BMC Public Health

A spatial-temporal statistical analysis of health seasonality: explaining HFMD infections within a children population along the Vietnamese south central coast

Phuong N. Truong[1]*, Thuong Vu Nguyen[2], Thao Thi Thanh Nguyen[2] and Alfred Stein[1]

Abstract

Background: Various neglected tropical diseases show spatially changing seasonality at small areas. This phenomenon has received little scientific attention so far. Our study contributes to advancing the understanding of its drivers. This study focuses on the effects of the seasonality of increasing social contacts on the incidence proportions at multiple district level of the childhood hand-foot-mouth disease in Da Nang city, Viet Nam from 2012 to 2016.

Methods: We decomposed the nonstationary time series of the incidence proportions for the nine spatial-temporal (S-T) strata in the study area, where S indicates the spatial and T the temporal stratum. The long-term trends and the seasonality are presented by the Fourier series. To study the effects of the monthly average ambient temperature and the period of preschooling, we developed a spatial-temporal autoregressive model.

Results: Seasonality of childhood hand-foot-mouth disease incidence proportions shows two peaks in all spatial strata annually: large peaks synchronously in April and small ones asynchronously during the preschooling period. The peaks of the average temperature are asynchronous with the seasonal peaks of the childhood hand-foot-mouth disease incidence proportions in the period between January and May, with the negative values of the regression coefficients for all spatial strata, respectively: $\beta_{T_11}^{S_1} = -0.18 \pm 0.07; \beta_{T_11}^{S_2} = -0.25 \pm 0.09; \beta_{T_11}^{S_3} = -0.14 \pm 0.05$. The increasingly cumulative preschooling period and the seasonal component of the incidence proportions are negatively correlated in the period between August and December, with the negative values of the regression coefficients for all temporal strata, respectively: $\beta_{T_32}^{S_1} = -0.40 \pm 0.01; \beta_{T_32}^{S_2} = -0.29 \pm 0.00; \beta_{T_32}^{S_3} = -0.25 \pm 0.01$.

Conclusions: The study shows that social contact amongst children under five years of age is the important driving factor of the dynamics of the childhood hand-foot-mouth disease outbreaks in the study area. The preschooling season when children's contact with each other increases stimulates the geographical variation of the seasonality of childhood hand-foot-mouth disease infections at small areas in the study area.

Keywords: Health seasonality, NTDs, Childhood HFMD, STAR, Fourier analysis, Disease, dynamics, health geography, small area analysis.

* Correspondence: p.truong@utwente.nl
[1]Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, The Netherlands
Full list of author information is available at the end of the article

Background

Neglected tropical diseases (NTDs) have shown evidently seasonal patterns with various timings, amplitudes and types [1–3]. NTDs remain socio-economic burdens in developing countries, especially in low-income communities in Asia and Africa. The considerable seasonality of these NTDs is affected by seasonal changes of both environmental and social risk factors [4, 5]. Understanding the regularity or irregularity of NTDs' outbreaks is necessary to obtain optimal control over the disease spread [6], to reduce their burdens and to reach the UN sustainable development goals for health [7].

Amongst others, the infections of *Coxsackievirus A16* (CA16) and human *Enterovirus 71* (EV71) that mainly cause hand-foot-mouth disease (HFMD) in children under 5 years of age (U5s) have been recognized as an emerging public health problem in North East and South East Asian countries [8, 9]. So far, studies of the seasonal variations of childhood HFMD infections have mainly focused on climatic driving factors such as temperature, humidity and rainfall [9]. The effects of seasonally changing social contacts on the dynamics of childhood HFMD for small areas have been overlooked, even though HFMD viruses are mainly transmitted by directly physical contacts between infected and non-infected children [10]. Additional research has been done for northeastern Asian countries [11–13], whereas some studies were also done for southeastern Asian regions [14, 15]. Despite the increasing application of mathematical models to understand the epidemiology of infectious diseases, spatial-temporal statistical methods (STS) have not yet been well recognized and leveraged for studying seasonality of NTDs, i.e. to quantitatively explain the spatial-temporal dynamics of the disease outbreaks [16], especially for small geographical areas.

In this research, we study the seasonality of childhood HFMD in Viet Nam and its geographical variations at small areas as affected by seasonally changing weather and the social contacts. The study of Horby et al. in Viet Nam in 2011 [17] shows that the physical contact amongst U5s themselves is the most intensive, especially during the period at preschool. Hence, the social contact amongst U5s is measured by the annually cumulated period of time they spend at preschool since their first day at preschool. We distinguish between the trend and the seasonality. The trend establishes the long-term changing pattern of the mean level, whereas the seasonality represents yearly periodic variations. We propose a two stage STS analysis to extract the seasonal patterns and to identify the driving factors of the incidences of childhood HFMD infections in Da Nang city, Viet Nam as an application. In so doing, we aimed to identify and understand the effects of seasonally changing social contacts of U5s on the dynamics of HFMD outbreaks at multiple district level.

Methods

Study area and data

Da Nang city (Fig. 1), the biggest city on the south central coast of Viet Nam has an area of 1,285 km^2 and a population of more than 1 million people in 2016 [18]. Approximately 8% of Da Nang city's population consists of U5s [18]. According to the guideline from the Vietnamese Ministry of Health on diagnosis and treatment of HFMD, a confirmed HFMD case is defined when a child has a positive RT-PCR assay for CA16 or EV71. Weekly counts of new cases, i.e. the incidences from seven mainland districts of the city from 2012 to 2016 were obtained from the published reports of Da Nang city Preventive medicine center (https://ksbtda-nang.vn/). Because these weekly counts have many small number of cases (< 5 cases per week), statistics based upon them are not reliable. To avoid this small number problem, the monthly count of new cases that is the sum of the weekly counts of new cases from this month was analyzed in this study. Note that the first and last weeks of the months can include 1 to 3 days of the previous or the next months. The demographical data were obtained from the demographic statistics of Da Nang city. The monthly population of U5s was interpolated from the yearly data with the assumption of a constant monthly growth rate. The monthly meteorological data from 2012 to 2016 were obtained from the Viet Nam institute of meteorology, hydrology and climate change (http://www.imh.ac.vn).

