Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Theoretical Biology 554 (2022) 111279

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Spatiotemporal pattern recognition and dynamical analysis of COVID-19 in Shanghai, China

Haonan Zhong, Kaifa Wang, Wendi Wang *

School of Mathematics and Statistics, Southwest University, Chongqing 400715, PR China

ARTICLE INFO

Keywords:
Spatial statistics
Autocorrelation analysis
Dynamic model
Daily reproduction number

ABSTRACT

Shanghai suffered a large outbreak of Omicron mutant of COVID-19 at the beginning of March 2022. To figure out the spatiotemporal patterns of the epidemic, a retrospective statistical investigation, coupled with a dynamic model, is implemented in this study. The hotspots of SARS-CoV-2 transmissions are identified, and strong aggregative effects in the decay stage are found. Besides, the visualization of disease diffusion is provided to show how COVID-19 disease invades all districts of Shanghai in the early stage. Furthermore, the calculations from the dynamic model manifest the effect of detections to suppress the epidemic dissemination. These results reveal the strategies to improve the spatial control of disease.

1. Introduction

In early March 2022, the Omicron variant of SARS-CoV-2 struck Shanghai of China, impacting over millions of population and eventually causing the lockdown of city till early June, 2022, which resulted in huge casualties and economic losses to China (Shanghai Municipal Health Commission, 2022). The epidemic caused by Omicron variant has two remarkable characteristics of high infectivity and high concealment, which quickly invaded most administrative districts in Shanghai and generated numerous asymptomatic cases, increasing the difficulty of detection. (The administrative division map is presented in Fig. 1.) The epidemic was eliminated on 31th May, and the public traffic system was allowed to restart on 1st June (The State Council Information Office of the People's Republic of China, 2022). The recognition of spatiotemporal patterns of this pandemic process is helpful for future disease prevention and control.

Spatial statistical analysis and dynamical modeling are powerful to find out the main features of spatial spread of epidemics. For example, the speeding-up pattern and spatiotemporal correlation of geographic spread of the early geographic spread of COVID-19 in China were obtained in Xue et al. (2021) and Gao et al. (2022). The transmission ability of COVID-19 in the regional difference of China was studied by Hu et al. (2020). The influences of direct and indirect infections on the epidemic were considered by Zhong and Wang (2020). An age-structured epidemic model was constructed by Duan et al. (2022) to find that shortening the diagnosis period can result in an enormous selective pressure on the evolution of SARS-CoV-2. A multi-scale model study was given by Li and Xiao (2021) to investigate the interaction of SARS-CoV-2 transmission and information dissemination dynamics during the outbreak of emerging infectious diseases. A probability model was proposed to retrospectively quantify the confidence of giving the end-of-outbreak declaration during the COVID-19 epidemic in Wuhan (Yuan et al., 2022). Recently, the epidemiological characteristics of SARS-CoV-2 infection outbreak in Shanghai in the spring of 2022 were studied by Xian et al. (2022) and Liu et al. (2022), where the infection force of the virus strain during the pandemic was analyzed and the time-varying reproduction number was estimated.

In this study, we collect the whole epidemic data of daily new asymptomatic cases and confirmed cases of various districts in Shanghai from the initial phase to the elimination phase of outbreak. On the basis of the data from each district, we identify the spatiotemporal pattern of the epidemic through two geographic statistics of Moran's I and Getis Ord G statistics, which can judge the spatial distribution pattern of the epidemic in different stages, and then present the diffusion patterns in import and diffusion stage. We also propose a difference model to investigate the daily reproduction number of each district, and implement the comparison between the statistical results and dynamical results. The results are helpful to improve the spatial prevention and control of epidemic diseases.

The organization of this paper is arranged as follows. The materials and methods are given in Section 2, including data collection, spatial geographic statistical methods and dynamic model formulation. The results of statistics and dynamics are given in Section 3. In Section 4, we present conclusions and draw some suggestions.

* Corresponding author.
E-mail address: wendi@swu.edu.cn (W. Wang).

https://doi.org/10.1016/j.jtbi.2022.111279
Received 29 June 2022; Received in revised form 5 September 2022; Accepted 14 September 2022
Available online 20 September 2022

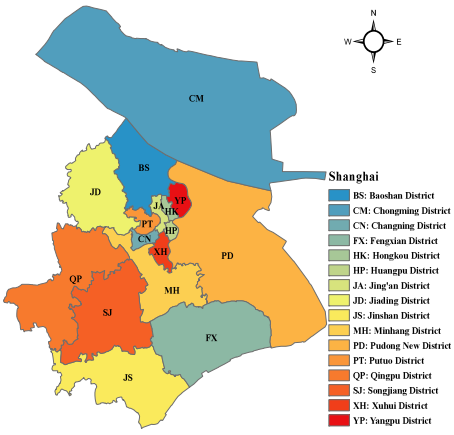

Fig. 1. Administrative division map of Shanghai. There are 16 administrative districts in Shanghai and the central downtown area consists of HP, XH, CN, PT, JA, HK, YP.

2. Materials and methods

2.1. Data collection and stages classification

Shanghai suffered a severe epidemic resulted from Omicron variant of SARS-CoV-2, which began on 1st March, 2022 and gradually faded away in late May. Considering the public traffic system in Shanghai was allowed to restart on 1st June, 2022 (The State Council Information Office of the People's Republic of China, 2022), the end time of epidemic could be set on that day. There are multifold epidemic data released in the official website of Shanghai Municipal Health Commission (Shanghai Municipal Health Commission, 2022). We extract the daily new confirmed cases and the daily new asymptomatic cases of the whole Shanghai from 1st March to 31th May, 2022. Besides, we record the daily epidemic data of each district of Shanghai according to the real time big data report of epidemic situation for novel coronavirus pneumonia (Anon, 2022).

Based on the development process of this epidemic, the outbreak in Shanghai is approximately divided into three stages: import and diffusion stage (IDs), outbreak and epidemic stage (OEs) and decline and elimination stage (DEs). Using Monte-Carlo Sampling Method, we estimate the switch times t_1 between IDs and OEs, and the switch time t_2 between OEs and DEs.

2.2. Spatial geographic statistical methods

Spatial autocorrelation analysis is applied to represent the relationships among adjacent regions with respect to the same attributes. Attributes of similar high or low values result in positive autocorrelation, while opposing high and low attribute values result in negative autocorrelation (Liu et al., 2008). In this study, we adopt two kinds of autocorrelation analysis methods: Moran's I and Getis Ord G Index analysis. ArcGIS Desktop 10.5 software is used to for analysis and visual presentation, including figures and video generation.

