Science of the Total Environment 648 (2019) 550–560

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Exploring spatiotemporal nonstationary effects of climate factors on hand, foot, and mouth disease using Bayesian Spatiotemporally Varying Coefficients (STVC) model in Sichuan, China

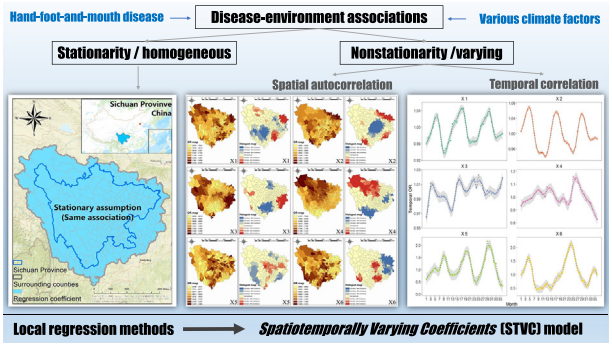
Chao Song [a,b,c], Xun Shi [b,*], Yanchen Bo [d], Jinfeng Wang [c,e], Yong Wang [c], Dacang Huang [c]

[a] School of Geoscience and Technology, Southwest Petroleum University, Chengdu, Sichuan 610500, China
[b] Department of Geography, Dartmouth College, Hanover, NH 03755, USA
[c] State Key Laboratory of Resources and Environmental Information System (LREIS), Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[d] State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[e] University of Chinese Academy of Sciences, Beijing 100049, China

HIGHLIGHTS

- The Bayesian STVC model is proposed for exploring the spatiotemporal nonstationary local exposure-response relationships.
- Both climate and socioeconomic factors were found to be associated with HFMD occurrence under spatiotemporal scales.
- Local temporal seasonal trends and spatial hot spots of HFMD occurrence and HFMD-climate associations were detected.
- The logistic STVC model spatialized odds ratio indicator into local ORs for mapping disease-environment associations.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 25 May 2018
Received in revised form 1 August 2018
Accepted 8 August 2018
Available online 09 August 2018

Editor: SCOTT SHERIDAN

Keywords:
HFMD epidemics
Disease-climate associations
Bayesian STVC model
Spatiotemporal nonstationarity
OR spatialization and mapping
Local regression

ABSTRACT

Background: Pediatric hand, foot, and mouth disease (HFMD) has generally been found to be associated with climate. However, knowledge about how this association varies spatiotemporally is very limited, especially when considering the influence of local socioeconomic conditions. This study aims to identify multi-sourced HFMD environmental factors and further quantify the spatiotemporal nonstationary effects of various climate factors on HFMD occurrence.

Methods: We propose an innovative method, named *spatiotemporally varying coefficients* (STVC) model, under the Bayesian hierarchical modeling framework, for exploring both spatial and temporal nonstationary effects in climate covariates, after controlling for socioeconomic effects. We use data of monthly county-level HFMD occurrence and data of related climate and socioeconomic variables in Sichuan, China from 2009 to 2011 for our experiments.

Results: Cross-validation experiments showed that the STVC model achieved the best average prediction accuracy (81.98%), compared with ordinary (68.27%), temporal (72.34%), spatial (75.99%) and spatiotemporal (77.60%) ecological models. The STVC model also outperformed these models in the Bayesian model evaluation. In this study, the STVC model was able to spatialize the risk indicator odds ratio (OR) into local ORs to represent spatial and temporal varying disease-climate relationships. We detected local temporal nonlinear seasonal trends and

* Corresponding author at: Department of Geography, Dartmouth College, Hanover, NH 03755, USA.
E-mail address: xun.shi@dartmouth.edu (X. Shi).

https://doi.org/10.1016/j.scitotenv.2018.08.114

spatial hot spots for both disease occurrence and disease-climate associations over 36 months in Sichuan, China. Among the six representative climate variables, temperature (OR = 2.59), relative humidity (OR = 1.35), and wind speed (OR = 0.65) were not only overall related to the increase of HFMD occurrence, but also demonstrated spatiotemporal variations in their local associations with HFMD.

Conclusion: Our findings show that county-level HFMD interventions may need to consider varying local-scale spatial and temporal disease-climate relationships. Our proposed Bayesian STVC model can capture spatiotemporal nonstationary exposure-response relationships for detailed exposure assessments and advanced risk mapping, and offers new insights to broader environmental science and spatial statistics.

1. Introduction

Pediatric hand, foot, and mouth disease (HFMD) is an emerging worldwide infectious disease that occurs mainly in children under five years old and can lead to death (Koh et al., 2016). HFMD is caused by a group of enteroviruses, such as enterovirus 71 (EV71) and coxsackievirus A16 (CV-A16) (Huang et al., 2018), and is readily transmitted through coughs, sneezes or contacts with infected feces or contaminated surfaces (Zhao et al., 2018). Environmental factors, such as climate, may affect virus survival and transmission environment, also human activity and contact frequency, which may greatly determine the eventual disease occurrences (Cheng et al., 2018; Shi and Kwan, 2015; Suminski et al., 2008). In China, HFMD is recognized as a leading infectious disease after it has been made statutorily notifiable in May 2008 (Xing et al., 2014). The morbidity of HFMD in China increased drastically from 37.6/100,000 persons in 2008 to 139.6/100,000 persons in 2013 (Liu et al., 2015).

Extensive studies suggest that HFMD is associated with environmental factors pertaining to climate, e.g. ambient temperature (Xiao et al., 2017; Zhu et al., 2015), humidity (Cheng et al., 2018; Onozuka and Hashizume, 2011), precipitation (Chen et al., 2014; Wang et al., 2011a), wind speed (Ma et al., 2010; Zhang et al., 2016c), air pressure (Wang et al., 2015), and sunshine (Zhang et al., 2016a) in many Asian countries (Lee et al., 2015), including China (Bo et al., 2014; Hu et al., 2012; Wang et al., 2011a), Japan (Onozuka and Hashizume, 2011), Vietnam (Nguyen et al., 2017), and South Korea (Kim et al., 2016). Climate environmental conditions not only affect the reproduction and transmission of the viruses causing HFMD, but also influence the behavior of children which may increase their contacts (Bélanger et al., 2009; Suminski et al., 2008), thus indirectly increasing opportunities of viral transmission among young children (Huang and Wang, 2018).

Recent studies further quantified the nonlinear risk effects of climate factors on HFMD (Chen et al., 2014; Wang et al., 2017; Xiao et al., 2017; Zhang et al., 2016c). For instance, Wang et al. found that a higher risk of HFMD was associated with temperatures ranging from 70 °F to 80 °F (Wang et al., 2011b); Xiao et al. found that the HFMD-temperature relationship was a nonlinear inverted V-shape curve (Xiao et al., 2017); Du et al. found that temperatures > 24.85 °C and a relative humidity between 80.59% and 82.55% lead to higher risk of HFMD (Du et al., 2016); Zhang et al. found that the HFMD risk peak occurred at humid values between 15 and 20 and 30–35 (Zhang et al., 2016b); Wang et al. found a nonlinear (inverted U) association between temperature and HFMD in summer, with a maximum morbidity at 27 °C (Wang et al., 2017).

However, most of those studies assume that the disease-environment association is stationary (Chen et al., 2010), i.e., the exposure-response association is homogeneous across the entire area and time period. In reality, the disease-environment association, especially for a communicable disease, is more likely to be nonstationary (varying), either spatially (Osei and Stein, 2017), temporally (Lee and Sheridan, 2018), or both, i.e., the associations vary among different spatial units and time frames. This kind of nonstationarity in HFMD-climate associations has been discussed either only in the spatial (Hu et al., 2012; Wang et al., 2016) or the temporal dimension (Liu et al., 2016; Yu et al., 2014) so far, and no studies have systematically investigated this problem within a spatiotemporal framework. In addition, while socioeconomic conditions may influence the climate effects on HFMD (Bo et al., 2014; Huang et al., 2014; Xu, 2017), few studies have incorporated socioeconomic factors in the HFMD-climate association assessment, especially from a spatiotemporal perspective.

Ecological regression models have commonly been used in environmental epidemiology to capture the risk association between disease and relevant covariates (Wakefield, 2006). These are fixed effect models that only estimate a single coefficient to represent the exposure-response association for each covariate based on the implicit assumption of stationarity. Given the complexity of the relationship between infectious diseases and environmental factors, nonstationarity impacts of neighborhood covariates are plausible (Leyk et al., 2012), especially for a large geographical region. A varying coefficients model that accounts for nonstationarity in the covariates will be able to give richer and more informative details (Hastie and Tibshirani, 1993). When geographic information on neighborhoods is available, the spatial nonstationary impacts of the covariates could be taken into account by the spatially varying coefficients (SVC) models (Gelfand et al., 2003). The common stationary regressions are called simple “global” models (Brunsdon et al., 1996), correspondingly, the nonstationary SVC models are called as “local” analysis methods in general (Wolf et al., 2018). Some studies have provided excellent overviews of SVC development (Banerjee et al., 2014; Finley et al., 2007). SVC models could readily be formulated as the latent random variable model within a Bayesian hierarchical modeling (BHM) framework (Lindgren and Rue, 2015). BHM is a multi-level statistical model in a hierarchical form that estimates the parameters of the posterior distribution using the Bayesian statistics method (Allenby and Rossi, 2006) and has been widely used in environmental epidemiology (Blangiardo et al., 2013).