Spatial-temporal heterogeneity analysis

To analyze the spatial-temporal heterogeneity and identify the predominant risk factors of HFMD, we compared the q-statistic of the nine spatial-temporal stratification to that of without stratification: [19]; where: N is the

$$q = 1 - \frac{\sum_{h=1}^{H} N_h \sigma_h^2}{N\sigma^2}$$

total number of HFMD cases in the study area and σ^2 is the corresponding total variance, N_h and σ_h^2 are those of various spatial-temporal strata that are considered in the statistical test, H is the total number of these strata. The q-statistic for heterogeneity analysis regarding covariates was provided by the `geodetector` library of R [19, 20].

Spatial-temporal stratification

To evaluate the effects of U5s as the population at risk on the magnitude of the seasonality of the childhood HFMD incidences, we stratified the districts by the density of U5s. This classification resulted into three spatial strata (Fig. 1): S_1 is an urban stratum with the highest density of U5s (> 500 per km^2), S_3 is the rural stratum having the lowest density (< 100 per km^2) and S_2 is located in between.

We transformed the monthly counts of HFMD incidences to the monthly HFMD incidence proportions in order to correct the effects of the internal heterogeneity of
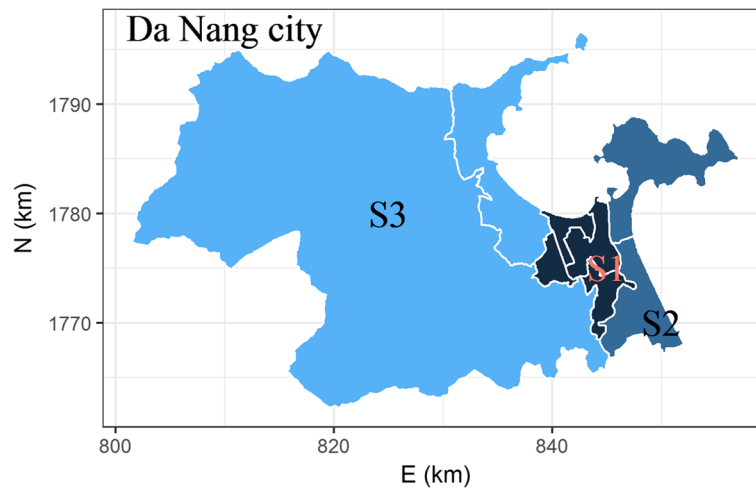
Fig. 1 Spatial strata in the mainland of Da Nang city, Viet Nam: S_1 is the most densely populated districts, S_3 has the lowest population density as compared to S_1 and S_2

the population at risk within the spatial strata. Let $i = 1:12$ index the months, $j = 1:5$ the years, $d = 1:7$ the districts and $s = 1:3$ the spatial strata. IR_{ij}^s is used to denote the incidence proportion at spatial stratum s in month i of year j. To account for unequal numbers of days in months, the IR_{ij}^s were adjusted as: $IR_{ij}^s = \dfrac{365.25 \times \sum\limits_s I_{ij}^d}{12 \times D_i \times \sum\limits_s N_{ij}^d \times 1,000}$ [21], where I_{ij}^d is the monthly incidence proportion, N_{ij}^d is the population of U5s, D_i is the number of days per month, $\sum\limits_s I_{ij}^d$ is the sum of HFMD incidences of the districts and $\sum\limits_s N_{ij}^d$ is the total population of U5s belonging to the stratum s. This adjustment removes the false variations of IR_{ij}^s attributed to the variations of the at-risk populations and the length of time unit, partly due to the use of the weekly reported data. IR_{ij}^s therefore equals the number of infected U5s per 1,000 U5s at risk per month, provided that removing the infected children from the population at risk only has a marginal correction effect.

Modelling seasonality

The seasonality of IR_{ij}^s was diagnosed by using the seasonal index (SI). The original value of IR_{ij}^s for each stratum s in month i in year j was represented as the percentage of the mean value of the IR_{ij}^s for stratum s in year j. SI of stratum s in month i equals the median of the percentages from all 5 years: $SI_i^s = median\left(\dfrac{IR_{ij}^s}{12^{-1}\sum\limits_{i=1}^{12} IR_{ij}^s} \times 100, j = 1:5\right)$ [22]. The median allows for the inclusion of the effects of

extreme values over the years. The standard error of SI was derived from the bootstrap confident interval assuming a normal distribution. A value of $SI_i^s > 100\%$ indicates that the incidence proportion is above the yearly average incidence proportion and vice versa.

