

Comparison of different predictive models on HFMD based on weather factors in Zibo city, Shandong Province, China

Original Paper

*Xiujun Li and Feng Cui contributed equally to this work as co-corresponding authors.

Cite this article: Liu LL, Hu YC, Qi C, Zhu YC, Li CY, Wang L, Cui F, Li XJ (2022). Comparison of different predictive models on HFMD based on weather factors in Zibo city, Shandong Province, China. *Epidemiology and Infection* **150**, e10, 1–8. <https://doi.org/10.1017/S0950268821002508>

Received: 27 May 2021

Revised: 24 October 2021

Accepted: 10 November 2021

Keywords:

Generalised additive model; hand-foot-and-mouth diseases; predict; random forest regression; support vectors regression

Author for correspondence:

Xiujun Li,

E-mail: xjli@sdu.edu.cn

L. L. Liu¹, Y. C. Hu², C. Qi¹, Y. C. Zhu¹, C. Y. Li¹, L. Wang³, F. Cui^{3,*} and X. J. Li^{1,*} 

¹Department of Biostatistics, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan 250012, Shandong, China; ²Institute of Clinical Trials and Methodology 90 High Holborn London, University College London, London, UK and ³Zibo Center for Disease Control and Prevention, Zibo 255000, Shandong, China

Abstract

The early identification and prediction of hand-foot-and-mouth disease (HFMD) play an important role in the disease prevention and control. However, suitable models are different in regions due to the differences in geography, social economy factors. We collected data associated with daily reported HFMD cases and weather factors of Zibo city in 2010~2019 and used the generalised additive model (GAM) to evaluate the effects of weather factors on HFMD cases. Then, GAM, support vectors regression (SVR) and random forest regression (RFR) models are used to compare predictive results. The annual average incidence was 129.72/100 000 from 2010 to 2019. Its distribution showed a unimodal trend, with incidence increasing from March, peaking from May to September. Our study revealed the nonlinear relationship between temperature, rainfall and relative humidity and HFMD cases and based on the predictive result, the performances of three models constructed ranked in descending order are: SVR > GAM > RFR, and SVR has the smallest prediction errors. These findings provide quantitative evidence for the prediction of HFMD for special high-risk regions and can help public health agencies implement prevention and control measures in advance.

Background

Hand, foot and mouth disease (HFMD) is a serious infectious disease caused mainly by an enterovirus, threaten mostly children under 5 years old [1–3]. Weather variables, as important predictors of HFMD, which could influence the transmission and reproduction of enterovirus directly or indirectly, have been pointed out in prior study [3–7], but results vary in different regions. For instance, some studies showed that temperature, relative humidity, precipitation, wind speed and sunshine all have an important influence on HFMD cases in Shandong province, Shanghai and Shenzhen [8–10]. However, Huang *et al.* found that wind speed and precipitation have no relationship with HFMD cases in Guangzhou city [11], and Liu *et al.* found that relative humidity has no effect on HFMD cases in Jiangsu Province [12]. In addition, some scholars revealed that temperature, humidity and wind speed have a positive correlation with HFMD, while rainfall and sunshine are negatively correlated with HFMD in studies of Chongqing, Gansu and Shanghai [10, 13, 14]. On the contrary, a study in Hunan province showed a positive effect between rainfall and HFMD incidence, while wind speed has a negative impact on HFMD incidence [15]. The discrepancies of these findings could be likely attributed to the diversities of geographic, socioeconomic and population variables or statistic methods and data sources.

Previous studies attempt to apply various traditional models to predict HFMD incidence, such as SARIMA model, Generalised additive model (GAM), as well as combined models [16]. Currently, the application of machine learning provided new ideas for disease warning measures by applying mathematical models after statistical sorting the historical information of data and applied for the prediction of infectious disease [17–19], such as HFMD, Dengue [20]. However, no studies have proven that machine learning models have been superior to traditional models in prediction, and traditional predictive models are still used. It remains unclear which type models would be better for HFMD epidemic prediction, which hinders HFMD control and prevention [21–23]. In addition, a unified model cannot be found suitable for all regions due to these discrepancies of data source, methods and variables in regions [24]. Therefore, it is urgent to establish a predictive model based on the actual situation of a specific region to predict more accurately.

Zibo city is a high-risk area of HFMD in Shandong Province. The annual average incidence was 127.9/100 000 from 2010 to 2019, exceeding the average level of Shandong Province (93.70/100 000) [17]. Thus, this study aims to explore the optimal model for HFMD prediction based on meteorological factors in Zibo city, helping the formulation and implementation of