Most Bayesian SVC models in practice only consider nonstationary spatial effects for only the intercept (unobserved residuals) (Blangiardo et al., 2013), but not the covariates (observed independent variables). More importantly, Bayesian SVC models are not spatiotemporal forms, and to our best knowledge, no such Bayesian spatiotemporal local regression model concerning nonstationarity has been proposed (Wolf et al., 2018).

To address the problems discussed above, we proposed a new method, named *spatiotemporally varying coefficients* (STVC) model, under the BHM framework, to quantify and characterize the spatiotemporally varying (nonstationarity) associations between exposures (climate variables) and response (HFMD occurrence). In this study, we use the county-level disease data in conjunction with the six aforementioned climate factors, as well as fourteen candidate socioeconomic factors in the Sichuan Province of China for a 36-month period from the years 2009 to 2011. We develop this Bayesian STVC model as an extended local regression method to the common global ecological regression approaches and SVC models, under a spatiotemporal form.

Our objective with this paper is three-fold. Firstly, to use the STVC model we developed to detect spatiotemporally nonstationary associations between environmental factors and disease outcomes, and compare it with the other four conventional models under a cross-validation design. Secondly, to identify HFMD risk factors accounting

for both climate and socioeconomic aspects under spatiotemporal scales. Thirdly, to quantify spatially and temporally varying disease distribution and disease-climate associations of HFMD at the local scale.

2. Material and methods

2.1. Study area and data

Sichuan Province is located in southwest China between 26.40°N and 33.68°N latitude, and 98.31°E and 107.99°E longitude. There are 135 county-level areal units in Sichuan Province. To reduce the edge effect, we included counties bordering Sichuan, into the study scope. As a result, our study area is comprised in total, of 243 county-level areal units (Fig. 1).

For the study area, we acquired county-level monthly data including HFMD occurrence, climate and socioeconomic variables for 36 months from the years 2009 to 2011. Our HFMD occurrence data in children aged between 0 and 9 years were from the China Information System for Disease Control and Prevention (CISDCP). The monthly climate data we used in this study was based on the raw data collected from 727 climate stations throughout China from the China Climate Data Sharing Service System (Bo et al., 2014). Data of yearly socioeconomic variables obtained from the China County Statistical Yearbook, China Statistical Yearbook for Regional Economy, and China City Statistical Yearbook (Song et al., 2018b). We included a total of six climates and fourteen socioeconomic variables as potential environmental-related factors of HFMD for this study (Additional file Table S1). We performed z-score standardization to the potential twenty factors to make them dimensionless.

2.2. Statistical methods

2.2.1. Odds ratio

Odds ratio (*OR*) is widely used to measure the risk of disease exposure to a determinant in epidemiology (Bland and Altman, 2000). The logistic regression model is the primary method to generalize *OR* beyond two binary variables (Harrell, 2001). For disease occurrence data of this case, the form of an ordinary logistic ecological regression

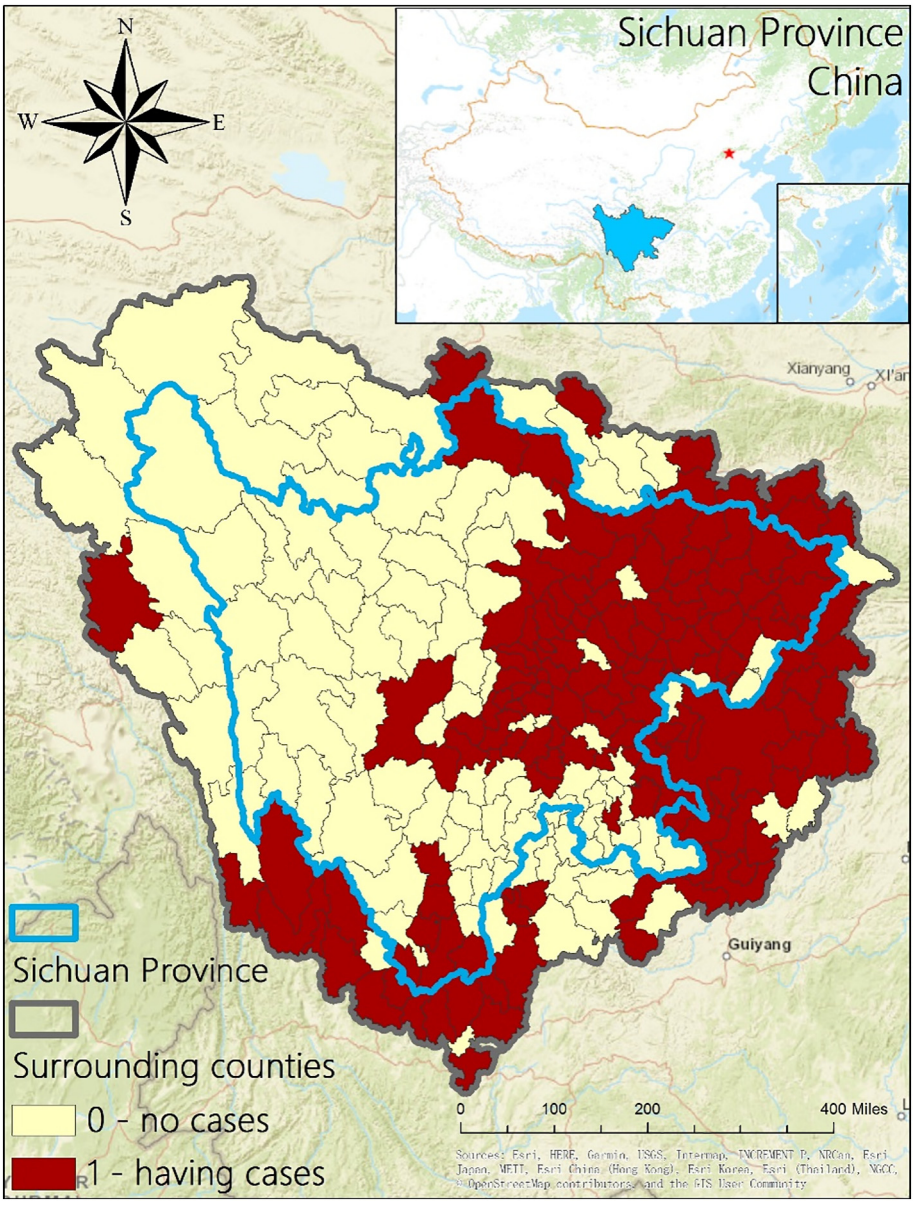


Fig. 1. Geographical distribution of reported HFMD occurrence in the study area in June 2009 (Blue border: Sichuan Province of China; Gray border: counties surrounding Sichuan). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

model is given as follows:

$$In\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \sum_k^m \beta_k C_k + \sum_j^n \beta_j SE_j \quad (1)$$

where i is the index of the records; P_i is the expected value of the dependent variable (probability of the occurrence of a disease); C_k is the k-th climate independent covariate; SE_j is the j-th socioeconomic independent covariate; β_0 is the intercept term; β_k and β_j are the estimated coefficients representing the exposure-response relationships between the dependent and independents.

Risk indicator *OR* could be obtained directly by the logistic regression model using $OR = e^{\beta}$. In epidemiology, an *OR* value higher than one indicates that the exposure variable is a positively correlated risk factor, lower than one means a negatively correlated risk factor and equal to one means an unrelated factor. The traditionally OR is an overall risk indicator for the entire study area, seldom spatialized for disease risk mapping (Best et al., 2005), not to mention for mapping disease-covariates relationships. One goal of this study is to localize the overall OR into a localized risk indicator to characterize the disease occurrence distribution and disease-environment relationships at both spatial and temporal scales. We refer to the traditional OR as "overall *OR*" to distinguish it from "local *OR*" in this paper.

2.2.2. General spatiotemporal model

We denote the county-level areal units as $i = 1, \ldots, I$ ($I = 243$) and the months as $t = 1, \ldots, T$ ($T = 12$). A general form of a spatiotemporal model with a logistic distribution for independent is decomposed additively into components regarding space and time as follows:

$$\eta_{it} = In\left(\frac{P_{it}}{1-P_{it}}\right) = \beta_0 + \mu_i + \nu_i + \gamma_t + \varphi_t \quad (2)$$

where η_{it} is the structured additive linear predictor; P_{it} is the probability of disease occurrence in space i and time t; β_0 quantifies the intercept fixed effect; terms μ_i, ν_i, γ_t and φ_t represent main spatial and temporal random effects.

The spatial components include two random effects: one assumes a Gaussian exchangeable prior to model unstructured heterogeneity, which is $\nu_i \sim N(0, \delta_\nu^2)$, and the other one assumes an intrinsic conditional autoregressive (CAR) prior for the spatially structured variability (Besag, 1974), which is formulated as follows:

$$\mu_i | \mu_{j \neq i} \sim N\left(\frac{1}{m_i}\sum_{i\sim j}\mu_i, \frac{\sigma^2}{m_i}\right) \quad (3)$$

where i~j indicates that areas i and j are neighbors, m_i is the number of areas that share boundaries with the i-*th* area, and σ^2 is the variance component. Spatial dependence in μ_i assumes a CAR prior that extends the well-known Besag model (Besag, 1974) with a Gaussian distribution and implies that each μ_i is conditional on the neighbor μ_j with variance dependent on the number of neighboring counties m_i of county i.

