Response to Reviewer 2 Comments

Thank you very much for giving us an opportunity to revise our manuscript. We appreciate the editor and reviewers very much for their constructive comments and suggestions on our manuscript. We have studied the comments carefully and made corrections. Hope meet with approval.

**Point 1**
Weather data of 2009 is not representative of 2009-2012 period. See for example,

- [https://ggweather.com/enso/oni.htm](https://ggweather.com/enso/oni.htm)
- It is possible this has played a role in such low RR for the environmental factors.

**Response 1**

- We updated the meteorological data of the article. The current data is complete data for 2009-2012.
- There were 26 surveillance stations within Guangdong province.
- The meteorological surveillance stations were mapped in Figure 2 of the article.
- All data was obtained from the China Meteorological Data Sharing Service System.
- The monthly county level meteorological variables were calculated using ordinary spatial kriging methods based on the 26 surveillance stations.

**Point 2**
Model is potentially ill-posed with identifiability and singularity problems.

**Minor point**

**Point 2.1**
This is an empirical model based on assumptions and is not a truth of nature. As such all terms, “followed” should be replaced by “assumed to follow” or “assumed”.

**Response 2.1**
We modified them in the article.

**Point 2.2**
Provide a justification for CAR model (or why not a SAR model), for example. Did your exploratory analysis suggest it? The justification could simply be a reference relevant for infectious diseases spatial epidemiology.

**Response 2.2**
Spatially structured random effect is commonly assumed to follow a conditional autoregressive prior structure in infectious diseases spatial epidemiology. Related references have been added.

• Zhang, X.; Xu, C.; Xiao, G. Space-time heterogeneity of hand, foot and mouth disease in children and its potential driving factors in Henan, China. BMC INFECT DIS 2018, 18, 638.

Point 2.3
Figures 4 and 5 caption needs clarity. It is not immediately obvious what is the variable that has been depicted in Figure 5. All we see is possible heterogeneity in space-time. Heterogeneity is not necessarily the same as spatio-temporal interaction. I suggest that the authors argue with example from the plots wht they mean by spatio-temporal interaction, here.

Response 2.3
In the article, we supplemented the description of spatiotemporal interaction and heterogeneity with q-statistic.

Point 2.4
There are several grammatical flaws.

Response 2.4
We re-examined the article and the modified the syntax errors we found.

Major issues

Point 2.5
The authors want to separately model the year effect and month effect (seasonality). For this they have suggested an additive combination of month and year effect- assuming orthogonality between these two effects. This however ignores the interaction between months and year. This is important because months are a subset of years. The correct framework for such modelling is SARIMA models that allows for serial and seasonal correlations. The data should have only one unit of time, either month or years, usually the smaller of the two units. Else singularity – via multicollinearity (for example) - is a common issue. The orthogonality assumption might still be considered if the authors had used monthly weather patterns for all months included in the study (and not just 2009).

Response 2.5
• Thank you for your reminder. Months are subsets of years, so it was not appropriate to separately model the year effect and month effect. In the new model, we removed the year component and only kept the smaller unit month component, and modified the spatiotemporal interaction item to the interaction between the space item and the month item.
• We calculated the variance inflation factor (VIF) for all candidate variables to assess the multicollinearity; variables with a VIF less than 5 were selected for inclusion in the model. Table 2 showed the results of the multicollinearity analysis.

Point 2.6
Is there a possibility of identifiability between $v_i$ and $\mu_i$? A CAR model already includes an independent Gaussian component!

**Response 2.6**
The Besag, York and Mollie (BYM) model was used to model the common spatial component, which consisted of two components: spatially structured random effect $\mu_i$ and spatially unstructured random effect $\nu_i$. The spatial structured effect was used to explain the phenomena that neighbouring regions tend to have similar overall disease risks, and spatial unstructured effect was used to explain the phenomena having no spatial autocorrelation. In the spatial structure, a conditional autoregressive prior was commonly used, and spatially unstructured random effect was assumed to follow a Gaussian distribution. There is a “bym” in the R-INLA package to complete the identification of $\mu_i$ and $\nu_i$.


**Point 2.7**
What is the rationale for selecting random walks for month and year once the authors have already modelled month and year effect? I hope that the authors have considered the fact that including a random walk introduces an infinite variance non-stationarity on the process! Have the authors considered the implications of this in the context of HFMD prevalence/incidence? If the residual effects in your empirical model is a random walk it would often indicate that the proposed model is ill-posed. There are tests (unit root) that can be performed to assess if you really need a random walk assumption for a particular variable.

