Infectious Disease Modelling 10 (2025) 1190–1207

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

Development of two-dimension epidemic prediction model

Jianping Huang [a,b,*], Wei Yan [b], Han Li [b], Shujuan Hu [b], Zihan Hao [b], Licheng Li [b], Xinbo Lian [a,c], Danfeng Wang [a]

[a] *Collaborative Innovation Center for Western Ecological Safety, Lanzhou University, Lanzhou, 730000, China*
[b] *College of Atmospheric Sciences, Lanzhou University, Lanzhou, 730000, China*
[c] *College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, 730000, China*

ARTICLE INFO

Article history:
Received 18 May 2024
Received in revised form 20 May 2025
Accepted 26 June 2025
Available online 27 June 2025
Handling Editor: Dr. Jianhong Wu

Keywords:
Epidemic diffusion modelling
Spatial transmission dynamics
Spatial epidemic prediction
Statistical-dynamic analysis

ABSTRACT

Epidemic prediction is a crucial foundation of disease control policy-making. Owing to the high population connectivity of current epidemics, it is essential to capture the spatial transmission of infectious diseases. However, most models currently used in epidemic prediction are single-point models, and they can only capture the time-dynamic increase of cases in limited areas. In this study, we develop a two-dimension epidemic prediction model by introducing diffusion processes, which take spatial transmission epidemics into account. We utilize mathematical theorems to prove a well-posed solution of the model. In addition, we also consider various influencing factors that affect the spread of epidemics, and introduce multiple parameterization schemes. Results suggest that this two-dimension model provides more precise predict the spatial and temporal distribution of confirmed cases. The regional average prediction score of COVID-19 in July 2022 in Lanzhou is 76.5 % and COVID-19 from May 1st to May 31st, 2023 in China is 70.7 %, respectively. Our results offer a scientific foundation for further study on the prediction of spatial epidemics, which contributes to an in-depth understanding of epidemic dynamics and provides valuable reference for the formulation of public health strategies and policies.

1. Introduction

Infectious diseases have posed a significant threat to human health and social development (Sun et al., 2016). In the context of climate change, the changing development pattern of infectious diseases will directly or indirectly affect the outbreak and spread of many infectious diseases (Wu et al., 2014). In the spring of 2009, a novel influenza A (H1N1) virus emerged in Mexico, resulting in approximately 18,000 deaths globally (Wang et al., 2014). After that, the Ebola virus outbreak in West Africa between 2013 and 2016 led to the infection of over 28,000 people, with a death toll exceeding 11,000 individuals (Garske et al., 2017). In 2012, Middle East respiratory syndrome (MERS) coronavirus (MERS-CoV) was detected in Saudi Arabia, resulting in more than 500 deaths in less than 3 years (Corti et al., 2015). In March 2020, the World Health Organization (WHO) officially declared the outbreak of COVID-19 as a global pandemic, with over 775 million confirmed

---

* Corresponding author. Collaborative Innovation Center for Western Ecological Safety, Lanzhou University, Lanzhou, 730000, China.
E-mail address: hjp@lzu.edu.cn (J. Huang).
Peer review under the responsibility of KeAi Communications Co., Ltd.

https://doi.org/10.1016/j.idm.2025.06.009

cases as of March 31, 2024 (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports). On July 23, 2022, the WHO declared the monkeypox outbreak was a global health emergency, with a cumulative total of 85,000 confirmed cases (Nuzzo et al., 2022; Zhang et al., 2024). Although there have been many significant achievements in the prevention and control of infectious diseases, these diseases still pose unpredictable risks (Heesterbeek et al., 2015).

Epidemiological model is a crucial approach for predicting the spread of infectious diseases (Chen et al., 2020). It unveils dynamic mechanisms of transmission and epidemiological characteristics of infectious diseases, and guides effective prevention and control strategies (Li et al., 2017). The susceptible-infectious-removed (SIR) and susceptible-exposed-infectious-removed (SEIR) epidemiological models are the two most widely-used dynamic models (Wu et al., 2020). Confronted with high population connectivity, localized outbreaks are possible to evolve into global pandemics (Huang, Huang, et al., 2020). Consequently, predicting the distribution of infectious diseases is of vital importance, whereas spatial epidemic modeling is particularly effective in describing the dynamic evolution of epidemics (Earn et al., 2000). Two-dimension epidemiological model enables the estimation of the spatial distribution of epidemics and the transmission across various spatial and temporal scales, thereby helping to make public health policies (Sun et al., 2007).

A precise prediction system should include the influence of environmental factors, such as temperature, humidity, and air quality (Lian et al., 2020; Liu et al., 2021). Chowdhury et al. (2018) found that there was a significant positive correlation between increasing temperature and the incidence of malaria, and enteric fever increased with rainfall. In 2020, Huang, Huang, et al. (2020) pointed out that approximately 73.8 % of the COVID-19 cases were concentrated in regions with absolute humidity between 3 g/m^3 to 10 g/m^3. It means that humid conditions may affect the spread of COVID-19 pandemic. Ma et al. (2020) pointed out the daily mortality of COVID-19 is negatively with absolute humidity. In addition, infectious diseases vary with geographic location, and local ecological changes can lead to the increase of epidemic intensity (Wilson, 2010). For example, the fungus Coccidioides immitis, which triggered epidemics in Central and South America, Southwestern United States and Mexico, multiplies on the surface of arid and semi-arid areas characterized by alkaline soil, hot summers and short moist winters (Flynn et al., 1979).

In addition, a reliable and accurate prediction system also needs to consider the impact of human factors, such as population density, population mobility and control measures. With the rapid development of transportation systems, transportation networks accelerate the spread of disease (Xu et al., 2013). When people move from one region to another, they serve as a part of global dispersal process (Wilson, 2010). High-speed rail and aviation associated with Wuhan led to a 25.4 % and 21.2 % increase in the number of new confirmed cases of COVID-19, respectively (Zhu & Guo, 2021). And international travelers result in dengue's rapid and large-scale spread by carrying the disease into non-endemic regions (Liebig et al., 2019). The spread of COVID-19 in China suggests that timely government control can prevent hundreds of thousands of confirmed cases (Tian et al., 2020). Thus, control measures and travel restrictions are essential in the early stage of epidemic outbreaks, which may influence the spread speed and scale of diseases (Liu et al., 2024). Wang et al. (2020) suggest that self-quarantine at an outbreak's early stage, is conducive to prevent the transmission of infection.

Based on the second version of the Global Prediction System for Epidemiological Pandemic (GPEP-2) developed by Huang et al. (2020, 2023a, 2023b), a two-dimension diffusion model was developed in this study. This two-dimension diffusion model characterizes the spread of epidemics in different regions. We combine three parameterization schemes in the prediction system, including control measures, vaccination, and seasonal changes in infection rates, to simulate the spatio-temporal evolution of actual epidemics. The model further improves prediction accuracy, and provides a foundation for high-resolution accurate prediction and early warning of global epidemics. Analyzing the diffusion of epidemic in different regions promotes information sharing and cooperation among regions, thus providing relatively more scientific reference for the development of control measures and resource allocation.

2. Data

Epidemic data of COVID-19 in Lanzhou City, Gansu Province and Suzhou City, Jiangsu Province were obtained from the Health Commission of Lanzhou City (https://wjw.lanzhou.gov.cn/) and the Chinese Center for Disease Control and Prevention (CDC, https://www.chinacdc.cn/), respectively. The time range of epidemic date in Lanzhou is from July 6th to July 28th, 2022 and in Suzhou is from February 14th to March 7th, 2022. Daily data collected include the number of newly confirmed cases, asymptomatic cases, asymptomatic cases transferred to confirmed cases, recovered cases, and death cases. The number of confirmed cases includes the number of people who transferred from asymptomatic to confirmed cases, and we can get the number of new infected cases by subtracting the number of asymptomatic transferred cases. The epidemic data in China after 2023 are derived from the national fever clinic monitoring data provided by the CDC Bureau of the Chinese Health Commission, including the total number of fever clinics in all regions of the country, the number of positive fever clinics and the proportion of positive fever clinics monitoring data.