The CAR prior model assumes that the disease occurrence risk in a spatial area is derived from nearby geographical neighbors (spatial autocorrelation). That is, the closer the space distances, the more similar disease occurrence risk is in these spatial areas.

The temporal components also include two random effects: The term ϕ_t is the unstructured time effect, which is specified using an independent mean-zero normal prior to the unknown variance σ_ϕ^2. The term γ_t represents the structured time effect and is modeled dynamically through a neighboring structure. Here, the random walk (RW) dynamic model is used as a prior for the structured time effect (Blangiardo et al., 2013), whose prior density π is written as follows:

$$\pi\left(\gamma_t \middle| \sigma_\gamma^2\right) \propto \exp\left(-\frac{1}{2\sigma_\gamma^2}\sum_{t=2}^{T}(\gamma_t - \gamma_{t-1})^2\right) \quad (4)$$

Similar to CAR, the RW prior model assumes that the temporal variation of disease occurrence risk is influenced by adjacent time points (temporal correlation), which is expressed as a smoothly varying nonlinear curve.

2.2.3. STVC model and local OR

Based on the general spatiotemporal model above, we proposed an innovative local regression method, named *spatiotemporally varying coefficients* (STVC) model to further consider the nonstationary random effects, i.e., spatially and temporally local characteristics of the association between exposures (e.g., climate factors) and response (e.g., disease occurrence). Utilizing this STVC model by assuming a logistic likelihood function for independent data distribution, researchers can further obtain spatially and temporally localized OR values between exposures and response.

The structured additive predictor η_{it} in the STVC model for HFMD case of this study is formulated as follows:

$$\eta_{it} = \beta_0 + \xi_i + \psi_t + \sum_k^m f\left(\mu_{k,i} C_{k,it}\right) + \sum_k^m f\left(\gamma_{k,t} C_{k,it}\right) + \sum_j^n \beta_j SE_j \quad (5)$$

where $C_{k,it}$ denotes the k climate factors (main covariates with random effect) in the i-th area for the t-th month; *SE* represents the n socioeconomic factors (control covariates with fixed effect); β_0 *and* β_j are fixed intercept and covariates coefficients, respectively; ξ_i and ψ_t are the main spatial and temporal random effects, respectively. Similar to their counterparts in the Eq. (2), they can be seen as varying spatial and temporal intercepts.

Function $f(\,)$ denotes the sub-level latent Gaussian models that are used to approximate the nonstationary exposure-response relationships between climate factors and disease occurrence. The CAR model, same as Eq. (3), is considered as the nonstationary spatial model for m climate factors. The spatial nonstationary random effects of climate are expressed as $\sum_k^m f(\mu_{k,i} C_{k,it})$, where $\mu_{k,i}$ are the varying spatial coefficients that represent the local spatial disease-climate relationships. Similarly, the RW model, same as Eq. (4), is considered as the nonstationary temporal model for m climate factors. The temporal nonstationary random effects of climate are expressed as $\sum_k^m f(\gamma_{k,t} C_{k,it})$, where $\gamma_{k,t}$ are the varying temporal coefficients that represent the local temporal disease-climate relationships.

The common *OR* value of each covariate can be obtained directly by using $OR = e^{\beta}$ from an ordinary logistic regression model, where β is the estimated coefficient of the environmental factor. Similarly, for k main covariates (climate factors) with the nonstationary assumption, the spatially local *ORs* of each covariate are obtained with $OR_{k,i} = e^{\mu_{k,i}}$, correspondingly, the temporally local *ORs* of each covariate are obtained with $OR_{k,t} = e^{\gamma_{k,t}}$. Also, we could obtain the local *OR*s for the spatial and temporal intercepts with $OR_i = e^{\xi_i}$ and $OR_t = e^{\psi_t}$, respectively. Note that $OR_{k,i}$ and $OR_{k,t}$ are for the k-th covariate-disease risk associations, whereas OR_i and OR_t are only for the risk of the disease itself.

With local *OR*, we could measure the localized risk associations between disease and environmental factors in each geographical area and time frame. The interpretation of local *OR* is similar to that of the traditional overall *OR*. An *OR* for a spatial or temporal structure > 1 indicates that the spatial or temporal structure of HFMD occurrence is positively associated with this environmental factor. The higher the *OR*, the higher the risk. For disease mapping, a spatially local *OR* greater than one indicates that the spatial unit is a high-risk area, an *OR* value less than one indicates that the spatial unit is a low-risk area, and an *OR* equal to one means that the risk in this areal unit is at an average level.

Essentially, compared with the general spatiotemporal model in Eq. (2), the core difference of the STVC model is that it further incorporates the nonstationary effects of both spatial and temporal dimensions in the observable explanatory variables, such as climate factors. The general spatiotemporal model in Eq. (2) here can also be seen as a type of varying coefficient model only for intercepts. Components μ_i and ν_i are varying intercepts across spatial areas, and components γ_t and φ_t are varying intercepts across temporal periods. However, these varying intercepts in general spatiotemporal model can only represent the occurrence risk of the disease itself (Blangiardo et al., 2013), not the risk associations between covariates and disease, which could be detected by the STVC model.

2.2.4. Model implementation

To evaluate the performance of the STVC model, we compared it with the other four regression models, including the ordinary ecological model considering only the covariates effects; the temporal ecological model considering the covariates and the overall temporal effects; the spatial ecological model considering the covariates and the overall spatial effects; and the general spatiotemporal ecological model considering the covariates, and the overall spatial and temporal effects.

The ordinary ecological regression model is (herein referred to as model 1) is given by Eq. (1).

The temporal ecological regression model is as follows (herein referred to as model 2):

$$\eta_{it} = \beta_0 + \sum_k^m \beta_k C_k + \sum_j^n \beta_j SE_j + \gamma_t + \varphi_t \tag{6}$$

The spatial ecological regression model is as follows (herein referred to as model 3):

$$\eta_{it} = \beta_0 + \sum_k^m \beta_k C_k + \sum_j^n \beta_j SE_j + \mu_i + \nu_i \tag{7}$$

The general spatiotemporal ecological regression model is as follows (herein referred to as model 4):

$$\eta_{it} = \beta_0 + \sum_k^m \beta_k C_k + \sum_j^n \beta_j SE_j + \mu_i + \nu_i + \gamma_t + \varphi_t \tag{8}$$

The STVC regression model (herein referred to as model 5) is given by Eq. (5).

2.2.5. Covariates selection

We used three criteria to select socioeconomic variables for modeling. First, we calculated the variance inflation factor (VIF) for all candidate variables to assess the multicollinearity (Vatcheva et al., 2016). The larger the VIF, the more severe the multicollinearity. Normally, the variable selection considers VIF < 10 as the screening standard (Bo et al., 2014). Second, we used the forward stepwise regression to exclude variables that were not statistically significant (Song et al., 2018a), in which we set 0.05 and 0.1 as the alpha cut. Finally, we retain a covariate in the Bayesian model unless its removal would increase deviance information criterion (DIC) value by 30 units or more (Burnham and Anderson, 2004).

2.3. Model evaluation and cross-validation

We ran cross-validation to evaluate the performance of the five Bayesian models aforementioned. Specifically, we randomly sampled 10%, 15%, and 20% of the existing data to create three test sets, and used the rest of the data as the training sets for modeling. We evaluated each model from three aspects as follows.

(1) Bayesian model complexity and fitness

The Deviance Information Criterion (*DIC*) is a well-known comparison criterion for Bayesian models, defined as follows (Spiegelhalter et al., 2002):

$$\mathrm{DIC} = \overline{D} + p_D \tag{9}$$

where $\overline{D}$ is the mean of model posterior deviance and p_D is the effective number of parameters. A large $\overline{D}$ indicates a great error in the model. A large p_D indicates a high complexity of the model. Models with smaller *DIC* have better fitness, i.e., are better supported by the data. *DIC* and p_D are the smaller, the better.

(2) Bayesian model predictive quality

The Conditional Predictive Ordinate (*CPO*) is defined as a cross-validated predictive density at a given observation and can be used to access predictive measures (Held et al., 2010). For continuous distributions, it is defined as follows:

$$CPO_i = p\left(y_i^* | y_f\right) \tag{10}$$

where y_i^* is the predicted value and y_f is the sample of observations y, which is used to fit the model and to estimate the posterior distribution of the parameters. In practice, the cross-validated logarithmic score (*LS*) computed from *CPO* is widely used to evaluate the predictive quality for Bayesian models. A smaller *LS* indicates a better prediction of a Bayesian model. *LS* is calculated as follows:

$$LS = -\frac{1}{n}\sum_{i=1}^{n} \log(CPO_i) \tag{11}$$

(3) Actual prediction accuracy

For logistic regression models, the confusion matrix is commonly employed to obtain actual prediction accuracy (*PA*) of the model. If we use "a" to denote the number of samples whose true values are 1 (e.g., having-cases) and the model-predicted results are also 1; "b" to denote the number of samples whose true values are 1, but the model-predicted results are 0 (e.g., no-cases); "c" to denote the number of samples whose true values are 0, but the model-predicted results are 1; and "d" to denote the number of samples whose true values are 0, and the model-predicted results are also 0, then the prediction accuracy of the regions with 1 as the true value is $PA(1) = \frac{a}{a+b}$, the regions with 0 as the true value is $PA(0) = \frac{d}{c+d}$, and the overall accuracy is $PA(1,0) = \frac{a+d}{a+b+c+d}$. These three indicators are the larger, the better.