**Response 2.7**
We have considered the fact that including a random walk introduces an infinite variance non-stationarity on the process. For an 1 order random walk, its variance is $t \times \sigma^2$, when $t \to \infty$, its variance will become infinite. However, in reality, researches, time is limited. In our study, the length of time was 4 years, so during this time, we considered that the variance of the random walk will not be so large, so we put random walk into the model. We read the relevant references and found that random walk model was also used in related studies.


**Point 2.8**
How does the estimation procedure differentiate (identifiability) between the additive components, $\gamma$ and $\varphi$? The same goes for $\tau$ and $\omega$.

**Response 2.8**
In the previous model, I defined $\gamma$ and $\varphi$ as the structured random effect and unstructured random effect of the year, assumed to follow 1 order random walk and Gaussian distribution, respectively. The corresponding R program was as follows:
Now the model in the article has been revised according to the comments of the reviewers, removing the yearly effect and retaining only the monthly effect in temporal component.

**Point 2.9**  
Could you please mention the total number of parameters and what was the total sample size?  
**Response 2.8**  
There was an intercept, six regression coefficients, spatial component (structured and unstructured), temporal component (structured and unstructured), and spatiotemporal interaction component in the article. A total of 12 parameters. There were 911640 reported HFMD cases in Guangdong Province from 2009-2012.

**Point 2.10**  
What are the spatial and temporal resolution of all explanatory factors? That is, do you have all variables available at county level? If not please discuss the limitations of your model how are the data aggregated. Biases caused due to data aggregation are well studied. Please refer to relevant articles.  
**Response 2.10**  
We have obtained average GDP data at county level, but in terms of annual units. Monthly meteorological data was obtained from 26 stations within Guangdong province, county level meteorological data was calculated using ordinary spatial kriging methods. The corresponding limitations have been added to the article.

**Point 2.11**  
Deviance information criterion of an ill posed model is of little value.  
**Response 2.11**  
We modified the model and recalculated the DIC value.

**Point 3**  
The study lacks novelty and reads like a replication study- in line with other provinces within China.

**Point 3.1**  
Also this reviewer could not be sure if proposing such a complex model has enhanced our understanding of the epidemiology of the disease given the numerous epidemiology studies on HFMD already reported from China (and the neighbourhood) on the same disease from other provinces.  
**Response 3.1**  
Our research incorporated spatiotemporal effect in the regression that controlled uncertainties resulted from residual confounding, which made the estimation of the impact of meteorological variables more accurate. Through the study of spatial effect and temporal effect, the hot spots and high-risk time of HFMD were identified. Through spatiotemporal interaction detection, it was found that the spatial effect of HFMD had positive nonlinear enhancement with the
temporal effect. After adding the spatiotemporal interaction component to the model, the goodness of fit of the model was greatly increased. In the spatiotemporal model with no interaction term (model 4), DIC value was 202885.9. When the interaction component was added, DIC value was reduced by 77.8% to 45036.5. Therefore, adding interaction component can improve the accuracy of the model, which is more important in the prediction model and was not considered in many previous related studies. It was a comprehensive study of HFMD in terms of spatiotemporal effects and meteorological variables.

**Point 3.2**  
Given the low risk ratios (RR) for the environmental factors one might produce the same maps simply by interpolating a monthly model of counts based at district level. That is, on the one hand the model seems ill-posed and on the other it would have questionable predictive impact.  

**Response 3.2**  
When the spatiotemporal effect was added to the model, as a control of confounding factors, the relative risk of meteorological variables would be less than that without spatiotemporal effect, and this was more accurate results. Although the regression coefficients of meteorological variables were not high, meteorological variables also had an impact on HFMD. For example, when monthly average temperature increased 1 ℃, the relative risk of HFMD would increase 4.5% (95%CI: 2.1%-6.9%) in Guangdong. Therefore, we think that models incorporating meteorological variables will be more advantageous than models that simply use the reported case.

**Point 4**  
In summary the Introduction starts the well on the epidemiology side. But then ends with this statement- “Traditional methods, such as regression, time series analysis, spatiotemporal scan statistics, etc., 79 cannot be used to address the aforementioned knowledge gaps. Therefore, this study chose the 80 Bayesian spatiotemporal model, which added the spatiotemporal effect to the generalized additive 81 model, to identify spatiotemporal variations and the effects of potential predictors [20].” It is an unsubstantiated statement and does not follow scientific rigour.  

**Response 4**  
We deleted this unsubstantiated statement.