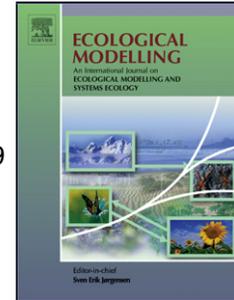


Journal Pre-proof

Identifying the Spatio-temporal Risk Variability of Avian Influenza A H7N9 in China

Ping Zhang, Jianwen Wang, Peter M. Atkinson



PII: S0304-3800(19)30315-1
DOI: <https://doi.org/10.1016/j.ecolmodel.2019.108807>
Article Number: 108807
Reference: ECOMOD 108807
To appear in: *Ecological Modelling*
Received Date: 31 May 2018
Revised Date: 8 August 2019
Accepted Date: 28 August 2019

Please cite this article as: Zhang P, Wang J, Atkinson PM, Identifying the Spatio-temporal Risk Variability of Avian Influenza A H7N9 in China, *Ecological Modelling* (2019), doi: <https://doi.org/10.1016/j.ecolmodel.2019.108807>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.

Identifying the Spatio-temporal Risk Variability of Avian Influenza A H7N9 in China

Ping Zhang^{1*}, Jianwen Wang¹, Peter M. Atkinson²

¹ *College of Geo-Exploration Science and Technology, Jilin University, Changchun, 130061, China*

* zp@jlu.edu.cn

² *Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster, LA1 4YW, UK*

* Corresponding author: zp@jlu.edu.cn

Highlights

- Spatial random effects on the relative risk of human H7N9 infection in four seasons
- Different covariates had a significant influence in different seasons
- Parameter settings in the Bayesian hierarchical model had an influence

Abstract

The outbreak of H7N9 epidemic in human has seasonal changes. However, up to now there is no research on the spatial-temporal variation characteristics of the relative risk of human H7N9 infection, and the covariate combination that has a greater impact on the relative risk of human H7N9 infection in different seasons. This study used China as the study area to predict the seasonal relative risk of human H7N9 infection through a Bayesian hierarchical conditional autoregressive model (BHCAR), which including five covariates (population density, number of live poultry markets, average precipitation, average temperature, and average relative humidity), seasonal random effects, and spatial random effects. Moreover, the sensitivity of the Bayesian hierarchical model (BH) to predict the seasonal relative risk of human H7N9 infection by changing the parameter settings of the BH prior distribution was analyzed. It was found that the relative risk of human H7N9 infection in spring and winter had spatial random effects, but not in summer and autumn. In spring, autumn and winter, the combination of population density and the number of live poultry markets had a significant influence on the relative risk of human H7N9 infection. In summer, however, the relative risk of human H7N9 infection was largely affected by population density,

the number of live poultry markets, average precipitation and average temperature. Further, the standard deviation of the normal distribution to which the covariate coefficient in the BH was subject seemed to have an influence on the prediction and fitting effect of the seasonal relative risk of human infection with H7N9.

Keywords: BHCAR; Human infection with H7N9; Spatial-temporal risk variation; Sensitivity analysis

1. Introduction

Since February 19, 2013, when the first patient infected with the novel influenza A H7N9 virus from an avian source showed symptoms, 1567 laboratory-confirmed human infections have been reported to the World Health Organization (WHO) in China as of 2 March 2018, including at least 615 deaths (WHO, 2018). The outbreaks of influenza A H7N9 infection in humans normally has taken on a seasonal trend with five major epidemic waves so far (WHO, 2018). The H7N9 virus is a low pathogenic avian influenza A virus, it can cause an asymptomatic infection in birds (Chen et al., 2013; Liu et al., 2013), but it can normally cause severe respiratory illness in humans (Tang and Chen, 2013). Another immediately notable feature of H7N9 is the rapid accumulation of laboratory-confirmed cases of infection in human beings, even though phylogenetic and epidemiological evidence suggests that transmission is mainly zoonotic (Cowling et al., 2013).

So far there is no evidence of efficient or sustained human-to-human H7N9 virus transmission (Fang et al., 2013), but the H7N9 virus can occasionally transmit from person to person (Qi et al., 2013; Poovorawan, 2014). In particular, three factors of avian influenza A H7N9 make it possible to be the next pandemic strain of influenza, which including the virus has the ability to cause disease in humans; there is little immunity to the virus within the population; the virus has the capacity for sustained human-to-human transmission (Tanner et al., 2015). If H7N9 becomes transmissible in wild migratory birds like H5N1, it will likely spread to many countries. Therefore, the spatial transmission pattern of novel H7N9 poses an epidemic threat for China, and the globe, especially for human beings (Zhang et al., 2015). Thus, the avian influenza

A H7N9 is an unusually dangerous for humans.

Some methods have been applied to map the spatial risk distribution of human infections with avian influenza A H7N9 in China, such as boosted regression tree models (BRT) (Fang et al., 2013; Gilbert et al., 2014; Li et al., 2015), ecological and case-control methods (Ge et al., 2017), ecological niche modeling (ENM) (Xu et al., 2016), Logistic regression model (Fuller et al., 2014; Dong et al., 2015), species distribution models (SDMs) (Bui et al., 2017), Bayesian model (Hu et al., 2015). Among these, most of which were spatially predictive of human infection with avian influenza A H7N9 (Fang et al., 2013; Fuller et al., 2014; Gilbert et al., 2014; Dong et al., 2015; Li et al., 2015; Xu et al., 2016; Bui et al., 2017; Ge et al., 2017), very few research work explored the spatio-temporal variability of the risk of human infection with avian influenza A H7N9 (Hu et al., 2015). So it is still unclear how the probabilities of occurrence of human infection with avian influenza A H7N9 varying spatially and temporally in China and which underlying risk factors are involved.

Through the numerous statistical techniques that are now available and increased acceptance of statistical methodology (Unkel et al., 2012), statistical models have become increasingly popular in public health. Among these, BH are especially popular in the analysis of spatial and spatio-temporal infectious diseases (Marínez-Beneito et al., 2008; Banks et al., 2012; Zou et al., 2012; Zhang et al., 2014; Zou et al., 2014). BH allow for multiple sources of data, decomposing complex problems into a subset of simpler problems governed by simple rules of probability, and accounting for parameter uncertainty; specifically, they have the ability to consider scientifically meaningful structures and *a priori* knowledge regarding the model parameters (Arab et al., 2008; Zhang et al., 2014). Although some work was done to illustrate the spatio-temporal distribution of the risk of H7N9 infection in China with a BHCAR (Hu et al., 2015), this was done for only one city. This paper had two objectives. One was to apply a BHCAR to identify the spatial-temporal variation in relative risk of human H7N9 incidence by season at the county level for China, and to find the optimal combination of relative variables (influence factors) fitting the relative risk of human infections in H7N9. The other one was to conduct the sensitivity analysis of the parameters of the

BH to see how different parameter settings affect the results of the relative risk of human H7N9 infection.

2. Methods

2.1. Data sources

Data on laboratory-confirmed avian influenza A H7N9 human cases reported in China from 19 February 2013 to 16 February 2015 were provided by the World Health Organization (WHO, 2018), social media and Zhang et al. (2015). Basic geographic data were provided by the Data Sharing Infrastructure of Earth System Science (2015). Population data at the county level were provided by Demographic Statistics of the County in China of 2000 and 2010 (Public Security Bureau, 2000, 2010). Monthly average precipitation and temperature data were obtained from the WorldClim-Global Climate Data (2011), which were provided at a 30 second resolution and in grid format. Monthly average relative humidity data were provided by National Meteorological Information Center (2018), which was a text format of the climatological annual data set of meteorological stations in China. The locations of live poultry markets in 2012 in mainland China were obtained from Baidu Map (2016), which provides a Chinese location-based network digital map search service.

2.2. Data analysis & data processing

The `seasonplot()` function in R (Hyndman and Athanasopoulos, 2018) was used here to plot the number of H7N9 human cases that have been reported in China between February 2013 and December 2014. The horizontal axis of Fig. 1 was the month, and the vertical axis was the number of cases of human H7N9 infection in each month. It clearly showed the potential seasonal patterns of changes in the number of H7N9 human cases in China in the year of 2013 and 2014 (Fig. 1). The seasonal patterns differed in winter and spring, while summer and autumn showed similar trends (Fig. 1). There was a sharp fall in the winter of 2014, while a sharp rise in that of 2013. By contrast, there was a sharp rise followed by a sharp decline in the spring of 2013, while a slow rise followed by a slow decline in that of 2014 (Fig. 1). In the

summer and autumn of 2013 and 2014, however, the number of human cases of H7N9 was relatively small, showing neither increasing nor decreasing trend (Fig. 1).

Further, most human infections occurred in spring and winter, and less cases occurred in summer and autumn (Fig. 2). This may confirm the seasonal characteristics of human infection with H7N9 influenza virus, and winter and spring were the peak seasons for human infection with H7N9. Besides, most of the H7N9 human cases were clustered in the east and southeast regions of China, and the remaining cases occurred in the neighboring regions of those. The spatial distribution of H7N9 human cases had a tendency to spread from east and southeast to the adjacent and other areas (Fig. 2).