2.4. Bayesian inference

Our models are formalized within a Bayesian framework by extending the concept of hierarchical structure. For each spatiotemporal hierarchical Bayesian model, there are three levels, namely, data distribution, spatiotemporal process, and parameter, with each level further containing some sub-levels (Bakka et al., 2018). Particularly, for the data distribution level, we employed the logistic likelihood function. For the spatiotemporal process level, we combined CAR and RW sub-models to account for the spatial and temporal random effects (Martínez-Bello et al., 2018). For the parameter level, we specified the non-informative priors for the parameters and their variance

Table 1
Model evaluation results of the five alternative Bayesian models in three cross-validation experiments.

	10% Cross-validation			15% Cross-validation			20% Cross-validation		
	DIC	*pD*	*LS*	*DIC*	*pD*	*LS*	*DIC*	*pD*	*LS*
Model 1	8010.98	6.02	0.5750	7597.18	6.02	0.5774	7203.89	6.01	0.5817
Model 2	7324.76	35.86	0.5258	6936.34	35.77	0.5272	6584.97	35.52	0.5318
Model 3	6523.59	190.52	0.4673	6225.77	188.67	0.4722	5896.36	187.99	0.4752
Model 4	6125.04	196.21	0.4394	5840.08	195.11	0.4436	5537.17	194.02	0.4469
Model 5	5191.31	408.24	0.3741	4962.17	301.11	0.3779	4761.72	321.22	0.3848
	PA(*1*)	*PA*(*0*)	*PA*(*1,0*)	*PA*(*1*)	*PA*(*0*)	*PA*(*1,0*)	*PA*(*1*)	*PA*(*0*)	*PA*(*1,0*)
Model 1	55.18%	75.77%	64.73%	70.33%	67.70%	69.08%	70.33%	71.65%	70.99%
Model 2	61.45%	81.06%	70.54%	73.44%	70.42%	72.01%	73.46%	75.48%	74.48%
Model 3	76.14%	71.87%	74.16%	80.82%	72.23%	76.74%	79.48%	74.71%	77.07%
Model 4	78.55%	73.26%	76.10%	80.82%	75.32%	78.21%	79.22%	77.78%	78.49%
Model 5	80.48%	81.06%	80.75%	83.93%	79.67%	81.91%	83.66%	82.89%	83.27%

Note: (model 1: ordinary; model 2: temporal; model 3: spatial; model 4: spatiotemporal; model 5: STVC). *DIC*: deviance information criterion; *pD*: effective number of parameters; *LS*: logarithmic score; *PA*(*1*): prediction accuracy for have-cases regions; *PA*(*0*): prediction accuracy for no-cases regions; *PA*(*1,0*): prediction accuracy for all-cases regions.

components, which allowed the observed data to have the greatest influence on posterior distributions (Schrödle and Held, 2011). We employed a proximate Bayesian inference approach, named Integrated Nested Laplace Approximation (INLA) to estimate posterior parameters of the presented Bayesian models in R software (Lindgren and Rue, 2015). A major advantage of using the INLA approach is a relatively short computation time with accurate parameter estimates (Rue et al., 2017). The core codes for fitting general spatiotemporal models have been published (Blangiardo et al., 2013; Schrödle and Held, 2011; Ugarte et al., 2014).

3. Results

3.1. Model evaluation and comparison

Table 1 shows the evaluation results of the five alternative logistic regression models in three cross-validation experiments. With the lowest *DIC*, *LS* and the highest *PA* values, the new STVC model (model 5) is the best model in terms of model fitness and predictive ability, when compared with the other four models. Although the average complexity (p_D) of the STVC model is almost 57 times that of the traditional ecological logistic regression model (model 1), model 5 shows the best model fitness and predictive ability. Not only that the STVC model performed best in the Bayesian model evaluation, but also achieved the best in the cross-validation experiments on average prediction accuracy (81.98%), compared with ordinary (68.27%), temporal (72.34%), spatial (75.99%), and spatiotemporal (77.60%) ecological models. On average, of the three cross-validation experiments, model 5 was able to correctly predict about 80% of HFMD cases in all three types of regions. In addition, the spatial information in model 3 had a better performance than the temporal information in model 2. The combination of spatial and temporal information in model 4 further improve the performance, demonstrating the effect of the spatiotemporal integration. Most importantly, model 5 surpassed model 4 with an average prediction accuracy increase of 4.38% for the all-cases region, reflecting the benefits of further incorporating the spatial and temporal nonstationary characteristics of covariates (climate variables).

3.2. Environmental factors for HFMD

We first used the common OR to detect the overall associations between environmental factors and HFMD after controlling the spatiotemporal disease occurrence effects. A multicollinearity evaluation of the six climate factors shows that the VIFs are all less than five units (supplementary file Table S2), indicating that the degree of confounding effects of the climate factors is acceptable. In addition, socioeconomic covariates selection results accounting for multicollinearity, significance, and DIC are summarized in supplementary file S2 (Tables S2, S3, and S4). Table 2 summarizes overall OR values and the statistics for the posterior estimated parameters of the selected covariates. When all covariates are standardized, the OR here can be interpreted as a relative indicator, i.e., the higher the OR value, the riskier the covariate is.

Among the six climate variables, the HFMD occurrence risk increased as temperature (OR = 2.59), relative humidity (OR = 1.35), and sunshine hours (OR = 1.06) increased. The disease occurrence risk decreased as wind speed (OR = 0.95) and air pressure (OR = 0.65) increased. These results suggest that hot and humid climate conditions may be an ideal environment for increasing HFMD occurrence risk. The OR of precipitation is close to 1, suggesting that it was not related to HFMD on the scale of the entire region. With regards to the five selected socioeconomic variables, we found that the HFMD occurrence risk increased under higher economic development, demonstrated by the enterprise number density (OR = 3.43), per capita household savings (OR = 2.44), and per capita gross domestic product (OR = 1.41). The children gender ratio (OR = 1.35) indicates a higher risk for boys than girls. The per capita industrial output (OR = 0.71) is the only socioeconomic covariate that showed a negative value associated with HFMD.

3.3. Temporally varying HFMD-climate associations

We then used the optimal model 5 (STVC model) to detect the HFMD occurrence risk and disease-environment (HFMD-climate) associations on both spatial and temporal scales. Fig. 2 shows the trend of temporally local OR of the disease itself in the entire study area. There were very obvious seasonal variations during the 36 months from 2009 to 2011, with three high peaks in summer (around May) and three low peaks in winter (around November). Fig. 3 further illustrates these temporally varying HFMD-climate associations for six climate

Table 2
Odds ratio (OR) values and estimated posterior parameters of the climate and socioeconomic factors on HFMD occurrence.

Covariates	OR	Mean	SD	0.025 CI	0.975 CI
Temperature	2.59	0.9508	0.2858	0.3980	1.5201
Relative humidity	1.35	0.2978	0.1401	0.0232	0.5729
Sunshine hours	1.06	0.0621	0.1207	−0.1749	0.2988
Air pressure	0.95	−0.0423	0.2439	−0.5203	0.4374
Wind speed	0.65	−0.4259	0.1250	−0.6720	−0.1812
Precipitation	0.99	−0.0010	0.0007	−0.0023	0.0004
Enterprise number density	3.43	1.2326	0.2641	0.7166	1.7532
Per capita household savings	2.44	0.8921	0.2068	0.4850	1.2971
Per capita gross domestic product (GDP)	1.41	0.3465	0.1672	0.0183	0.6748
Children gender ratio	1.35	0.3015	0.1152	0.0761	0.5281
Per capita industrial output values	0.71	−0.3486	0.2210	−0.7832	0.0844

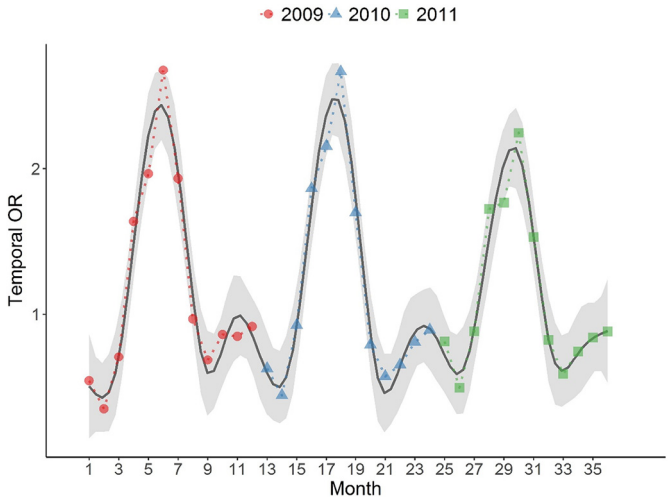


Fig. 2. Main temporal risk trend of HFMD occurrence during 36 months from the years 2009 to 2011 in Sichuan, China.

factors. Different from the overall OR in Table 2, the effects of the six climate factors on HFMD varied during the 36 months. Variations in temperature (×1), relative humidity (×2), and wind speed (×5) were relatively similar to that of the temporal OR of HFMD itself (Fig. 2), with three seasonal high peaks during three years. The other three climate factors did not have this similarity: precipitation (×3), air pressure (×4) and sunshine hours (×6) were related to some high risks in spring and winter.