2.2.1. Global and local Moran's I

The Moran's I takes values in $[-1, 1]$. When I is close to 1, around 0 or close to -1, it implies the aggregated, random or dispersed distribution in the interest region, respectively. Suppose that the number of spatial locations is n, and X_i is the value of location i. The formula for calculating the global Moran's I is

$$I = \frac{n}{S_0}\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(X_i-\bar{X})(X_j-\bar{X})}{\sum_{i=1}^{n}(X_i-\bar{X})^2},\ (i\neq j),$$

where

$$\bar{X} = \sum_{i=1}^{n} X_i,\ S_0 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}$$

and $W = \{w_{ij}\}$ is the link matrix of distance. If there is no spatial autocorrelation, the expectation and variance of I satisfy $E(I) = -1/(n-1)$ and

$$Var(I) = \frac{\frac{1}{2}(n-1)n^2 s_{w1} - (n-1)n s_{w2} - 2(\sum_{i\neq j} w_{ij})^2}{(n+1)(n-1)^2(\sum_{i\neq j} w_{ij})^2},$$

where $s_{w1} = \sum_{i\neq j}\sum_j (w_{ij}+w_{ji})^2$ and $s_{w2} = \sum_k(\sum_i w_{ik} + \sum_j w_{kj})^2$. For Moran's I, a significance test of normal distribution can be performed through the following equation

$$Z = \frac{I - E(I)}{\sqrt{Var(I)}} \sim N(0,1).$$

In order to find out whether there is a spatial autocorrelation in the local space, i.e. hotspot, we use the local Moran's I, which was proposed by Anselin (1995). It is a local indicator of spatial association (LISA) to manifest the state of each region, and the computation formula is given by

$$I_i = \frac{Z_i}{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sum_{i\neq j}^{n} w_{ij}Z_j,$$

where $Z_i = X_i - \bar{X}, Z_j = X_j - \bar{X}$. A significance test of normal distribution of the local Moran's I can be performed by

$$Z_i = \frac{I_i - E(I_i)}{\sqrt{Var(I_i)}} \sim N(0, Var(I_i)),$$

where $E(I_i) = -\sum_j^n w_{ij}/(n-1)$, $s_y = \sum_j^n (y_j-\bar{y})^2/[\sum_j^n (y_j-\bar{y})^2]^2$ and

$$Var(I_i) = \frac{(n-s_y)}{n-1}\sum_{j\neq i}^{n} w_{ij}^2 + \frac{(2s_y-n)}{(n-1)(n-2)}\sum_{k\neq i}^{n}\sum_{h\neq i}^{n} w_{ik}w_{ih} - [E(I_i)]^2.$$

2.2.2. Local Getis Ord G index

Local Getis Ord G index was proposed by Geits and Ord (1992) and Ord and A. Geits (1995). It is used to find the hotspots of disease incidence in an interesting area and the computation formula is

$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij}X_j}{\sum_{j=1}^{n} X_j}.$$

Similarly, a significance test of G_i^* can be performed by

$$Z_{G_i^*} = \frac{G_i^* - E(G_i^*)}{\sqrt{Var(G_i^*)}} \sim N(0, Var(G_i^*)),$$

where

$$E(G_i^*) = \frac{1}{n}\sum_{j=1}^{n} w_{ij}, Var(G_i^*) = \frac{\sum_{j=1}^{n} w_{ij}(n-\sum_{j=1}^{n} w_{ij})Var(X)^2}{n^2(n-1)\bar{X}^2}.$$

2.3. Difference equation model

Suppose that $A_i(t), C_i(t),\ i = 1, 2, \ldots, 16$, are the increments of reported asymptomatic cases and confirmed cases of ith district on the tth day. In light of the regulations on epidemic prevention (Shanghai Municipal Health Commission, 2022), the reported cases were immediately quarantined and acquired treatment in medical institutions. We suppose that the daily recovery rates of asymptomatic and confirmed cases r^A, r^C are invariant over time for simplifying simulation. Therefore, we get the total numbers of reported asymptomatic and confirmed cases of ith district on the tth day:

$$\tilde{A}_i(t) = \sum_{k=0}^{t} e^{-r^A(t-k)}A_i(k),\quad \tilde{C}_i(t) = \sum_{k=0}^{t} e^{-r^C(t-k)}C_i(t). \tag{1}$$

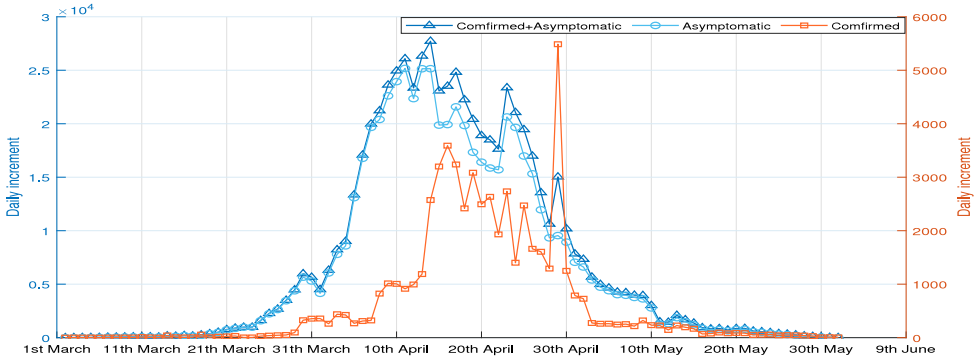

Fig. 2. Illustration of the daily new increment of confirmed cases, asymptomatic cases and their summation in Shanghai. The asymptomatic cases accounts for the majority of positive cases and the confirmed cases accounts for a few. The data begins on 1st March and ends on 31th May.

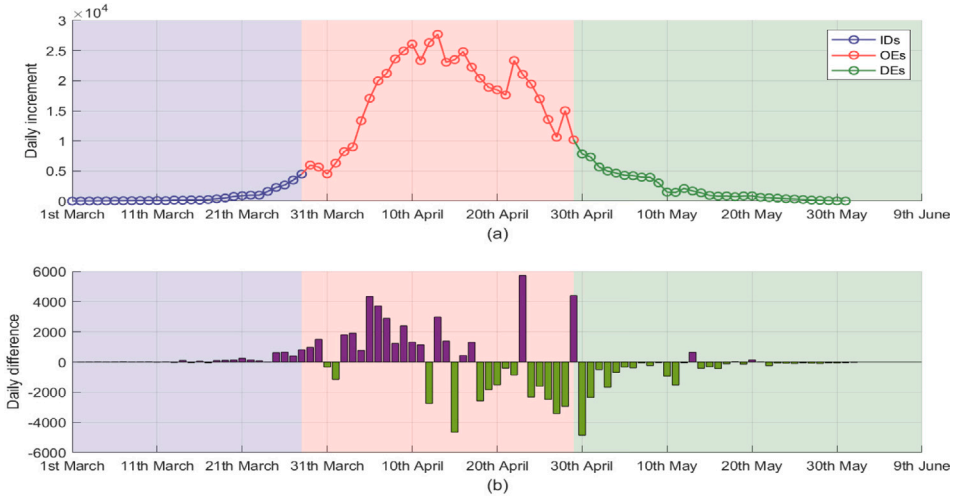

Fig. 3. Illustrations of stage classification for the whole outbreak in Shanghai. Panel (a) presents the three stages including import and diffusion stage (IDs), outbreak and epidemic stage (OEs) and decay and elimination stage (DEs) of total positive cases in Shanghai. The switch times are 28th March and 29th April. Panel (b) presents the difference values of daily new increment of total cases in Shanghai.