For the monthly average precipitation and temperature data, they were extracted within the China boundary first, and then converted from radian unit into length unit, resampled to a 100 m cell size. Finally, monthly average precipitation and temperature at the county scale were obtained through the zonal statistics tool in ArcGIS using the county as the defined zones. For the monthly average relative humidity data, semi-variance analysis was carried out on the meteorological stations and then the monthly average relative humidity surfaces in China at the resolution of 100m grid were obtained by Kriging. Using the tool of Zonal Statistics as Table in ArcGIS as well, the monthly average relative humidity at the county scale were obtained. Based on the population statistics of each county in 2000 and 2011, the population statistics of each county in China in 2013 and 2014 were deduced, and then combined with the area of each county, the population density of each county in China in 2013 and 2014 were calculated. In this study, four seasons were divided by astronomical means: spring from March to May, summer from June to August, autumn from September to November, and winter from December to February. The seasonal average temperature, precipitation, relative humidity and population density were obtained by adding the monthly average temperature, precipitation, relative humidity and population density of the corresponding months, and then taking the average. .

2.3. *Monthly incidence rate*

The incidence rate is the number of new cases per population at risk in a given

time period (Rothman et al., 2008). In epidemiology, incidence rate can be replaced by the attack rate, which is the biostatistical measure of frequency of morbidity, or speed of spread, in an at risk population. It is used in hypothetical predictions and during actual outbreaks of disease. Population at risk is defined as one that has no immunity to the attacking pathogen which can be either a novel pathogen or an established pathogen. It is used to project the number of cases to expect during an epidemic. Here, monthly incidence rate is to be used, calculated as follows:

$$I = \frac{C_m}{P_m} \times K \quad (1)$$

where, I represents the monthly incidence rate, m stands for a certain month, C_m represents the number of new H7N9 cases in the month, P_m is the average population number at risk for the month, and K could be 100%, 1,000 ‰, 10,000/10,000, 100,000/100,000 and so on. Here K is assigned with 100,000/100,000. P_m in Equation (1) for the average population number for the month can be calculated through the following two equations.

$$P_{t_0} \times (1 + x_a)^{t_1 - t_0} = P_{t_1} \quad (2)$$

$$P_{t_1} \times (1 + x_b)^{12} = P_{t_1+1} \quad (3)$$

where, P_{t_0} represents the average population number for the start year, P_{t_1} is the average population number for the terminal year, x_a and x_b denotes the annual population growth rate and the monthly population growth rate, respectively, t_0 and t_1 represents the start year and the terminal year, respectively.

2.4. Bayesian hierarchical modeling

It was assumed that there was a true (unobserved) process underlying the observed epidemic of H7N9 infection in human cases, which we incorporate into the

framework of a BH. BH modelling is a flexible approach, which can accommodate both spatial effects and temporal dynamics in a unified framework, as well as incorporate situation-specific characteristics such as covariate information for certain populations or for H7N9 infection in humans (Zou et al., 2012). It was also assumed that the spatial component of the model is required due to dependence among adjacent counties and the temporal aspect is a direct result of a plausible Markov structure. Typically, BH consist of three components: the data model (i.e. the conditional distribution of the data given hidden processes and parameters); the process model (i.e. the conditional distribution of the hidden processes given the parameters); and the parameter model (i.e., the prior distribution of the parameters) (Zhuang and Cressie, 2014).

Data model

Here, the data model relates the number of H7N9 cases in each county and month to the probability of occurrence. The observed number of H7N9 cases at the county level is assumed to occur independently and follow a Poisson distribution.

$$y_{ij} \sim \text{Poisson}(\theta_{ij} e_i) \quad (4)$$

where, y_{ij} is the observed number of H7N9 cases in county j for month i ; e_{ij} is the expected incidence rate of H7N9 infection in humans in county j for month i ; θ_{ij} is the relative risk of H7N9 infection in humans in county j for month i .

Process model

The process component of the model relates the probability of observing a given number of H7N9 cases y_{ij} to the linear trend in time and three random effect terms that capture the non-spatial, spatial, and spatio-temporal patterns in the H7N9 case data. To model the Poisson distributed outbreak probability, a standard log-linear transform is used. Thus, the probability of observing a given number of H7N9 cases in each county and in each month is modelled similarly to Fransworth et al. (2009). This can be written as follows:

$$\text{Model 1: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + v_{ij} + \gamma_{ij}$$

$$\text{Model 2: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \alpha_{3ij}\beta_{3ij} + \gamma_{ij}$$

$$\text{Model 3: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \alpha_{3ij}\beta_{3ij} + \alpha_{4ij}\beta_{4ij} + \gamma_{ij}$$

$$\text{Model 4: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \alpha_{3ij}\beta_{3ij} + \alpha_{4ij}\beta_{4ij} + \alpha_{5ij}\beta_{5ij} + \gamma_{ij}$$

$$\text{Model 5: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \gamma_{ij}$$

$$\text{Model 6: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \alpha_{3ij}\beta_{3ij} + \gamma_{ij}$$

$$\text{Model 7: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \alpha_{3ij}\beta_{3ij} + \alpha_{4ij}\beta_{4ij} + \gamma_{ij}$$

$$\text{Model 8: } \log(\theta_{ij}) = \alpha_0 + \alpha_{1ij}\beta_{1ij} + \alpha_{2ij}\beta_{2ij} + \alpha_{3ij}\beta_{3ij} + \alpha_{4ij}\beta_{4ij} + \alpha_{5ij}\beta_{5ij} + \gamma_{ij}$$

where, v_{ij} is an independent random effect term associated with season i in county

j ; γ_{ij} is the spatial random effect term for season i in county j ; α_0 is the intercept;

$\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and α_5 represent the coefficients for the five covariates (seasonal average precipitation, seasonal average temperature, seasonal average relative humidity, population density and number of live poultry markets) for the season i in the county j , respectively; $\beta_1, \beta_2, \beta_3, \beta_4$ and β_5 represent the five covariates respectively.

Parameter model

All parameters within BH are stochastic and are assigned appropriate probability distributions (Lawson, 2013). Hence, a single parameter value is simply one possible realization of the possible set of values for the parameter, the probability of which, before integration, is defined by the prior distribution (Lawson, 2013). Note that one interpretation of prior distributions is that they provide additional 'data' for a problem and so they can be used to improve estimation or identification of parameters.

Here, the prior distributions for all model parameters in the hierarchy have also been specified since our model is fully Bayesian. The independent random effect v_{ij} ,

corresponds to a latent process operating independently in each county and year.

$$v_{ij} \sim N(0, \sigma_v^2 I), \text{ for } i=1,2,3 \dots n, \text{ and } j=1,2,3 \dots k \quad (5)$$

where, I is an $n_i \times n_j$ indicator matrix, n is the total number of county in China, i is the season, j is the county.

This component models the overall unstructured heterogeneity in the data by assuming no relationship among neighboring counties or years, but with a variance that is common to all counties and years. The spatial component, modeled by γ_{ij} describes the unobserved transmission process among counties in each year contributing to the observed spatial structure of outbreaks each year. Here γ_{ij} is an intrinsic Gaussian CAR model. For each season i and county j , the CAR model states that γ_{ij} is related to the γ terms for the neighboring counties, and each neighbouring county is independent of all other counties outside the local neighborhood. Specifically, let the set of neighbors of county j be denoted by $j+$. Then, for each season i in county j , the conditional relationship can be written as follows:

$$\gamma_{ij} \sim N\left(\frac{1}{n_{j+}} \sum_{i \in j+} \gamma_i, \frac{\sigma_\gamma^2}{n_{j+}}\right) \quad (6)$$

where, n_{j+} is the number of neighbors of county j and σ_γ^2 is the variance common to all counties. Thus, the conditional mean of γ_{ij} is simply the average value of its

neighbors γ_{ij+} , with conditional variance $\frac{\sigma_\gamma^2}{n_{j+}}$ inversely proportional to the number of

neighbors. In the conditional mean in equation (6), the neighboring counties are equally weighted so that all neighbors of county j influence it equally.

Spatial variation in our model is limited to counties sharing a border. The spatial unit of this study adopts the vector format of China's counties. The county as an

administrative unit does vary in size, but it is consistent with the real space shape and size of each county. Using the county as a spatial unit to study the spatial random effects, the research results are more realistic.

However, there are no a priori restrictions on specifying the neighborhood structure or county weights. The linear trend in time, βt_i , is the rate ratio between two consecutive years and provides the estimate of the rate at which the number of H7N9 cases changes between years. Let $\beta \sim N(0, 10000)$ be the prior distribution for the time trend parameter. Finally, the space-time interaction random effect $\gamma_{ij} t_i$ models the interaction between space and time during the course of the epidemic.