3.4. Spatially varying HFMD-climate associations

Fig. 4(a) shows the local OR risk map representing the spatial HFMD occurrence distribution of the disease itself. In Fig. 5, we used six local OR risk maps representing the spatially varying HFMD-climate associations for the six climate factors. We also obtained the hotspot maps based on the corresponding OR risk maps to show which regions have significant clusters of high-risk hot spots and low-risk cold spots, as seen in Figs. 4(b) and 5. Supplemental file S3 includes the details of the hot spot analysis. Note that "not significant" regions in the hot spot map do not necessarily indicate absence or presence of risk, just that the risks in these regions are not significant enough to form a cluster.

All of the local OR maps in Figs. 4(a) and 5 showed prominent spatial aggregation characteristics, which suggests that spatial autocorrelation is useful when applied to both disease and disease-climate associations in modeling. For risk of HFMD itself in Sichuan Province of China, we identified three high-risk hot spots in which officials would need to focus more on in practice, as well as three low-risk cold spots shown in Fig. 4(b). Comparing hot spot maps between Figs. 4(b) and 5, we discovered: for disease hot spot I, ×1 and ×2 were significantly contributing positive risk factors (red); ×6 was significantly contributing negative risk factor (blue); the other climate factors did contribute various risks at the local scale but were not significant enough to form hot spots. Similarly, for disease hot spot II, ×5 and ×6 contributed

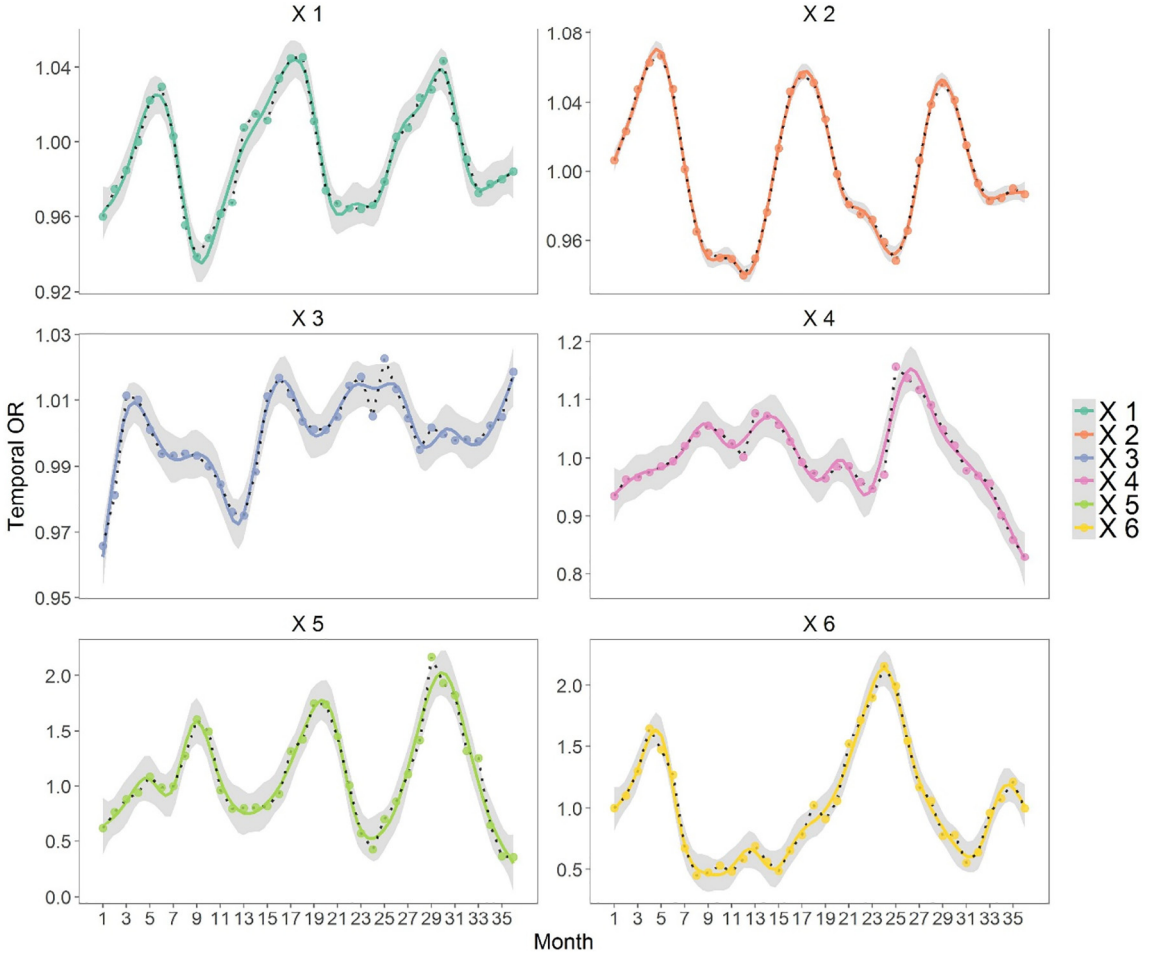


Fig. 3. Temporal nonstationary climate effects on HFMD occurrence during 36 months from the years 2009 to 2011: ×1 temperature, ×2 relative humidity, ×3 precipitation, ×4 air pressure, ×5 wind speed, and ×6 sunshine hours.

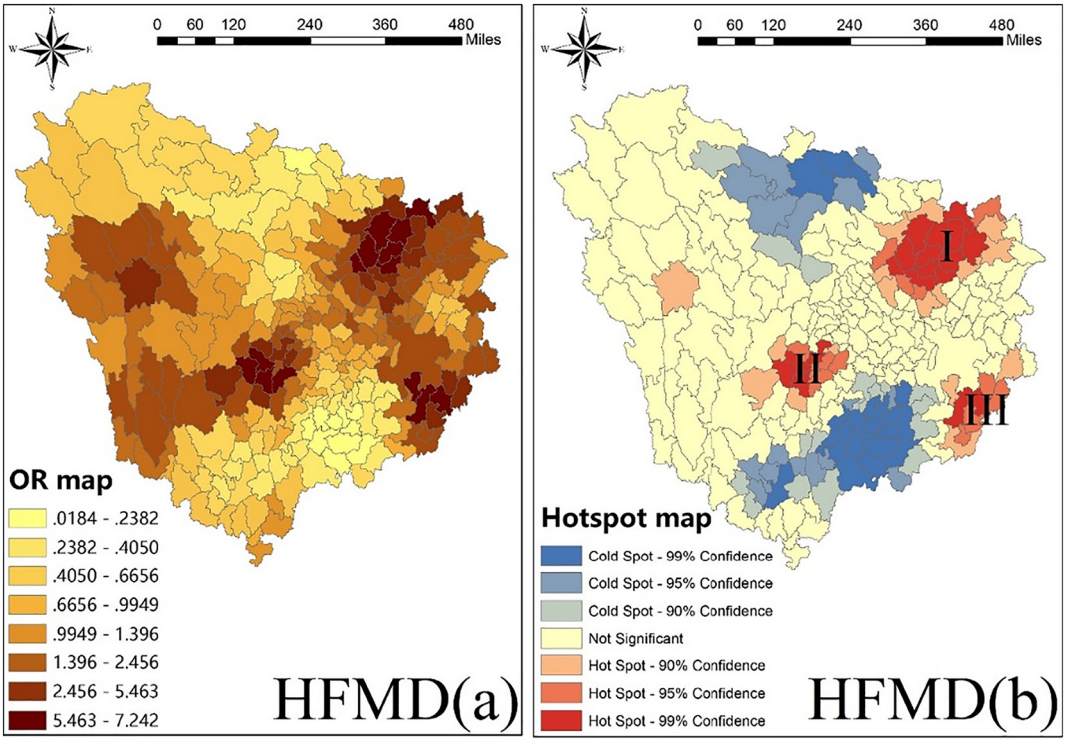


Fig. 4. Spatial local odds ratio (OR) map of the main HFMD occurrence risk (a) and its hotspot map (b) at the county level in Sichuan, China.

significantly positive risk, while ×2 contributed significantly negative risks. Lastly, for disease hot spot III, ×1, ×2, and ×3 contributed significantly positive risk, while ×5 contributed significantly negative risks. With local OR maps, we could further analyze the differences in the contribution of these environmental factors at each monitoring site, which is of great significance for the prevention and control of local disease transmission.

4. Discussion

Our results suggest that both climate and socioeconomic factors were associated with HFMD occurrence under spatiotemporal scales in Sichuan, China. We also showed that socioeconomic variables were confounding factors in detecting associations between HFMD and the climate.