Suppose that $a_i(t)$ and $b_i(t)$ are the proportions of unreported asymptomatic and confirmed cases with respect to $\tilde{A}_i(t), \tilde{C}_i(t)$, and β is the infection probability of each contact with an asymptomatic or symptomatic individual. Since all the reported cases were immediately isolated and did not cause any infection, we get the infection force of ith district on the tth day:

$$f_i(t) = \beta \frac{\tilde{S}_i(t)}{\tilde{N}_i} \left(a_i(t)\tilde{A}_i(t) + b_i(t)\tilde{C}_i(t) \right),$$

where $\tilde{S}_i(t)$ and $\tilde{N}_i$ are the susceptible population of ith district on the tth day and the total population of ith district. Since $\tilde{S}_i(t)$ is very close to $\tilde{N}_i$ in the outbreak in Shanghai, that is, $\tilde{S}_i(t)/\tilde{N}_i \approx 1$, the infection force is further simplified as

$$f_i(t) = \beta \left(a_i(t)\tilde{A}_i(t) + b_i(t)\tilde{C}_i(t) \right) = \beta_i^A(t)\tilde{A}_i(t) + \beta_i^C(t)\tilde{C}_i(t),$$

where $\beta_i^A(t)$, $\beta_i^C(t)$ are the effective transmission rates of asymptomatic and symptomatic individuals, respectively.

Assume that Λ^A, Λ^C are the transition ratio matrices of asymptomatic cases and confirmed cases with the elements λ_{ij}^A, λ_{ij}^C, $(i \neq j)$ representing the immigration ratios of asymptomatic cases and confirmed cases from j district to i district and λ_{jj}^A, λ_{jj}^C representing the emigration ratios of j districts satisfying

$$\sum_{i \neq j}^{16} \lambda_{ij}^A = -\lambda_{jj}^A, \quad \sum_{i \neq j}^{16} \lambda_{ij}^C = -\lambda_{jj}^C.$$

Suppose p, q, $(p + q = 1)$ are the proportions of positive cases flowing into the asymptomatic cases and confirmed cases respectively. Then we get following difference equation model for asymptomatic and confirmed cases:

$$\begin{aligned} A(t+1) &= pIF + e^{-r^A}\Lambda^A A(t), \\ C(t+1) &= qIF + e^{-r^c}\Lambda^C C(t), \end{aligned} \tag{2}$$

where I is a 16×16 identity matrix and

$$F = [f_1(t), f_2(t), \ldots, f_{16}(t)]^T,$$

$$A(t) = [A_1(t), A_2(t), \ldots, A_{16}(t)]^T,$$

$$C(t) = [C_1(t), C_2(t), \ldots, C_{16}(t)]^T.$$

2.4. Daily reproduction number

Based on model (2), we can further estimate the daily reproduction number $\mathcal{R}_i(t)$ on the tth day of the ith district in Shanghai. By definition, the daily reproduction number is the infected numbers in susceptible population by one positive case during its whole infectious period. At time t, if one asymptomatic case or one confirmed case is invaded in the ith district in Shanghai, i.e.,

$$(A(t), C(t)) = (e_i, 0) \quad \text{or} \quad (A(t), C(t)) = (0, e_i). \tag{3}$$

Here e_i denotes a 16×1 vector with the ith element being 1 and others being 0. In order to obtain the number of new infections caused by this invader, taking (3) as initial conditions of model (2) respectively,

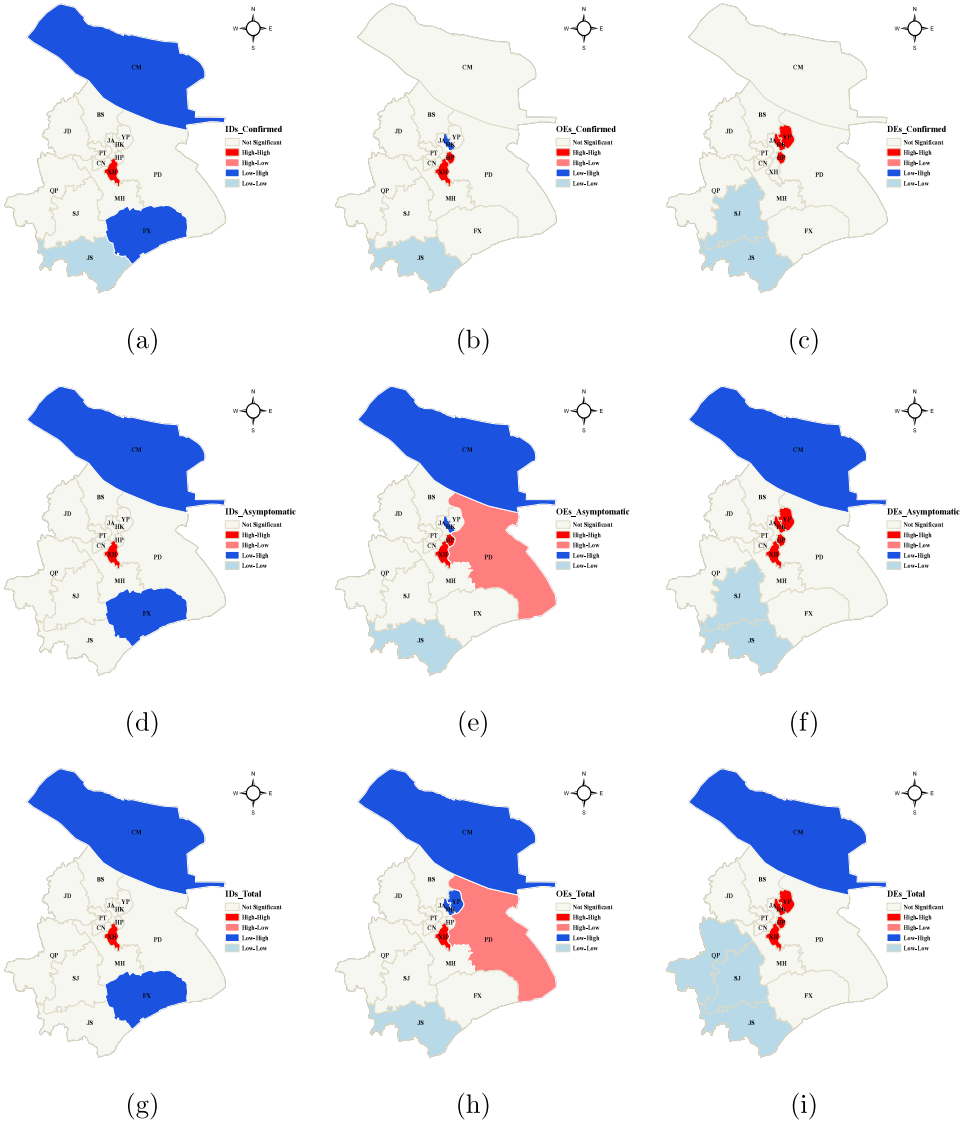

Fig. 4. Local indicator spatial association (LISA) clustering maps for three stages of the outbreaks in Shanghai.