All models and Markov chain Monte Carlo (MCMC) procedure for each model were run using the OpenBUGS 3.2.3 software downloaded from the OpenBugs website (<http://www.openbugs.net/>). The MCMC procedure for each model was run for 10,000 iterations after a burn-in period of 10,000 iterations to ensure convergence of all model parameters. In this study, convergence diagnostics, autocorrelation statistics were done in OpenBugs. By observing the iterative history graphs and Gibbs dynamic sampling trace graphs of related variables in OpenBugs after updating the BHCAR, it was found that the iterative history curves and Gibbs dynamic sampling trace curves tended to be stable, indicating that the convergence was good. In the autocorrelation function graphs of related variables, the results of autocorrelation function of related variables were soon close to 0, indicating that Monte Carlo chain had tended to converge. Part of the iterative history graphs, Gibbs dynamic sampling trace graphs, and the autocorrelation analysis graphs were included in the appendix of this paper.

2.5. Deviance Information Criterion (DIC)

The deviance information criterion (DIC) proposed by Spiegelhalter et al. (2002) is used as a method of model comparison for complex hierarchical models fitted within a Bayesian framework, which is a Bayesian generalization of the Akaike information criterion (AIC) (Fung et al., 2014). Using DIC aims to identify models that best explain the observed data, but with the expectation that they are likely to minimize

uncertainty about observations generated in the same way (Spiegelhalter et al., 2002). In brief, a measure of model fit is used in conjunction with a measure of model complexity to form the criteria upon which model comparisons can be made (Deeth et al., 2015)

Adapting the notation of Spiegelhalter et al. (2002), model fit is estimated by the posterior mean deviance, $\overline{D(\theta)}$, where $D(\theta) = -2\ln(f(D|\theta))$ is the deviance. Models that are more complex are likely to have a lower deviance and, therefore, a better (i.e. lower) measure of model fit. To account for this, a penalty term for the level of model complexity is used in the calculation of the DIC (Deeth et al., 2015). The model complexity is measured by estimating the effective number of parameter, p_D , which is given as:

$$p_D = \overline{D(\theta)} - D(\hat{\theta}) \quad (7)$$

$$DIC = \overline{D(\theta)} + p_D$$

where $D(\hat{\theta})$ is the deviance calculated at $\hat{\theta}$, a point estimate of θ . Often the posterior mean of the model parameters is used for $\hat{\theta}$. Then the DIC is calculated as a sum of model fit and model complexity.

Smaller values of DIC indicate a better-fitting model. However, note that the DIC is not intended for identification of the 'correct' model, but rather merely as a method of comparing a collection of alternative formulations (all of which may be incorrect), as with other model choice criteria (Zhu and Carlin, 2000). Here, the values of DIC were calculated in the OpenBUGS software.

3. Results

3.1. Seasonal spatio-temporal risk variability of human H7N9 infection

Because the cases number of human infected with H7N9 was relatively small in summer and autumn (Fig. 2) and its incidence rate was relatively low, the BHCAR cannot be operated in summer and autumn. However, in spring and winter, the cases number of human infected with H7N9 was relatively high (Fig. 2) and its incidence rate

was relatively high, so the BHCAR can be operated normally in spring and winter (Fig. 3). Under the influence of spatial random effect, the relatively high risk areas of the human infection with H7N9 were scattered in the central and eastern regions in spring, but showed different spatial distribution patterns in winter (concentrated in the south and southeast, or scattered throughout the country, or scattered in the south of the Yangtze river) (Fig. 3).

In the absence of spatial random effect, the relatively high risk areas of human H7N9 infection were scattered in the spring, and the high risk areas might be in the northeast, north and northwest of China (Fig. 4). The relative risk of human H7N9 infection in summer was very low, with only an occasional high risk area near Shanghai (Fig. 5). Although the relative risk of human H7N9 infection in autumn was relatively low, it was higher than that in summer. The high-risk areas only appeared sporadically in the west, central and south, while the eastern part of Taiwan island and the southern part of Hainan island were relatively at greater risk (Fig. 6). The relative risk of human H7N9 infection in winter varied from high to low, with high risk areas appearing sporadically in central and western China, or in patches in the west and north (Fig. 7).

3.2. Effects of covariates on the spatio-temporal risk variability of human H7N9 infection

If the spatial random effect was taken into account, the DIC value in the model fitting of the relative risk of human H7N9 infection in model 1 with two covariates in spring and winter were the minimum (141.4 and 165.0, respectively) (Table 1), indicating that the population density and the number of live poultry markets had the largest impact on the relative risk of human H7N9 infection in spring and winter. If the spatial random effect was not taken into account, the DIC value of model 7 with four covariates in the model fitting of the relative risk of human H7N9 infection in summer was the minimum (25.07) (Table 1), indicating that the population density and the number of live poultry markets, the average temperature and precipitation had the largest impact on the relative risk of human H7N9 infection in summer.

Similarly, the DIC value in the fitting of the assessment of the relative risk of human H7N9 infection in model 5 with two covariates in autumn was the minimum (21.85) (Table 1), indicating that the population density and the number of live poultry market had the largest impact on the relative risk of human H7N9 infection in autumn. In addition, the results of this study also found that the model fitting effect with spatial random effect was not as good as the model without spatial random effect (Table 1), which may imply that the relative risk of human infection with H7N9 was not strongly spatially random.

3.3. Sensitivity analysis based on different prior distributions

Based on the prior distribution of different parameters, the sensitivity analysis of the prediction results of the relative risk of human H7N9 infection by the BH was conducted in this study. In this sensitivity analysis, the standard deviation of the covariate coefficients in the BH were changed to conform to the normal distribution, and the covariate coefficients were subject to the normal distribution of different shapes, so as to see the change in the relative risk of human infection with H7N9. Here, three different standard deviations were used, so that the normal distribution curves had three different shapes. The first standard deviation was 0.0005 (Table 1) based on Lawson's study (Lawson, 2013). Based on the first standard deviation, the other two different standard deviations were set as well, one larger than the first standard deviation (1), and the other smaller than the first standard deviation ($1.0E-8$) (Table 2), so that the normal distribution curve was one flat, and one thin and tall.

The standard deviation of the normal distribution to which the covariate coefficient was subject seemed to have an influence on the prediction and fitting effect of the relative risk of human infection with H7N9. When the standard deviation set to 1, the DIC value of the relative risk of human infection with H7N9 in spring was the lowest (133.0); when the standard deviation set to 0.0005, the DIC value of the relative risk of human infection with H7N9 in summer was the lowest (25.07); and when the standard deviation was set $1.0E-8$, the DIC value of relative risk of human infection with H7N9 in autumn and in winter were the minimum (21.74 and 163.3, respectively)

(Table 2). But how about the size of the standard deviation affects the BH fitting effect, this had not yet been determined in this study.

4. Discussion

Here, a BH with five covariates was used to predict the relative risk of human H7N9 infection in the four seasons. The results showed that the combination of two covariates (population density and the number of live poultry markets) had a great influence on the relative risk of human H7N9 infection in spring, autumn and winter. In summer, the combination of four covariates (population density, number of live poultry markets, average precipitation and average temperature) had a greater impact on the relative risk of human H7N9 infection. This is inconsistent with previous studies (Fang et al., 2013; Fuller et al., 2014), so it is debatable which covariates have a truly statistically significant relationship to the relative risk of human H7N9 infection. However, another result was obtained in this study, that is, the combination of covariates that had an impact on the relative risk of human H7N9 infection was different in different seasons, and the combination of covariates that had an impact on the relative risk of human H7N9 infection throughout the year was determined, which was not found by previous studies. This result may vary with the number of model covariates. In future research, other covariates should be explored in the Bayesian model framework.

Besides, the CAR allows disease risk to spread smoothly within a cluster with a disjoint multiplicative jump between clusters. Removing the CAR component of the model would assume a constant disease risk within a cluster, which is unlikely to be true in general (Anderson et al., 2014). BHCAR can only be running when human H7N9 cases number was more and its incidence rate was higher like in spring and winter time here, it showed that the spatial distribution of the relative risk of human H7N9 infection in spring and winter had spatial random effects. But when human H7N9 infection cases number was less and its incidence rate was lower in summer and autumn, the BHCAR cannot be running, it illustrated the spatial distribution of the relative risk of human H7N9 infection in summer and autumn had no spatial random

effects. This study could reveal something in spatial correlation for the seasonal relative risk of human H7N9, but not all. As far as the method of this study is concerned, it cannot be compared with field data. Further, it was also found that the operation of the BHCAR was greatly affected by the parameter setting. It can only run under the setting of some parameters, rather than under the setting of any parameters. For example, it only worked when the standard deviation of the normal distribution of the covariate coefficients was 0.0005, but when the standard deviation was set to 1 or $1.0E-8$, the BHCAR cannot run at all.

The BHCAR used here indicates that environmental factors, seasonal random effects and spatial random effects together affect the spatio-temporal dynamic changes of diseases. But for how the environmental factors (including human population densities et al) would "quantitatively" determine spatial-temporal dynamics of disease in-depth analysis, it can be considered that environmental factors may interact with the time and place of disease onset and change with the change of time and space. It may be done in future research. All the environmental factor coefficients in eight models of this study are assumed to follow the normal distribution. By changing the standard deviation of the normal distribution, the model fitting results of the seasonal relative risk of human H7N9 infection can be obtained, thus influencing the spatio-temporal dynamic change of the relative risk of human H7N9 infection.