With regards to the climate as a variable, we found, consistent with the previous studies, that a hot and humid climate might be an ideal environment for HFMD transmission (Cheng et al., 2018; Onozuka and Hashizume, 2011; Xiao et al., 2017; Zhu et al., 2015). Moreover, we found that increased sunshine hours were also positively correlated with HFMD occurrence. One explanation is that more sunshine hours will motivate children to spend more time outdoors, which can facilitate contacts among them via air and body, in turn, the spread of the virus. In addition, more sunshine hours can increase the surface temperature, which is good for breeding and transmission of the virus. Wind speed and air pressure were not positively correlated with HFMD occurrence.

With regards to socioeconomic conditions, most variables that we tested were positively associated with the HFMD occurrence in Sichuan, China, i.e., the enterprise number density, per capita household savings, and per capita GDP, indicating that the economic development was positively correlated with the occurrence of HFMD. This finding is consistent with other studies of HFMD using different socioeconomic variables, such as GDP (Huang et al., 2014; Li et al., 2018), enterprise number density and per capita fixed assets investment (Song et al., 2018a), and an increase of migratory workers from rural areas (Zeng et al., 2013). In urban areas, the higher population density leads to an easy spread of the virus (Li et al., 2018). Most children in the developed regions of China go to daycares centers or kindergartens, whereas children in undeveloped areas usually stay at home where there is less of a chance to be in contact with HFMD-infected children (Xu, 2017). Our finding that boys had a higher risk than girls is also consistent with other studies, e.g., (Xu, 2017; Zeng et al., 2013) and (Wang et al., 2011b). The per capita industrial output was the only socioeconomic covariate that was negatively correlated with HFMD occurrence, possibly because the pollution in those industrial areas reduced the outdoor activities of children.

Going beyond the overall disease-climate associations aforementioned, a more important contribution of this study is that we have detected the local-scale spatial and temporal variations in the HFMD-climate associations, i.e., spatiotemporal nonstationary varying random effects, in the Sichuan Province of China.

In the temporal dimension, our results indicated that the temporal disease-climate associations between HFMD and air temperature, relative humidity as well as wind speed were very similar to the temporal trend of HFMD itself. This suggests that these three climate variables are better explanatory factors for HFMD occurrence at a localized temporal scale relating to the disease throughout 36 months of three years, with the highest risk peak during summer around May, as consistent with other studies (Li et al., 2014; Xing et al., 2014). In addition, we found that there was a second peak in winter (Ma et al., 2010) of each year in Fig. 1, and upon further examination (Fig. 3), we found that the precipitation, air pressure and sunshine hours might be related to that increase. Temporally varying disease-climate associations indicate that both hot and humid environments in the summer, and cold and dry environments in the winter may contribute to HFMD occurrence risk increase.

In the spatial dimension, we created maps that exhibit clusters (spatial autocorrelation) of HFMD itself and its associations with the climate in Sichuan Province. The nonstationary HFMD-climate associations were more complicated in the spatial dimension than the temporal dimension. Our findings extend the results of a recent study that explored the spatially nonstationary impact of air temperature on HFMD (Liao

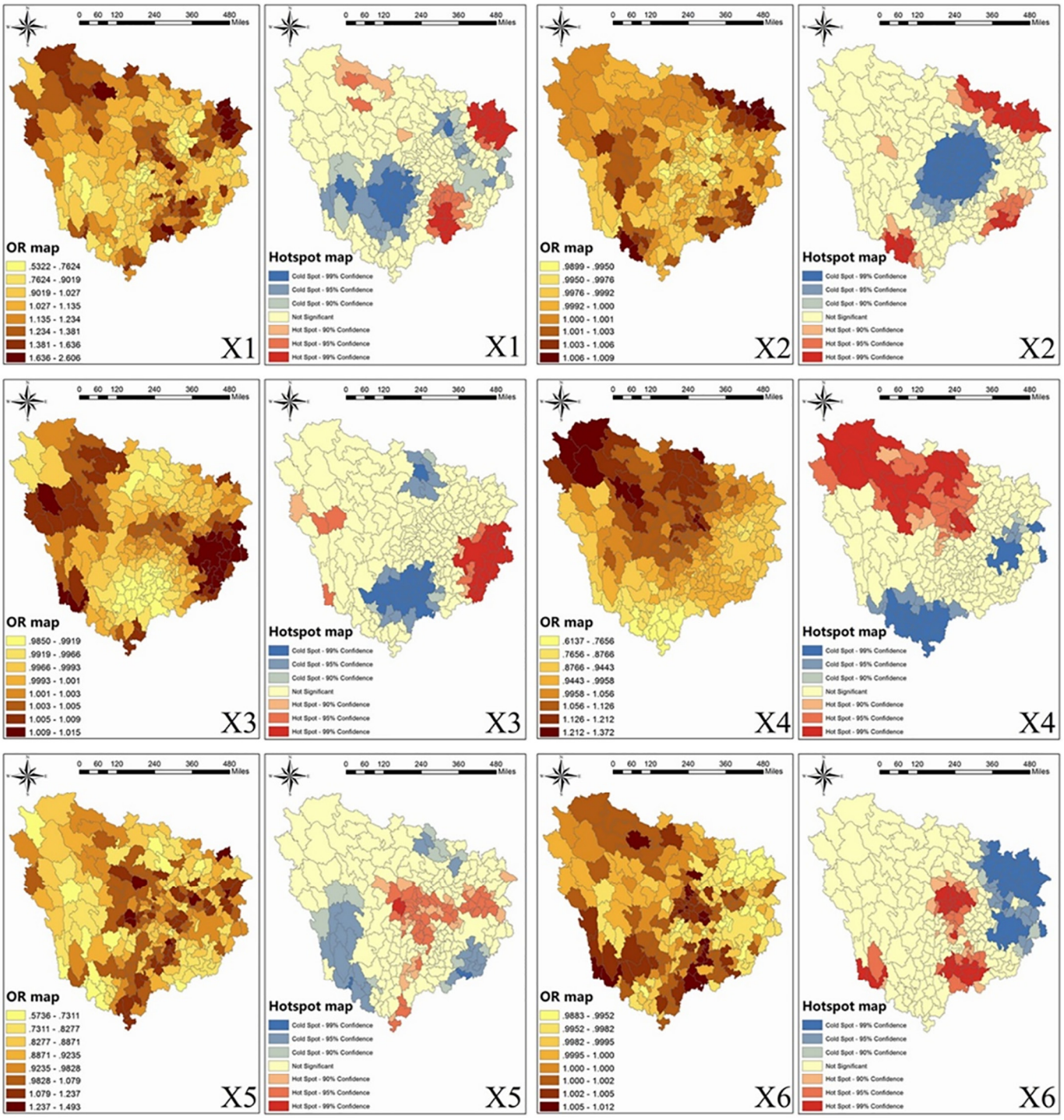


Fig. 5. Local odds ratio (OR) maps (brown) of spatially varying nonstationary climate effects on HFMD occurrence and their corresponding spatial hot spot maps (red and blue) at the county level in Sichuan, China: ×1 temperature, ×2 relative humidity, ×3 precipitation, ×4 air pressure, ×5 wind speed, and ×6 sunshine hours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

et al., 2016) by including six climate factors rather than only air temperature, and also by on a spatiotemporal scale. In addition, the spatially nonstationary random effects of covariates (varying coefficients) embedded in the STVC model could capture local spatial relationships between climate and disease. Such information cannot be achieved by models that do not incorporate spatially nonstationary effects (Li et al., 2018). The results showed important implications of strong spatial patterns of HFMD occurrence risk and disease-climate associations in Sichuan, China. Our risk maps reveal hot spot counties to which particular governmental interventions about HFMD should be applied, and where particular attention to the local effects of climate on the disease is needed.

In summary, the main contribution of this study is to propose the STVC model for local regression modeling and demonstrate its advantages in both environmental science and spatial statistics. We summarized the advantages of the STVC model in three aspects as follows.

I. Methodological development

The core novelty of the proposed STVC model is the incorporation of the nonstationary random effects in both spatial and temporal dimensions, which is an important extension and improvement of the former regression models. Specifically, the modifications to the former three categories of regression models are summarized as below:

Firstly, the STVC model with spatiotemporally nonstationary assumption is an alternative to the common ecological regression approaches and nonlinear evaluation models, such as generalized linear model (GLM) or generalized additive model (GAM), where covariates effects are assumed to be homogenous (stationarity) across space and over time (Wakefield, 2006); Secondly, the STVC model is an important extension of the SVC model into the spatiotemporal form (Wolf et al., 2018); Thirdly, the STVC model is an improvement to the general spatiotemporal models that only incorporate spatial and temporal random effects in the intercept component (Blangiardo et al., 2013), not the observed independent covariates.

The STVC model implements the nonstationarity assumption by obtaining local spatial and temporal coefficients for observable explanatory variables, which seems to be more reasonable and realistic (Leyk et al., 2012; Osei and Stein, 2017), as the model's fitness and prediction accuracy got improved.

II. Applications of the STVC model

With regards to disease mapping, the STVC model is effective in localizing the associations between disease outcome and covariates, while the traditional disease mapping only considers the spatial distribution of the disease itself (Best et al., 2005). To our best knowledge, this study is the first that produces spatialized local OR values that can be directly utilized in mapping disease-covariate associations in environmental epidemiology.