we can get the number of infections in each district after that. Since the infectious durations of asymptomatic and confirmed cases are $1/r^A$ and $1/r^C$ respectively, by taking the sum of abovementioned infections in all 16 districts during the period, we get the daily reproduction numbers $\mathcal{R}_i^A(t)$ or $\mathcal{R}_i^C(t)$, resulted from one asymptomatic case or one confirmed case,

$$\begin{aligned}\mathcal{R}_i^A(t) &= \sum_{j=1}^{16} \tilde{A}_j(t+[1/r^A]) + \sum_{j=1}^{16} \tilde{C}_j(t+[1/r^A]),\ (A(t),C(t)) = (e_i,0),\\ \mathcal{R}_i^C(t) &= \sum_{j=1}^{16} \tilde{A}_j(t+[1/r^C]) + \sum_{j=1}^{16} \tilde{C}_j(t+[1/r^C]),\ (A(t),C(t)) = (0,e_i),\end{aligned} \tag{4}$$

where $[\cdot]$ is the rounding operator, $\tilde{A}_j(k)$ and $\tilde{C}_j(k)$, which are defined in (1), denote the total numbers of asymptomatic and confirmed cases of jth district on the kth day respectively.

3. Results

3.1. Stage classification

The stage classification is based on the total daily positive cases in Shanghai. We draw the line chart of the daily new increment of confirmed cases, asymptomatic cases and total cases in Fig. 2.

Based on the values of the total cases in Fig. 2 and the Monte-Carlo Sampling Method, we obtain the estimated switch times:

$$t_1 = 28 \pm 5,\ t_2 = 61 \pm 1.$$

For the convenience of analysis, we set $t_1 = 28$ and $t_2 = 60$, which indicate that the corresponding switch times are 28th March and 29th April. The line chart for the differences of daily new increment of total cases is shown in Fig. 3.

3.2. Spatiotemporal pattern

We now present the global and local Moran's I and local Getis Ord G Index, which are used to judge the spatial distribution patterns of the epidemic in the aforementioned three stages. By using the software ArcGIS Desktop 10.5, we get the summations of confirmed, asymptomatic and total cases in IDs, OEs and DEs. Based on the summations, we get the global Moran's I, Z-score and P-value related to confirmed, asymptomatic and total positive cases in IDs, OEs and DEs respectively, which are shown in Table 1.

From Table 1, we see that the full global Moran's Is in DEs of confirmed, asymptomatic and total cases pass the significance test with $P < 0.01$ and all Z-scores are positive. It implies there is an

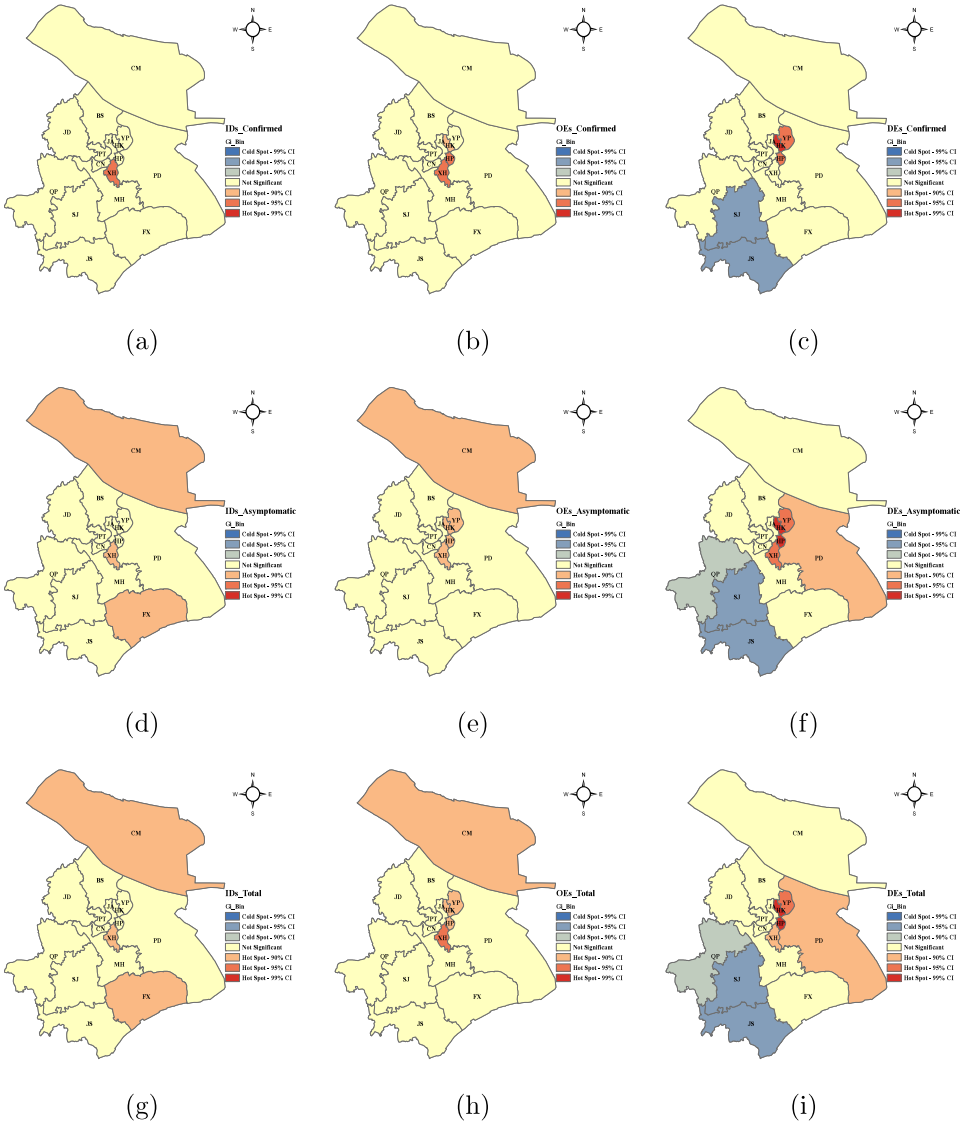

Fig. 5. Getis hotspot maps for three stages of the outbreaks in Shanghai.