5. Conclusion

In this study, a BHCAR was applied to explore the seasonal and spatial variation of the relative risk of human H7N9 infection, and the influence of different covariance combinations in different seasons on the relative risk of human H7N9 infection. From the perspective of ecology, this study obtained the seasonal variation characteristics of relative risk of human H7N9 infection, that is, the relative risk of human H7N9 infection in spring and winter has a spatial random effect, while the relative risk of human H7N9 infection in summer and autumn does not have a spatial random effect. In summer, population density, the number of live poultry markets, average precipitation and average temperature have a significant impact on the relative risk of

human H7N9 infection, while in spring, autumn and winter, only population density and the number of live poultry markets have a significant impact. These results provide a reasonable theoretical basis for the effective seasonal prevention and control of human H7N9 infection and the maintenance of normal health level of the human population.

By setting different parameters in the prior distribution of the BH, the sensitivity analysis was also conducted here to explore the changes in predicting the relative risk of human H7N9 infection with the BH. It was found that the standard deviation of the normal distribution to which the covariate coefficient was subject seemed to have an influence on the prediction and fitting effect of the relative risk of human infection with H7N9. Therefore, the setting of prior distribution parameters of the BH has an impact on the prediction and fitting results of the relative risk of human H7N9 infection. In order to achieve the optimal prediction and fitting, it is necessary to find the right parameters as far as possible.

Acknowledgements

This study was supported by the 13th five-year science and technology program of Jilin provincial education department (grant No. JJKH20180162KJ). We would like to thank the reviewers and editors for their good comments and suggestions on the revision of this paper, which has greatly helped to improve the quality of this paper.

Appendix

References

- Anderson, C., Lee, D., Dean, N., 2014. Identifying clusters in Bayesian disease mapping. *BIostatistics*. 15(3), 457-469.
- Arab, A., Hooten, M. B., Wikle, C. K., 2008. Hierarchical spatial models. in: Shekhar, S. and Xiong, H. (Eds.), *Hierarchical Spatial Models in Encyclopedia of GIS*. Springer, New York, pp. 425.
- Baidu Map, 2016. [http:// map.baidu.com/](http://map.baidu.com/) (accessed 1 March 2015).
- Banks, D., Datta, G., Karr, A., Lynch, J., Niemi, J., Vera, F., 2012. Bayesian CAR models

- for syndromic surveillance on multiple data streams: theory and practice. *INFORM FUSION*. 13, 105-116.
- Bui, C. M., Gardner, L., MacIntyre, R., Sarkar, S., 2017. Influenza A H5N1 and H7N9 in China: A spatial risk analysis. *PLOS ONE*. 12(4), e0174980.
- Chen, Y., Liang, W., Yang, S., Wu, N., Gao, H., Sheng, J., Yao, H., Wo, J., Fang, Q., Cui, D., Li, Y., Yao, X., Zhang, Y., Wu, H., Zheng, S., Diao, H., Xia, S., Zhang, Y., Chan, K. H., Tsoi, H. W., Teng, J. L., Song, W., Wang, P., Lau, S. Y., Zheng, M., Chan, J. F., To, K. K., Chen, H., Li, L., Yuen, K. Y., 2013. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. *Lancet*. 381(9881), 1916-1925.
- Cowling, B. J., Jin, L., Lau, E. H. Y., Liao, Q., Wu, P., Jiang, H., Tsang, T. K., Zheng, J., Fang, V. J., Chang, Z., Ni, M. Y., Zhang, Q., Ip, D. K. M., Yu, J., Li, Y., Wang, L., Tu, W., Meng, L., Wu, J. T., Luo, H., Li, Q., Shu, Y., Li, Z., Feng, Z., Yang, W., Wang, Y., Leung, G. M., Yu, H., 2013. Comparative epidemiology of human infections with avian influenza A H7N9 and H5N1 viruses in China: a population-based study of laboratory-confirmed cases. *Lancet*. 382, 129-137.
- Data Sharing Infrastructure of Earth System Science, 2015. <http://www.geodata.cn/> (accessed 1 March 2015).
- Deeth, L. E., Deardon, R., Gillis, D. J., 2015. Model choice using the deviance information criterion for latent conditional individual-level models of infectious disease spread. *Epidemiol. Methods*. 4(1), 47-68.
- Dong, W., Yang, K., Xu, Q-L., Yang, Y-L., 2015. A predictive risk model for A (H7N9) human infections based on spatial-temporal autocorrelation and risk factors: China, 2013-2014. *Int J Env Res Pub He*. 12, 15204-15221.
- Fang, L. Q., Li, X. L., Liu, K., Li, Y. J., Yao, H. W., Liang, S., Yang, Y., Feng, Z. J., Gray, G. C., Cao, W. C., 2013. Mapping spread and risk of avian influenza A (H7N9) in China. *Sci Rep*. 3, 2722.
- Fransworth, M. L. and Ward, M. P., 2009. Identifying spatio-temporal patterns of transboundary disease spread: examples using avian influenza H5N1 outbreaks. *Vet Res*. 40, 20.

- Fuller, T., Havers, F., Xu, C., Fang, L-Q., Cao, W-C., Shu, Y., Widdowson, M-A., Smith, T. B., 2014. Identifying areas with a high risk of human infection with the avian influenza A (H7N9) virus in East Asia. *J Infection*. 69, 174-181.
- Fung, T., Wang, J. J. J., Seneta, E., 2014. The deviance information criterion in comparison of normal mixing models. *INT STAT REV*. 82(3), 411-421.
- Ge, E., Zhang, R., Li, D., Wei, X., Wang, X., Lai, P-C., 2017. Estimating risks of inapparent avian exposure for human infection: avian influenza virus A (H7N9) in Zhejiang Province, China. *Sci. Rep.* 7, 40016.
- Gilbert, M., Golding, N., Zhou, H., William Wint, G. R., Robinson, T. P., Tatem, A. J., Lai, S., Zhou, S., Jiang, H., Guo, D., Huang, Z., Messina, J. P., Xiao, X., Linard, C., Van Boeckel, T. P., Martin, V., Bhatt, S., Gething, P. W., Farrar, J. J., Hay, S. I., Yu, H., 2014. Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. *Nat. Commun.* 5, 4116.
- Hu, W. B., Zhang, W. Y., Huang, X. D., Clements, A., Mengersen, K., Tong, S., 2015. Weather variability and influenza A (H7N9) transmission in Shanghai, China: A Bayesian spatial analysis. *Environ Res.* 136, 405-412.
- Hyndman, R.J., Athanasopoulos, G., 2018. *Forecasting: principles and practice*, 2nd ed. OTexts: Melbourne, Australia. OTexts.com/fpp2 (accessed 25 June 2019).
- Lawson, A. B., 2013. Chapter 2. Bayesian Inference and Modelling. in Lawson, A. B. (Eds.), *Bayesian Disease Mapping: Hierarchical Modelling in Spatial Epidemiology* (Second Edition). Chapman and Hall/CRC, pp. 19-34.
- Li, X-L., Yang, Y., Sun, Y., Chen, W-J., Sun, R-X., Liu, K., Ma, M-J., Liang, S., Yao, H-W., Gray, G. C., Fang, L-Q., Cao, W-C., 2015. Risk distribution of human infections with avian influenza H7N9 and H5N1 virus in China. *Sci. Rep.* 5, 18610.
- Liu, D., Shi, W., Shi, Y., Wang, D., Xiao, H., Li, W., Bi, Y., Wu, Y., Li, X., Yan, J., Liu, W., Zhao, G., Yang, W., Wang, Y., Ma, J., Shu, Y., Lei, F., Gao, G. F., 2013. Origin and diversity of novel avian influenza A H7N9 viruses causing human infection: phylogenetic, structural, and coalescent analyses. *Lancet*. 381(9881), 1926-1932.
- Marínez-Beneito, M. A., Conesa, D., López-Quílez, A., López-Maside, A., 2008. Bayesian Markov switching models for the early detection of influenza epidemics. *Stat Med*.

27, 4455-4468.

National Meteorological Information Center, 2018. <http://www.data.cma.cn/> (accessed on 27 December 2018).

OpenBugs website, 2016. <http://www.openbugs.net/> (accessed on 2 January 2016).

Poovorawan, Y., 2014. Epidemic of avian influenza A (H7N9) virus in China. *Pathog. Glob. Health.* 108, 169-170.

Public Security Bureau, 2000. *Demographic Statistics of the County in China 2000.* Beijing: Mass Press.

Public Security Bureau, 2010. *Demographic Statistics of the County in China 2010.* Beijing: Mass Press.

Qi, X., Qian, Y. H., Bao, C. J., Guo, X. L., Cui, L. B., Tang, F. Y., Ji, H., Huang, Y., Cai, P. Q., Lu, B., Xu, K., Shi, C., Zhu, F. C., Zhou, M. H., Wang, H., 2013. Probable person to person transmission of novel avian influenza A (H7N9) virus in Eastern China 2013; epidemiological investigation. *BMJ.* 347, f4752.

Rothman, K. J., Greenland, S., Lash, T. L., 2008. *Modern Epidemiology*, third ed. Lippincott Williams & Wilkins.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *J R STAT SOC B.* 64(4), 583-639.