Moreover, most previous environmental studies treat the exposure-response relationship as being constant across space and over time (Nieuwenhuijsen, 2015), the STVC model offers a new approach to capture spatiotemporal local-scale variations in associations between environmental factors and health outcomes. This approach can be generalized and applied to other applications in the broader environmental science, especially for large-scale geographical datasets.

III. Method versatility and scalability

The BHM-based STVC model is relatively more convenient to be converted into other types to adapt different research objects compared with frequency domain statistics (Wolf et al., 2018), due to the advantages of Bayesian statistics and hierarchical modeling framework (Blangiardo et al., 2013). For instance, the STVC model is not limited to logistic distribution for an independent; instead, it can be applied to a broad variety of likelihood models for different data distributions, such as Gaussian, Poisson, binomial, and negative binomial (Bakka et al., 2018).

There are some limitations to our study. Firstly, the socioeconomic data used in this study do not contain any temporal changes, as the data are the summation of one year. Secondly, we did not include such environmental factors as land cover, land use (He et al., 2015), and air pollution (Yang et al., 2018; Yin et al., 2018), which could potentially influence HFMD. Thirdly, there might be unreported HFMD cases, but we were not able to obtain tangible information about underreporting (Hu et al., 2012), which is a potential bias in modeling. Finally, our proposed STVC model was trained using only the HFMD data in the Sichuan province of China, and should in the future be applied to the larger study area, different data distributed assumptions and other infectious diseases.

5. Conclusions

In this study, we proposed a new local regression approach, named *spatiotemporally varying coefficients* (STVC) model under the BHM framework to detect the nonstationary exposure-response relationships between the dependent and independent variables in both spatial and temporal dimensions, and successfully applied it to HFMD case. Our study expanded the limited knowledge of the complex local-scale nonstationary associations between climate factors and HFMD, revealing the temporally nonlinear (seasonal trends) and spatially clustered (hot spots) disease-climate relationships in Sichuan, China. To our best knowledge, this is the first study that investigates both spatial and temporal nonstationary effects of climate on HFMD. We found that air temperature, relative humidity, and wind speed were not only overall related to the increase of HFMD occurrence, but also demonstrated spatiotemporal variations in their local associations with HFMD in Sichuan, China. These findings can help deepen the understanding of complex associations between HFMD and climate, and provide support for county-level HFMD prevention and control. In addition, the STVC model with logistic distribution has solved the OR spatialization problem for advanced mapping on disease-environment associations. More importantly, the proposed Bayesian STVC model should be a new alternative and sophisticated category method for local regression modeling in spatial statistics and Geostatistics. The STVC model could be an efficient way to capture spatiotemporal nonstationary exposure-response relationships for detailed exposure assessments and advanced risk mapping with large-scale geospatial data, not only important to environmental epidemiology, but also offer new insights to the broader environmental science.

Funding

The work was supported by the National Natural Science Foundation of China (grant number 41701448), the Young Scholars Development Fund of Southwest Petroleum University (grant number 201699010064), a grant from State Key Laboratory of Resources and Environmental Information System, and the State Key Laboratory of Remote Sensing Science.

Acknowledgments

The authors thank Henry Chung (Michigan State University, US) and Thom Goodwin for proofreading our paper carefully, María Dolores Ugarte (Public University of Navarre, Spain) for her help with coding. The authors also acknowledge colleagues in the Spatial Information Technology and Big Data Mining Research Center in Southwest Petroleum University for their assistant in this study. Lastly, we would like to thank the editor and three anonymous reviewers for their constructive comments and valuable suggestions in improving this manuscript.

Competing interests

The authors declare that they have no competing interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2018.08.114.

References

Allenby, G.M., Rossi, P.E., 2006. Hierarchical bayes models. The Handbook of Marketing Research: Uses, Misuses, and Future Advances, pp. 418–440.

Bakka, H., Rue, H., Fuglstad, G.-A., Riebler, A., Bolin, D., Krainski, E., et al., 2018. Spatial Modelling With R-INLA: A Review. arXiv Preprint. arXiv:1802.06350.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2014. Hierarchical Modeling and Analysis for Spatial Data. CRC press.

Bélanger, M., Gray-Donald, K., O'Loughlin, J., Paradis, G., Hanley, J., 2009. Influence of weather conditions and season on physical activity in adolescents. Ann. Epidemiol. 19, 180–186.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. J. R. Stat. Soc. Ser. B Methodol. 192–236.

Best, N., Richardson, S., Thomson, A., 2005. A comparison of Bayesian spatial models for disease mapping. Stat. Methods Med. Res. 14, 35–59.

Bland, J.M., Altman, D.G., 2000. The odds ratio. BMJ 320, 1468.

Blangiardo, M., Cameletti, M., Baio, G., Rue, H., 2013. Spatial and spatio-temporal models with R-INLA. Spat. Spatio-Temporal Epidemiol. 7, 39–55.

Bo, Y., Song, C., Wang, J., Li, X., 2014. Using an autologistic regression model to identify spatial risk factors and spatial risk patterns of hand, foot and mouth disease (HFMD) in mainland China. BMC Public Health 14, 358.

Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. 28, 281–298.

Burnham, K.P., Anderson, D.R., 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociol. Methods Res. 33, 261–304.

Chen, V., Wu, P., Yang, T., Su, H., 2010. Examining non-stationary effects of social determinants on cardiovascular mortality after cold surges in Taiwan. Sci. Total Environ. 408, 2042–2049.

Chen, C., Lin, H., Li, X., Lang, L., Xiao, X., Ding, P., et al., 2014. Short-term effects of meteorological factors on children hand, foot and mouth disease in Guangzhou, China. Int. J. Biometeorol. 58, 1605–1614.

Cheng, Q., Bai, L., Zhang, Y., Zhang, H., Wang, S., Xie, M., et al., 2018. Ambient temperature, humidity and hand, foot, and mouth disease: a systematic review and meta-analysis. Sci. Total Environ. 625, 828–836.

Du, Z., Zhang, W., Zhang, D., Yu, S., Hao, Y., 2016. The threshold effects of meteorological factors on hand, foot, and mouth disease (HFMD) in China, 2011. Sci. Rep. 6, 36351.

Finley, A.O., Banerjee, S., Carlin, B.P., 2007. spBayes: an R package for univariate and multivariate hierarchical point-referenced spatial models. J. Stat. Softw. 19, 1.

Gelfand, A.E., Kim, H.-J., Sirmans, C., Banerjee, S., 2003. Spatial modeling with spatially varying coefficient processes. J. Am. Stat. Assoc. 98, 387–396.

Harrell, F.E., 2001. Ordinal Logistic Regression. Regression Modeling Strategies. Springer, pp. 331–343.

Hastie, T., Tibshirani, R., 1993. Varying-coefficient models. J. R. Stat. Soc. Ser. B Methodol. 757–796.

He, Y., Bo, Y., de Jong, R., Li, A., Zhu, Y., Cheng, J., 2015. Comparison of vegetation phenological metrics extracted from GIMMS NDVIg and MERIS MTCI data sets over China. Int. J. Remote Sens. 36, 300–317.

Held, L., Schrödle, B., Rue, H., 2010. Posterior and cross-validatory predictive checks: a comparison of MCMC and INLA. Stat. Modell. Regress. Struct. 91–110.

Hu, M., Li, Z., Wang, J., Jia, L., Liao, Y., Lai, S., et al., 2012. Determinants of the incidence of hand, foot and mouth disease in China using geographically weighted regression models. PLoS One 7, e38978.

Huang, D., Wang, J., 2018. Monitoring hand, foot and mouth disease by combining search engine query data and meteorological factors. Sci. Total Environ. 612, 1293–1299.

Huang, J., Wang, J., Bo, Y., Xu, C., Hu, M., Huang, D., 2014. Identification of health risks of hand, foot and mouth disease in China using the geographical detector technique. Int. J. Environ. Res. Public Health 11, 3407–3423.

Huang, J., Liao, Q., Ooi, M.H., Cowling, B.J., Chang, Z., Wu, P., et al., 2018. Epidemiology of recurrent hand, foot and mouth disease, China, 2008–2015. Emerg. Infect. Dis. 24, 432.

Kim, B.I., Ki, H., Park, S., Cho, E., Chun, B.C., 2016. Effect of climatic factors on hand, foot, and mouth disease in South Korea, 2010-2013. PLoS One 11, e0157500.

Koh, W.M., Bogich, T., Siegel, K., Jin, J., Chong, E.Y., Tan, C.Y., et al., 2016. The epidemiology of hand, foot and mouth disease in Asia: a systematic review and analysis. Pediatr. Infect. Dis. J. 35, e285.

Lee, C., Sheridan, S., 2018. A new approach to modeling temperature-related mortality: non-linear autoregressive models with exogenous input. Environ. Res. 164, 53–64.

Lee, C., Tang, J., Hwang, J., Shigematsu, M., Chan, T., 2015. Effect of meteorological and geographical factors on the epidemics of hand, foot, and mouth disease in island-type territory, East Asia. Biomed. Res. Int. 2015.