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

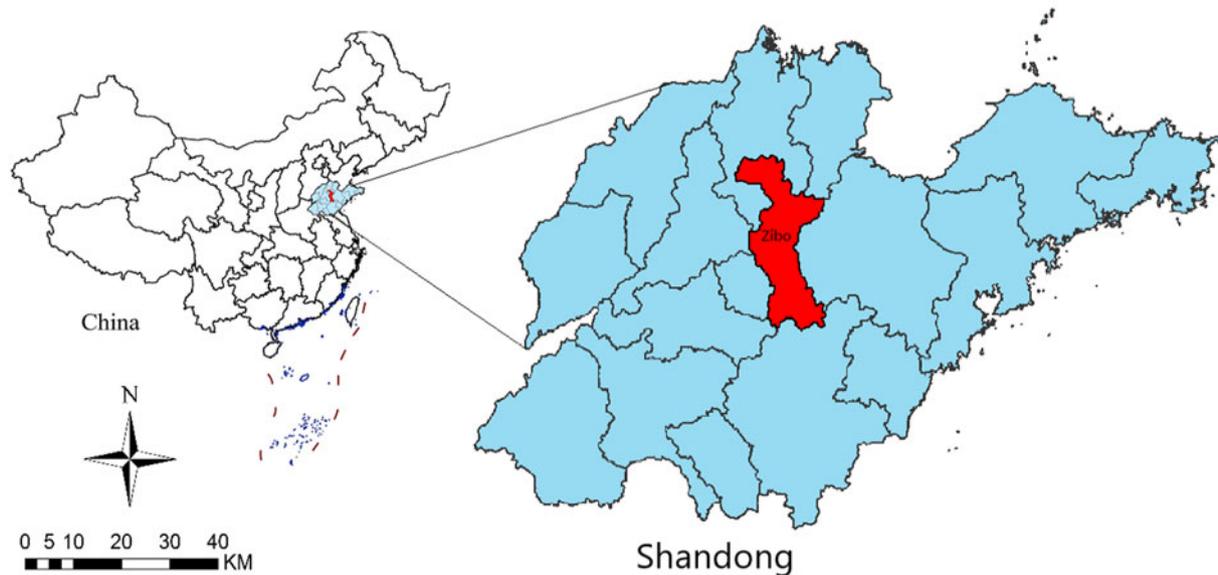


Fig. 1. Geolocation of Zibo city in Shandong Province, China.

preventive and control measures of HFMD and providing new objective information for the associated department.

Materials and methods

Study area

Zibo is a central city in the Shandong Province of China, located between latitude 35°55' N and 37°17' N and longitude 117°32' E and 118°31' E (Fig. 1). Zibo city has a temperate monsoon climate. By 2018, the city consists of five districts and three counties, covering an area of 5965 square kilometres, with a permanent resident population of 4.702 million [25].

Data collection

We collected daily reported HFMD cases during 1 January 2010 and 31 December 2019 from the Diseases Reporting Information System of the Zibo Center for Disease Control and Prevention. Reported cases were divided into two epochs (2010–2018, 2019), the former was defined as a training set to construct models, and the latter as a test set to verify the accuracy of predictive models. Patients were identified according to the diagnostic criteria defined by the National Health Commission of the People's Republic of China (<http://www.moh.gov.cn>).

Daily climatic data were collected from the Chinese Meteorological Data Sharing System (<http://cdc.nmic.cn>) and it is publicly accessible, including daily mean, maximum, minimum temperature (°C), precipitation (mm), mean air pressure (k pa), wind speed (m/s), relative humidity (%) and duration of sunshine (h) respectively.

Model construction

GAM with a Quasi-Poisson family was used to fit the association of weather factors and the number of HFMD [20]. We used categorical indicator variables to control the effects of the weekend and public holidays, and penalised smoothing splines functions to adjust for seasonality and long-term time trends in morbidity

[14]. Considering HFMD incubation period range from 3 to 10 days, a lag of up to 14 days was chosen for this study. Each lag day was modelled, and the optimal lag model used to analyse the relationship between meteorological factors and HFMD was selected by comparing the R^2 of models [26]. Akaike's information criterion for quasi-Poisson (QAIC) and previous similar studies were adopted to choose the df for climate factors [7, 27].

$$\log(u_t) = \beta_0 + ns(\text{weather}, df) + \dots + ns(\text{time}, 8 \times 10) + as.factor(\text{weekend}) + as.factor(\text{holiday}) \quad (1)$$

where: u_t is the expected number of HFMD on t day, β_0 is the intercept, $ns()$ present penalised smoothing splines.

Support vectors regression (SVR) model maps the nonlinear samples from the original space to the high-dimensional space, where conducts linear regression for finding an appropriate regression hyperplane, that is, the sample fits the optimal regression function, which is used for regression prediction. For the sample (x_i, y_i) , SVR allows a difference closed to ϵ between predicted value $f(x_i)$ and observed value y_i , which is equivalent to building an interval band with a width of 2ϵ , if the training sample falls into this interval band, it is considered to be predicted correctly [18, 28]. Thus, the SVR problem can be formalised as:

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + \mathcal{C} \sum_{i=1}^m l_i |f(x_i) - y_i| \quad (2)$$

where, \mathcal{C} is the constant of regularisation, l_i is the ϵ -insensitive loss function, ω , b is model parameters to be determined, m is the number of HFMD cases.