For the analysis of the dynamics, we applied to IR_{ij}^s an additive conceptual model with three main components: a trend (μ), a seasonality (ξ) and random noises (ε):

$$IR_{ij}^s = \mu\left(IR_{ij}^s\right) + \xi\left(IR_{ij}^s\right) + \varepsilon\left(IR_{ij}^s\right). \qquad (1)$$

We used the Fourier series to represent different oscillations of $\mu(IR_{ij}^s)$ and $\xi(IR_{ij}^s)$ in (1) by a linear combination of sinusoidal functions with different frequencies, magnitudes and phases [23]:

$$\mu\left(IR_{ij}^s\right) = \mu_0^s + \sum_{p=1}^{P} A_{\mu p}^s \sin 2\pi \frac{k_p}{60} t_{ij} + \sum_{p=1}^{P} B_{\mu p}^s \cos 2\pi \frac{k_p}{60} t_{ij}, k_p = 1:P,$$

$$\qquad (2)$$

$$\xi\left(IR_{ij}^s\right) = \xi_0^s + \sum_{q=1}^{Q} A_{\xi q}^s \sin 2\pi \frac{k_q}{12} t_{ij} + \sum_{q=1}^{Q} B_{\xi q}^s \cos 2\pi \frac{k_q}{12} t_{ij}, k_q = 1:Q.$$

$$\qquad (3)$$

Here P and Q are the number of the sine and cosine functions representing the trend and seasonal components, respectively, μ_0^s and ξ_0^s are the intercepts defining the baselines of the trend and seasonal components, $A_{\mu p}^s$, $B_{\mu p}^s$, $A_{\xi q}^s$ and $B_{\xi q}^s$ are the Fourier coefficients and t_{ij} indexes time. The amplitude of the trend equals: $C_{\mu p}^s = sign(B_{\mu p}^s)(A_{\mu p}^{s\,2} + B_{\mu p}^{s\,2})^{1/2}$; that of the seasonal component equals: $C_{\xi q}^s = sign(B_{\xi p}^s)(A_{\xi q}^{s\,2} + B_{\xi q}^{s\,2})^{1/2}$, where `sign` is the function extracting the sign of a real number [24].

Iterative weighted least square estimation was used to fit (2) and (3) by minimizing the generalized cross-

validation value: $GCV = RSS + \lambda \times PEN_L$, where RSS is the weighted residual sum of squares, λ is the smoothing parameter and PEN_L is a penalty term as a measure of Fourier series roughness [25].

Identifying driving factors of seasonality

To contrast the effects of preschooling period as a proxy measure of the social contacts and the weather season on the seasonality of the disease, we further stratified the seasonal components $\xi(IR_{ij}^s)$ into three temporal strata: T_1 is the dry and preschooling season from January to May, T_2 is the dry and summer holiday season from June to July and T_3 is the rainy and preschooling season from August to December. Hence, for all j and s, $\xi(IR_{ij}^s)$ belong to T_1 for $i =$ 1:5, to T_2 for $i = 6{:}7$ and to T_3 for $i = 8{:}12$, respectively.

To relate $\xi(IR_{ij}^s)$ of T_1, T_2 and T_3 to the potential risk factors, a spatial-temporal regression model was deployed:

$$\xi_T\left(IR_{ij}^s\right) = \beta_{T0}^s + \sum_{n=1}^{N}\beta_{Tn}^s X_{Tnij}^s + \beta_{T(N+1)}^s P_{Ti} + \varepsilon_{Tij}^s, T = T_1 : T_3,$$
(4)

where N is the total number of the weather covariates, β_{T0}^s and β_{Tn}^s are the regression parameters, X_{Tnij}^s are the weather covariates, ε_{Tij}^s are the spatial-temporal autoregressive regression residuals, P_{Ti} are the number of months at preschools, cumulated from August current year to May next year. Notice that P_{Ti} did not vary with s and j as they were obtained by: $P_{Ti} = \begin{cases} i + 5 \text{ for } i \in T_1 \\ 0 \text{ for } i \in T_2 \\ i-7 \text{ for } i \in T_3 \end{cases}$.

To account for spatial-temporal autocorrelation remaining in the residual, we applied the Spatial-Temporal AutoRegressive model (STAR) [26] for the error term ε_{Tij}^s in (4):

$$\varepsilon_{Tij}^s = \gamma_{T1}^s \varepsilon_{T(i+1)j}^s + \gamma_{T2}^s \varepsilon_{T(i-1)j}^s + \sum_{m=1}^{M}\rho_{Tm}^s \varepsilon_{Tijm}^s$$
$$+ \omega_{Tij}^s, \text{ with } \omega_{Tij}^s \sim N\left(0, \left(\sigma_{T\omega}^s\right)^2\right)$$
(5)

where M is the total number of the first-order spatial neighbors of S_s, γ_{T1}^s and γ_{T2}^s are the temporal autoregressive coefficients, ρ_{Tm}^s are the spatial autoregressive

coefficients, $\left(\sigma_{T\omega}^s\right)^2$ is the variance of the i.i.d. random error ω_{Tij}^s. The model parameters were estimated by maximum likelihood (ML) [27]. Schwarz's Bayesian criterion (BIC) was used to indicate the fitness of the models. The Shapiro–Wilk test [28] was used for normality test. The analyses and modellings were executed in R. 3.4.4, using mainly `stat4` library [29] and `fda` library [30].

Results
Seasonality at multiple district level

The seasonality of IR_{ij}^s was evidenced by the variations of the corresponding SI_i^s (Table 1). For all spatial strata, the maximum $SI_i^s > 100\%$ fell into two periods, between April and May and between September and October. These indicate that, for example for S_2 in April, the median incidence proportion was 1.66 times higher than the average incidence proportion of the year. The minimum $SI_i^s < 60\%$ were in the period between December and January. These show that, for example for S_2 in January, the median incidence proportion was 2.5 times lower than the average incidence proportion of the year.

As outcomes from the iterative fitting processes, the optimally fitted trends of S_1 and S_2 had two sine and two cosine functions ($P = 2$) and the optimal $\log\lambda = 4$ corresponding to the minimum values of the GCV equal to 2.28 and 3.33, respectively. For S_3, the fitted model included four of each of the trigonometric functions ($P = 4$) with $\log\lambda = 2$ and the minimum GCV equal to 6.1. The fitted models for the seasonality (3) have five of each sine and cosine functions ($Q = 5$) with $\log\lambda = 0$ and the minimum GCV equal to 0.87, 2.6 and 4.1 for the S_1, S_2 and S_3, respectively. The trends show a gradual decrease from 2012 to 2015, whereas from the end of 2015 onwards, they increase (Fig. 2). The magnitude is highest in S_2, being approximately double to those of S_1 and S_3.

