistical Analysis

A stepwise regression analysis with a forward selection procedure was conducted using the SPSS version 22.0 software [42], which started with no variables in the model and added the variable that most improved the model each time iteratively ($p < 0.05$), until no more improvement was achieved by adding new variables. A final set of model predictors of HB cases was selected based on the overall adjusted R-square value and further used to establish the ANN model.

2.5. ANN Model

ANN models simulate human brain functioning to predict values of independent variables when the relationship between independent variables and dependent variables is non-linear. A key difference among various types of ANN models is the training algorithm used for determining the weights for neurons between different layers. The most commonly used training algorithm is called Back Propagation, a multilayer feed-forward network trained by an error back-propagation algorithm, which trains ANN models to minimize the difference between the model output and the measured

target [22]. The back-propagation ANN model used in this study consists of three layers: an input layer, a hidden layer, and an output layer; each layer contains multiple neurons. An artificial neuron in a typical ANN architecture receives input neurons via connections and sends calculated results via connections to output layer (Figure 2). Each connection has an unknown parameter, called weight, which is kept adjusted until the overall performance of the ANN model meets the requirements of prediction.

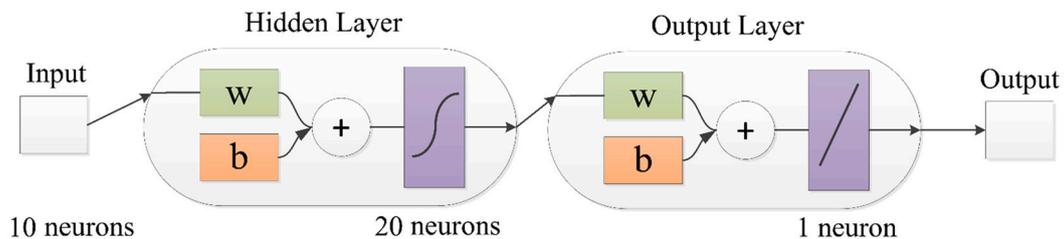


Figure 2. Structure of an artificial neural network (ANN) model. Weight matrix vector and threshold vector are denoted by w and b , respectively.

2.5.1. Model Development

In this study, we used an ANN toolbox in the MATLAB program (version 2013a) to train a feed-forward ANN model using the error back-propagation algorithm, and to integrate environmental variables for the prediction of HB outbreaks [43]. The input layer consisted of the variables selected from the statistical analysis of different environmental predictors of HB. The number of hidden neurons was adjusted to 20. A log-transformation was conducted to shrink the discrepancy of HB cases among different areas, and the output layer comprised the raw yearly log-transformed numbers of HB cases at the county level as the target.

For model development, data for 2006–2008 (89 groups for each year) were combined and then randomly split into three subsets for training (60% of data, 161 groups), validating (20%, 53 groups), and testing (20%, 53 groups). The training set was used to design the model, in which the weights of connections were estimated, while the validation set was used to estimate the mean square error (MSE) of the model. Training and validating were iteratively conducted until the MSE reached its minimum value. Then, the unbiased generalization estimate of performance, i.e., linear correlation coefficient (LCC), was calculated based on the testing set, and the entire process of model development was conducted to optimize LCC.

2.5.2. Model Assessment

To assess the model's ability to predict the HB outbreaks, we used 178 groups of data from 2009–2010 (89 groups for each year), which were left out in the model development phase to test whether the prediction results were identical with the reported numbers. A range of statistical measures were calculated to compare prediction and observation in both model development and assessment stages, including the linear correlation coefficient (LCC), root mean square error (RMSE), standard deviation of bias (SDB), arithmetic mean (AM) and median.

2.6. Sensitivity Analysis

Sensitivity analysis of the model was conducted by the changing one-factor-at-a-time (OAT) method, to demonstrate the extent to which each input variable influenced the output. To that end, each input variable included in the ANN model was alternately changed from -20% to $+20\%$ (increasing by $\pm 5\%$ at a time). Sensitivity was measured by the percent change in the model output.

3. Results

3.1. Predictors of HB

As shown in Table 1, the most significant environmental variables selected for the final regression model using the stepwise regression analysis procedure were: sheep density in 2010 (Sheep10, heads/km²), max temperature of warmest month (Bio5, °C), mean temperature of wettest quarter (Bio8, °C), minimum EVI (EVImin), minimum LST (LSTmin, °C), precipitation of wettest month (Bio13, mm), mean LST (LSTmean, °C), accessibility (Access), mean EVI (EVImean), minimum MIR (MIRmin), and precipitation of warmest quarter (Bio18, mm).

Table 1. Summary of the stepwise regression analysis of environmental factors affecting the annual incidence of HB in counties of Inner Mongolia (note: the variables were entered if $p \leq 0.05$, and removed if $p \geq 0.10$).

Model	Variables Entered	Variables Removed	R Square	Significance of F Value Change
1	Sheep10		0.129	0.000
2	Bio5		0.237	0.000
3	Bio8		0.283	0.000
4	EVImin		0.296	0.004
5	LSTmin		0.333	0.000
6	Bio13		0.386	0.000
7	LSTmean		0.419	0.000
8	Access		0.435	0.000
9	EVImean		0.445	0.006
10		Bio8	0.443	0.313
11	MIRmin		0.455	0.002
12	Bio18		0.460	0.044

3.2. Model Development

We next compared the predicted and raw log-transformed numbers of HB cases at county level during 2006–2008 (Figure 3). In general, ANN model predictions were consistent with the raw numbers with a high LCC of 0.9134. We conducted a comparison of statistical measures between predicted and raw yearly log-transformed numbers of HB cases during 2006–2008 (Table 2). As Table 2 shows, both AM and median of the predicted numbers were close to those of the raw numbers, which implied similarity between predictions and reports. This was further confirmed by a low RMSE of 0.3664. A low SDB (0.3656) revealed that errors in the prediction of log-transformed numbers of HB cases tended to be uniform among different counties.