Table 1
The results of global Moran's I analysis in three different stages.

Phase	IDs(Confirmed)	OEs(Confirmed)	DEs(Confirmed)
Moran's *I*	−0.073	−0.040	**0.292**
Z-score	−0.051	0.310	**2.694**
P-value	0.959	0.757	**0.007**
Phase	IDs(Asymptomatic)	OEs(Asymptomatic)	DEs(Asymptomatic)
Moran's *I*	−0.064	−0.051	**0.316**
Z-score	0.023	0.215	**2.934**
P-value	0.981	0.829	**0.003**
Phase	IDs(Total)	OEs(Total)	DEs(Total)
Moran's *I*	−0.065	−0.050	**0.319**
Z-score	0.018	0.219	**2.948**
P-value	0.986	0.826	**0.003**

aggregated distribution in various districts of Shanghai in the decay phase. Furthermore, the clustering maps of LISA and local Getis Ord *G* index in Figs. 4 and 5 show that the COVID-19 spread in Shanghai has a strong aggregation in two dimensions of time and space.

More specifically, in Fig. 4, the maps of (a)–(c) show that the confirmed cases are mainly concentrated in the downtown districts over all the three phases. Note that the Omicron variant of SARS-CoV-2 generates a large fraction of asymptomatic cases which account for over 90%. The maps of (d)–(f) and (g)–(i) have similar patterns. These indicate that in the first two phases, the total positive cases are concentrated in downtown districts, Chongming and Fengxian Districts. But in the decay phase, the epidemic mainly spreads in downtown districts and Pudong New District; the surrounding districts control the epidemic more quickly than the downtown districts, such as Qingpu, Songjiang and Jinshan Districts. In Fig. 5, the pattern of each map remains similar to that in Fig. 4. It implies the two statistic outcomes are compatible with each other, and have a good reliability.

3.3. Dissemination pattern in IDs

Based on the 28 days' data in the initial phase, we show the diffusion process of districts with positive cases in Shanghai in Fig. 6. Fig. 6 only presents the diffusion maps of the days when a district is reported with at least one positive case, and the integral diffusion maps are shown in Supplementary video file (ShangHai_InitialDiffusion_Movie.avi). The initial areas of the outbreak in Shanghai are Putuo and Jiading Districts. The general tendency of the diffusion is that the Omicron variant first struck the surrounding districts, then invaded the central downtown area and finally infected Chongming District. The intrusion order of

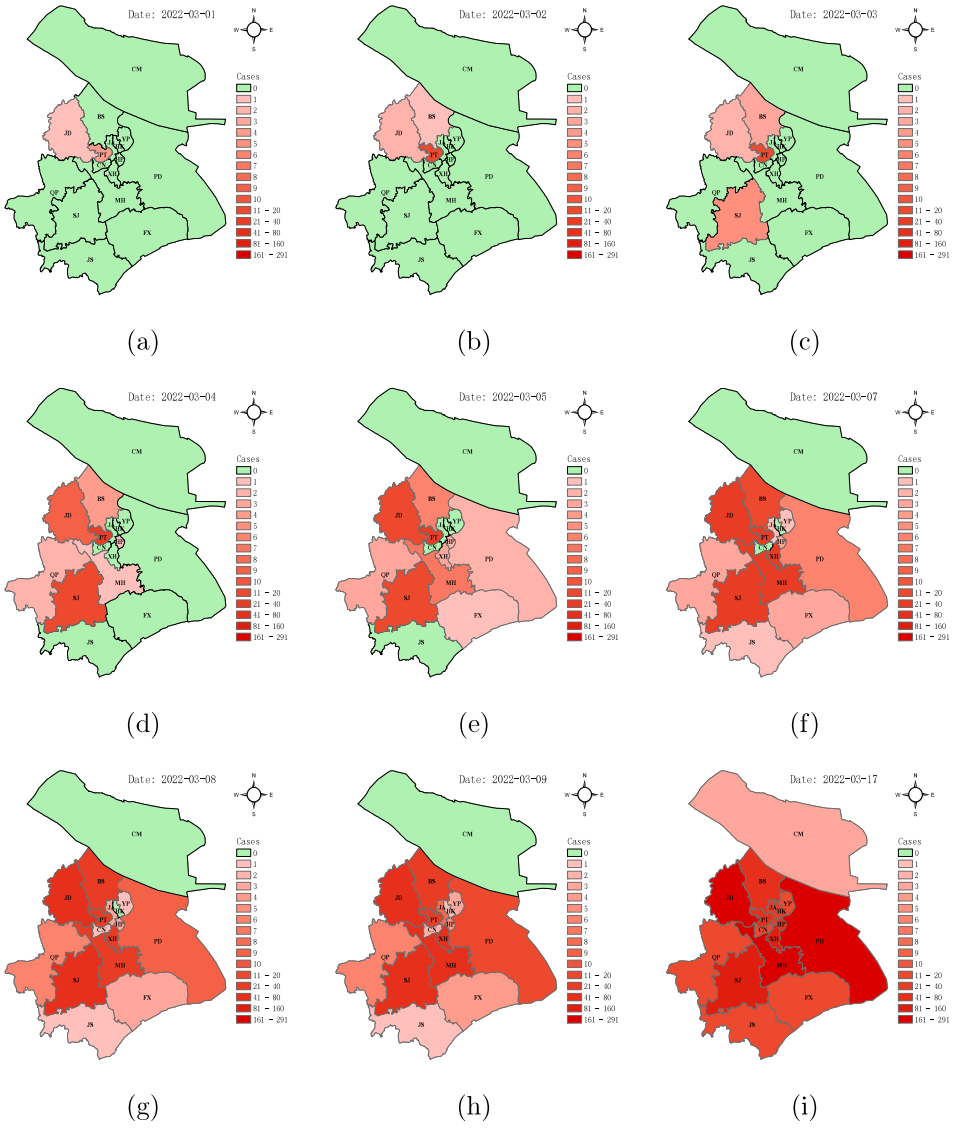

Fig. 6. Diffusion maps of districts with positive cases in Shanghai. The color shades indicate the cumulative value of positive cases in each district.

Omicron variant manifests that the prevention and control measures of central downtown area are much more effective than that of surrounding areas, and the district with the longer boundary burdens relatively greater pressure of epidemic prevention and control. Indeed, the invasion time to central downtown was not very early, but it remains a hotspot in the maps in Figs. 4 and 5. This implies that the Omicron variant is much easier to diffuse in the areas with high density of population and high activity of economy. Moreover, Chongming District, which was the last one invaded by the epidemic disease, could be attributed to the Yangtze river, which limits the population mobility as a natural barrier.