Fig 3 presents the seasonal patterns of IR_{ij}^s. In T_1, all the peaks fell within April as indicated by the maximum SI_i^s in Table 1. In T_3, the peaks occurred in August (S_3), September (S_2) and October (S_1). All the deepest troughs were in January, whereas the small troughs occurred in June. The maximum amplitude of seasonality was equal to 4.56 infected U5s per 1,000 U5s at risk per month.

Table 1 Seasonal indices of IR_{ij}^s for 3×3 spatial-temporal strata in the mainland of Da Nang city

| Strata		T_1					T_2		T_3				
Month		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
SI % (Std. Error)	S_1	44.6 (19.5)	52.5 (15.4)	123.0 (41.9)	138.6 (38.1)	106.9 (52.8)	83.0 (27.4)	108.1 (23.0)	105.0 (17.3)	136.9 (31.7)	158.1 (41.7)	88.1 (36.2)	55.1 (15.5)
	S_2	41.0 (25.0)	72.9 (26.4)	94.6 (41.9)	165.7 (43.3)	107.0 (45.5)	73.5 (27.4)	121.7 (21.4)	105.0 (10.4)	138.1 (36.2)	136.3 (27.4)	89.7 (38.2)	54.7 (19.56)
	S_3	49.4 (25.7)	67.1 (36.5)	122.1 (32.0)	142.5 (54.5)	137.8 (43.7)	86.5 (28.4)	93.8 (26.7)	124.4 (18.1)	131.6 (38.6)	121.8 (29.4)	70.7 (16.9)	52.3 (21.0)
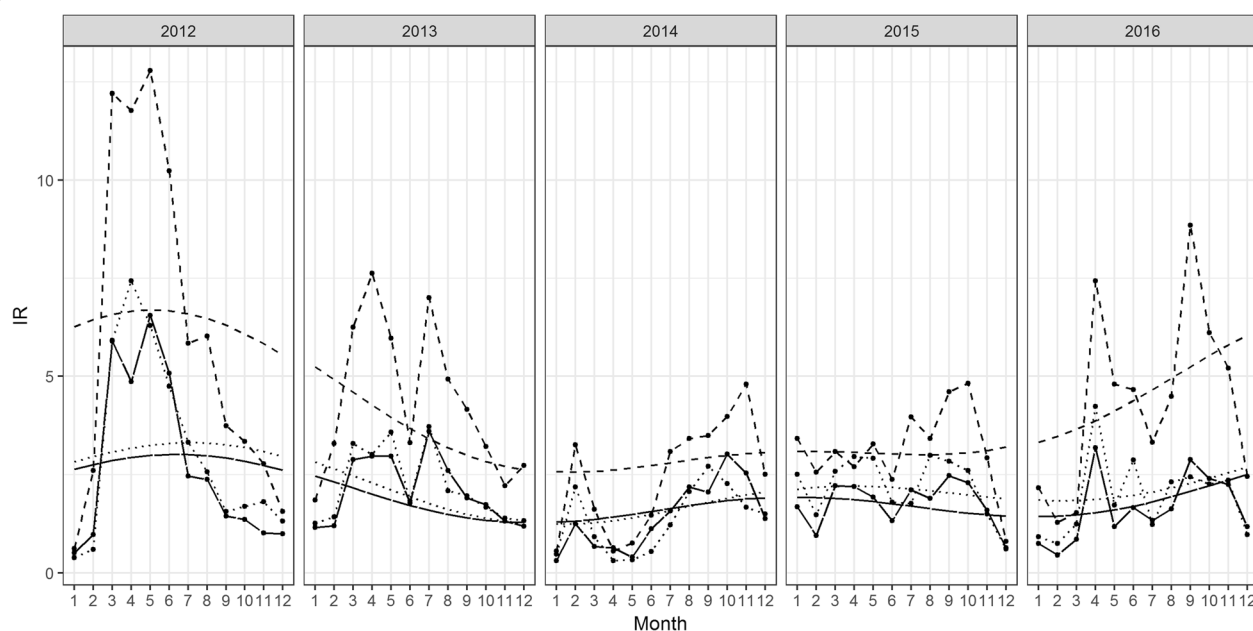
Fig. 2 Time plot of IR_{ij}^s and the temporal trends for S_1 (long-dashed line), S_2 (dashed line) and S_3 (dotted line). IR_{ij}^s in the dry season had larger magnitude than in the rainy season

Driving factors of seasonality

The monthly relative humidity and sunny hours in Da Nang city in the study period were significantly correlated at lag zero with the average temperature (r_{lag0} = − 0.82 and 0.84, respectively). Meanwhile, $\xi(IR_{ij}^s)$ from all three