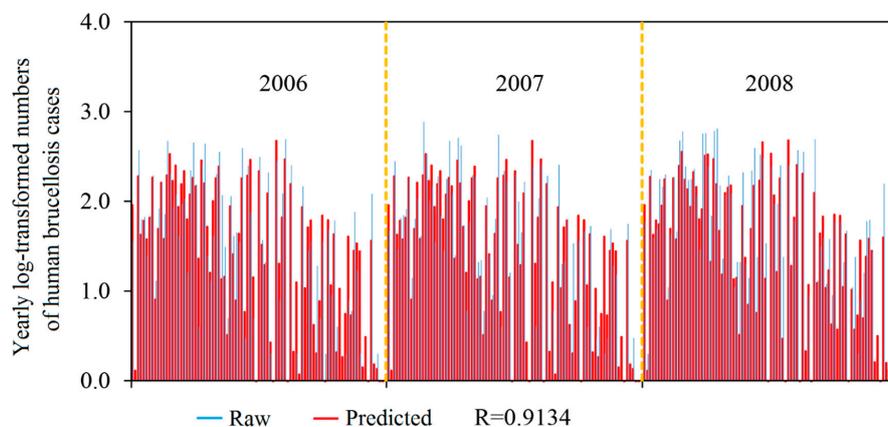


Figure 3. Comparison between predicted (red) and raw yearly log-transformed numbers of human brucellosis cases (blue) at the county level in the model development phase. X-axis denotes counties, with 89 counties listed horizontally in the same order in all three panels (2006, 2007 and 2008).

Table 2. Comparison of predicted and raw yearly log-transformed numbers of human brucellosis cases in the model development phase (2006–2008) in Inner Mongolia, China.

Statistical Measure	Raw	Predicted
Linear correlation coefficient (LCC)	NA	0.9134
Root mean square error (RMSE)	NA	0.3664
Standard deviation of bias (SDB)	NA	0.3656
Arithmetic mean (AM)	1.3998	1.3745
Median	1.5682	1.5826

NA: Not applicable.

Overall, the ANN model did not show any tendency to either overestimate or underestimate the average log-transformed number of HB cases in the model development phase. The percent errors in the prediction were between -50% and 50% in 71 (79.8%) out of 89 counties (Figure 4), with the spatial distribution shown in Figure S1a. The reported averages of HB cases in 12 counties (Alxa Right Banner, Alxa Left Banner, Baiyun Mining District, Dalad Banner, Dongsheng City, Ejin Banner, Otog Banner, Otog Front Banner, Urat Rear Banner, Uxin Banner, Ejin Horo Banner, and Jungar Banner) were either 0 or 1, thus the percent error was not applicable in these counties. The percent errors in the remaining six counties (Haggin Rear Banner, Liangcheng County, Linhe District, Ningcheng County, Tumed Right Banner, and Tumed Left Banner) were -75.80% , 126.96% , -57.88% , 60.62% , 269.56% , and 102.16% , respectively.

3.3. Model Assessment

To assess the performance of the ANN model, the predicted and raw log-transformed numbers of HB cases per county were compared in 2009 and 2010 (Figure 5). In general, ANN model predictions were consistent with the raw numbers of HB cases with a high LCC of 0.8930, which was close to but slightly lower than the LCC in the model development phase (0.9134). Compared to 2006–2008, the reported numbers of HB cases increased in 2009 and 2010, particularly in Horqin Right Front Banner. A large-scale outbreak of HB was reported during 2010 in Horqin Right Front Banner where 1490 people were infected. The total number of HB cases in this county increased from 202 in 2006 to 1490 in 2010.

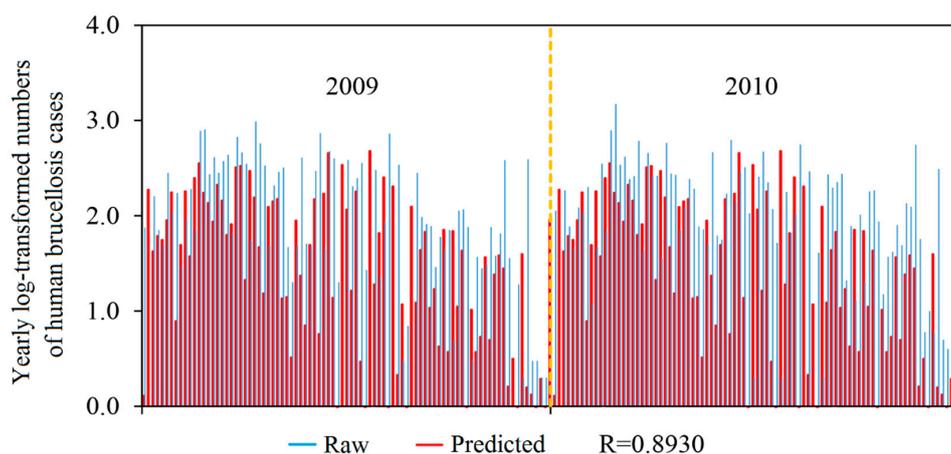


Figure 5. Comparison between predicted (red) and raw yearly log-transformed numbers of human brucellosis cases (blue) at the county level in the model assessment phase. X-axis denotes counties, with 89 counties listed horizontally in the same order in both panels (2009 and 2010).