3. Two-dimension epidemic model

3.1. Model design

The two-dimension model equations, which incorporate the diffusion process, are based on the GPEP-2 proposed by Huang et al. (2020, 2023a, 2023b). The GPEP-2 model defines seven disease states, including susceptible (S), protected

(P), exposed (E, infected cases in a latent period), infected (I, infected cases that have not been quarantined), quarantined (Q, confirmed and quarantined cases), recovered (R), and mortality (M) cases. Then, by introducing the different diffusion terms, we have the following two-dimension model equations:

$$\frac{\partial S(x,y,t)}{\partial t} = -\frac{\beta(x,y,t)I(x,y,t)S(x,y,t)}{N(x,y,t)} - \alpha(x,y,t)S(x,y,t) + \nabla\cdot(D_S(x,y)\nabla(S(x,y,t))), \tag{1}$$

$$\frac{\partial E(x,y,t)}{\partial t} = \frac{\beta(x,y,t)I(x,y,t)S(x,y,t)}{N(x,y,t)} - \gamma(x,y,t)E(x,y,t) + \nabla\cdot(D_E(x,y)\nabla(E(x,y,t))), \tag{2}$$

$$\frac{\partial I(x,y,t)}{\partial t} = \gamma(x,y,t)E(x,y,t) - \delta(x,y,t)I(x,y,t) + \nabla\cdot(D_I(x,y)\nabla(I(x,y,t))), \tag{3}$$

$$\frac{\partial Q(x,y,t)}{\partial t} = \delta(x,y,t)I(x,y,t) - \lambda(x,y,t)Q(x,y,t) - \kappa(x,y,t)Q(x,y,t), \tag{4}$$

$$\frac{\partial R(x,y,t)}{\partial t} = \lambda(x,y,t)Q(x,y,t) + \nabla\cdot(D_R(x,y)\nabla(R(x,y,t))), \tag{5}$$

$$\frac{\partial M(x,y,t)}{\partial t} = \kappa(x,y,t)Q(x,y,t), \tag{6}$$

$$\frac{\partial P(x,y,t)}{\partial t} = \alpha(x,y,t)S(x,y,t) + \nabla\cdot(D_P(x,y)\nabla(P(x,y,t))). \tag{7}$$

where the sum of seven categories in all grids was equal to total population (N) at any time, and it can be expressed as follows:

$$S(x,y,t) + E(x,y,t) + I(x,y,t) + Q(x,y,t) + R(x,y,t) + M(x,y,t) + P(x,y,t) = N(x,y,t). \tag{8}$$

The above presented model contains the following assumptions:

(1) Assuming that the population is diffusion mixed during the epidemic, and $D(x,y)$ represents the diffusion rate.
(2) It is evident in the equations that local quarantined cases and mortality cases cannot move to other places, which means that they are no longer able to spread viruses. In contrast, the other five categories including susceptible, protected, exposed, infected, and recovered exhibit different diffusion abilities as follows:

$$D_S > D_E > D_P > D_R > D_I, \tag{9}$$

The coefficients $\alpha, \beta, \gamma, \delta, \lambda$ and κ in equations (1)–(7) represent protection rate, infection rate, inversion of the average latent period, rate at which infected people enter quarantine, time-dependent recovery rate, and time-dependent mortality rate, respectively. The sum of N in equation (8) is assumed to be constant, which means that birth and death rates are not considered. Compared to traditional SEIR model, the improved model equations (1)–(7) introduce two new states of diseases: protected (P) and quarantined (Q) cases. Considering the increasing of people's self-protection awareness during the epidemic, there is a group of individuals (i.e., protected cases) who have a far lower chance of being infected with viruses than the susceptible individuals. Besides, it is impossible to detect and quarantine all confirmed cases due to the complexity of reality, and “Q” only represents isolated confirmed cases.

3.2. Process of epidemic diffusion

An epidemic process refers to a process that pathogens are discharged from infectious sources, invade susceptible population through certain transmission routes, form new infections, and subsequently propagate. Unlike individual phenomenon in the process of infection, an epidemic process represents a collective phenomenon occurring within populations. An epidemic process must have three basic links, including sources of infection, routes of transmission and susceptible population. These three interconnected components collectively impact the development of epidemics. The absence of any one of these links would hinder the transmission of infectious diseases among people. In addition, the epidemic intensity of infectious diseases is also constrained by natural and social factors. However, it is difficult to describe such a complex process with mathematical formulas. For simplicity, we will focus on elucidating the diffusion process.

The diffusion rate $D(x,y)$ in equation (9) is assumed as follows:

$$D(x,y) = D_0(x,y) + D_1(x,y) + D_2(x,y) + D_3(x,y) + D_4(x,y), \tag{10}$$

where $D_0(x, y)$ represents the impact of pollution on the epidemic process of infectious diseases. The role of respiratory viruses in the pathogenesis of severe respiratory infections is an issue of great importance. Airborne transmission, including droplet, droplet nucleus and dust, is the key to respiratory infections (Tellier et al., 2019). The main pathway of exposure from air pollution is through the respiratory tract. Droplet transmission occurs when an infected person exhales, sneezes or coughs, releasing pathogens which are then inhaled by others. People living in high concentration of pollution are reduced the resistance to viral and bacterial infections, resulting the increase of epidemic diffusion scale. Thus, we used the concentration of $PM_{2.5}$ as an alternative indicator in this model.

The $D_1(x, y)$ in equation (10) represents the impact of natural factors, which is mainly related to local climate and geographical conditions. For airborne infectious diseases, there is obvious seasonality with a high incidence in winter and spring, and their outbreaks are cyclical. Furthermore, natural factors can alter epidemic characteristics of infectious diseases by influencing human living habits and body resistance. For example, people stay indoors more during hot summers and cold winters, and there may be an increase in the incidence of certain infectious diseases. This trend could be attributed to factors such as proximity of individuals and reduced ventilation (Lian et al., 2023).

The $D_2(x, y)$ in equation (10) represents the impact of residential environment on the epidemic process of infectious diseases. In crowded residential areas, close contact between people is relatively more frequent, which facilitates the transmission of viruses and bacteria. Moreover, regions with crowded living conditions and high population density are prone to a high incidence of infectious diseases, as large population size and high concentration of people lead to an increased risk of epidemic transmission.

The $D_3(x, y)$ in equation (10) indicates the impact of economic and medical conditions on the transmission of infectious diseases. In certain regions, limited health education leads to inadequate disease awareness, and the lack of knowledge in turn hinders epidemic control. As a result, it is easy to control infectious diseases in regions with superior medical facilities and systems, while it is difficult to control diseases in areas with scarce medical resources. These regions may face issues such as inadequate vaccine and medication supplies, a shortage of medical personnel, and outdated health facilities. These issues exacerbate the epidemic together.

The $D_4(x, y)$ in equation (10) represents the impact of local population mobility and transportation conditions on the epidemic process of infectious diseases. In areas with high local population mobility, the probability of contact between susceptible and infected individuals increases, and it favors the spread of infectious diseases. It is worth noting that large-scale cross regional population mobility during holidays accelerates the spread and prevalence of infectious disease. Additionally, the mobility can lead to cross-regional transmission of viruses or other pathogens, particularly to areas that originally had low rates of infectious diseases. It thereby exacerbates the spread of epidemics in broad regions (Li et al., 2023).

4. Well-posedness of solution

It can be observed from equations (1)–(7) that only equations (1)–(3) are coupled by S, E and I, and the last four equations are single-variation partial differential equations essentially. Consequently, the well-posedness of equations (1)–(3) and four others will be discussed separately in the following.

Assuming that equations (1)–(3) satisfy the homogeneous Dirichlet boundary condition (non-homogeneous boundary conditions and other boundary conditions have similar conclusions), which is

$$S(x, y, t) = E(x, y, t) = I(x, y, t) = 0, (x, y) \in \partial\Omega, t > 0, \tag{11}$$

with the following initial condition,

$$S(x, y, 0) = S_0(x, y), E(x, y, 0) = E_0(x, y), I(x, y, 0) = I_0(x, y), (x, y) \in \Omega, \tag{12}$$

where $\Omega \in R^2$ is the bounded region in two-dimension plane, and $\partial\Omega$ is the boundary. $S_0(x, y)$, $E_0(x, y)$, and $I_0(x, y)$ are the bounded nonnegative functions in Ω. According to Theorem 1.1 in William E Fitzgibbon's research (Fitzgibbon et al., 2021), the well-posedness of systems (1)–(3) can be achieved as the following theorem 1 (Supplementary Material).