spatial strata had significantly lower correlation with the rainfall than with the average temperature at lag zero (r_{lag0} = − 0.13, − 0.10, − 0.23 versus 0.74, 0.74, 0.63 for S_1, S_2, S_3, respectively). The average temperature was apparently the sole explanatory weather variable included or
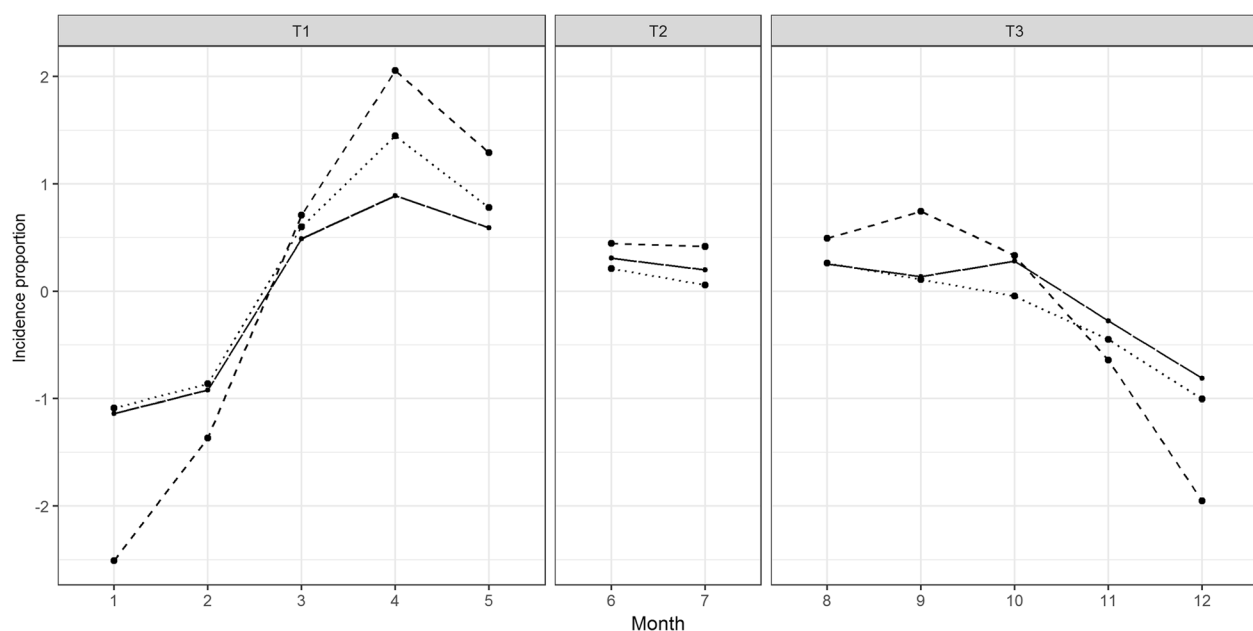

Fig. 3 Seasonality of IR_{ij}^s from the mainland of Da Nang city estimated from the 5 year time series for S_1 (long-dashed line), S_2 (dashed line) and S_3 (dotted line). The highest peaks occurred in the middle of the dry season (April). The other peaks occurred after the children went to preschools

$N = 1$ in (4). The maximum average temperature of over 30 °C occurred in May and June; the minimum of approximately 22 °C occurred in January every year (Fig. 4). The $\xi_T(IR_{ij}^s)$ all satisfied the Shapiro-Wilk normality test (*p*-value < 0.05), thus satisfying the specifications of (4) and (5) without any need for transformation.

The results of the *q*-statistic show that the spatial-temporal stratification into nine strata reflecting the influences of the density of U5s and the monthly average temperature yielded a significant homogeneity within the strata. The *q*-statistic of the nine spatial-temporal strata was equal to 0.21 (*p*-value = 0.00), while that of the combinations of the months versus the districts was not significant, *q*-statistic = 0.26 (p-value = 0.24).

To compare the effects of the average temperature and the preschooling period (Fig 4), the parameters in (4) and (5) were consequently estimated for two sub-models: 1) with preschooling period or $\beta_{T(N+1)}^s = \beta_{T2}^s$ and 2) without preschooling period or $\beta_{T(N+1)}^s = 0$. Accordingly, every spatial-temporal stratum had two sets of estimated parameters, except for those in T_2 because children did not attend preschools in this period. The estimated results are presented in Tables 2 and 3. The effects of the spatial-temporal autocorrelations were taken into account in these estimations, resulting in estimated parameters as presented in Tables 4 and 5 in Appendix. The corresponding BIC values of the selected models were all minimized.

T_2 is the transitional period between the dry and rainy season and also the period of summer holidays. Thus, the interest of studying the driving factor of the seasonality laid in T_1 and T_3. Table 2 shows that in the sub-model one, the average temperature was negatively correlated with $\xi(IR_{ij}^s)$ in T_1. This negative β_{T1}^s indicates a mismatch between the peak of the seasonality of the disease incidences and the average temperature. In sub-model two as can be seen in

Table 3, the β_{T2}^s had negative values for all strata in T_3; whereas, the β_{T1}^s were minimal. The average temperature therefore had no significant effect in T_3.

Discussion

In this section, we discuss our main findings, the consistency of the seasonal patterns extracted from IR_{ij}^s using our methods compared to other methods and we also highlight the limitations of this study.

In this study, we have decomposed the monthly time series of the adjusted incidence proportions of childhood HFMD infections for the spatial-temporal strata of the mainland of Da Nang city to reveal the trend and the seasonality. The yearly larger outbreaks simultaneously happened in all three spatial strata in April. The variation of the seasonality of the second smaller outbreaks in T_3 is intriguing. This variation suggests the effects of the season of increasing social contacts among U5s when they go to preschool on the onset of HFMD infection.

We have shown statistical evidences of these effects during the preschooling period. The positive values of the spatial-temporal regression coefficients imply the simultaneity of the peaks and vice versa [21]. The mismatch between the maximum incidence proportions and the maximum average temperature indicates that the increase of the average temperature was the driving factor, not its maximum values. In the last two quarters of all years, the average temperature was not the predominant risk factor. The contrary was evident in T_1. At small areas of multiple districts in the study area, the increasing social contact has been shown to be the important driving factor of the geographical variation of the seasonality of the childhood HFMD outbreaks.