We further compared statistical measures between the predicted and raw log-transformed numbers of HB cases during 2009–2010 (Table 3). Both AM and median of predicted log-transformed numbers of HB cases were underestimated by the model in the assessment phase. There was a decrease in LCC (0.9134 to 0.8930) and an increase in RMSE (0.3664 to 0.5603) in the model assessment phase, compared to the model development phase, thus indicating larger discrepancies between the predicted and raw numbers of HB cases in the model assessment phase.

Table 3. Comparison of predicted and raw yearly log-transformed numbers of human brucellosis cases in the assessment phase (2009–2010) in Inner Mongolia, China.

Statistical Measure	Raw	Predicted
Linear correlation coefficient (LCC)	NA	0.8930
Root mean square error (RMSE)	NA	0.5603
Standard deviation of bias (SDB)	NA	0.3793
Arithmetic mean (AM)	1.8190	1.4067
Median	1.9370	1.5993

NA: Not applicable.

The percent errors of the predicted and raw numbers of HB cases during 2009–2010 were compared. The percent errors were between -50% and 50% in 67 out of 89 counties (Figure 6), $>50\%$ in Uxin Banner (92.72%), and $<-50\%$ in the remaining 18 counties. The spatial distribution of the percent errors in the prediction can be found in Figure S1b. As mentioned above, the ANN model had a tendency to underestimate the average yearly numbers of HB cases in the model assessment phase, which was different from the tendency observed in the model development phase.

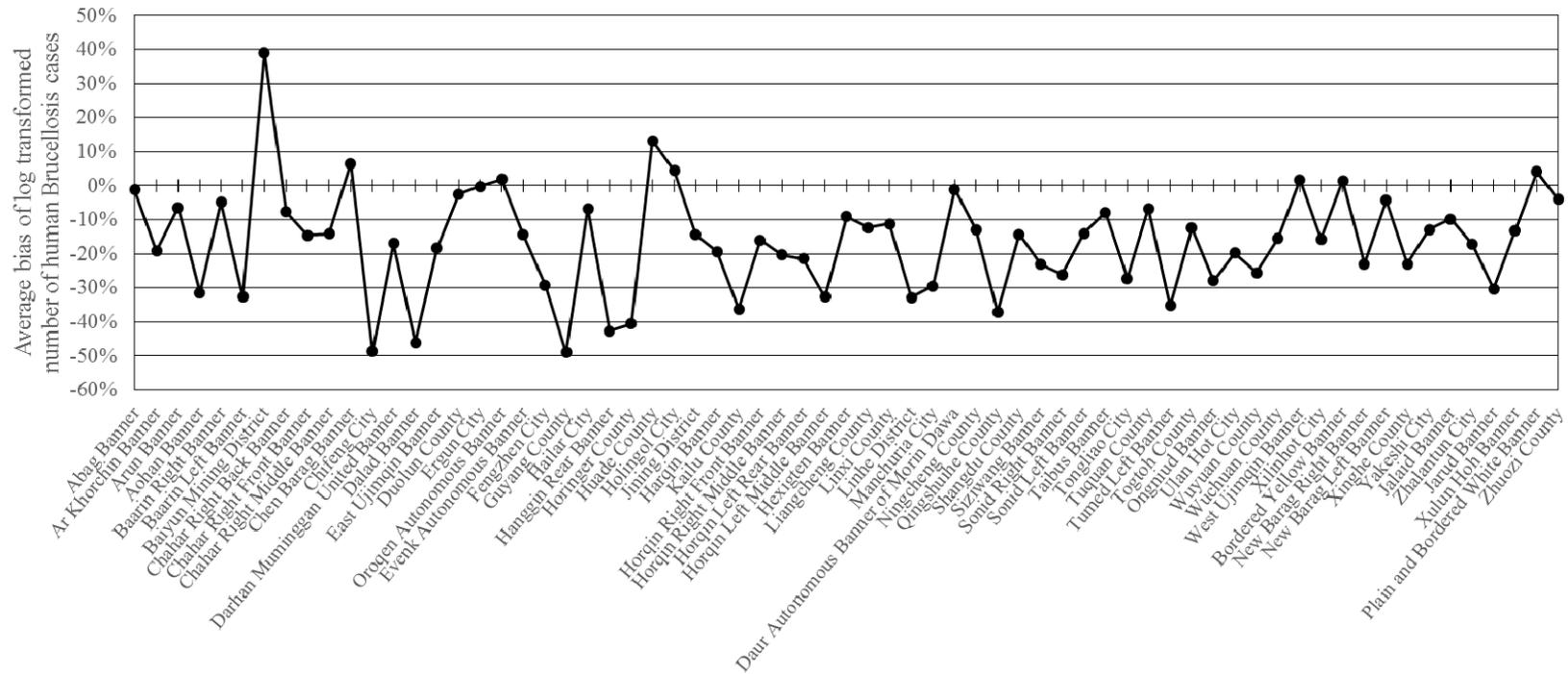


Figure 6. Percent errors of the log-transformed numbers of human brucellosis cases during 2009–2010 in 67 out of 89 counties in Inner Mongolia, China. The percent errors in the remaining 22 counties, outside of the −50% to 50% error range, were described in the text, but left out in this figure for the sake of clarity.

3.4. Model Estimation

The spatial distribution of predicted and raw log-transformed numbers of HB cases was, to a large degree, consistent with each other for the years between 2006 and 2010 (Figure 7). It was also observed that the central counties have tended to cluster since 2006, forming a hot spot (colored in red) which gradually expanded over years, and, in 2010, covered Darhan Muminggan United Banner in the west to East Ujimqin Banner, West Ujimqin Banner, Barrin left Banner, and Ongniud Banner in the east. In addition, counties in northeastern Inner Mongolia formed another hotspot from Oroqen Autonomous Banner in the north to Jarud Banner in the south. The predicted and real locations of these two hotspots were quite similar, which indicated an acceptable performance of the model prediction in terms of identifying the clustering patterns.