Random forest was applied to deal with classification and regression problems. If the CART tree is a regression tree, then the adopted principle is the least mean square deviation [26]. That is, for any partition feature A , the data was divided into sets D_1 and D_2 by the partition points, calculating the minimum mean square deviation of D_1 and D_2 sets and the corresponding feature and eigenvalue partition points. Random forest regression

Table 1. Summary statistics for daily HFMD cases and weather factors in Zibo city, China, 2010–2019 ($n = 56\ 058$)

Variables	Mean (s.d.)	Min	P_5	Median	Max	P_{95}
Number of cases	14.26 (17.50)	0.00	0.00	8.00	117.00	52.00
Maximum temperature (°C)	18.74 (10.65)	−8.00	1.00	20.60	38.10	33.10
Minimum temperature(°C)	8.82 (10.58)	−17.70	−8.00	9.80	28.50	23.80
Mean temperature (°C)	13.29 (10.52)	−13.10	−3.60	14.80	32.00	27.52
Relative humidity (%)	6.75 (1.80)	0.00	3.30	6.40	10.00	9.10
Sunshine duration (h)	6.34 (3.74)	0.00	0.00	7.30	13.20	11.50
Precipitation (mm)	981.34 (8.34)	962.00	968.30	918.50	1003.80	994.70
Wind speed (m/s)	1.70 (0.77)	0.00	0.70	1.60	6.90	3.20
Pressure (kpa)	2.0 (8.03)	0.00	0.00	0.00	142.40	11.90

s.d., standard deviation, P_x , interquartile range.

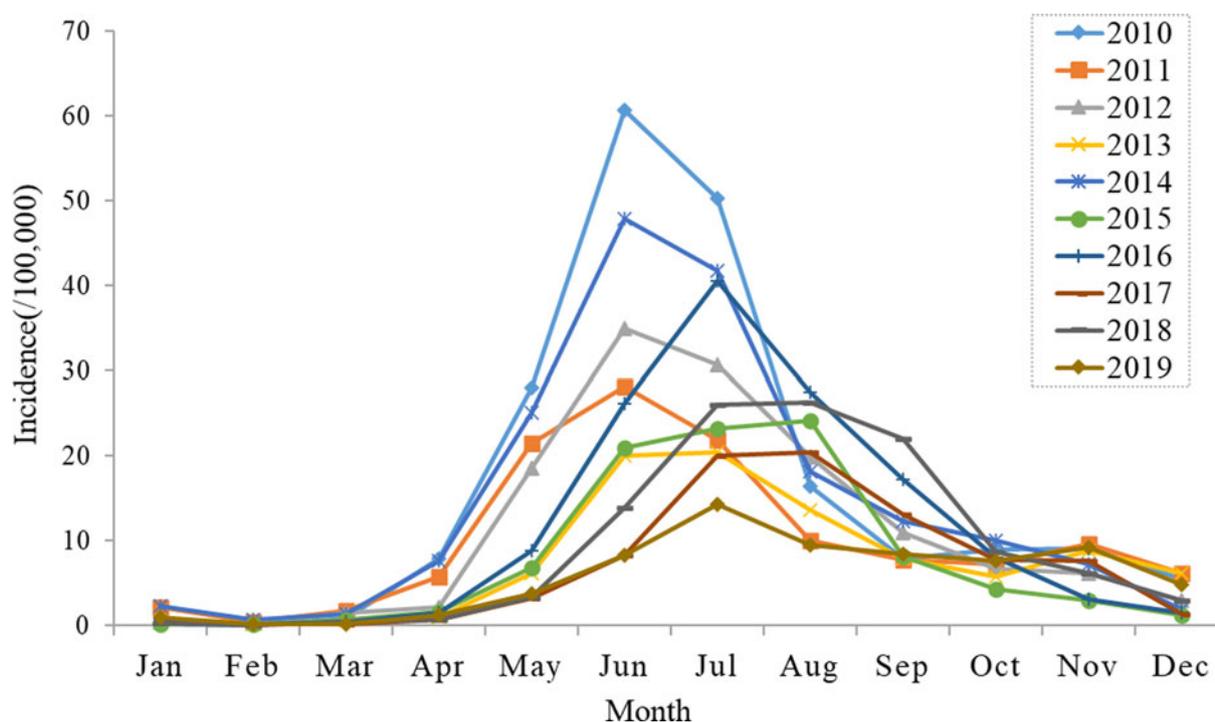


Fig. 2. Time series diagram of yearly HFMD cases from 2010 to 2019 in Zibo city, Shandong Province.

(RFR) run efficiently on high-dimension data sets, in addition, it is more accurate and robust to noise [19]. The main step of obtaining an optimal model is to determine two parameters: mtry and decision tree (ntree) [29]. When used in regression, mtry is usually 1/3 of all factors in the database. The prediction of CART trees is based on the average of leaf nodes, so the prediction of random forest is the average of the predicted values of all trees. The function is as follows:

$$\min_{A,s} \left[\min_{c_1} \sum_{x_i \in D_1(A,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in D_2(A,s)} (y_i - c_2)^2 \right] \quad (3)$$

where, c_1 is the sample output mean of D_1 dataset, and c_2 is the sample output mean of D_2 dataset. The predictive validity of the

models was evaluated by the criteria of the root mean square error (RMSE), mean absolute error (MAE) and R^2 .

Data collection and statistical analyses were performed by excel 2016 and R 4.0.1 software respectively in the study. We used the package ‘e1071’ for support vector regression [30], ‘randomForest’ for the RFR [31] and ‘mgcv’ for GAMs [32] respectively.

Results

Descriptive analysis

Daily HFMD counts and weather variables during the study period are shown in Table 1. Daily average cases are 14.26 and the daily maximum count reached 117 cases.