The estimated spatial-temporal regression parameters of (4) in T_3 are interestingly more informative than those in T_1 because all the smaller peaks occurred in T_3. As the
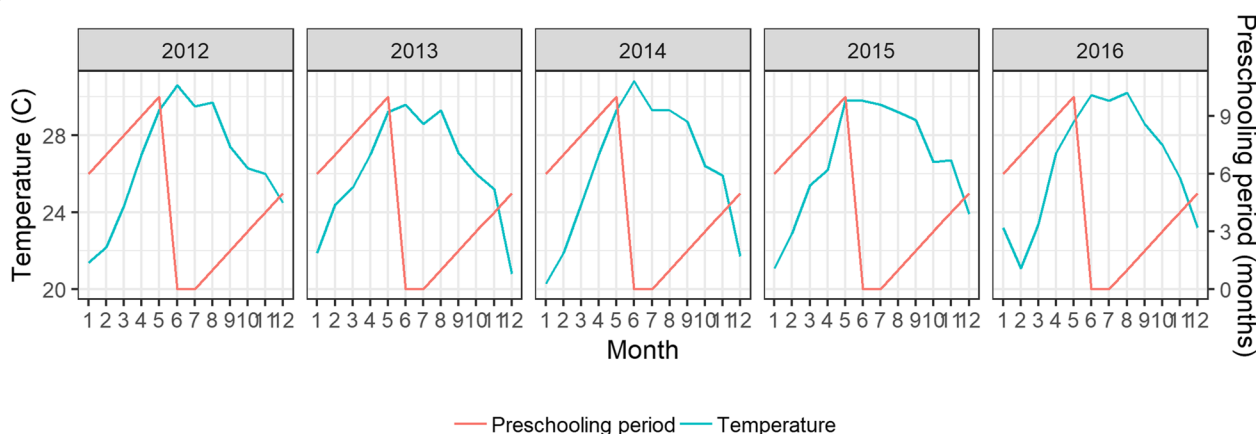

Fig. 4 Temporal variation of the driving factors of the seasonality of HFMD incidences: the monthly average ambient temperature (cyan line) and the cumulative preschooling period (red line)

Table 2 ML estimates of the spatial-temporal regression coefficients between the $\xi(IR_{ij}^s)$ and the average temperature, *without* the cumulative preschooling period included

Strata	T_1			T_2			T_3		
Parameter (Std. Error)	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3
β_{T0}^s	5.44 (2.08)	7.12 (2.45)	4.10 (1.47)	3.25 (0.73)	8.83 (0.00)	6.55 (0.00)	1.57 (0.65)	−2.81 (0.08)	0.69 (0.35)
β_{T1}^s (Temperature)	−0.18 (0.07)	− 0.25 (0.09)	− 0.14 (0.05)	− 0.10 (0.02)	− 0.00 (0.00)	−0.74 (0.00)	− 0.12 (0.03)	0.02 (0.00)	− 0.08 (0.01)
BIC	−23.28	−9.63	−45.38	−15.73	− 95.88	−95.23	−41.90	−69.34	− 66.72

period at the preschools in T_3 cumulatively increased from August to December, the negative values of β_{T2}^s indicate the occurrence of the smaller peaks at the beginning of the preschooling period. Recall from the study of Horby et al. [17] that the physical contacts of U5s are most intense within this age group in Viet Nam. Given that the physical contact is the prominent means of passing HFMD viruses, the nature, frequency and duration of the contacts of those children during the preschooling period explain the occurrence and the geographical dynamics of these smaller outbreaks.

The trends show the occurrences of the large outbreaks at the beginning and at the end of the observed period. Study of a longer period of time would show a long-term oscillation of the large outbreaks with the frequency of every two to 3 years in this city. The possible reasons can be the required lapse of time to cumulate the critical susceptible population [31] or the existence of multiple viral strains in the study area [32, 33].

The decomposed trends and seasonal components were highly correlated with those fitted by the method of seasonal and trend decomposition using Loess [34]. The correlation coefficients of the trends from both methods at lag zero varied from 0.85 to 0.90. Those of the seasonality varied from 0.83 to 0.97. The similar results from both decomposition methods imply the consistency of the seasonality derived from the Fourier decomposition method. This method provides a semi-parametric approach to extract different components of a time series including the long-term and short-term irregular and regular fluctuations, taking into account effects of possible confounders [35, 36]. Using this

approach instead of the common seasonal ARIMA, the drawbacks of stationarization by differencing and statistical transformation can be avoided [37]. Moreover, Fourier decomposition in combination with the spatial-temporal autoregressive model allow both spatial and temporal autocorrelation existing in the data to be included into the model calibration. This reduces the biases in the estimation of the associations between the incidence proportions and the risk factors. Notwithstanding, from a statistical point of view, the small number of spatial strata in the study area places one of the limitations of the study to include the effects of the spatial auto-correlation. In addition, the separation the temporal trend from the time series was based mainly upon expert judgments. In other words, the efficiency of the Fourier decomposition in many cases relies on the expert's familiarity with the understudied phenomena.

Conclusions

HFMD has been becoming one of the most important pediatric NTDs in the Northeastern and Southeastern Asian countries. By applying spatial-temporal statistical analyses, the results have shown that at multiple district level, the social contact is the important driving factor of the spatial variation of the disease's outbreaks. This study provides statistical evidences of the effects of the seasonality of increasing social contact amongst U5s on geographical dynamics of HFMD infection outbreaks. Our findings contribute to extend the understanding of the underlying driving factors of the disease dynamics at small areas. This contribution is necessary to inform the next insightful research into the spatial-temporal dynamics of the transmission of the NTDs within the local population.