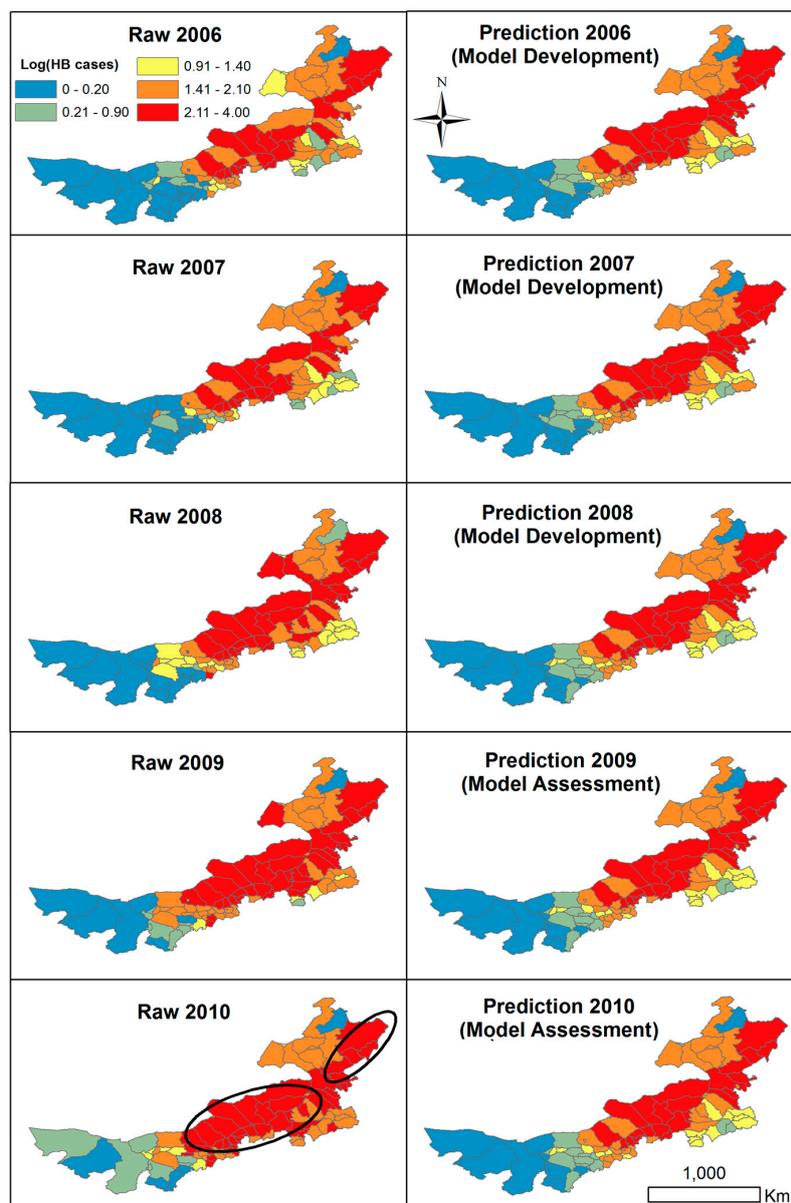


Figure 7. Spatial distribution of predicted and raw yearly log-transformed numbers of human brucellosis (HB) cases during 2006–2010 in Inner Mongolia, China. The location of the major and minor hot spots is marked by the large and small black circles, respectively.

A major cold spot was observed to the west of central Inner Mongolia and successfully captured by the model (Figure 7, counties colored in blue and green). Although this cluster has been gradually shrinking since 2006, this trend has not been reflected in predictions over years. The model also failed to capture some abrupt changes. For example, in New Barag Right Banner, the increase in the yearly log-transformed numbers of HB cases during 2008–2009 was not accurately predicted. However, prediction was accurate again in 2010 when the numbers of HB cases decreased. Likewise, predictions were lower in Sonid Right Banner during 2008–2010, an area with an increasing reported number of HB cases since 2008. HB cases were also underestimated in southeastern Inner Mongolia, including Ar Khorchin Banner, Naiman Banner, Kailu County, Horqin Left Middle Banner, Horqin Left Rear Banner, Tongliao City, and Hure Banner. On the other hand, HB was not prevailed in Alxa Right Banner in the west and Genhe City in the north, and model predictions for these two counties were also close to zero. Hence, although the model had difficulties in the prediction in counties with high values and abrupt changes, it still accurately predicted HB cases in counties with low reported numbers and could reproduce similar spatial patterns of HB cases in major counties of Inner Mongolia.

3.5. Sensitivity Analysis

There was no direct evidence implying that the correlations between individual predictor variables and predicted yearly log-transformed numbers of HB cases were linear. The increase in predicted yearly log-transformed numbers of HB cases was associated with a decrease in precipitation of wettest month (Bio13, mm), accessibility (Access) and MIR, and with an increase in precipitation of warmest quarter (Bio18, mm) and sheep density (Sheep10, heads/km²) (Figure 8). Changes in other predictor variables, i.e., maximum temperature of warmest month (°C) (Bio 5), LST (°C) and EVI, resulted in the mixed changes in model predictions.