Theorem 1. *Based on the definite solution conditions of* (11) *and* (12), *there is a unique global weak solution* (*solution in the square integrable space*) *for systems* (1) - (3), *and the solutions* $S(x, y, t)$, $E(x, y, t)$ *and* $I(x, y, t)$ *are uniformly bounded on* $t \in \mathbb{R}_+$.

In addition, based on systems (1)–(3) discussed above and assuming the other four single variable equations (4)–(7) satisfied the same initial condition and boundary condition like (11) and (12), equations (4) and (6) are initial value problems for first-order linear ordinary differential equations essentially. It is easy to use the method of constant variation to obtain the expression of the solution, and it can be proven that the solution has an upper bound controlled by a known function (initial value and external forcing term), indicating the existence and uniqueness of the solutions (Supplementary Material). Equations (5) and (7) have similar forms and are special cases of equation systems (1)–(3). Therefore, according to Theorem 1, the definite solution problem of equations (5) and (7) is well-posedness.

The above discussion indicates that the two-dimension epidemic model (1)–(7) we constructed is well-posedness under given initial and boundary conditions. The detailed proof process can be found in the Supplementary Material.

5. Parameterization schemes

Parameterization is a mathematical process that expresses the state quantities of a system, process or model as a function of some independent variables called parameters. This method was utilized in atmospheric sciences initially (Stensrud, 2009). Considering the unpredictability of the real world, where various situations may arise unpredictably, integrating parameterization into epidemic models enables a relatively more comprehensive representation of infectious disease transmission. Through different parameterization schemes, we can get comparatively better prediction results in different situations of the real world. As shown in Fig. 1, we have constructed the following parameterization schemes in the two-dimension model.

5.1. Parameterization scheme for vaccination

Vaccines are crucial tools in pandemic response and protection against severe diseases, which also provide protection against infection and transmission (Zou et al., 2022). When a community has a high vaccination rate, the spread of viruses is significantly reduced, as most people viruses encounter are immune. The more people who are vaccinated, the lower the risk of virus exposure for those who cannot be vaccinated. Therefore, constructing a vaccine parameterization scheme is helpful for predicting infectious diseases. At the onset of the epidemic, we assume that vaccinations have been completed, which can already achieve the simulation effect. $P(0)$ and $S(0)$ represent the initial number of protected and susceptible individuals, respectively. In this parameterization scheme, vaccination transforms the susceptible individuals of $S(0)$ into protected group of $P(0)$. The parameterization scheme is constructed as follows:

$$P(0) = S(0) \cdot Vr \cdot Ve, \tag{13}$$

$$S(0) = S(0) - S(0) \cdot Vr \cdot Ve, \tag{14}$$

where *Vr* represents vaccination rate, and *Ve* represents protection rate of vaccines.

5.2. Parameterization scheme for control measures

Owing to the continuous mutation of viruses and the difficulty of large-scale vaccination, control measures such as reducing human interactions are still the most effective methods to prevent the development of epidemics. We developed a parameterized scheme of different control measures to accurately simulate and predict the outbreak of epidemics (Huang et al., 2020, 2021). This parameterization contains four coefficients, including *Days_con*, *E*0, *Infection_attenuation_rate* and *Diffusion_attenuation_rate*. Since the first confirmed case is often detected several days after infection, we assume that *Days_con* represents the time from the beginning of epidemics to the time when government departments start taking control measures. Furthermore, there are already several infected people at the first time of the government report, we therefore use *E*0 to represent the number of initially exposed cases. *Infection_attenuation_rate* represents attenuation rate of

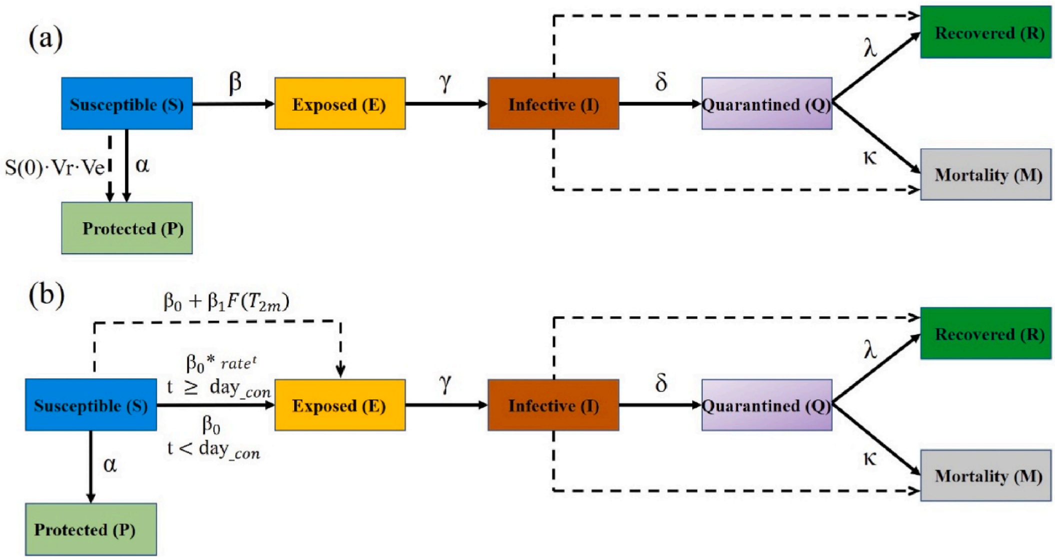

Fig. 1. Schematic diagram of the parameterization scheme of vaccination (a), control measures and seasonal change in temperature (b).

infection, and *Diffusion_attenuation_rate* represents attenuation rate of diffusion. Once government initiates strong control measures, the infection rate and diffusion rate of epidemics decline exponentially over time. The parameterization scheme is constructed as follows:

$$\beta(t) = \begin{cases} \beta_0 & t < Days\_con \\ \beta_0 \cdot Infection\_attenuation\_rate^t, & t \geq Days\_con \end{cases}, \tag{15}$$

$$D(t) = \begin{cases} D & t < Days\_con \\ D \cdot Diffusion\_attenuation\_rate^t, & t \geq Days\_con \end{cases}, \tag{16}$$

where β_0 and D represent historical infection rate or base infection rate, and basic diffusion rate, respectively.

5.3. Parameterization scheme for seasonal changes

Many environmental factors that vary with the seasons also affect the development of epidemics, and temperature is a major factor in the environment. In the real world, there is a basic infection rate β_0 and basic diffusion rate D_1 that influencing the diffusion of epidemic. In addition, the viruses can maintain different activities at different temperature, resulting the infection rate change with season. In this study, we constructed a parameterization scheme for the effect of temperature on the infection rate to incorporate the factor of temperature into the model. The actual infection rate $\beta(t)$ and diffusion rate $D_1(t)$ will change with temperature in different seasons. Then the parameterization scheme is constructed as follows:

$$\beta(t) = \beta_0(t) + \beta_1 F(T_{2m}), \tag{17}$$

$$D_1(t) = D_1 + a_1 F(T_{2m}), \tag{18}$$

where $F(T_{2m})$ is the probability distribution function (PDF) of local temperature at 2m above ground level obtained by Huang et al. (2020). They found that 60 % of the confirmed COVID-19 cases occurred in regions where the air temperature ranged from 5 °C to 15 °C. The parameter β_1 and parameter a_1 are obtained through nonlinear fitting. As the probability of virus transmission varies in different temperature regions, we calculated global distribution of annual PDF of temperature using reanalysis data from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR).

6. Determination of model coefficients

On account of the complexity of the real world as well as the concealment and mutation of viruses, it is difficult to obtain an accurate set of infectious disease model parameters. Many researchers used empirical parameters for sensitivity experiments, aiming at exploring the impact of different parameters on the development trend of infectious diseases. Nevertheless, statistical-dynamic forecasting method has been used widely in atmospheric sciences (Huang & Chou, 1990; Huang & Wang, 1991), and it improves the accuracy of weather and climate prediction. We can obtain parameters that are comparatively closer to the real world, and simulate the epidemic development trend through this method.