Table 3 ML estimates of the spatial-temporal regression coefficients between the $\xi(IR_{ij}^s)$ and the average temperature *including* the cumulative preschooling period

Strata	T_1			T_3		
Parameters (Std. Error)	S_1	S_2	S_3	S_1	S_2	S_3
β_{T0}^s	−5.16 (0.61)	8.00 (3.73)	−5.57 (0.55)	0.8 (0.03)	0.64 (0.03)	0.65 (0.09)
β_{T1}^s (Average temperature)	0.08 (0.03)	0.03 (0.04)	0.07 (0.02)	−0.00 (0.00)	−0.00 (0.00)	− 0.01 (0.00)
β_{T2}^s (Preschooling period)	0.32 (0.05)	−1.25 (0.53)	0.39 (0.05)	−0.40 (0.01)	−0.29 (0.00)	− 0.25 (0.01)
BIC	−50.00	− 68.79	−57.86	− 189.86	− 98.63	− 166.29

Appendix

Table 4 ML estimates of the spatial and temporal auto-regression coefficients of the sub-model one

Strata	T_1			T_2			T_3		
Parameters (Std. Error)	S_1	S_2	S_3	S_1	S_2	S_3	S_1	S_2	S_3
γ_{T1}^s	0.04 (0.06)	0.08 (0.05)	0.23 (0.04)	−0.37 0.70	−0.17 (0.00)	0.00 (0.00)	0.07 (0.06)	2.26 (0.06)	0.22 (0.05)
γ_{T2}^s	0.42 (0.14)	0.38 (0.05)	−0.09 (0.02)	0.26 0.28	−0.19 (0.00)	1.50 (0.02)	−0.55 (0.12)	0.87 (0.02)	0.38 (0.03)
ρ_{T1}^s	−0.61 (0.18)	−0.61 (0.21)	1.03 (0.04)	1.70 (0.00)	1.05 (0.00)	−0.51 (0.02)	0.16 (0.05)	−1.26 (0.11)	0.35 (0.07)
ρ_{T2}^s	1.54 (0.21)	1.62 (0.18)	0.09 (0.02)	0.18 (0.00)	0.27 (0.00)	0.00 (0.00)	1.28 (0.18)	−1.05 (0.10)	0.06 (0.02)
$\sigma_{T\omega}^s$	0.09 (0.01)	0.12 (0.02)	0.04 (0.01)	0.04 (0.00)	0.00 (0.00)	0.00 (0.00)	0.06 (0.01)	0.04 (0.01)	0.04 (0.01)

Table 5 ML estimates of the spatial and temporal auto-regression coefficients of the sub-model two

Strata	T_1			T_3		
Parameter (Std. Error)	S_1	S_2	S_3	S_1	S_2	S_3
γ_{T1}^s	−0.25 (0.04)	0.15 (0.03)	1.61 (0.06)	−0.51 (0.01)	−1.68 (0.04)	−0.06 (0.01)
γ_{T2}^s	0.07 (0.02)	0.04 (0.01)	8.12 (0.02)	−0.04 (0.02)	0.04 (0.01)	−0.21 (0.01)
ρ_{T1}^s	−0.07 (0.03)	0.21 (0.09)	5.82 (0.01)	−0.24 (0.02)	−0.71 (0.07)	0.23 (0.01)
ρ_{T2}^s	0.96 (0.04)	0.84 (0.09)	6.62 (0.00)	2.79 (0.06)	8.84 (0.01)	0.21 (0.00)
$\sigma_{T\omega}^s$	0.05 (0.01)	0.035 (0.01)	0.06 (0.01)	0.00 (0.00)	0.02 (0.00)	0.01 (0.00)

Author details
[1]Faculty of Geo-Information Science and Earth Observation, University of Twente, Enschede, The Netherlands. [2]Pasteur Institute, Ho Chi Minh City, Viet Nam.