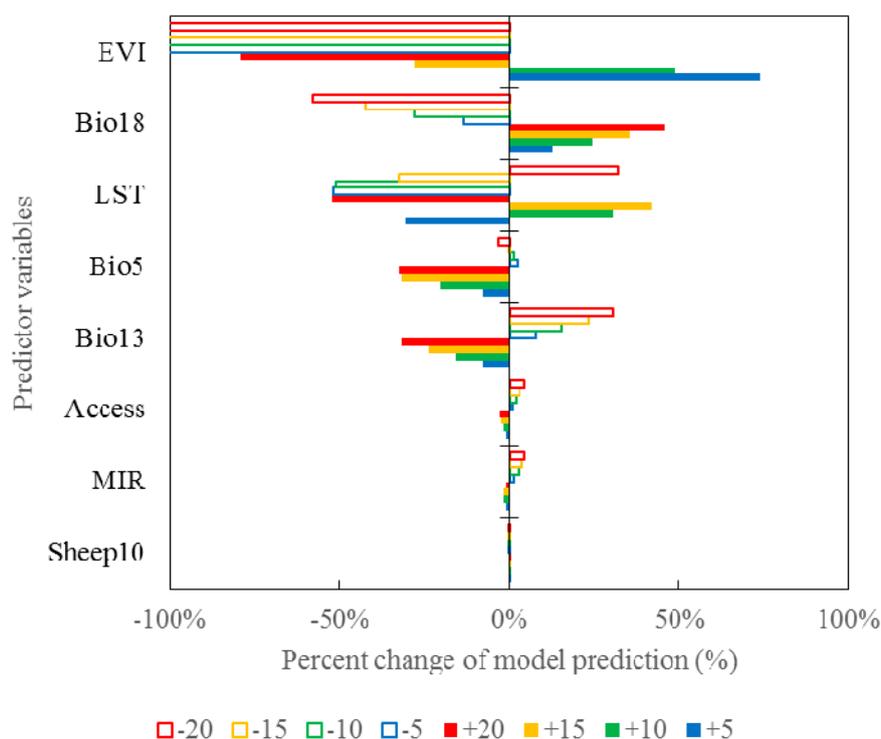


Figure 8. Sensitivity of ANN model to changes in predictor variables. Horizontal bars indicate the percent changes of model prediction caused by changes in the corresponding predictors. Positive values indicate overestimation while negative values indicate underestimation. The larger percent change of model prediction caused by changing a certain predictor, the more important that predictor is for model prediction.

Changes in EVI had the greatest influences on model predictions. Changes in Bio18, LST, Bio5 and Bio13 also affected model predictions significantly, for example, a change by 20% in Bio18 could lead to around 50% of the changes in prediction. The model prediction was not significantly sensitive to changes in Access, MIR and Sheep10, where the changes in model predictions were less than 5%. Changes beyond 100% resulting from changes in EVI (the most sensitive predictor variable) were left out in Figure 8, for clearer data visualization of the other seven variables. When EVI was decreased by 20%, model predictions were about three times lower than the original predictions.

The 2-D scatterplots were generated to visualize the suitable environment for HB outbreaks. For example, over 50% of HB cases occurred in regions with the precipitation of warmest quarter (Bio18) ranging 92–339 mm and the mean EVI ranging 0.09–0.23 (Figure 9a), and where the maximum temperature of warmest month (Bio5) and mean LST were 22.7–30.4 and 3.2–23.2 °C, respectively (Figure 9b).