Firstly, we refined the latest epidemic data into high resolution, and collected comparable data from similar outbreak scenarios to invert basic coefficients in the epidemic prediction model. Then, we set an initial value and a bound for each coefficient in the two-dimension epidemic model. In this step, we used the date collected from similar epidemic outbreaks to invert basic coefficients, and basic coefficients are used as the initial value in the epidemic model of the latest outbreak. In the next step, the latest epidemic data and coefficient optimization algorithms (such as least squares) were used to invert each coefficient in real time. In the above two inversion, the minimum variance sum of obtained time series parameter and actual data was adjusted through iterative calculation, so that the initial values of the coefficients were adjusted and kept close to the real value. Finally, we got a set of coefficients that were close to the real world, which were substituted into the epidemic model to predict the development of epidemic trends. The initial values of coefficients and initial number of individuals in various disease states may affect the accuracy of inverted coefficients. It is necessary to combine empirical assumptions and change the initial value many times to perform inversion for accurate and stable results. In addition, it is useful to empirically determine initial values of coefficients in the epidemic model, and use parallel methods in program algorithms to improve the speed of inversion and prediction. Adopting efficient numerical methods and optimization algorithms, it is possible to significantly enhance the computational speed and predictive accuracy of the model. The integration of these methods ensures the effectiveness and stability of the model in practical applications, which is crucial for epidemic analysis and management.

7. Process of system-prediction

The prediction process includes data collection, data processing, basic coefficients inversion and assimilation, simulation and prediction, as well as accuracy verification. Fig. 2 is a schematic diagram of the system-prediction process. Through

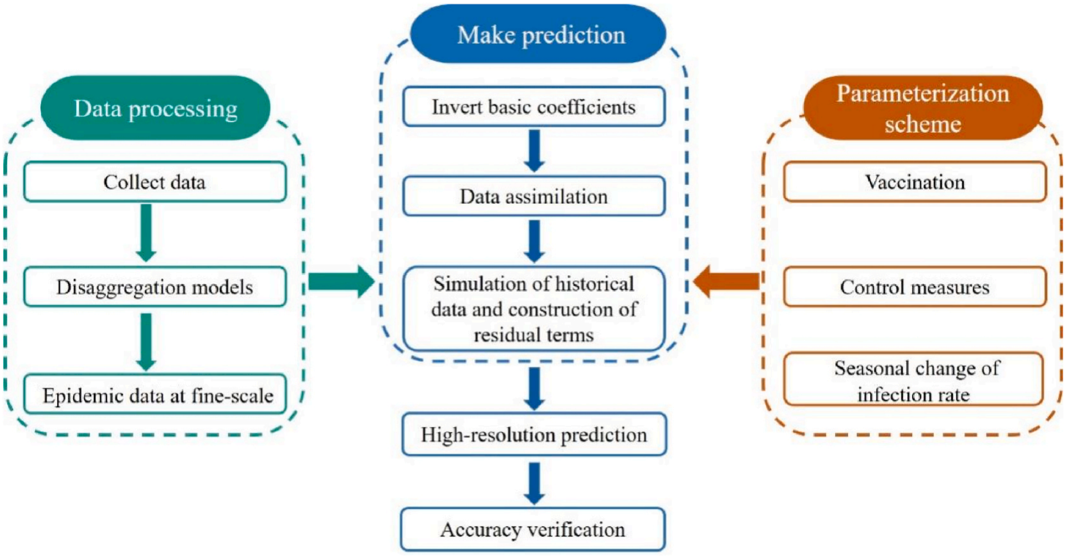

Fig. 2. Schematic diagram of two-dimension diffusion system prediction process.

precise data handling and model validation, this prediction process provides a solid foundation for dynamic analysis of epidemics.

Data collection: When an epidemic breaks out in a country or region, epidemic data are collected from the website of the local Health Commission. Social data include population size (https://hub.worldpop.org/), gross domestic product (Zhao et al., 2017) and urbanization levels (https://lbs.amap.com/). The meteorological data, such as 2m temperature and dew-point temperature data, are sourced from the ERA5 reanalysis dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF, https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=form).

Data processing: Owing to the epidemic data we collected are always in the city-level or province-level, we need to refine and process it to fine-scale level. The process primarily involves spatially epidemic data and covariates like population, temperature, and dew point temperature as inputs (see Fig. 3).

The number of daily cases in a given region is relatively small compared to the total population. Therefore, we assume that the number of cases y_i in a given region follows a Poisson distribution:

$$y_i \sim Pois(\mu_i), \tag{19}$$

where μ_i is expected confirmed cases, μ_i is equal to the sum of the incidence numbers at all grid points in the region($\mu_i = \sum_{j=1}^{n_i} pop_{i,j}\beta_{i,j}$), n_i is the total number of grid points in region i, and $pop_{i,j}, \beta_{i,j}$ denotes the population size and incidence rate in grid $j = 1, \ldots, n_i$, respectively.

Then, the logarithmic value of the incidence rate $\beta_{i,j}$ is considered to have a functional relationship with the covariate:

$$\log\left(\beta_{i,j}\right) = \beta_0 + \sum_{q}^{Q} \beta_q \mathrm{cov}(q)_{i,j} + \xi, \tag{20}$$

where β_0 represents intercept, $\mathrm{cov}(q)_{i,j}$ denotes covariate values in grid j of region i, with coefficients $\beta_q(q = 1, \ldots, \mathrm{Q})$, Q is the number of covariates, and ξ is a Gaussian random field.

Basic coefficients inversion: Similar epidemic data and coefficient optimization algorithm are used to invert basic coefficients of the two-dimension epidemic model.

Multi-source dynamic ensemble prediction (MDEP): When there is only one infectious source, the traditional infectious prediction models can achieve reliable prediction results. However, due to the complexity of the real world and the concealment of viruses, there is often more than one infectious source, that is, the development of the epidemic will change over time, and initial prediction results cannot represent real development. Therefore, Huang et al. (2023b) introduced an ensemble prediction method of weather forecasting (Zhu, 2005) into the infectious disease model, dynamically adjusting the model parameters based on the true development of epidemics. By summing up initial prediction results with dynamic prediction results, the integration result can be used as the final prediction result, thus obtaining accurate prediction results.

Simulation and prediction: Fitting simulate data with reported data, we invert basic coefficients and parameters in the parameterization scheme in the previous step. Besides, we use corresponding parameterization schemes for different regions. Then, we obtain predicted results through integration.