Tang, R. B., Chen, H. L., 2013. An overview of the recent outbreaks of the avian-origin influenza A (H7N9) virus in the human. *J. Chin. Med. Assoc.* 76, 245-248.

Tanner, W. D., Toth, D. J. A., Gundlapalli, A. V., 2015. The pandemic potential of avian influenza A (H7N9) virus: a review. *Epidemiol. Infect.* 143, 3359-3374.

Unkel, S., Farrington, C. P., Garthwaite, P. H., 2012. Statistical methods for the prospective detection of infectious disease outbreaks: a review. *J. R. Statist. Soc. A.* 175, 49-82.

World Health Organization. Global Health Observatory (GHO) data, 2018. <http://www.who.int/gho/en/> (accessed 20 April 2018).

WorldClim - Global Climate Data, 2011. <http://www.worldclim.org/> (accessed 1 March 2015).

Xu, M., Cao, C., Li, Q., Jia, P., Zhao, J., 2016. Ecological Niche Modeling of risk factors

- for H7N9 human infection in China. *Int J Env Res Pub He.* 13, 600.
- Zhang, Y., Arab, A., Cowling, B. J., Stoto, M. A., 2014. Characterizing influenza surveillance systems performance: application of a Bayesian hierarchical statistical model to Hong Kong surveillance data. *BMC PUBLIC HEALTH.* 14, 850.
- Zhang, Y., Shen, Z., Ma, C., Jiang, C., Feng, C., Shankar, N., Yang, P., Sun, W., Wang, Q., 2015. Cluster of human infections with avian influenza A (H7N9) cases: a temporal and spatial analysis., *Int. J. Environ. Res. Public Health.* 12, 816-828.
- Zhu, L. and Carlin, B. P., 2000. Comparing hierarchical models for spatio-temporally misaligned data using the deviance information criterion. *Stat Med.* 19, 2265-2278.
- Zhuang, L. L. and Cressie, N., 2014. Bayesian hierarchical statistical SIRS models. *Stat Methods Appl.* 23, 601-646.
- Zou, J., Karr, A. F., Banks, D., Heaton, M., Datta, G., Lynch, J., Vera, F., 2012. Bayesian methodology for the analysis of spatial-temporal surveillance data. *Stat Anal Data Min.* 5, 194-204.
- Zou, J., Karr, A. F., Datta, G., Lynch, J., Grannis, S., 2014. A Bayesian spatio-temporal approach for real-time detection of disease outbreaks: a case study. *BMC MED INFORM DECIS.* 14, 108.

Fig. 1. Seasonal plot of monthly cases number of human infection with H7N9 influenza virus in China during the period of Feb 2013- Dec 2014.

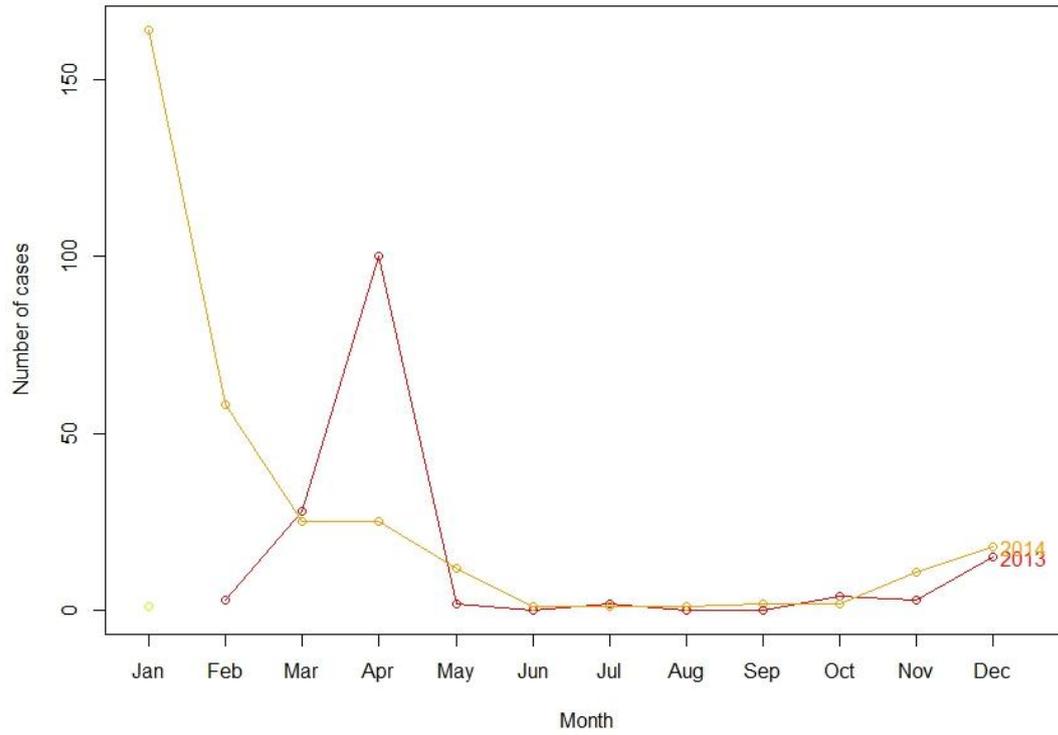
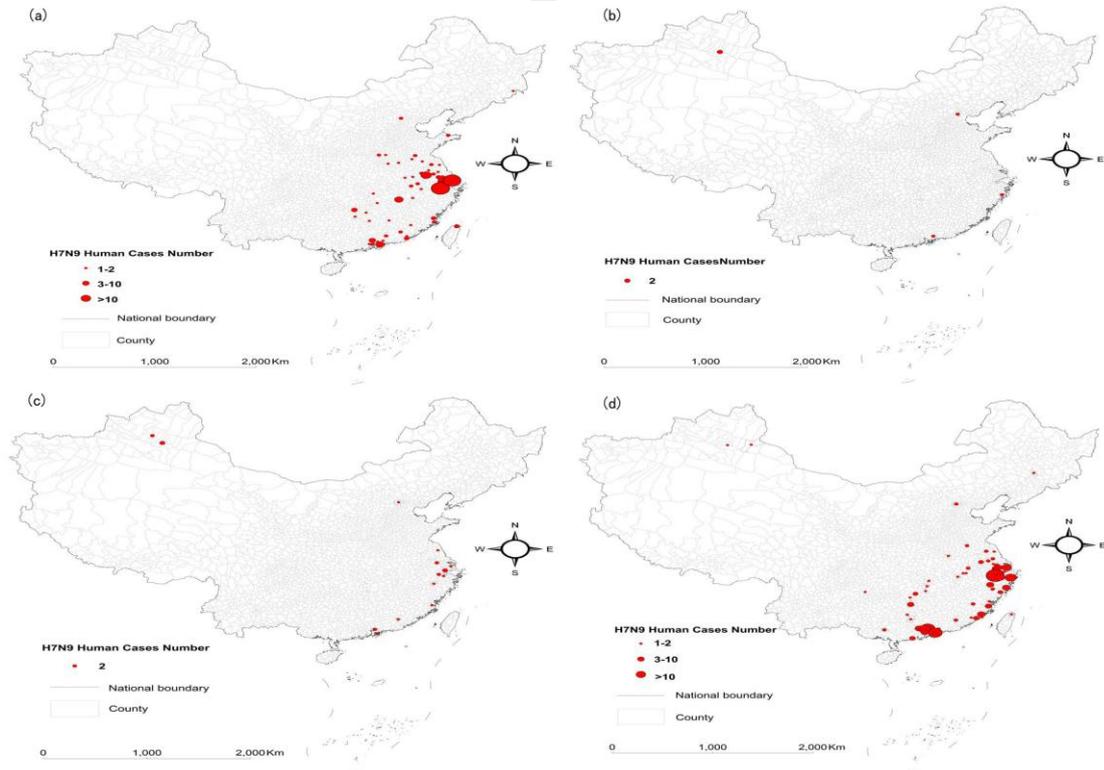


Fig. 2. Seasonal spatial distribution of the cases number of human infection with H7N9 influenza virus in China during the period of 1 Jan 2013 – 31 Dec 2014. (a) Spring; (b) Summer; (c) Autumn; (d) Winter.