Figure 2 shows monthly HFMD incidence distribution, and Figure 3 displays the time series of daily HFMD cases and weather

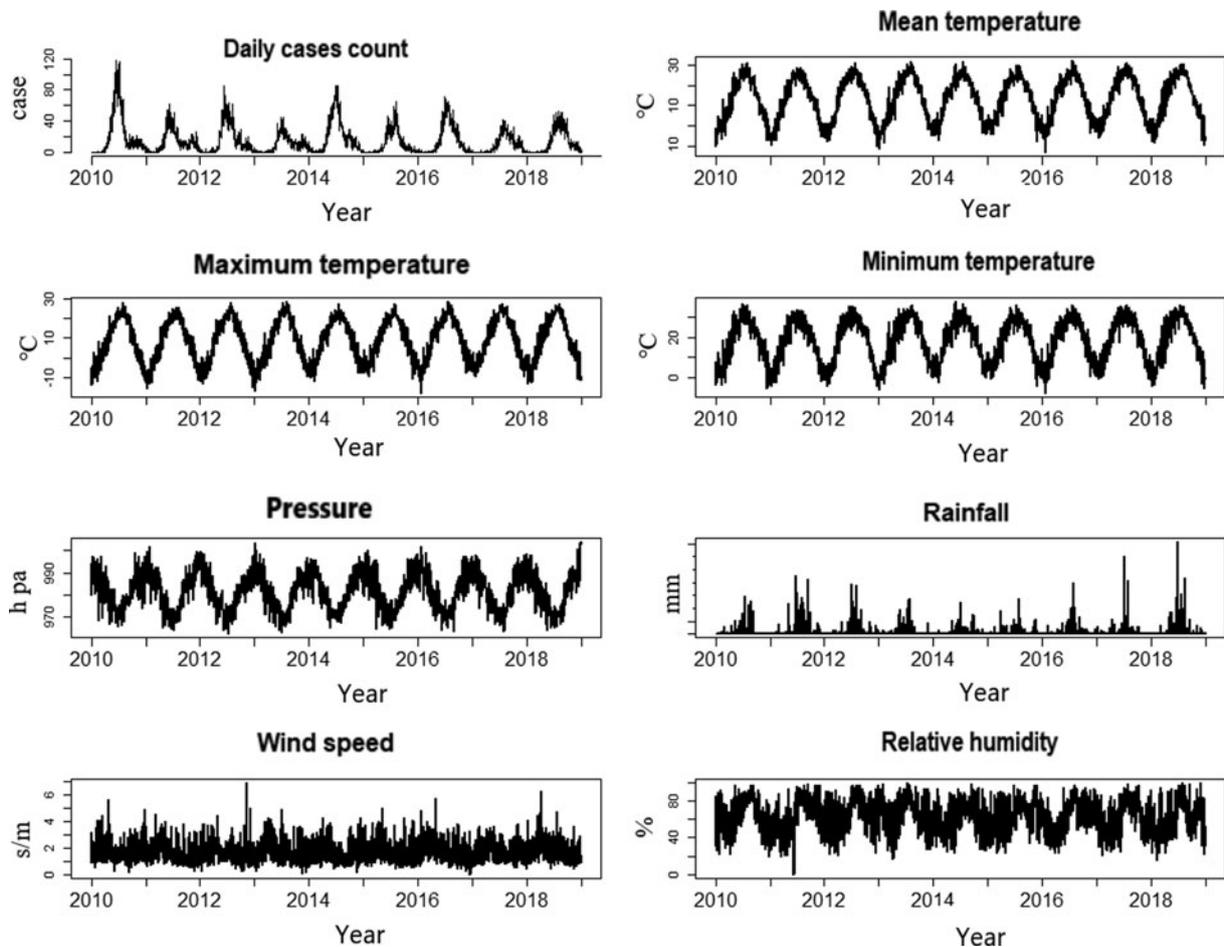


Fig. 3. Time series of yearly weather variables and HFMD cases during 2010–2019 in Zibo city, Shandong Province.

Table 2. Spearman's correlation coefficients between meteorological factors and daily HFMD cases in Zibo city, China, 2010–2019

Variables	PRE	PRS	WIN	TEM	RHU	SSD	T _{min}	T _{max}
PRE	1							
PRS	−0.27	1						
WIN	0 ⁿ	−0.11	1					
TEM	0.19	−0.87	(−0.01) ⁿ	1				
RHU	0.52	−0.33	−0.27	0.35	1			
SSD	−0.43	−0.12	0.22	0.20	−0.54	1		
T _{min}	0.29	−0.85	(−0.01) ⁿ	0.98	0.45	0.06	1	
T _{max}	0.11	−0.86	(−0.02) ⁿ	0.98	0.27	0.3	0.93	1
Number of cases	0.15	−0.50	(−0.07) ⁿ	0.60	0.30	0.09	0.61	0.58

Rⁿ: No Statistical Significance.

PRE, precipitation. PRS, pressure.

WIN, wind speed. RHU, relative humidity.

SSD, Sunshine duration. TEM, mean temperature.

T_{min}, daily minimum temperature. T_{max}, daily maximum temperature.

variables in Zibo city, 2010–2019. We found that the onset of HFMD presents an unimodal trend, increasing from March, concentrating from May to September (Fig. 2), and the onset of HFMD presents obvious seasonality and cycle occur almost every 2–3 years (Fig. 3). In addition, HFMD cases and climate vectors except for wind speed (Fig. 3).