References
1. Sultan B, Labadi K, Guégan JF, Janicot S. Climate drives the Meningitis epidemics onset in West Africa. PLoS Med. 2005;2:e6. https://doi.org/10.1371/journal.pmed.0020006.
2. Do TTT, Martens P, Luu NH, Wright P, Choisy M. Climatic-driven seasonality of emerging dengue fever in Hanoi, Vietnam. BMC Public Health. 2014;14:1078. https://doi.org/10.1186/1471-2458-14-1078.
3. Kim BI, Ki H, Park S, Cho E, Chun BC. Effect of climatic factors on hand, foot, and mouth disease in South Korea, 2010–2013. PLoS One. 2016;11:e0157500. https://doi.org/10.1371/journal.pone.0157500.
4. Pascual M, Dobson A. Seasonal patterns of infectious diseases. PLoS Med. 2005;2:e5. https://doi.org/10.1371/journal.pmed.0020005.
5. Altizer S, Dobson A, Hosseini P, Hudson P, Pascual M, Rohani P. Seasonality and the dynamics of infectious diseases. Ecol Lett. 2006;9:467–84.
6. Grassly NC, Fraser C. Seasonal infectious disease epidemiology. Proc R Soc B Biol Sci. 2006;273:2541–50.
7. United Nations. SDGs-Goal 3: Ensure healthy lives and promote well-being for all at all ages. http://www.un.org/sustainabledevelopment/health. Accessed 18 Jan 2018.
8. Hotez PJ, Bottazzi ME, Strych U, Chang LY, Lim YAL, Goodenow MM, AbuBakar S. Neglected tropical diseases among the association of Southeast Asian Nations (ASEAN): Overview and update. PLoS Negl Trop Dis. 2015;9:e0003575. https://doi.org/10.1371/journal.pntd.0003575.
9. Koh WM, Bogich T, Siegel K, Jin J, Chong EY, Tan CY, Chen MIC, Horby P, Cook AR. The epidemiology of hand, foot and mouth disease in Asia: a systematic review and analysis. Pediatr Infect Dis J. 2016;35:e285–300. https://doi.org/10.1097/INF.0000000000001242.
10. Frydenberg A, Starr M. Hand, foot and mouth disease. Aust Fam Physician. 2003;32:594–5.
11. Liu W, Ji H, Shan J, Bao J, Sun Y, Li J, Bao C, Tang F, Yang K, Bergquist R, Peng Z, Zhu Y. Spatiotemporal dynamics of hand-foot-mouth disease and its relationship with meteorological factors in Jiangsu province, China. PLoS One. 2015;10:e0131311. https://doi.org/10.1371/journal.pone.0131311.
12. Wang J, Cao Z, Zeng DD, Wang Q. Assessing local risk factors of Beijing hand-foot- mouth disease in China. Online J Public Health Inform. 2017;9(1):e009. https://doi.org/10.5210/ojphi.v9i1.7759.
13. Xu C, Zhang X, Xiao G. Spatiotemporal decomposition and risk determinants of hand, foot and mouth disease in Henan, China. Sci Total Environ. 2019;657:509-16.
14. NikNadia NMN, Sam IC, Rampal S, WanNorAmalina WMZ, NurAtifah G, Verasahib K, Ong CC, MohdAdib M, Chan YF. Cyclical patterns of hand, foot and mouth disease caused by Enterovirus A71 in Malaysia. PLoS Negl Trop Dis. 2016;10:e0004562. https://doi.org/10.1371/journal.pntd.0004562.
15. Phung D, Nguyen HX, Nguyen HLT, Do CM, Tran QD, Chu C. Spatiotemporal variation of hand-foot-mouth disease in relation to socioecological factors: a multiple-province analysis in Vietnam. Sci Total Environ. 2018;610:983–91.
16. Sofianopoulou E, Pless-Mulloli T, Rushton S, Diggle PJ. Modeling seasonal and spatiotemporal variation: the example of respiratory prescribing. Am J Epidemiol. 2017;186:101–8.
17. Horby P, Thai PQ, Hens N, Yen NTT, Mai LQ, Thoang DD, Linh NM, Huong NT, Alexander N, Edmunds WJ, Duong TN, Fox A, Hien NT. Social contact patterns in Vietnam and implications for the control of infectious diseases. PLoS One. 2011;6:e16965. https://doi.org/10.1371/journal.pone.0016965.
18. Da Nang Statistics Office. Da Nang statistical yearbook 2016. Da Nang: Statistical publishing house; 2017.
19. Wang JF, Zhang TL, Fu BJ. A measure of spatial stratified heterogeneity. Ecol Indic. 2016;67:250–6.
20. Wang JF, Li XH, Christakos G, Liao YL, Zhang T, Gu X, Zheng XY. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun region. China Int J Geogr Inf Sci. 2010;24:107–27.
21. Barnett AG, Dobson AJ. Analysing seasonal health data. Berlin: Springer; 2010.
22. Dodge Y. Seasonal index. In: Dodge Y, editor. The concise encyclopedia of statistics. New York: Springer New York; 2008. p. 476–9.
23. Bloomfield P. Fourier analysis of time series: an introduction. 2nd ed. New York: Wiley; 2000.
24. Naumova EN, MacNeill IB. Seasonality assessment for biosurveillance systems. In: Auget JL, Balakrishnan N, Mesbah M, Molenberghs G, editors. Advances in statistical methods for the health sciences: applications to cancer and aids studies, genome sequence analysis, and survival analysis. Boston: Birkhäuser Boston; 2007. p. 437–50.
25. Ramsay JO, Hooker G, Graves S. Smoothing: computing curves from noisy data. In: Ramsay JO, Hooker G, Graves S, editors. Functional data analysis with R and MATLAB. New York: Springer New York; 2009. p. 59–82.
26. Pace RK, Barry R, Clapp J, Rodriquez M. Spatio-temporal autoregressive models of neighborhood effects. J Real Estate Finance Econ. 1998;17:15–33.
27. Cole SR, Chu H, Greenland S. Maximum likelihood, profile likelihood, and penalized likelihood: a primer. Am J Epidemiol. 2014;179:252–60.
28. Yap BW, Sim CH. Comparisons of various types of normality tests. J Stat Comput Simul. 2011;81:2141–55.
29. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2017. https://www.R-project.org
30. Ramsay JO, Wickham H, Graves S, Hooker G. fda: Functional Data Analysis. R package version 2.4.8. 2018. https://cran.r-project.org/web/packages/fda/index.html.
31. Anderson RM. Populations and infectious diseases: ecology or epidemiology? J Anim Ecol. 1991;60:1–50.
32. Sarma N. Relapse of hand foot and mouth disease: are we at more risk? Indian J Dermatol. 2013;58:78–9.
33. Thao NTT, Donato C, Trang VTH, Kien NT, Trang PMT, Khanh TQ, Nguyet DT, Sessions OM, Cuong HQ, Lan PT, Huong VTQ, van Doorn HR, Vijaykrishna D. Evolution and spatiotemporal dynamics of Enterovirus A71 subgenogroups in Vietnam. J Infect Dis. 2017;216:1371–9.

34. Cleveland RB, Cleveland WS, McRae JE, Terpenning I. STL: a seasonal-trend decomposition procedure based on loess. J Off Stat. 1990;6:3–73.
35. Diggle P. Time series: a biostatistical introduction. Oxford statistical science series. Oxford: Oxford University Press; 1990.
36. Stolwijk AM, Straatman H, Zielhuis GA. Studying seasonality by using sine and cosine functions in regression analysis. J Epidemiol Community Health. 1999;53:235–8.
37. Makridakis S, Hibon M. ARMA models and the box–Jenkins methodology. J Forecast. 1997;16:147–63.

Publisher's Note