Leyk, S., Norlund, P.U., Nuckols, J.R., 2012. Robust assessment of spatial non-stationarity in model associations related to pediatric mortality due to diarrheal disease in Brazil. Spat. Spatio-temporal Epidemiol. 3, 95–105.

Li, Y., Zhang, J., Zhang, X., 2014. Modeling and preventive measures of hand, foot and mouth disease (HFMD) in China. Int. J. Environ. Res. Public Health 11, 3108–3117.

Li, L., Qiu, W., Xu, C., Wang, J., 2018. A spatiotemporal mixed model to assess the influence of environmental and socioeconomic factors on the incidence of hand, foot and mouth disease. BMC Public Health 18, 274.

Liao, J., Qin, Z., Zuo, Z., Yu, S., Zhang, J., 2016. Spatial-temporal mapping of hand foot and mouth disease and the long-term effects associated with climate and socio-economic variables in Sichuan Province, China from 2009 to 2013. Sci. Total Environ. 563, 152–159.

Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with R-INLA. J. Stat. Softw. 63.

Liu, S.L., Pan, H., Liu, P., Amer, S., Chan, T.C., Zhan, J., et al., 2015. Comparative epidemiology and virology of fatal and nonfatal cases of hand, foot and mouth disease in mainland China from 2008 to 2014. Rev. Med. Virol. 25, 115–128.

Liu, L., Luan, R., Yin, F., Zhu, X., Lü, Q., 2016. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model. Epidemiol. Infect. 144, 144–151.

Ma, E., Lam, T., Wong, C., Chuang, S., 2010. Is hand, foot and mouth disease associated with meteorological parameters? Epidemiol. Infect. 138, 1779–1788.

Martínez-Bello, D.A., López-Quílez, A., Torres, A.P., 2018. Spatio-temporal modeling of Zika and dengue infections within Colombia. Int. J. Environ. Res. Public Health 15.

Nguyen, H.X., Chu, C., Nguyen, H.L.T., Nguyen, H.T., Do, C.M., Rutherford, S., et al., 2017. Temporal and spatial analysis of hand, foot, and mouth disease in relation to climate factors: a study in the Mekong Delta region, Vietnam. Sci. Total Environ. 581, 766–772.

Nieuwenhuijsen, M.J., 2015. Exposure Assessment in Environmental Epidemiology. Oxford University Press, USA.

Onozuka, D., Hashizume, M., 2011. The influence of temperature and humidity on the incidence of hand, foot, and mouth disease in Japan. Sci. Total Environ. 410, 119–125.

Osei, F., Stein, A., 2017. Diarrhea morbidities in small areas: accounting for non-stationarity in sociodemographic impacts using Bayesian spatially varying coefficient modelling. Sci. Rep. 7, 9908.

Rue, H., Riebler, A., Sørbye, S.H., Illian, J.B., Simpson, D.P., Lindgren, F.K., 2017. Bayesian computing with INLA: a review. Annu. Rev. Stat. Its Appl. 4, 395–421.

Schrödle, B., Held, L., 2011. Spatio-temporal disease mapping using INLA. Environmetrics 22, 725–734.

Shi, X., Kwan, M., 2015. Introduction: geospatial health research and GIS. Ann. GIS 21, 93–95.

Song, C., He, Y., Bo, Y., Wang, J., Ren, Z., Yang, H., 2018a. Risk assessment and mapping of hand, foot, and mouth disease at the county level in mainland China using spatiotemporal zero-inflated Bayesian hierarchical models. Int. J. Environ. Res. Public Health 15, 1476.

Song, C., Yang, X., Shi, X., Bo, Y., Wang, J., 2018b. Estimating missing values in China's official socioeconomic statistics using progressive spatiotemporal Bayesian hierarchical modeling. Sci. Rep. 8, 10055.

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. Ser. B (Stat Methodol.) 64, 583–639.

Suminski, R.R., Poston, W.C., Market, P., Hyder, M., Sara, P.A., 2008. Meteorological conditions are associated with physical activities performed in open-air settings. Int. J. Biometeorol. 52, 189–197.

Ugarte, M.D., Adin, A., Goicoa, T., Militino, A.F., 2014. On fitting spatio-temporal disease mapping models using approximate Bayesian inference. Stat. Methods Med. Res. 23, 507–530.

Vatcheva, K.P., Lee, M., McCormick, J.B., Rahbar, M.H., 2016. Multicollinearity in regression analyses conducted in epidemiologic studies. Epidemiology 6 (Sunnyvale, Calif.).

Wakefield, J., 2006. Disease mapping and spatial regression with count data. Biostatistics 8, 158–183.

Wang, J., Guo, Y., Christakos, G., Yang, W., Liao, Y., Li, Z., et al., 2011a. Hand, foot and mouth disease: spatiotemporal transmission and climate. Int. J. Health Geogr. 10, 25.

Wang, Y., Feng, Z., Yang, Y., Self, S., Gao, Y., Longini, I.M., et al., 2011b. Hand, foot and mouth disease in China: patterns of spread and transmissibility during 2008–2009. Epidemiology 22, 781 (Cambridge, Mass.).

Wang, H., Du, Z., Wang, X., Liu, Y., Yuan, Z., Liu, Y., et al., 2015. Detecting the association between meteorological factors and hand, foot, and mouth disease using spatial panel data models. Int. J. Infect. Dis. 34, 66–70.

Wang, J., Cao, Z., Zeng, D.D., Wang, Q., Wang, X., 2016. Assessment for spatial driving forces of HFMD prevalence in Beijing, China. Proceedings of the Second ACM SIGSPATIALInternational Workshop on the Use of GIS in Emergency Management. ACM, p. 6.

Wang, P., Zhao, H., You, F., Zhou, H., Goggins, W.B., 2017. Seasonal modeling of hand, foot, and mouth disease as a function of meteorological variations in Chongqing, China. Int. J. Biometeorol. 61, 1411–1419.

Wolf, L.J., Oshan, T.M., Fotheringham, A.S., 2018. Single and multiscale models of process spatial heterogeneity. Geogr. Anal. 50, 223–246.

Xiao, X., Gasparrini, A., Huang, J., Liao, Q., Liu, F., Yin, F., et al., 2017. The exposure-response relationship between temperature and childhood hand, foot and mouth disease: a multicity study from mainland China. Environ. Int. 100, 102–109.

Xing, W., Liao, Q., Viboud, C., Zhang, J., Sun, J., Wu, J.T., et al., 2014. Hand, foot, and mouth disease in China, 2008–12: an epidemiological study. Lancet Infect. Dis. 14, 308–318.

Xu, C., 2017. Spatio-temporal pattern and risk factor analysis of hand, foot and mouth disease associated with under-five morbidity in the Beijing–Tianjin–Hebei Region of China. Int. J. Environ. Res. Public Health 14, 416.

Yang, Y., Luo, L., Song, C., Yin, H., Yang, J., 2018. Spatiotemporal assessment of PM2. 5-Related economic losses from health impacts during 2014−2016 in China. Int. J. Environ. Res. Public Health 15.

Yin, H., Pizzol, M., Jacobsen, J.B., Xu, L., 2018. Contingent valuation of health and mood impacts of PM 2. 5 in Beijing, China. Sci. Total Environ. 630, 1269–1282.

Yu, L., Zhou, L., Tan, L., Jiang, H., Wang, Y., Wei, S., et al., 2014. Application of a new hybrid model with seasonal auto-regressive integrated moving average (ARIMA) and nonlinear auto-regressive neural network (NARNN) in forecasting incidence cases of HFMD in Shenzhen, China. PLoS One 9, e98241.

Zeng, M., Pu, D., Mo, X., Zhu, C., Gong, S., Xu, Y., et al., 2013. Children of rural-to-urban migrant workers in China are at a higher risk of contracting severe hand, foot and mouth disease and EV71 infection: a hospital-based study. Emerg. Microbes Infect. 2, e72.

Zhang, W., Du, Z., Zhang, D., Yu, S., Hao, Y., 2016a. Boosted regression tree model-based assessment of the impacts of meteorological drivers of hand, foot and mouth disease in Guangdong, China. Sci. Total Environ. 553, 366–371.

Zhang, W., Du, Z., Zhang, D., Yu, S., Huang, Y., Hao, Y., 2016b. Assessing the impact of humidex on HFMD in Guangdong Province and its variability across social-economic status and age groups. Sci. Rep. 6, 18965.

Zhang, Z., Xie, X., Chen, X., Li, Y., Lu, Y., Mei, S., et al., 2016c. Short-term effects of meteorological factors on hand, foot and mouth disease among children in Shenzhen, China: non-linearity, threshold and interaction. Sci. Total Environ. 539, 576–582.

Zhao, Q., Li, S., Cao, W., Li, D., Qian, Q., Ren, H., et al., 2018. Modeling the present and future incidence of pediatric hand, foot, and mouth disease associated with ambient temperature in mainland China. Environ. Health Perspect. 126 (Online).

Zhu, L., Yuan, Z., Wang, X., Li, J., Wang, L., Liu, Y., et al., 2015. The impact of ambient temperature on childhood HFMD incidence in inland and coastal area: a two-city study in Shandong province, China. Int. J. Environ. Res. Public Health 12, 8691–8704.