Spearman correlation between weather factors and HFMD cases is depicted in Table 2. Variables positively correlated with the HFMD cases included precipitation, sunshine hours, mean temperature, maximum and minimum temperature and relative humidity, and variables negatively correlated included wind speed and air pressure. While there was no significant correlation between

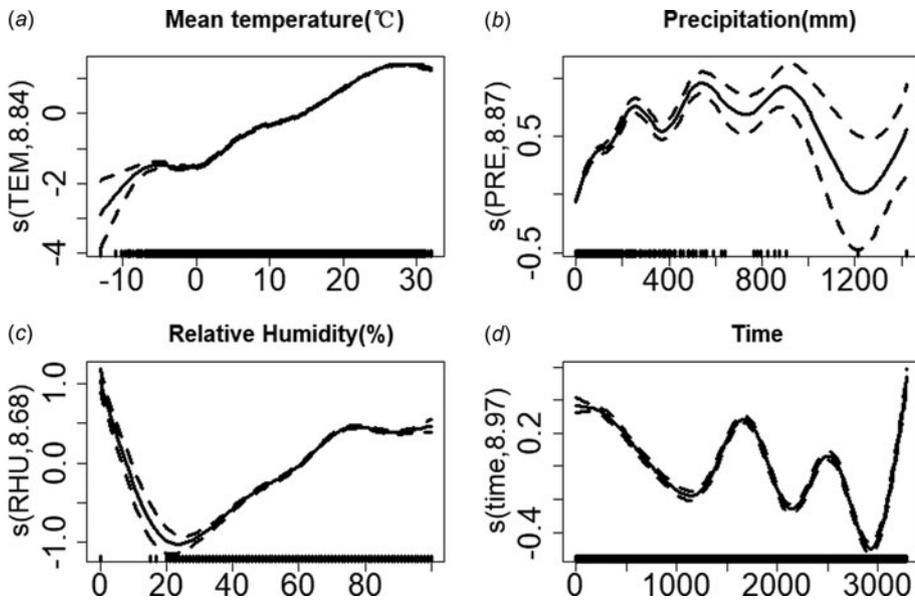


Fig. 4. Effects of meteorological factors on HFMD in Zibo city, Shandong Province, 2010–2019(A–D).

wind speed and sunshine hours and HFMD cases ($|r| < 0.1$). Taking the multicollinearity between variables into account, we can't take the variables that $|r|$ between variables are more than 0.8 into models. From Table 2, we find that there is multicollinearity between air pressure and the three temperature variables as well as among three temperature variables with each other ($|r| > 0.8$). Based on a previous study, we only take mean temperature into models [33]. In summary, weather variables taken into models are mean temperature, relative humidity and precipitation respectively.

Comparison of models

The independent variables included in SVR, RFR and GAM are the same, all including average temperature, relative humidity and precipitation respectively.

Figure 4 (the ordinate represents the smooth fitting value of weather factors to the incidence of HFMD) presents the nonlinear effects of mean temperature, precipitation and relative humidity on HFMD cases. We found the potential risks of HFMD increased as the temperature increased (below the threshold value of 28 °C), then became flatted. Remarkably, the data is relatively sparse and not significant (below -10 °C). As for relative humidity, there is a threshold at 77% about the effect of relative humidity on HFMD cases, when the risk of HFMD increases rapidly (less than 77%) and then decreases slightly and keep stable. While the effect of precipitation on HFMD cases presents 'M' shape (when precipitation < 600 mm), $s()$ increases rapidly, then falls (about 300 mm), then rise quickly (at 400 mm) and more than the previous peak.

Figure 5 shows the predictive results of the three models for HFMD cases in 2019 in Zibo city. The results of SVR prediction are closer to the observed values, predicting a large epidemic peak from May to August, and have the smallest predictive errors compared to GAM and RFR (Table 3). The accuracy of the three models was ranked in descending order as follow: SVR > GAM > RFR. The prediction results of SVR for HFMD are closest to the observed value, and the trend is roughly coincident to the original trend. By contrast, RFR and GAM can accurately predict the incidence trend of HFMD, but cannot accurately predict the specific

number of cases. Therefore, the SVR model was effective and applicable for the prediction of HFMD in Zibo city and RFR and GAM fail to be used to predict HFMD cases.

Discussion

HFMD has been increasingly recognised as a significant health problem, and climate change may increase the burden of HFMD. An early warning could help health agencies to release potential disease outbreaks through the combination of disease surveillance information and climate factors [34]. This study assessed the effect of temperature, precipitation and relative humidity on HFMD cases, and constructed the most appropriate predictive model for HFMD based on meteorological factors.