H7N9 Human Cases Number

1-2

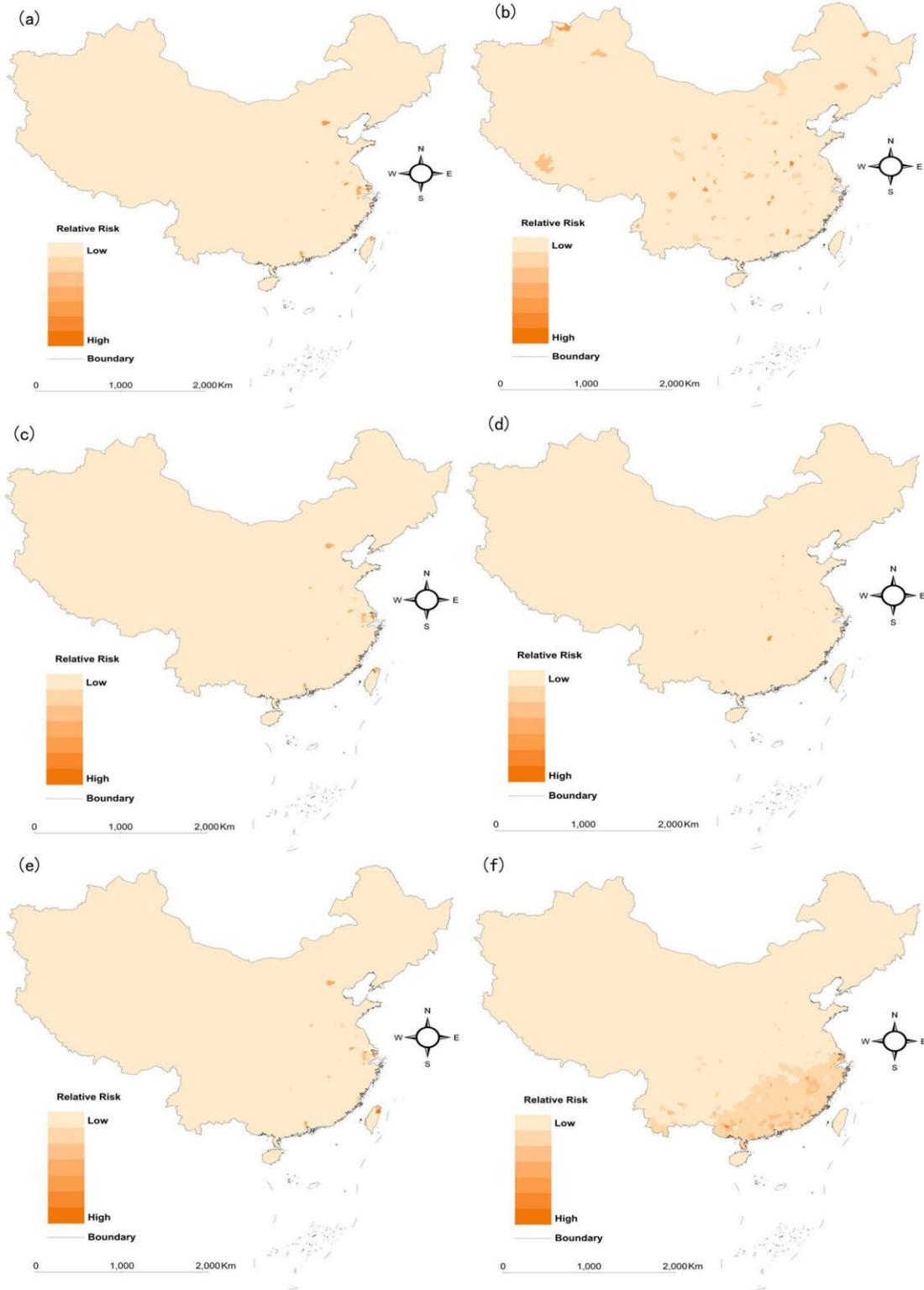
3-10

>10

National boundary

County

Fig. 3. Relative risk distribution of human infection with H7N9 influenza virus at county scale in China with Bayesian hierarchical CAR model. (a) Result of model 1 in spring with DIC being 141.4; (b) Result of model 1 in winter with DIC being 165.0; (c) Result of model 2 in spring with DIC being 146.6; (d) Result of model 2 in winter with DIC being 169.6; (e) Result of model 3 in spring with DIC being 147.4; (f) Result of model 3 in winter with DIC being 165.6.



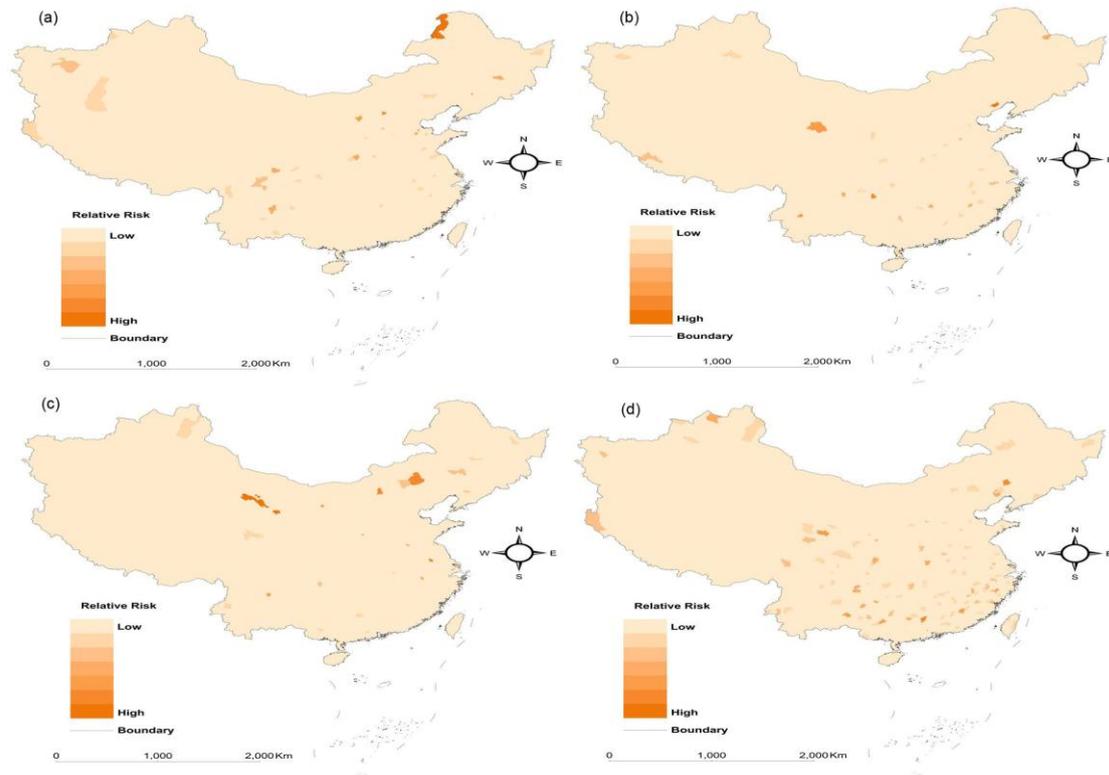
Relative Risk

Low

High

Boundary

Fig. 4. Relative risk distribution of human infection with H7N9 influenza virus at county scale in China in spring with Bayesian hierarchical model. (a) Result of model 5 with DIC being 133.0; (b) Result of model 6 with DIC being 137.4; (c) Result of model 7 with DIC being 138.2; (d) Result of model 8 with DIC being 135.9.



Relative Risk

Low

High

Boundary

Fig. 5. Relative risk distribution of human infection with H7N9 influenza virus at county scale in China in summer with Bayesian hierarchical model. (a) Result of model 5 with DIC being 25.66; (b) Result of model 6 with DIC being 26.05; (c) Result of model 7 with DIC being 25.07; (d) Result of model 8 with DIC being 25.1.