3.4. Dynamical results

First, we fix the transition matrices Λ^A and Λ^C in the difference equation model (2) by

$$\Lambda^A = h_0(t)Dist = \frac{c_0}{e^t+1}Dist,\ \Lambda^C = h_1(t)Dist = \frac{c_1}{e^t+1}Dist$$

where $Dist$ is an inverse distance matrix decided by the distance (km) between government offices of each district and $h(t)$ is a decreasing function depicting the limiting force of traffic flow. With the development of the epidemic and the strengthening of control measures, the daily undetected rate is getting smaller and smaller. We suppose the formations of daily undetected rate functions $a_i(t)$ and $b_i(t)$ are

$$a_i(t) = e^{-s_i^A \cdot t},\ b_i(t) = e^{-s_i^C \cdot t},$$

in which s_i^A and s_i^C are positive constants, indicating the control strength of the ith district (see Table 2).

From the real data, the range of the proportion of asymptomatic cases in positive cases is about $p \in [0.9, 0.95]$. Thus, $p = 0.91$ is taken in our model (2). The initial values $(A(0),\ C(0))$ of the difference equation model is set by the asymptomatic cases and confirmed cases on 1st March, 2022. In order to estimate the other parameters in the model, the adaptive Metropolis–Hastings algorithm is used to execute the Markov Chain Monte Carlo (MCMC) process (Haario et al., 2006) on the basis of the real data. Specifically, the algorithm runs for 10^6 iterations with a burn-in of 5×10^5 iterations, and the Geweke convergence diagnostic method is employed to assess the convergence of chains. As a result, the other parameters are estimated as following:

$$\beta = 2.41, r^A = 0.132, r^C = 2.8r^A, c_0 = 0.1, c_1 = 0.01,$$

Based on the estimated model parameters, the real data and simulation results of 16 districts of Shanghai are shown in Fig. 7, which capture well the basic characteristics of the outbreak. Indeed, they

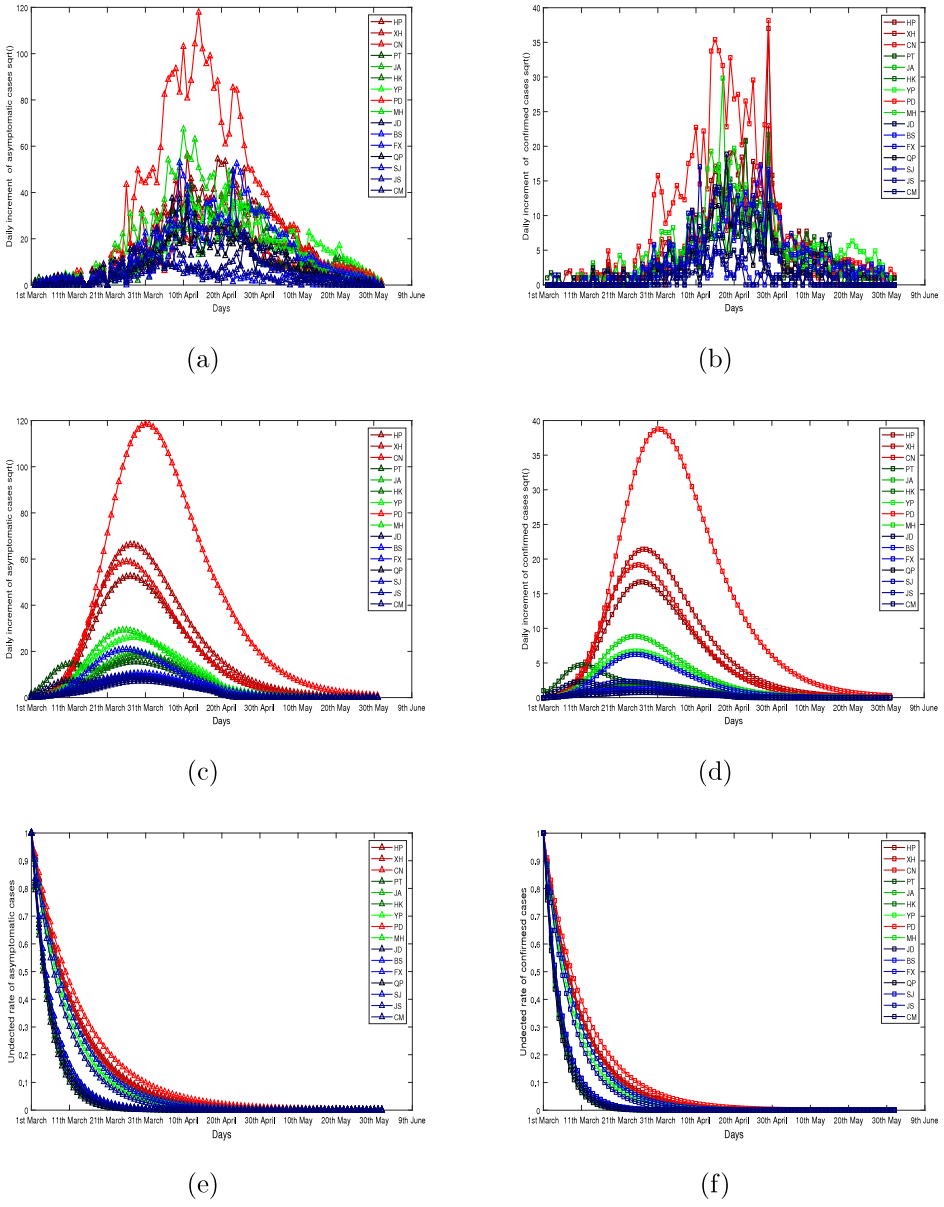

Fig. 7. Illustrations of the fitting result of model (2) in Shanghai. Panels (a), (b) indicate the real values of daily increment of asymptomatic (triangle marker) and confirmed (square marker) cases in each district in Shanghai, and panels (c), (d) indicate the corresponding fitting values of asymptomatic and confirmed cases, respectively. Panels (e), (f) indicate the daily undetected rate functions with respect to time of each district.

Table 2
Values of $\{s_i^A\}$ and $\{s_i^C\}$.

Districts	HP	XH	CN	PD	JA	HK	YP	PD
s_i^A	0.091	0.088	0.092	0.19	0.22	0.2	0.11	0.078
s_i^C	0.109	0.106	0.11	0.228	0.264	0.24	0.132	0.093
Districts	MH	JD	BS	FX	QP	SJ	JS	CM
s_i^A	0.1	0.2	0.18	0.1	0.23	0.18	0.12	0.21
s_i^C	0.12	0.24	0.216	0.12	0.276	0.216	0.144	0.252

fit well with the peak value, hitting peak time and the end time of the outbreak in Shanghai. Besides, the sort order of daily undetected rate functions of each district is consistent with that by our numerical simulations. For instance, the undetected rate function of Pudong New District is the largest and our simulation curves of positive cases predict this result. This means that shortening the time of contact tracing and promoting the daily detected rate are the very effective method of controlling the epidemic dissemination.