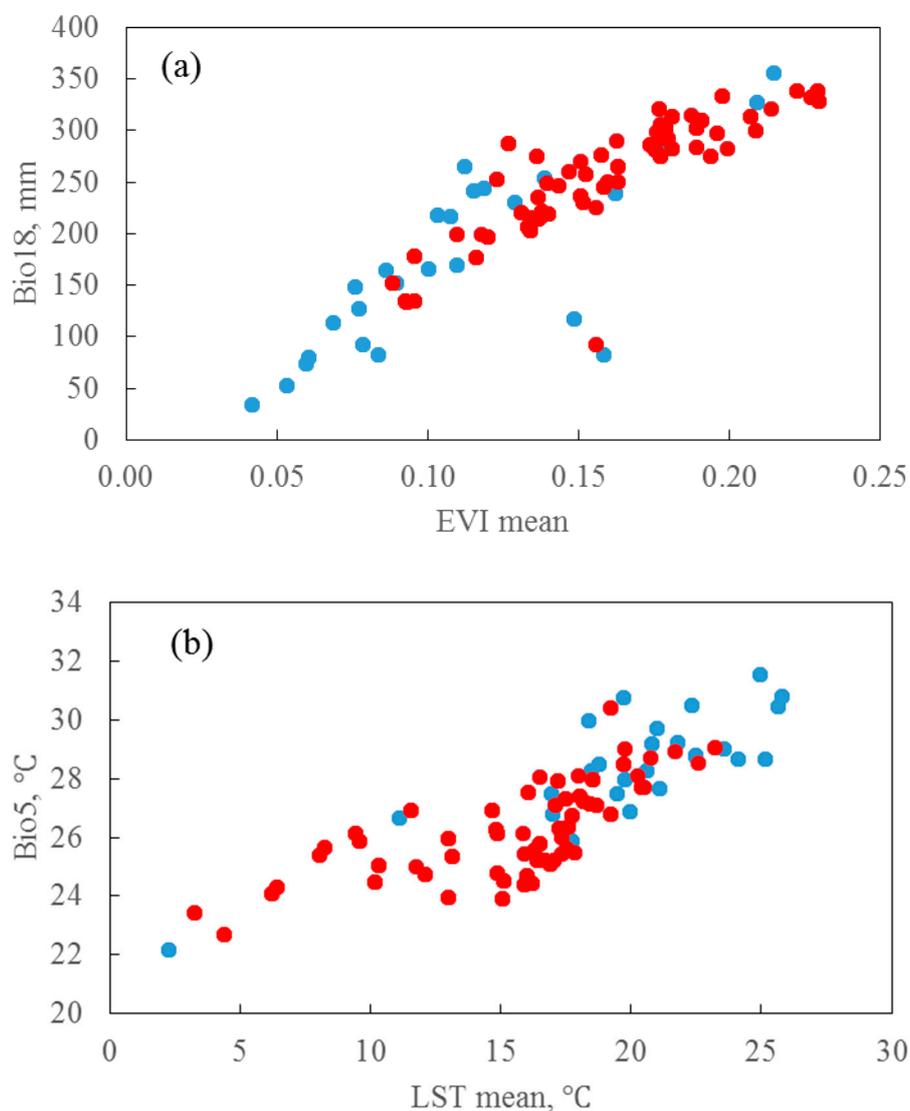


Figure 9. Modeled ecological niche of human brucellosis (HB), according to: (a) the precipitation of warmest quarter (Bio 18) and mean enhanced vegetation index (EVI mean); and (b) max temperature of warmest month (Bio 5) and mean land surface temperature (LST mean). All dots represent the ecological niche of HB with the suitable ecological niche highlighted in red, which explains over 50% of HB in historical records.

4. Discussion

The Chinese CDC defines brucellosis as a Class II reportable disease, which increasingly receives attention to its causes, predictions and prevention nationwide. Our study suggests that environmental factors regarding temperature and precipitation, such as LST and other climatic factors, are key drivers behind HB outbreaks in Inner Mongolia. In particular, according to the results of our study, the predicted numbers of HB cases decreased with an increase in the amount of precipitation during the wettest month. A possible reason is that sufficient water supply may increase the resistance of animals to diseases [44], and hence fewer humans are exposed to animal infection. On the other hand, an increase in the predicted numbers of HB cases was observed when precipitation increased over the entire warmest quarter. A possible explanation might be a higher risk of human exposure to the contaminated soil and water by the bacteria *Brucella*, originally found in animal wastes and bodies, through the accelerated runoff and circulation of water among land, air and surface which could be triggered by heavy and/or long precipitation.

The association between HB incidence and other predictor variables was complex. For example, a decrease in EVI was associated with smaller numbers of HB cases. This association may be because shrinking vegetation coverage reduces food supply to domestic animals, and reduction in the animal population may lower the possibility of brucellosis transmission. Conversely, an increase in vegetation coverage results in more water and food supply, which may increase energy availability and lower potential competition for food. Such an environment promotes the development of healthy individuals with strong immune systems capable of efficiently protecting against the brucellosis infection. In addition, the land surface temperature in a specific range (3.2–23.2 °C) was also favorable for the spread of brucellosis. According to our results, a decrease in the maximum temperature of the warmest month on the predicted number of HB cases was negligible, while an increase in the value of this variable was associated with the reduction in the predicted number of HB cases, apparently because lower temperature in the warmest month favors the survival of animals and decreases disease susceptibility as a result. Consequently, decreased disease susceptibility can result in a reduction of the numbers of HB cases. These results could help to understand better the changes in the ecological niche of and outbreak patterns of HB under global climate change scenarios. However, further studies are needed to elucidate the complex effects of climate change on the spatial distribution and outbreak patterns of HB.

The areas affected by brucellosis expanded over the study period, with the number of the affected counties increased from 76 in 2006 to 86 in 2010. The central and northeastern parts of Inner Mongolia were widely affected by brucellosis, where Horqin Right Front Banner was particularly severely affected, with a total number of 1490 HB cases reported in 2010, the highest among all 89 counties ever reported. The counties with more than 500 HB reported cases in 2010 showed clustering, with major and minor hot spots observed in the northeastern and central parts of Inner Mongolia, respectively. The majority of HB cases were reported in the major hot spot that includes four counties (Horqin Right Front Banner, Horqin Right Middle Banner, and Zhalantun City), with the reported numbers of infections increasing particularly fast in Horqin Right Front Banner and Horqin Right Middle Banner at an annual growth rate of 65% and 55%, respectively. In the minor hot spot, the number of HB cases decreased somewhat in Sonid Left Banner, Siziwang Banner and Shangdu County since 2009. However, the average annual growth rate of the number of HB cases in Horing County was 139%, which implied that this region remained under a high risk of brucellosis outbreaks.

A major limitation of our model is the difficulty of making accurate predictions for counties with 500 HB cases or higher, apparently attributed to the underestimation of the numbers of HB in the assessment phase. Given both positive (in the warmest quarter) and negative impacts (in the wettest month) of precipitation on HB outbreaks, some interactions among different environmental drivers might exist but not examined in this study. It should also be noted that changes of predictor variables were associated with changes in the yearly number of HB cases, and thus our model was not necessarily capable to explain the spread of brucellosis outbreaks for individual months. To improve

our model further, longer-term and more complete data with information of reported HB cases per month need to be collected and analyzed in future studies. Results from some advanced time series analysis based on satellite data with finer temporal resolution, such as the Breaks For Additive Season and Trend (BFAST) model, should be integrated with the monthly reported HB cases. More recent environmental and climatic products, such as CHELSA [45] and free-of-charge satellite data (e.g., Sentinel and ASTER), may also be a better match for disease data and provide more accurate explanations. Finer-scale population grid data [46–49] can help to convert the risks of HB outbreaks to relevant population-based metrics for future public health planning.

Despite these limitations, the ANN model approach was able to reproduce and predict the yearly log-transformed numbers of HB cases with a relatively high accuracy, in both model development and assessment phases. By analyzing the sensitivity of model prediction to changes in each predictor variable, we estimated the relationships between predictor variables and the numbers of HB cases as well as the importance of each predictor variable, which can improve our understanding of brucellosis epidemic and provide practical guidelines for decision making with important public health implications. The prediction of outbreaks generated from our model might become a fundamental tool for planning, designing and implementing cost-effective control interventions against the HB epidemic.