Relative Risk

Low

High

Boundary

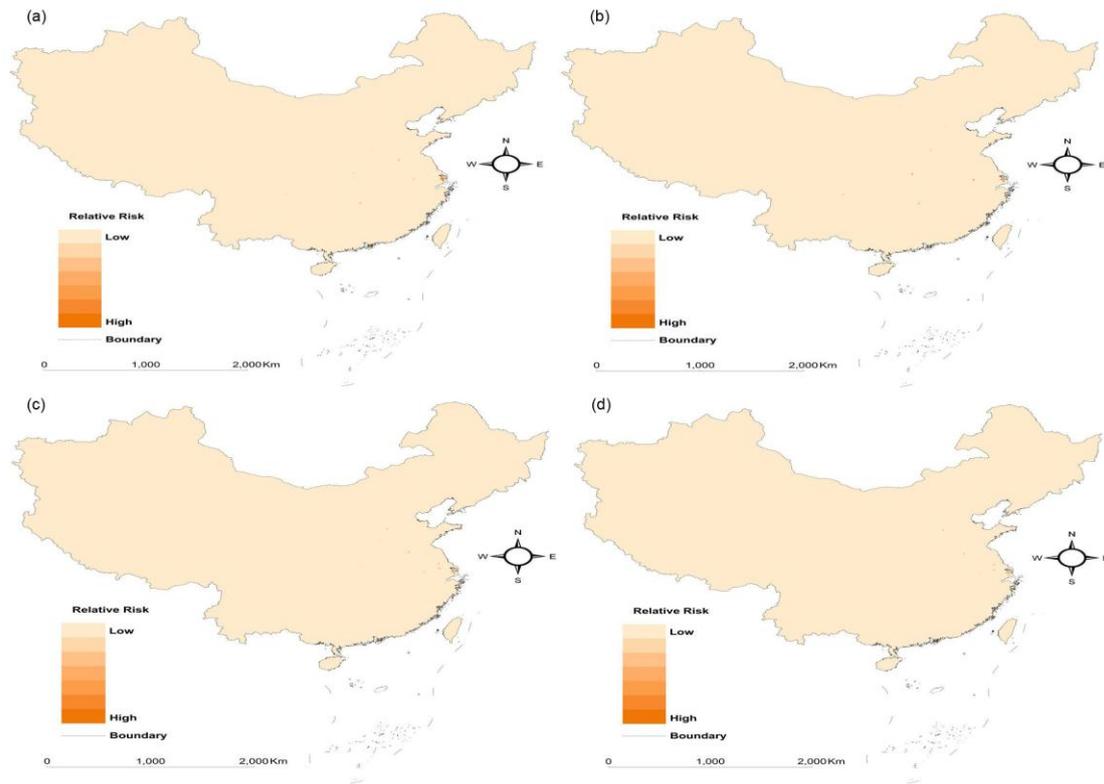
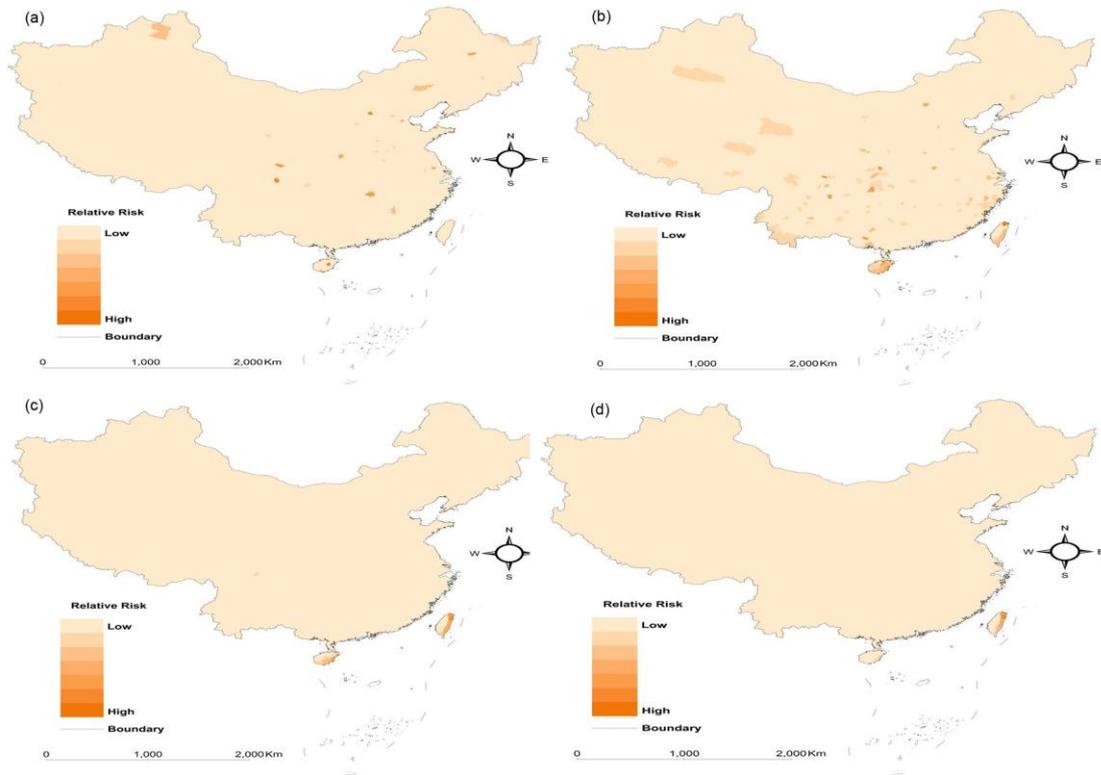


Fig. 6. Relative risk distribution of human infection with H7N9 influenza virus at county scale in China in autumn with Bayesian hierarchical model. (a) Result of model 5 with DIC being 21.74; (b) Result of model 6 with DIC being 23.48; (c) Result of model 7 with DIC being 22.94; (d) Result of model 8 with DIC being 25.13.



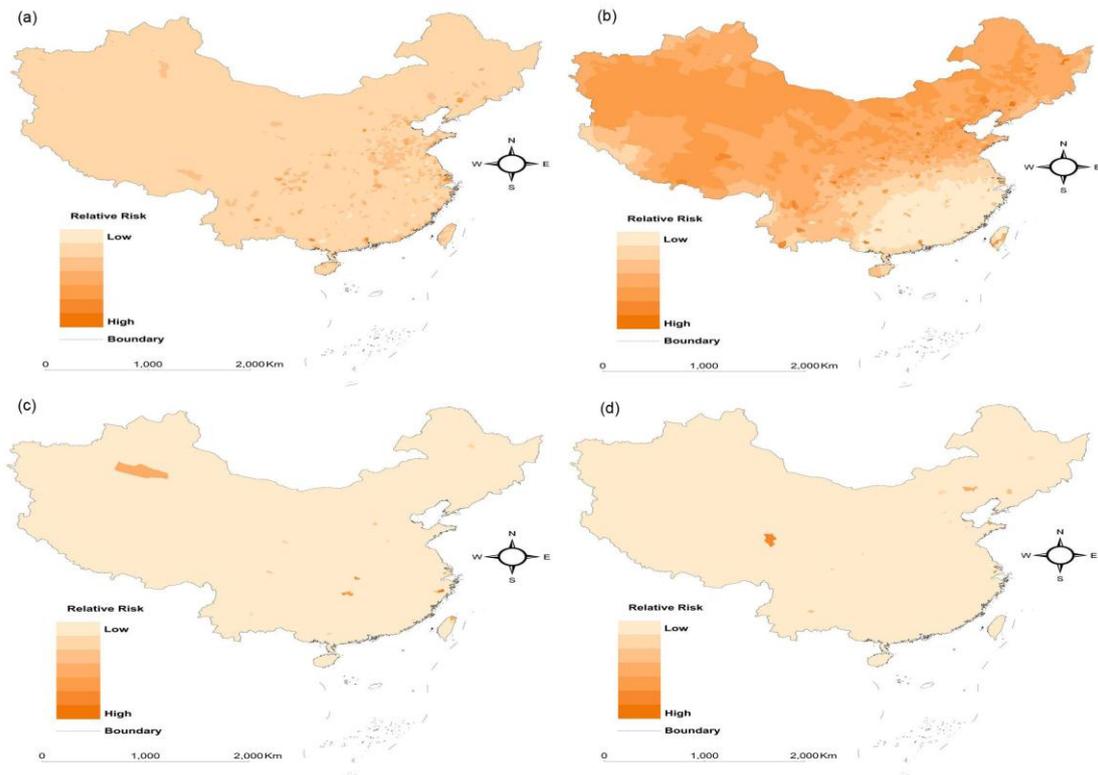
Relative Risk

Low

High

Boundary

Fig. 7. Relative risk distribution of human infection with H7N9 influenza virus at county scale in China in winter with Bayesian hierarchical model. (a) Result of model 5 with DIC being 163.3; (b) Result of model 6 with DIC being 168.4; (c) Result of model 7 with DIC being 172.9; (d) Result of model 8 with DIC being 166.2.



Relative Risk

Low

High

Boundary

Fig. A.1. History plot for the parameter of the relative risk of human infection with H7N9 for the first 15 iterations.

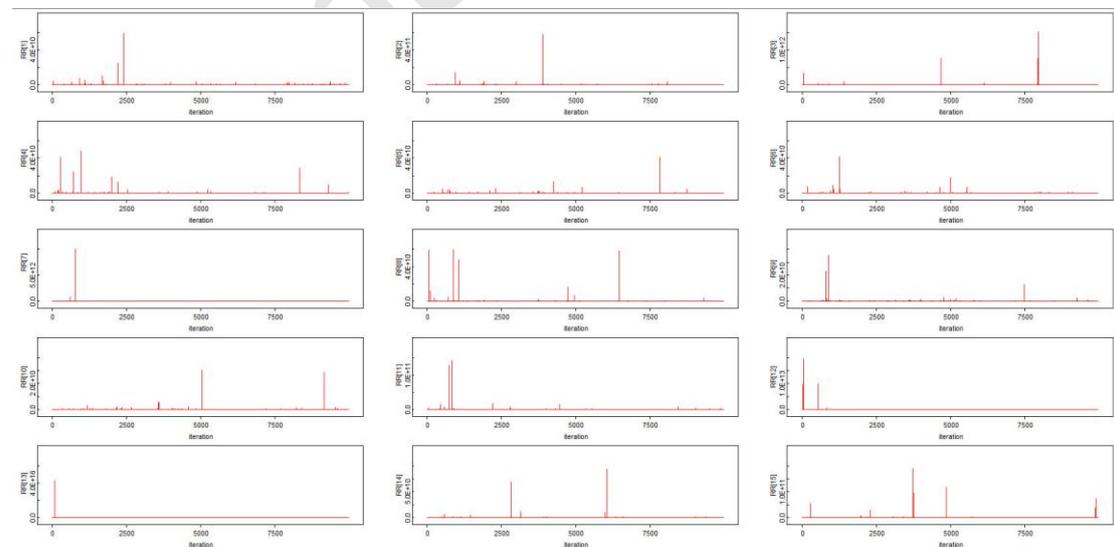


Fig. A.2. Trace plot for the parameter of the relative risk of human infection with H7N9

for the first 47 iterations.