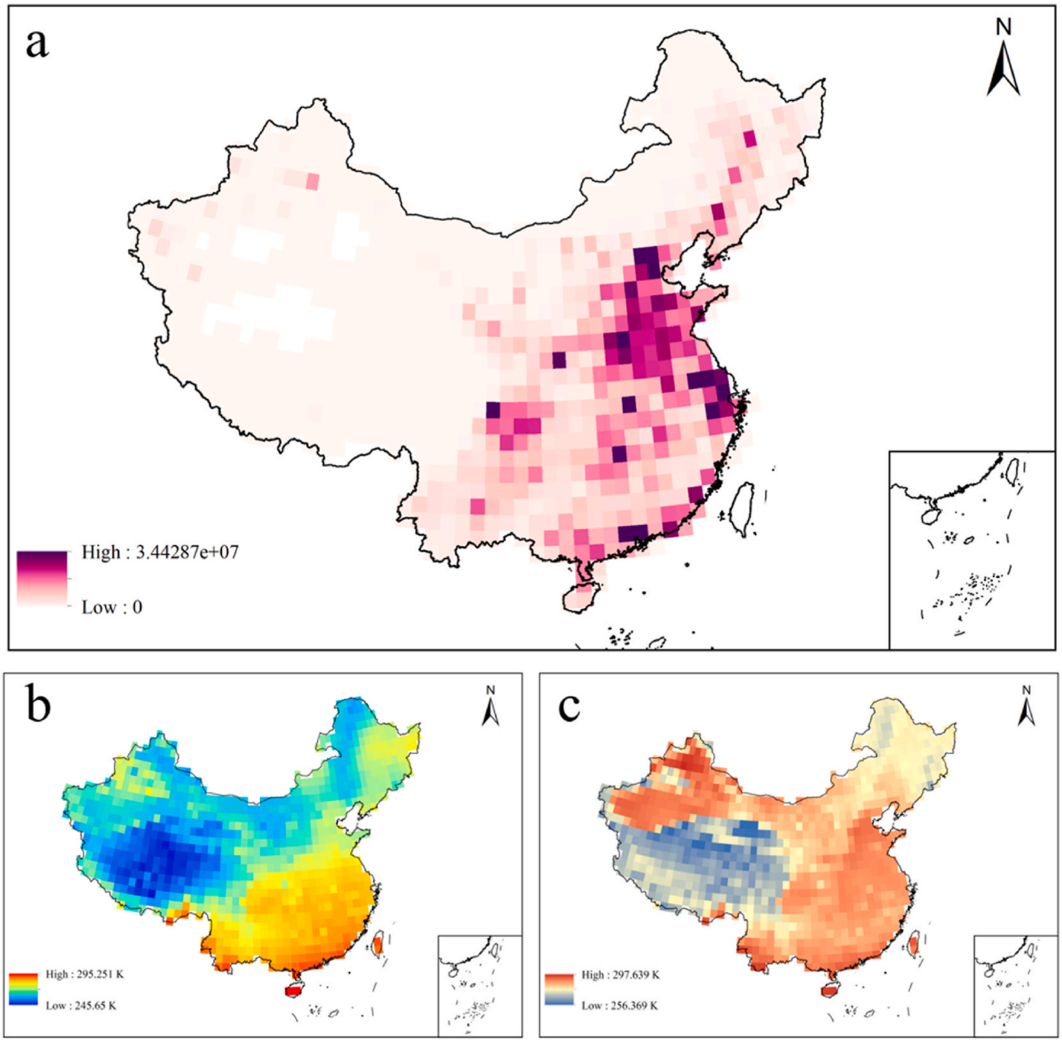

Fig. 3. Potential covariates used in disaggregation downscaling model. Population density (a, the blank space represents missing data, but it belongs to Chinese territory), 2m dewpoint temperature (b) and 2m temperature (c).

Accuracy verification: We compare prediction results with reported data to test prediction accuracy. By verifying the accuracy of the model, we can revise the algorithm well and improve the parameterization scheme, thus leading to a more accurate prediction system. We adopt the following two methods to evaluate the predictive ability of the model. The first method is based on relative error (RE). We assess the model's predictive accuracy at the macro level, calculating the discrepancy between predicted and actual numbers of infected individuals. This approach enables a rigorous assessment of the model's effectiveness in estimating the scale of infections.

$$RE = \frac{1}{T}\sum_{t=1}^{T}\left|\frac{R_s - R_t}{R_t}\right| \times 100\%, \tag{21}$$

$$MS = 1 - |RE|, \tag{22}$$

where MS (Macro Score) represents the prediction score of the sum of cumulative infected cases for epidemic prediction, and this score evaluates overall accuracy of predictions across various outbreaks. Additionally, RE denotes the relative error between cumulative predicted data and actual data for each outbreak, where “Rs” stands for predicted values, “Rt” represents actual values observed, and “T” denotes the number of samples considered in the analysis.

The second method is the bias score, which was previously used to measure the prediction bias of a model for a certain level of precipitation. On the grid points of the region, we set the value numerically equal to the ratio of the total number of grid points that meet the infection *MS* in the predicted region to the corresponding total number of grid points in the actual situation. This is used to evaluate the spatial prediction performance of the two-dimension diffusion model in the region.

$$BS = \frac{hits}{hits + misses} \times 100\%, \tag{23}$$

where *BS* (Bias Score) represents the prediction score at region level for epidemic prediction, *hits* and *misses* denote the number of grid points with *MS* greater than or equal to set score (60 %) on each grid point and the number of grid points with *MS* less than set score on each grid point, respectively.

8. Application of prediction modeling to COVID-19

8.1. Prediction of cases in Suzhou City

On February 14th, 2022, a local outbreak of COVID-19 occurred in Suzhou City, Jiangsu Province of China. On the first day, there were 8 confirmed and 4 asymptomatic infections. The local government implemented timely and proactive control measures to mitigate further spread of the epidemic. As of March 7th, the epidemic was under control and approaching its end, with a cumulative total of 141 confirmed cases.

We simulate this epidemic on an ideal grid, assuming Suzhou as a square area with [0, 52] × [0, 52]. We uniformly distribute the total population of Suzhou to the internal grid points of the square area except for the boundary, assuming that the 12 infected individuals on the first day are distributed on a grid point. Then we introduce the parameterization scheme of control measures into the model to simulate the evolution of the epidemic, and the simulation results are shown in Figs. 4 and 5. In terms of the temporal evolution trend of the number of confirmed cases illustrated in Fig. 4, the simulation results demonstrate a good concordance with the actual development. Therefore, it is credible to use diffusion models to simulate real development of infectious diseases. The model achieves reliable results in predicting the total number of infected individuals. As of March 7th, the average relative error between predicted results of cumulative infected cases and actual data of COVID-19 is 7.5 %. Additionally, analyzing the spatial distribution of confirmed cases, as shown in Fig. 5, we find that the COVID-19 first spread and developed in the initial grid. On account of the existence of diffusion terms in the model, the distribution of confirmed cases is spatially multi-point over time. Moreover, the number of confirmed cases at the initial infection spatial location shows a trend of increasing at first and then decreasing. If control measures and preventive measures are taken in the early stages of infection to control the spread of infectious sources, the extent and development intensity of infectious disease transmission will decrease significantly (S1, S2).

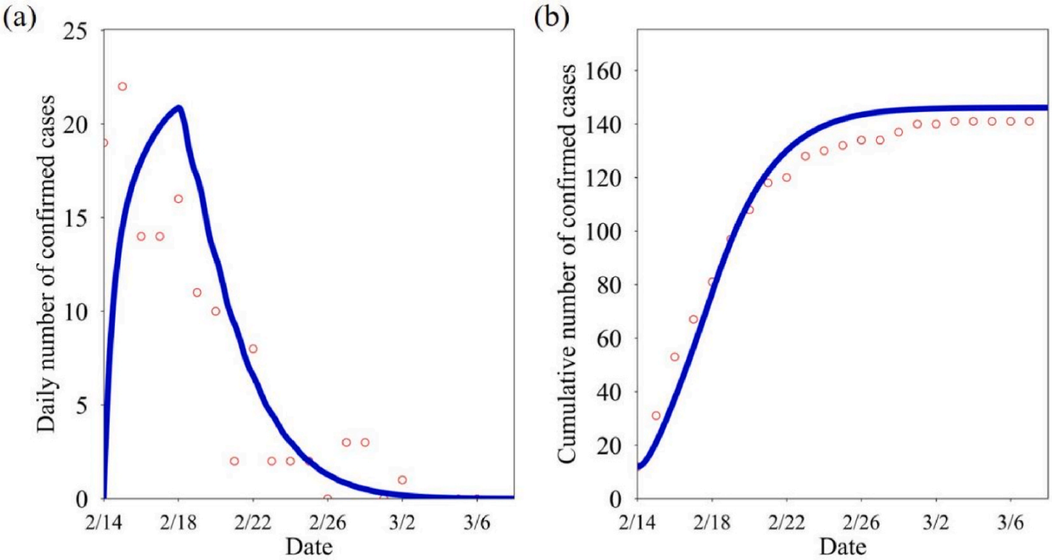

Fig. 4. Comparison of prediction results and reported data of daily (a) and cumulative (b) confirmed cases in Suzhou, Jiangsu Province from February 14th to March 7th in 2022. Red circles represent the curve of reported data of the epidemic, and solid blue lines represent the curve of prediction of the epidemic.