5. Conclusions

Our study presents an ANN model for predicting yearly numbers of HB cases at a county level in Inner Mongolia, China, which helps to understand better the environmental determinants of HB. The methodology presented in this paper can be used to analyze and predict the outbreaks of a broad range of climate-sensitive infectious diseases in other regions. Such information may become essential for successful implementation of surveillance, control, and prevention measures, and also for future public health planning and resource allocation.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/9/10/1018/s1, Figure S1: Percent errors of the predicted numbers of human brucellosis cases in 89 counties of Inner Mongolia, China, during: (a) 2006–2008; and (b) 2009–2010; Table S1: Description of remotely sensed, climatic, and socioeconomic variables used in the study.

Acknowledgments: All authors acknowledge funding support from the Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, Ministry of Education (No. JOF201704); the Key Laboratory of Spatial Data Mining & Information Sharing of Ministry of Education, Fuzhou University (No. 2017LSDMIS08); and the Key Laboratory for National Geographic State Monitoring (National Administration of Surveying, Mapping and Geoinformation) (No. 2015NGCM). We would also like to thank three anonymous reviewers for very helpful comments.

Author Contributions: P.J. conceived the study and designed the analyses. P.J. and W.G. undertook collection of the reports and extraction of data. J.W. and P.J. implemented the data processing and analysis. All authors contributed to the writing and editing of the report.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. 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](#)] [[PubMed](#)]
2. 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](#)]
3. Mufinda, F.C.; Boinas, F.; Nunes, C. Prevalence and factors associated with human brucellosis in livestock professionals. *Rev. Saude Publica* **2017**, *51*, 57. [[CrossRef](#)] [[PubMed](#)]
4. Sanjuan-Jimenez, R.; Colmenero, J.D.; Morata, P. Lessons learned with molecular methods targeting the bcbp-31 membrane protein for diagnosis of human brucellosis. *Clin. Chim. Acta* **2017**, *469*, 1–9. [[CrossRef](#)] [[PubMed](#)]
5. Franco, M.P.; Mulder, M.; Gilman, R.H.; Smits, H.L. Human brucellosis. *Lancet Infect. Dis.* **2007**, *7*, 775–786. [[CrossRef](#)]

6. 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](#)]
7. Cascio, A.; Iaria, C. Brucellar aortitis and brucellar spondylitis. *Lancet Infect. Dis.* **2015**, *15*, 145. [[CrossRef](#)]
8. Riethmuller, P.C.; Chalermkao, N. *Some Issues Associated with the Livestock Industries of the Asia-Pacific Region*; Food and Agriculture Organization of the United Nations, Regional Office for Asia and the Pacific Animal Production and Health Commission for Asia and the Pacific (APHCA): Bangkok, Thailand, 2002.
9. Shang, D.; Xiao, D.; Yin, J. Epidemiology and control of brucellosis in China. *Vet. Microbiol.* **2002**, *90*, 165–182.
10. 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](#)]
11. Liu, Z.G.; Di, D.D.; Wang, M.; Liu, R.H.; Zhao, H.Y.; Piao, D.R.; Tian, G.Z.; Fan, W.X.; Jiang, H.; Cui, B.Y.; et al. MLVA genotyping characteristics of Human *Brucella melitensis* isolated from Ulanqab of Inner Mongolia, China. *Front. Microbiol.* **2017**, *8*, 6. [[CrossRef](#)] [[PubMed](#)]
12. 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*, e004470. [[CrossRef](#)] [[PubMed](#)]
13. Merow, C.; Smith, M.J.; Silander, J.A. A practical guide to maxent for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **2013**, *36*. [[CrossRef](#)]
14. Royle, J.A.; Chandler, R.B.; Yackulic, C.; Nichols, J.D. Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods Ecol. Evol.* **2012**, *3*, 545–554. [[CrossRef](#)]
15. Warren, D.; Seifert, S. Ecological niche modelling in maxent: The importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* **2011**, *21*, 235–342. [[CrossRef](#)]
16. Merow, C.; Silander, J.A. A comparison of maxlike and maxent for modelling species distributions. *Methods Ecol. Evol.* **2014**, *5*, 215–225. [[CrossRef](#)]
17. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **1943**, *5*, 115–133. [[CrossRef](#)]
18. Rosenblatt, F. *The Perceptron: A Theory of Statistical Separability in Cognitive Systems*; Cornell Aeronautical Laboratory, Inc.: Buffalo, NY, USA, 1958.
19. Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **1980**, *36*, 93–202. [[CrossRef](#)]
20. Lippmann, R.P. An introduction to computing with neural nets. *IEEE ASSP Mag.* **1987**, *4*, 4–22. [[CrossRef](#)]
21. Werbos, P.J. Beyond regression: New tools for prediction and analysis in the behavioral sciences. *J. Shandong Inst. Commer. Technol.* **1975**, *29*, 65–78.
22. Zhang, Z.; Deng, Z.; Rusch, K.A. Modeling fecal coliform bacteria levels at gulf coast beaches. *Water Qual. Expo. Health* **2015**, *7*, 255–263. [[CrossRef](#)]
23. Baronchelli, A.; Ferrericancho, R.; Pastorsatorras, R.; Chater, N.; Christiansen, M.H. Networks in cognitive science. *Trends Cogn. Sci.* **2013**, *17*, 348–360. [[CrossRef](#)] [[PubMed](#)]
24. Wang, J.; Deng, Z. Modeling and prediction of oyster norovirus outbreaks along gulf of Mexico coast. *Environ. Health Perspect.* **2015**, *124*, 627–633. [[CrossRef](#)] [[PubMed](#)]
25. Bastiaansen, W.G.; Molden, D.J.; Makin, I.W. Remote sensing for irrigated agriculture: examples from research and possible applications. *Agric. Water Manag.* **2010**, *46*, 137–155. [[CrossRef](#)]
26. Jia, P.; Nie, Y.; Song, G. Detection of underground remains by remote sensing and geophysics. In Proceedings of the IEEE 18th International Conference on Geoinformatics, Beijing, China, 18–20 June 2010.
27. Jia, P.; Nie, Y.; Yang, L. Recognition and extraction of the ancient sites covered by thick vegetation in Hainan Province of China. *IEEE Geosci. Remote Sens. Symp.* **2010**. [[CrossRef](#)]
28. Zhu, J.; Jia, P.; Nie, Y. Analysis of the ancient river system in Loulan period in Lop Nur region. *Proc. SPIE* **2010**, *8203*, 820313.
29. 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](#)] [[PubMed](#)]
30. Jia, P.; Anderson, J.D.; Leitner, M.; Rheingans, R. High-resolution spatial distribution and estimation of access to improved sanitation in Kenya. *PLoS ONE* **2016**, *11*, e0158490. [[CrossRef](#)] [[PubMed](#)]
31. Jia, P.; Stein, A. Using remote sensing technology to measure environmental determinants of non-communicable diseases. *Int. J. Epidemiol.* **2017**, *46*, 1343–1344. [[CrossRef](#)] [[PubMed](#)]