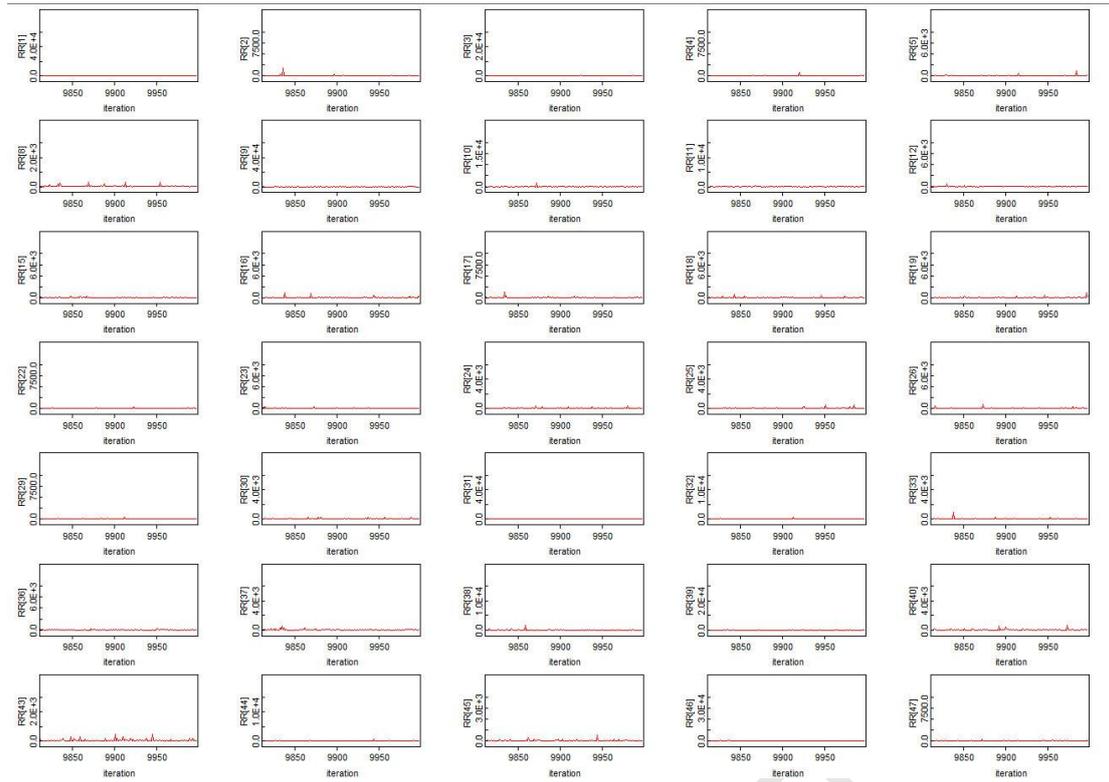


Fig. A.3. Autocorrelation plot for the parameter of the relative risk of human infection with H7N9 for the first 40 iterations.

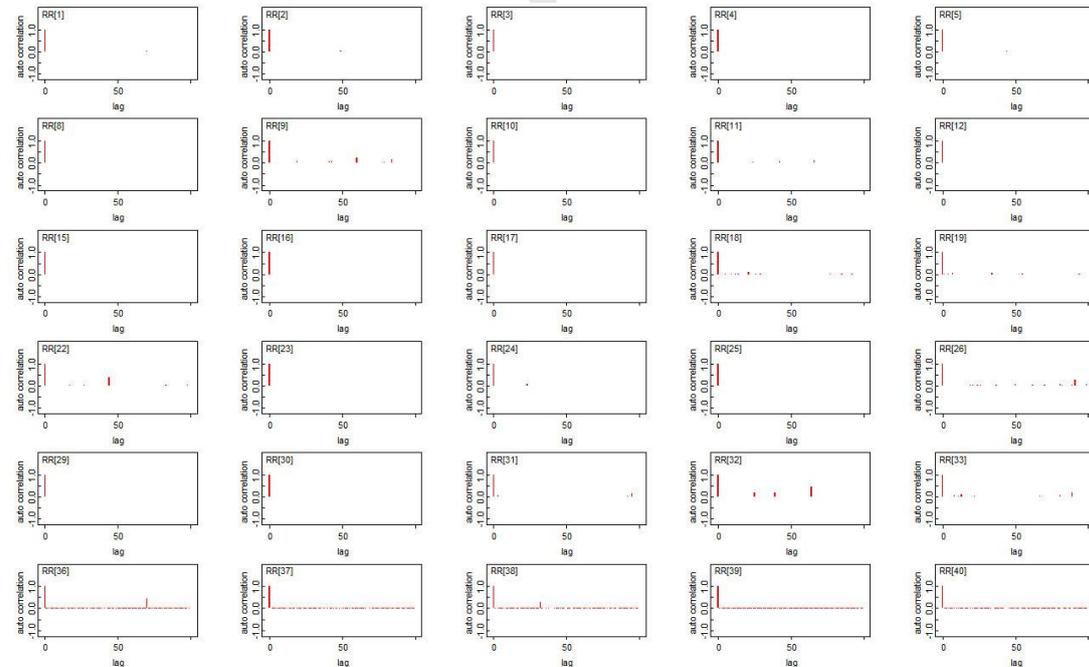


Table 1. The DIC, covariates and prior distribution of eight models.

Model	Covariates	CAR	Prior distribution	DIC
Model 1	Population density, live poultry market	With	$\alpha_1, \alpha_2 \sim N(0, 0.0005)$	141.4 (Spring); 165.0 (Winter)
Model 2	Population density, live poultry market, precipitation	With	$\alpha_1, \alpha_2, \alpha_3 \sim N(0, 0.0005)$	146.6 (Spring); 169.6 (Winter)
Model 3	Population density, live poultry market, precipitation, temperature	With	$\alpha_1, \alpha_2, \alpha_3, \alpha_4 \sim N(0, 0.0005)$	147.4 (Spring); 165.6 (Winter)
Model 4	Population density, live poultry market, precipitation, temperature, relative humidity	With	$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5 \sim N(0, 0.0005)$	145.9 (Spring); 164.3 (Winter)
Model 5	Population density, live poultry market	Without	$\alpha_1, \alpha_2 \sim N(0, 0.0005)$	133.1 (Spring); 25.66 (Summer); 21.85 (Autumn); 163.4 (Winter)
Model 6	Population density, live poultry market, precipitation	Without	$\alpha_1, \alpha_2, \alpha_3 \sim N(0, 0.0005)$	137.4 (Spring); 26.31 (Summer); 23.87 (Autumn); 168.4 (Winter)
Model 7	Population density, live poultry market, precipitation, temperature	Without	$\alpha_1, \alpha_2, \alpha_3, \alpha_4 \sim N(0, 0.0005)$	138.4 (Spring); 25.07 (Summer); 24.1 (Autumn); 173.8 (Winter)
Model 8	Population density, live poultry market, precipitation, temperature, relative humidity	Without	$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5 \sim N(0, 0.0005)$	136.2 (Spring); 25.1 (Summer); 25.81 (Autumn); 167.0 (Winter)

Table 2. The DIC and covariates of four Bayesian hierarchical models without CAR with two different prior distribution

Model	Covariates	CAR	Prior distribution	DIC
Model 5	Population density, live poultry market	Without	$\alpha_1, \alpha_2 \sim N(0, 1)$	133.0 (Spring); 25.8 (Summer); 23.03 (Autumn);

				163.6 (Winter)
Model 5	Population density, live poultry market	Without	$\alpha_1, \alpha_2 \sim N(0, 1.0E-8)$	133.1 (Spring); 25.79 (Summer); 21.74 (Autumn); 163.3 (Winter)
Model 6	Population density, live poultry market, precipitation	Without	$\alpha_1, \alpha_2, \alpha_3 \sim N(0, 1)$	137.3 (Spring); 26.05 (Summer); 24.52 (Autumn); 168.6 (Winter)
Model 6	Population density, live poultry market, precipitation	Without	$\alpha_1, \alpha_2, \alpha_3 \sim N(0, 1.0E-8)$	137.4 (Spring); 29.73 (Summer); 23.48 (Autumn); 168.4 (Winter)
Model 7	Population density, live poultry market, precipitation, temperature	Without	$\alpha_1, \alpha_2, \alpha_3, \alpha_4 \sim N(0, 1)$	138.2 (Spring); 26.35 (Summer); 25.38 (Autumn); 172.9 (Winter)
Model 7	Population density, live poultry market, precipitation, temperature	Without	$\alpha_1, \alpha_2, \alpha_3, \alpha_4 \sim N(0, 1.0E-8)$	138.4 (Spring); 26.56 (Summer); 22.94 (Autumn); 174.1 (Winter)
Model 8	Population density, live poultry market, precipitation, temperature, relative humidity	Without	$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5 \sim N(0, 1)$	136.5 (Spring); 26.78 (Summer); 25.22 (Autumn); 166.2 (Winter)
Model 8	Population density, live poultry market, precipitation, temperature, relative humidity	Without	$\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5 \sim N(0, 1.0E-8)$	135.9 (Spring); 29.06 (Summer); 25.13 (Autumn); 167.5 (Winter)