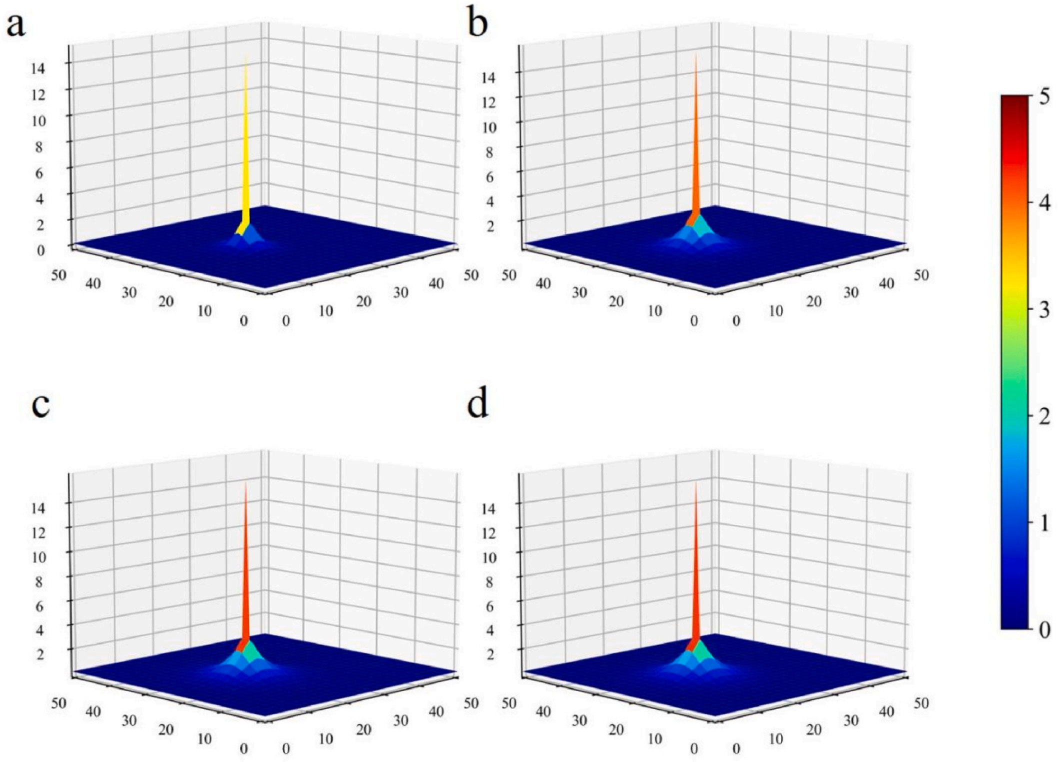

Fig. 5. Spatial distribution of the number of infected individuals simulated in an ideal grid. (a: February 18th; b: February 23rd; c: February 28th; d: March 7th).

8.2. *Prediction of cases in Lanzhou City*

A sudden outbreak of COVID-19 occurred in Lanzhou City, Gansu Province of China in July 2022. The Health Commission of Lanzhou City reported for the first time on July 8th, 2022, with 4 positive cases and 1 asymptomatic case, and the onset date of the first positive case is July 6th. As is well known, the onset date is several days earlier than the reported date. Thus, we collect daily number of newly diagnosed cases rather than reported cases in Lanzhou City from July 6th to July 28th in this research. The time evolution trend is shown in Fig. 6. Since the first positive case appeared on July 6th, the development of the epidemic in Lanzhou has been divided into multiple stages. Firstly, the development of the epidemic in Lanzhou was in the early stage, with no more than 100 daily new cases. Subsequently, there was a significant increase in daily new diagnosed cases from July 17th, reaching a maximum of 260 cases on July 21st. In the early stage of epidemic, we could discover one source of infection and made prediction based on limited data. However, there are multiple transmission chains in Lanzhou and the infection rate has changed over time. Therefore, the MDEP method must be employed for prediction.

The prediction results of the number of daily newly diagnosed cases is shown in Fig. 8a, which includes the first prediction results, the second prediction results and the dynamic integration results. The first wave of predictions yielded reliable results before July 15th, which is consistent with the actual development trend. Since July 16th, there has been a significant increase in the number of daily newly diagnosed cases, indicating the possibility of new potential infection sources. Therefore, we started dynamic prediction on July 16th. We conducted a new round of prediction, and integrated two prediction results to obtain final prediction results. The integration results are in good agreement with actual daily number of newly diagnosed cases, and the accuracy is significantly increased. Meanwhile, in terms of the evolution trend of the cumulative number of confirmed cases, the integration results are relatively consistent with actual development (Fig. 8b). The model can basically predict the development trend and achieve good results. The prediction score of MS (Fig. 10a) is 83.4 % on July 10th and 92.5 % on July 28th. The increase of prediction score indicates that the MDEP method can predict the impact of potential infection sources and the prediction results of cumulative cases become comparatively more accurate.

In terms of the spatial distribution of diagnosed cases, they are mainly concentrated in Chengguan District (Fig. 7). There were some diagnosed cases in Yuzhong County and Anning District over time, but no large-scale spread occurred. A sudden break occurred in Chengguan District on July 16th, with a significant increase in daily new diagnosed cases. We divide infection cases into grid points and use models for spatial prediction, and prediction results are shown in Fig. 9. The overall

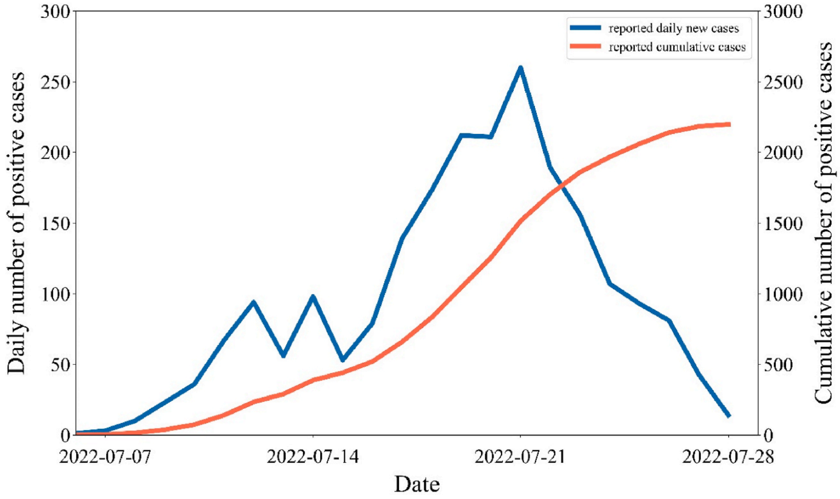

Fig. 6. Daily number of positive cases (blue line) and cumulative number of positive cases (orange line) of COVID-19 reported by fever clinics in China from March 1st to June 6th.

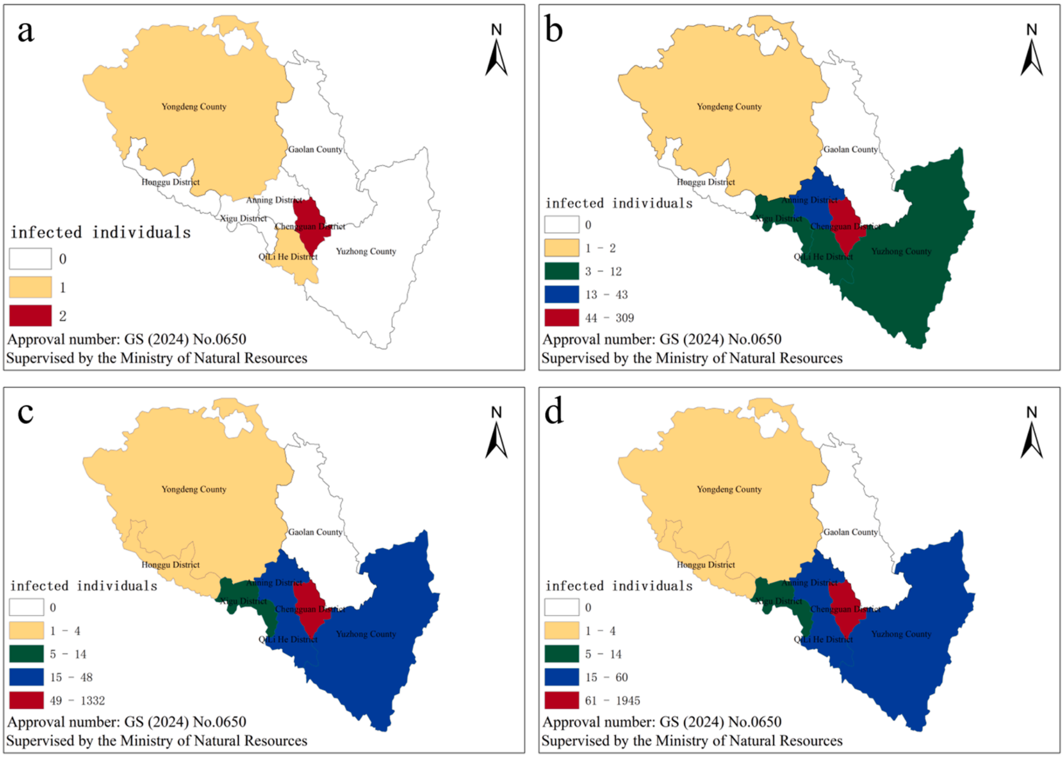

Fig. 7. Spatial distribution of cumulative number of positive individuals on July 7th (a), July 14th (b), July 21st (c), and July 28th (d).