3.5. Daily reproduction number and its comparison

Using model (2) and formula (4), we can calculate the daily reproduction number $\mathcal{R}_i(t)$ over the whole period of outbreak. First, according to the stage classification in Fig. 3 and the hotspot maps in Fig. 5, we have the following spatiotemporal classification

$$\text{Stages} = \{\text{IDs}, \text{OEs}, \text{DEs}\}, \quad \text{Regions} = \{\text{S}_1, \text{S}_2, \text{S}_3\}, \tag{5}$$

where S_1 = {XH, HP, HK, YP, PD} denotes the hotspot districts including the central downtown area and Pudong New District, S_2 = {CM, BS, JD, JA, PT, CN, MH, FX} denotes the non-significant districts, which surrounds the hotspot districts, S_3 = {QP, SJ, JS} denotes coldspot districts, which surrounds the non-significant districts. Therefore, based

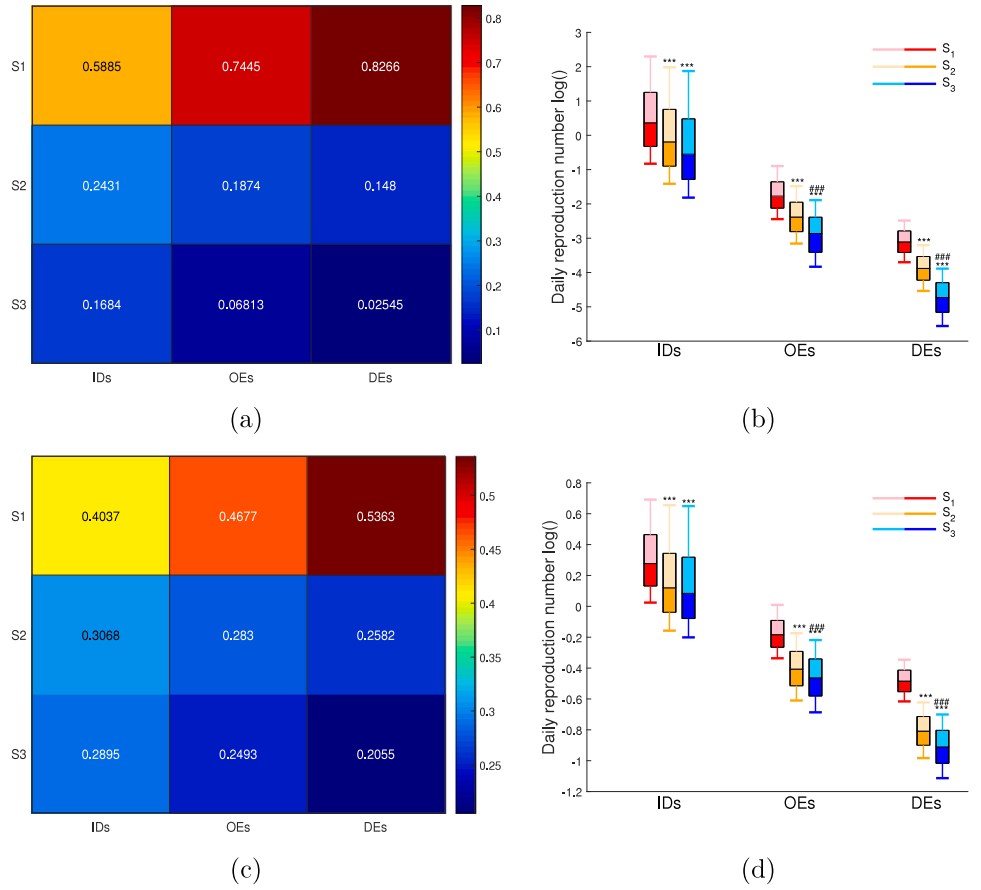

Fig. 8. Heatmaps and boxcharts of daily reproduction numbers under spatiotemporal classification (5). Panels (a) and (c) are the heat maps of daily reproduction number caused by importing one asymptomatic or confirmed case, where the values are normalized values according to the stages. Panels (b) and (d) are the boxcharts of daily reproduction number caused by importing one asymptomatic or confirmed case. In panels (b) and (d), *** and ### denote P-values < 0.001 comparing to S_1 and S_2.

on the mean operation, $\mathcal{R}_i(t)$ can be divided into 9 groups according to the spatiotemporal classification (5).

In order to highlight the comparison results, we draw the heatmaps and boxcharts for the 9 groups of $\mathcal{R}(t)$, which are presented in Fig. 8. The values in heatmaps are the normalized values of $\mathcal{R}(t)$ according to stages so as to avoid the extreme value impacts the dye of color blocks. The values in boxcharts are the log values so as to make all data visualized in the same panel. From Fig. 8, the central downtown area and Pudong New District have relatively larger reproduction numbers, and this phenomenon gets more significant in the decay stage of the outbreak. The simulation results are in good agreement with the LISA clustering map and the Getis hotspot map.

4. Conclusions and suggestions

In this study, we collect the epidemic data of various districts in Shanghai. The whole outbreak is classified into three stages including IDs, OEs and DEs. IDs is the period between 1st March and 28th March, OEs is the period between 29th March and 29th April, and DEs is the period between 30th April and 31th May. Through the geographical statistics, we investigate the spatiotemporal distribution pattern and find that the central downtown area and Pudong New District are the aggregative areas of infected cases, while the surrounding districts are of the less importance. In the study of epidemic dissemination in IDs, we find that the prevention and control measures of central downtown area are much more effective than that of surrounding areas and the district with the long boundary burdens a relatively greater pressure of epidemic prevention and control. Besides that, numerical simulations on the basis of the real data of positive cases in various districts are implemented according to the difference equation model. We find that the higher the daily undetected rate is, the more severe the epidemic in corresponding district is. Furthermore, we calculate the daily reproduction number of asymptomatic and confirmed cases in each district in the light of estimated parameters, which is consistent with the idea of Creswell et al. (2022). The heatmaps and boxcharts of the reproduction number fit well with the LISA clustering maps and Getis hotspot maps. Furthermore, we find that the time that the reproduction numbers declined to less than 1 was different in various districts. In particular, Pudong New District was the latest one that the reproduction number declined to less than 1, which was 8th April. This result is almost same as the previous study (Liu et al., 2022).

Our results indicate that the outbreak has a great aggregative effect. The central downtown area and Pudong New District are still the "high–high" type in the local spatial autocorrelation analysis in DEs, i.e., when the number of infected people in the region is large, the number of infected people in its neighboring regions is also large. Thus, these districts could be deemed as the high-burden type threatening the adjoined districts with higher incidence rates, and have a high risk of rebound and recurrence, which is partially verified in the epidemic notification in Shanghai on June 9, 2022 (Shanghai Municipal Health Commission, 2022).