The purpose of this study is to explore the influence of weather on HFMD and predict the incidence of disease on this basis. Therefore, other factors such as age, immunity of patients and the influence factors of virus infection are not considered, and only meteorological factors are taken into the model. Although the influence of meteorological factors on HFMD has been explored in published studies, the effects of these studies are inconsistent. Effect of temperature on HFMD presents an inverted V-shape in Guangzhou city and Beijing [35]. On the contrary, a linear correlation between temperature and HFMD cases was found in other studies of Guangzhou [34]. Additionally, the impact of wind speed on HFMD has always been controversial. Studies by WangHao *et al.* and Wuxinrui *et al.* found wind speed is a protective factor for HFMD [9, 15], the possible explanation is that wind speed increases the resistance to transmission of viruses, blocking the transmission of viruses, which reduces the chance of infected disease, while some studies show that wind speed is a risk factor for HFMD [10, 13, 14]. In this study, we observed that the effects of temperature, relative humidity and precipitation on HFMD presented nonlinear trends, and there are thresholds, which are consistent with studies in Guangdong province, Japan, Singapore [36–38]. These regional discrepancies could be likely attributed to the diversities of geographic, socio-economic and population variables or methods and data sources. Although there is no unified conclusion about biological mechanisms of climate factors on HFMD, acceptable explanations in

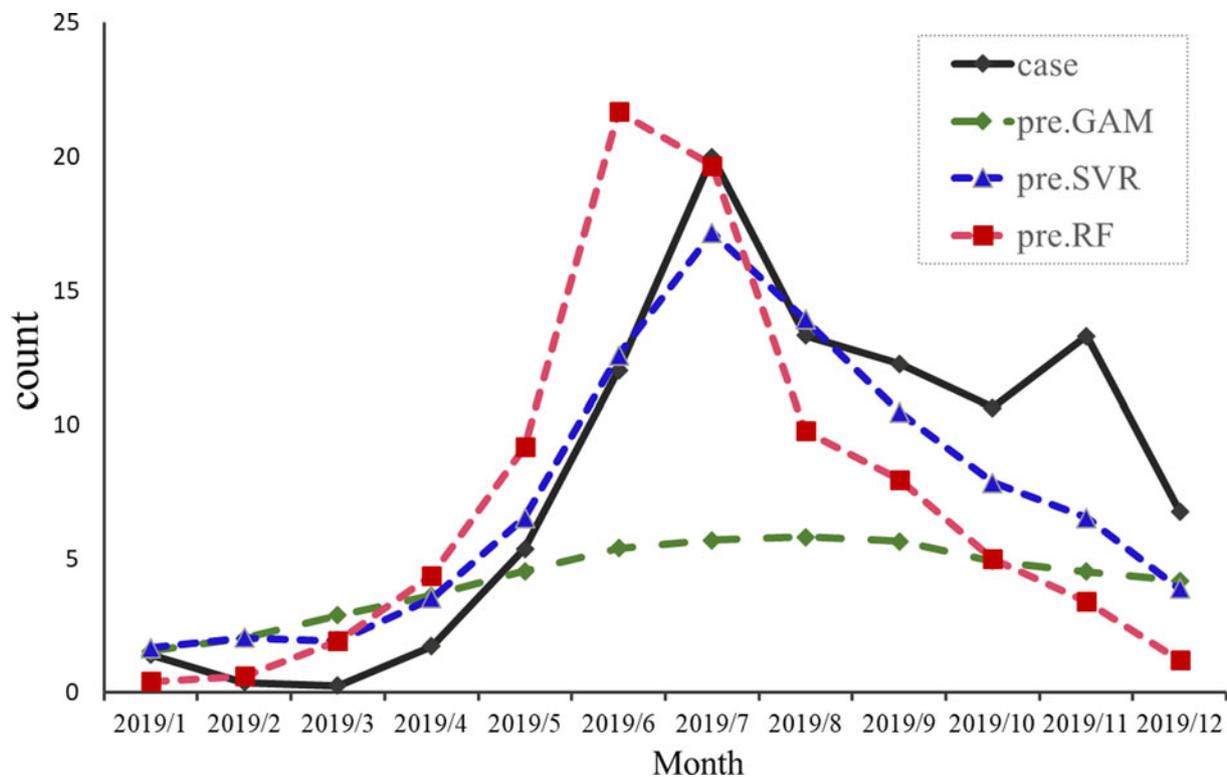


Fig. 5. Predictive results of three models in Zibo city, Shandong Province (2019) respectively.

Table 3. Comparison of the predictive performance of three different predictive models

Model	Fitted value (2010–2018)			Predictive value (2019)		
	MAE	MSE	RMSE	MAE	MSE	RMSE
GAM	7.29	125.72	11.21	5.41	55.29	7.43
SVR	9.41	182.37	13.50	1.46	5.44	2.33
RFR	14.83	547.06	23.38	7.99	113.96	10.68

published studies mainly include three aspects: survival and reproduction of pathogens, the host population and environmental factors [39]. Appropriate temperature and humidity are conducive to the propagation of the virus, increasing the probability of disease, and wind speed will accelerate the spread of the virus, thus leading to more infections. Therefore, related departments could use the impact of meteorological factors on HFMD in specific areas to predict the incidence, providing a scientific basis for the prevention and control of HFMD.