distribution of predicted results is almost consistent with that of actual development, and local prediction is also close to actual cases. As of July 28th, the predicted score of BS can reach 76.4 %, which means that good results have been achieved in both spatial distribution prediction and local prediction (Fig. 10b upper). When we set the MS of each grid is greater than 50 %, 60 % and 70 %, respectively. At first, the daily region-level prediction score shows different results. Nevertheless, the same daily BS were consistently generated since July 22nd (Fig. 10b bottom), which means the MDEP significantly increase the prediction score of epidemic prediction. The spatial prediction score shows a downward trend over time, indicating that the model can currently achieve good results in short-term prediction, while long-term prediction effect still needs to be

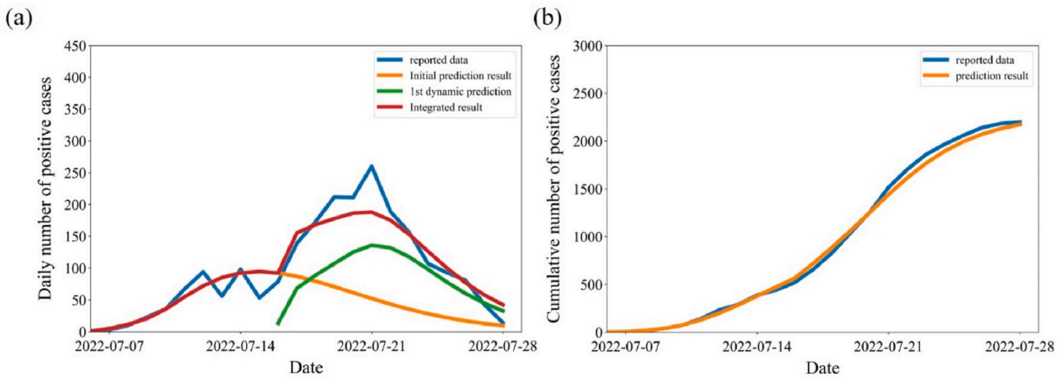

Fig. 8. Comparison of prediction results and reported data of daily (a) and cumulative (b) positive cases in Lanzhou City. Solid blue lines represent the curve of reported data of the epidemic.

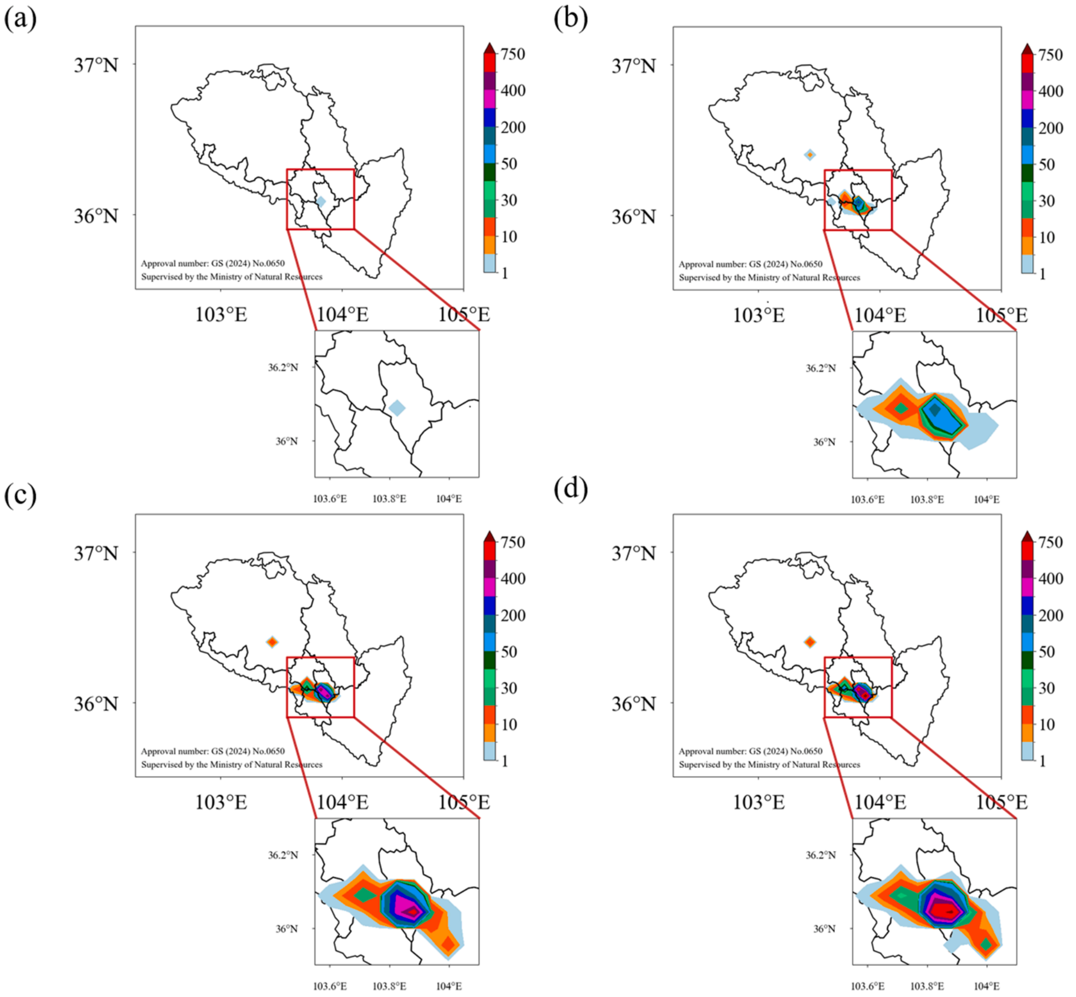

Fig. 9. Spatial distribution prediction of cumulative number of positive individuals on July 7th (a), July 14th (b), July 21st (c), and July 28th (d).

improved. In the future, long-distance diffusion and local development limitations will be introduced into the model to obtain relatively more accurate spatial prediction results.

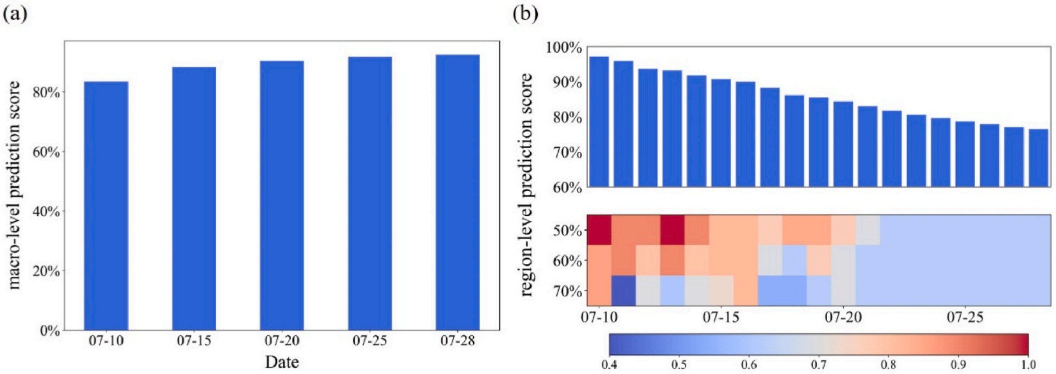

Fig. 10. Macro-level (a) and region-level (b: upper) prediction score of cumulative number of positive cases in Lanzhou and daily region-level prediction score (b: bottom) of the MS calculated by cumulative cases on each grid set as greater than 50 %, 60 % and 70 %, respectively.