Based on these study conclusions, four suggestions related to epidemic prevention and control are proposed and listed as following.

- Notice that the outbreak in Shanghai mainly gathers in the central downtown areas and Pudong New District, which has a great aggregative effect. The central downtown areas usually has a high population density, and Pudong New District is an important transportation hub in Shanghai, with a dense transportation

network. Thus, to better control the COVID-19 in an international metropolis like Shanghai, it is crucial to reduce the frequency of social contact and maintain social distance.

- According to the visualization of the initial diffusion process in Shanghai, we find that the virus first invaded its surrounding districts, then gradually invaded the central downtown area. However, the central downtown areas became soon a hotspot and remained for a long time. Therefore, for the international metropolis with epidemic situation, from the perspective of spatial control, we suggest to focus on the prevention and control of its core areas, which needs a long time to adhere to.
- Since more than 90% of the positive cases of Omicron variant infection are asymptomatic cases, and the concealment infection is significantly increased, it is advocated to strengthen personal health protection, choose off-peak travel and reduce the travel to densely populated areas.
- According to the different epidemic characteristics, it is more economic and feasible to adopt a dynamical partition prevention and control measures in corresponding districts because there are different process and elimination time of epidemic in various district.

We have two speculations that the dense traffic system and active economic vitality may contribute to the formation of hotspots for a large class of epidemic models. In addition, the invariance assumption of daily reduction rates r^A and r^C may be improved. We leave these as future researches.

CRediT authorship contribution statement

Haonan Zhong: Conceptualization, Data curation, Methodology, Formal analysis, Software, Visualization, Writing – original draft. **Kaifa Wang:** Conceptualization, Data curation, Methodology, Software, Visualization, Validation, Funding acquisition, Writing – original draft. **Wendi Wang:** Methodology, Supervision, Validation, Funding acquisition, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors are very grateful to the anonymous referees for their valuable comments and suggestions. This work was supported by the National Natural Science Foundation of China (grant numbers: 12071381, 12171396).

Appendix A. Supplementary data

Supplementary video file: ShangHai_InitialDiffusion_Movie.avi.

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jtbi.2022.111279.

References

Anon, 2022. Real time big data report of epidemic situation for novel coronavirus pneumonia. https://voice.baidu.com/act/newpneumonia/newpneumonia/?from=osari_aladin_banner. (Accessed 31 May 2022).

Anselin, L., 1995. Local indicators of spatial association - LISA. Geogr. Anal. 27 (2), 93–115. http://dx.doi.org/10.1111/j.1538-4632.1995.tb00338.x.

Creswell, R., Augustin, D., Bouros, I., et al., 2022. Heterogeneity in the onwards transmission risk between local and imported cases affects practical estimates of the time-dependent reproduction number. Phil. Trans. R. Soc. A 380, 20210308. http://dx.doi.org/10.1098/rsta.2021.0308.

Duan, X.C., Li, X.Z., Martcheva, M., Yuan, S., 2022. Using an age-structured COVID-19 epidemic model and data to model virulence evolution in Wuhan. China. J. Bio. Dyn. 16 (1), 14–28. http://dx.doi.org/10.1080/17513758.2021.2020916.

Gao, R., Yu, S.C., Wang, Q.Q., 2022. Spatiotemporal evolution of COVID-19 epidemic in the early phase in China. Chin. J. Epidemiol. 43 (3), 297–304. http://dx.doi.org/10.3760/cma.j.cn112338-20211217-00996.

Geits, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistic. Geogr. Anal. 24 (3), 189–206. http://dx.doi.org/10.1111/j.1538-4632.1992.tb00261.x.

Haario, H., Laine, M., Mira, A., Saksman, E., 2006. DRAM: efficient adaptive MCMC. Stat. Comput. 16 (4), 339–354. http://dx.doi.org/10.1007/s11222-006-9438-0.

Hu, Y., Wang, K.F., Wang, W.D., 2020. Analysis of transmissibility of COVID-19 and regional differences in disease control. ACTA Math. Appl. Sinica 43 (2), 227–237. http://dx.doi.org/10.12387/C2020019.

Li, T., Xiao, Y., 2021. Linking the disease transmission to information dissemination dynamics: An insight from a multi-scale model study. J. Theoret. Biol. 526, 110796. http://dx.doi.org/10.1016/j.jtbi.2021.110796.

Liu, X.N., Huang, F., Wang, P., 2008. Spatial Analysis Theories and Methods on GIS (Second Version). Science Press, Beijing, pp. 189–194.

Liu, Z.X., Zhu, W.L., Wang, W.B., 2022. Assessment on the outbreak caused by Omicron viriant in Shanghai based on time-varying reproduction number. Shanghai J. Prevent. Med. https://kns.cnki.net/kcms/detail/31.1635.R.20220507.1027.002.htm.

Ord, J.K., A. Geits, A., 1995. Local spatial autocorrelation statistics: distribution issues and an application. Geogr. Anal. 27 (4), 286–306. http://dx.doi.org/10.1111/j.1538-4632.1995.tb00912.x.

Shanghai Municipal Health Commission, 2022. https://wsjkw.sh.gov.cn. (Accessed 31 May 2022).

The State Council Information Office of the People's Republic of China, 2022. The 201st press conference on the prevention and control of COVID-19 in Shanghai. http://www.scio.gov.cn/xwfbh/gssxwfbh/xwfbh/shanghai/Document/1724946/1724946.htm. (Accessed 31 May 2022).

Xian, L., Lin, J., Yu, S., Zhao, Y., Zhao, P., Cao, G., 2022. Epidemiological characteristics of SARS-CoV-2 infection outbreak in Shanghai in the spring of 2022. Shanghai J. Prevent. Med. 34 (4), 294–299. http://dx.doi.org/10.19428/j.cnki.sjpm.2022.22058.

Xue, S., Hua, Y.W., Lei, Y.H., et al., 2021. Early signs of geographic spread of COVID-19: lessons learnt from outbreaks in Wuhan 2020 and Nanjing 2021. Inter. Health http://dx.doi.org/10.1093/inthealth/ihab080.

Yuan, B.Y., Liu, R., Tang, S.Y., 2022. A quantitative method to project the probability of the end of an epidemic: Application to the COVID-19 outbreak in Wuhan, 2020. J. Theor. Biol. 545, 111149. http://dx.doi.org/10.1016/j.jtbi.2022.111149.

Zhong, H.N., Wang, W.D., 2020. Mathematical analysis for COVID-19 resurgence in the contaminated environment. Math. Bio. Eng. 17 (6), 6909–6927. http://dx.doi.org/10.3934/mbe.2020357.