Based on the impacts of meteorological factors on HFMD, we constructed and compared three predictive models. The comparison of prediction may help epidemiologists choose the most appropriate model for a given situation. Although three models can solve nonlinear prediction problems, for the predictive efficiency, the order of the three predictive models is SVR > GAM > RFR. GAM excelled flexibly controlling the confounding factors affecting diseases, such as time series, weekend and holiday effect. The adjusted degree of freedom of factors, it has been widely used to predict a variety of diseases, but it is too complex in calculation and sensitive to the choice of parameters. Machine learning methods don't require complex calculations, and the choice of

parameters is less likely to cause a large change in error. In contrast, machine learning is less affected by parameter selection. SVR model is a new approach of machine learning provided better performance and greater potential in predicting infectious disease. It uses a nonlinear kernel function followed by linear regression in this feature space and can minimise the actual risk by seeking to minimise the structural risk and better solves problems such as small samples and nonlinearity, which are the most prominent feature of SVR [40]. Because it carries out linear regression in high-dimensional feature space by insensitive loss function and makes use of strong data inclusion to reduce the complexity of the model for a long time. It avoids the traditional process from induction to deduction, achieves the efficient inference from training sample to prediction sample, greatly simplifies the traditional regression problem. Although the SVR model algorithm is simple and has good robustness, it fails to implement for large-scale training sample, which will consume a lot of machine memory and operation time. However, due to fast training speed and two random principles, RFR can process a large amount of data and not easy to generate overfitting problems. In addition, it has a strong ability to make use of data, and obtain the

importance of eigenvalue, but takes a long time when the number of decision trees. In summary, for prediction of HFMD in 2019, Zibo city, the results of SVR are closest to the observed value, and has minimal predictive errors.

There are several limitations which should be noted. Firstly, we only consider meteorological factors as risk factors for HFMD, but other influencing factors, such as population density, economic conditions, geographical environment are not taken into consideration, which may decrease the predictive ability of the model [12]. Secondly, due to discrepancy in geographical location, living custom and other aspects, the results may only apply to HFMD in Zibo city, and cannot be extended to other regions with different weather patterns.

In summary, we compared three models based on the influence of weather factors on HFMD incidence, and concluded that the SVR is superior to GAM and RFR, and can be used for forecasting HFMD cases, providing scientific evidence for further application in HFMD prevention and control in Zibo city, Shandong Province.

Financial support. This study is supported by the State Key Research Development Program of China (2019YFC1200502) and the Project supported by the General Administration of Customs, P.R. China (2019HK125).

Conflicts of interest. All the authors have declared no competing interests.

Ethical standards. HFMD data were provided by the Zibo Center for Disease Control and Prevention. No informed consent was required because no individual-level analysis was performed. This study was supported by the State Key Research Development Program of China (2019YFC1200502) and the Project supported by the General Administration of Customs, P.R. China (2019HK125). The information contained in the patient records is anonymised and identified prior to analysis. Only aggregated data were analysed and reported.

Data availability statement. The data of this study are available from the Zibo Center for Disease Control and Prevention but restrictions apply to the availability of these data, which are used under licence for the current study, and so cannot be shared publicly.