32. Jia, P.; Joyner, T.A.; Sun, Y. Short-term associations between accumulated rainfall and atmospheric moisture during landfall of three Atlantic hurricanes. *Geogr. Bull.* **2014**, *55*, 49–62.
33. Xue, Y. The impact of desertification in the Mongolian and the Inner Mongolian grassland on the regional climate. *J. Clim.* **1996**, *9*, 2173–2189. [[CrossRef](#)]
34. 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](#)] [[PubMed](#)]
35. Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; Centro Internacional de Agricultura Tropical (CIAT). *Global Rural-Urban Mapping Project, Version 1 (GRUMPV1): Population Density Grid*; NASA Socioeconomic Data and Applications Center (SEDAC): Palisades, NY, USA, 2011.
36. Jia, P.; Sankoh, O.; Tatem, A.J. Mapping the environmental and socioeconomic coverage of the indepth international health and demographic surveillance system network. *Health Place* **2015**, *36*, 88–96. [[CrossRef](#)] [[PubMed](#)]
37. Nelson, A. Estimated Travel Time to the Nearest City of 50,000 or More People in Year 2000. Available online: <http://bioval.jrc.ec.europa.eu/products/gam/> (accessed on 10 January 2012).
38. Scharlemann, J.P.; Benz, D.; Hay, S.I.; Purse, B.V.; Tatem, A.J.; Wint, G.W.; Rogers, D.J. Global data for ecology and epidemiology: A novel algorithm for temporal Fourier processing MODIS data. *PLoS ONE* **2008**, *3*, e1408. [[CrossRef](#)] [[PubMed](#)]
39. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2005**, *25*, 1965–1978. [[CrossRef](#)]
40. Baptista-Rosas, R.C.; Hinojosa, A.; Riquelme, M. Ecological niche modeling of *Coccidioides* spp. in western North American deserts. *Ann. N. Y. Acad. Sci.* **2007**, *1111*, 35–46. [[CrossRef](#)] [[PubMed](#)]
41. Gonzalez, C.; Wang, O.; Strutz, S.E.; Gonzalez-Salazar, C.; Sanchez-Cordero, V.; Sarkar, S. Climate change and risk of leishmaniasis in North America: Predictions from ecological niche models of vector and reservoir species. *PLoS Negl. Trop. Dis.* **2010**, *4*, e585. [[CrossRef](#)] [[PubMed](#)]
42. IBM Corp. *IBM SPSS Statistics for Windows, Version 22.0*; IBM Corp: Armonk, NY, USA, 2013.
43. MathWorks. *Global Optimization Toolbox: User's Guide (2013a)*; MathWorks: Natick, MA, USA, 2013.
44. Grandin, T.; Deesing, M.J. Chapter 4—Genetics and behavior during handling, restraint, and herding. In *Genetics and the Behavior of Domestic*, 2nd ed.; Academic Press: San Diego, CA, USA, 2014; pp. 115–158.
45. Karger, D.N.; Conrad, O.; Böhrner, J.; Kawohl, T.; Kreft, H.; Soriaaauza, R.W.; Zimmermann, N.; Linder, H.P.; Kessler, M. Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **2017**, *4*, 170122. [[CrossRef](#)] [[PubMed](#)]
46. Jia, P.; Gaughan, A.E. Dasymeric modeling: A hybrid approach using land cover and tax parcel data for mapping population in alachua county, florida. *Appl. Geogr.* **2016**, *66*, 100–108. [[CrossRef](#)]
47. Jia, P.; Qiu, Y.; Gaughan, A.E. A fine-scale spatial population distribution on the high-resolution gridded population surface and application in Alachua County, Florida. *Appl. Geogr.* **2014**, *50*, 99–107. [[CrossRef](#)]
48. Gaughan, A.E.; Stevens, F.R.; Linard, C.; Jia, P.; Tatem, A.J. High resolution population distribution maps for southeast Asia in 2010 and 2015. *PLoS ONE* **2013**, *8*, e55882. [[CrossRef](#)] [[PubMed](#)]
49. Tatem, A.J.; Gaughan, A.E.; Stevens, F.R.; Patel, N.N.; Jia, P.; Pandey, A.; Linard, C. Quantifying the effects of using detailed spatial demographic data on health metrics: A systematic analysis for the AfriPop, AsiaPop, and AmeriPop projects. *Lancet* **2013**, *381*, S142. [[CrossRef](#)]