References

- Xing W *et al.* (2014) Hand, foot, and mouth disease in China, 2008–12: an epidemiological study. *Lancet Infectious Diseases* **14**, 308–318.
- Li R *et al.* (2017) Epidemiological characteristics and spatial-temporal clusters of mumps in Shandong Province, China, 2005–2014. *Scientific Reports* **7**, 46328.
- Du Z *et al.* (2019) Interactions between climate factors and air pollution on daily HFMD cases: a time-series study in Guangdong, China. *Science of the Total Environment* **656**, 1358–1364.
- Schmidt CW (2018) More cases of hand, foot, and mouth disease in China: a consequence of climate change? *Environmental Health Perspectives* **126**. doi: <http://dx.doi.org/10.1289/EHP3778>.
- Onozuka D and Hashizume M (2011) The influence of temperature and humidity on the incidence of hand, foot, and mouth disease in Japan. *Science of the Total Environment* **410–411**, 119–125.
- Xiao X *et al.* (2017) The exposure-response relationship between temperature and childhood hand, foot and mouth disease: a multicity study from mainland China. *Environment International* **100**, 102–109.
- Zhu L *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. *International Journal of Environmental Research and Public Health* **12**, 8691–8704.
- Zhang Z *et al.* (2016) Short-term effects of meteorological factors on hand, foot and mouth disease among children in Shenzhen, China: non-linearity, threshold and interaction. *Science of the Total Environment* **539**, 576–582.
- Wang H *et al.* (2015) Detecting the association between meteorological factors and hand, foot, and mouth disease using spatial panel data models. *International Journal of Infectious Diseases* **34**, 66–70.
- Qi H *et al.* (2018) Impact of meteorological factors on the incidence of childhood hand, foot, and mouth disease (HFMD) analyzed by Dlnms-based time series approach. *Infectious Diseases Poverty* **7**, 7.
- Huang Y *et al.* (2013) Effect of meteorological variables on the incidence of hand, foot, and mouth disease in children: a time-series analysis in Guangzhou, China. *BMC Infectious Diseases* **13**, 134.
- Liu W *et al.* (2015) Spatiotemporal dynamics of hand-foot-mouth disease and its relationship with meteorological factors in Jiangsu Province, China. *PLoS One* **10**, e0131311.
- Gou F *et al.* (2018) Different responses of weather factors on hand, foot and mouth disease in three different climate areas of Gansu, China. *BMC Infectious Diseases* **18**, 15.
- Wang P *et al.* (2017) Seasonal modeling of hand, foot, and mouth disease as a function of meteorological variations in Chongqing, China. *International Journal of Biometeorology* **61**, 1411–1419.
- Wu X *et al.* (2017) Spatio-temporal clustering analysis and its determinants of hand, foot and mouth disease in Hunan, China, 2009–2015. *BMC Infectious Diseases* **17**, 645.
- Zhan Z *et al.* (2019) Real-time forecasting of hand-foot-and-mouth disease outbreaks using the integrating compartment model and assimilation filtering. *Scientific Reports* **9**, 2661.
- Zhu L *et al.* (2016) Assessment of temperature effect on childhood hand, foot and mouth disease incidence (0–5years) and associated effect modifiers: a 17 cities study in Shandong Province, China, 2007–2012. *Science of the Total Environment* **551–552**, 452–459.
- Zou JJ *et al.* (2019) Application of a combined model with seasonal autoregressive integrated moving average and support vector regression in forecasting hand-foot-mouth disease incidence in Wuhan, China. *Medicine (Baltimore)* **98**, e14195.
- Cipullo S *et al.* (2019) Predicting bioavailability change of complex chemical mixtures in contaminated soils using visible and near-infrared spectroscopy and random forest regression. *Scientific Reports* **9**, 4492.
- Laureano-Rosario AE *et al.* (2018) Application of artificial neural networks for dengue fever outbreak predictions in the northwest coast of Yucatan, Mexico and San Juan, Puerto Rico. *Tropical Medicine and Infectious Disease* **3**, 5. doi: 10.3390/tropicalmed3010005.
- Hempstalk K *et al.* (2015) Machine learning algorithms for the prediction of conception success to a given insemination in lactating dairy cows. *Journal of Dairy Science* **98**, 5262–5273.
- Ghafouri-Kesbi F *et al.* (2017) Predictive ability of random forests, boosting, support vector machines and genomic best linear unbiased prediction in different scenarios of genomic evaluation. *Animal Production Science* **57**. doi: 10.1071/AN15538.
- Van Der Heide EMM *et al.* (2019) Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *Journal of Dairy Science* **102**, 9409–9421.
- Xiao D *et al.* (2015) Modeling and predicting hemorrhagic fever with renal syndrome trends based on meteorological factors in Hu County, China. *PLoS One* **10**, e0123166.
- Xu Q *et al.* (2017) Forecasting the incidence of mumps in Zibo city based on a Sarima model. *International Journal of Environmental Research and Public Health* **14**, 925. doi: 10.3390/ijerph14080925.
- Yu G *et al.* (2019) Short-term effects of meteorological factors and air pollution on childhood hand-foot-mouth disease in Guilin, China. *Science of the Total Environment* **646**, 460–470.
- Chen S *et al.* (2019) The application of meteorological data and search index data in improving the prediction of HFMD: a study of two cities in Guangdong Province, China. *Science of the Total Environment* **652**, 1013–1021.
- A Comparative Study between Least-Squares Support Vector Machines and Partial Least Squares in Simultaneous Spectrophotometric Determination of Cypermethrin, Permethrin and Tetramethrin.
- Wang Y *et al.* (2019) Calibration of a low-cost Pm2.5 monitor using a random forest model. *Environment International* **133**(Pt A), 105161.

30. **Sananmuang T *et al.*** (2021) Support vector regression algorithm modeling to predict the parturition date of small - to medium-sized dogs using maternal weight and fetal biparietal diameter. *Veterinary World* **14**, 829–834.
31. **Zhao D and Wu Q** (2018) An approach to predict the height of fractured water-conducting zone of coal roof strata using random forest regression. *Scientific Reports* **8**, 10986.
32. **Li R *et al.*** (2019) Climate-driven variation in mosquito density predicts the spatiotemporal dynamics of dengue. *Proceedings of the National Academy of Sciences of The United States of America* **116**, 3624–3629.
33. **Zhang D *et al.*** (2019) The relationship between meteorological factors and mumps based on boosted regression tree model. *Science of the Total Environment* **695**, 133758.
34. **Chen C *et al.*** (2014) Short-term effects of meteorological factors on children hand, foot and mouth disease in Guangzhou, China. *International Journal of Biometeorology* **58**, 1605–1614.
35. **Ma E *et al.*** (2010) Is hand, foot and mouth disease associated with meteorological parameters? *Epidemiology and Infection* **138**, 1779–1788.
36. **Urashima M *et al.*** (2003) Seasonal models of Herpangina and hand-foot-mouth disease to simulate annual fluctuations in urban warming in Tokyo. *Japanese Journal of Infectious Diseases* **56**, 48–53.
37. **Xu Z *et al.*** (2019) The effect of temperature on childhood hand, foot and mouth disease in Guangdong Province, China, 2010–2013: a multicity study. *BMC Infectious Diseases* **19**, 969.
38. **Hii YL *et al.*** (2011) Short term effects of weather on hand, foot and mouth disease. *PLoS One* **6**, e16796.
39. **Yi L *et al.*** (2019) The impact of climate variability on infectious disease transmission in China: current knowledge and further directions. *Environmental Research* **173**, 255–261.
40. **Zhang X *et al.*** (2014) Applications and comparisons of four-time series models in epidemiological surveillance data. *PLoS One* **9**, e88075.