8.3. Prediction of COVID-19 diffusion in China

Since the outbreak of COVID-19, the Chinese government always insists on prioritizing the physical health and life safety of the people, and has invested heavily in epidemic control, vaccine development and patient care. In the context of mass vaccination of the population and decreased pathogenicity of COVID-19, the Chinese government started adjusting control measures. Since January 8th, 2023, the management measures of COVID-19 in China have been changed from category B infectious disease with management of category A to category B disease with management of category B (http://www.nhc.gov.cn/xcs/). Although the pathogenicity of COVID-19 has declined, COVID-19 did not disappear completely. This means that people are likely to be infected if the protection of vaccines declines. From April to May in 2023, as human mobility rapidly increased during International Labour Day vacation, the spread of COVID-19 accelerated significantly, leading to a resurgence in the number of infections (Fig. 11). As nucleic acid testing is no longer routinely conducted, the quantification of positive cases is now primarily through fever clinics. The data show the development trend of COVID-19 infection in China, but they do not represent the actual number of confirmed cases.

We collect data on positive cases in fever clinics of different provinces across China from March 1st to June 6th, 2023, and the time evolution trend is shown in Fig. 11. From March 1st to April 20th, the incidence of positive cases remained relatively low and exhibited stable development, while the number of positive cases in fever clinics began to significantly increase from April 23rd. On May 20th, the daily number of new positive cases achieved its peak. By June 6th, the number of daily new positive cases was less than 15000 cases. The cumulative number of positive cases in fever clinics nationwide reached 870,948. In terms of spatial distribution, as shown in Fig. 12, the number of new confirmed cases has existed in fever clinics of all provinces in mainland China since April 23rd, 2023, and the epidemic was in the primary development stage. As of May 14th, 2023, the epidemic had widely spread across the country and the epidemic was in the primary outbreak stage. Compared with April 23rd, the number of new confirmed cases in the mid-May phase of the epidemic in China showed a significant increasing trend, with the number of confirmed cases increasing by nearly 12 times. During the outbreak stage in mid-May, the number of new confirmed cases in South China accounted for 59.31 % of all regions, while that in Northwest

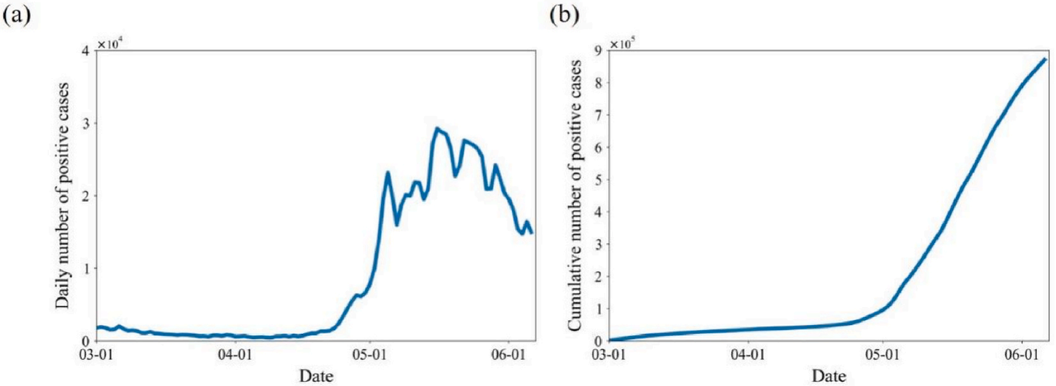

Fig. 11. Daily number of positive cases (a) and cumulative number of positive cases (b) of COVID-19 reported by fever clinics in China from March 1st to June 6th.

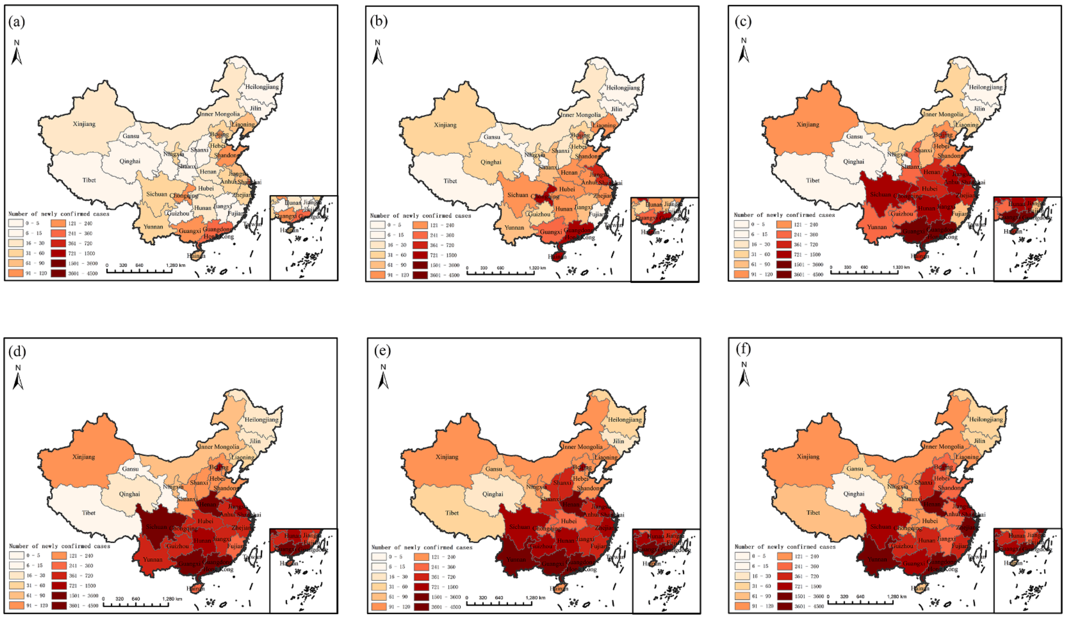

Fig. 12. Spatial distribution of cumulative number of positive individuals on April 23rd (a), April 30th (b), May 7th (c), May 14th (d), May 21st (e), and May 28th (f) in 2023.

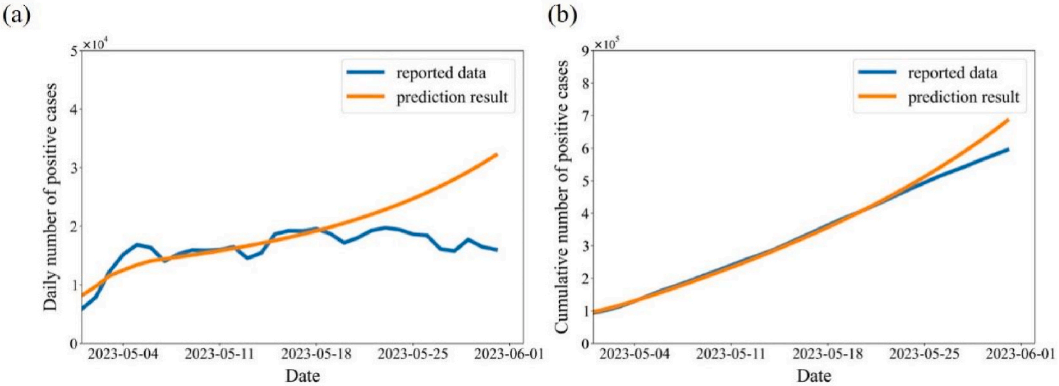

Fig. 13. Comparison of prediction results and reported data of daily (a) and cumulative (b) positive cases in China's fever clinics from May 1st to May 31st in 2023. Solid blue lines represent the curve of reported data of the epidemic, and solid orange lines represent the curve of prediction of the epidemic.

China was 1.84 %. It indicates that there is a spatial difference in the distribution of the number of confirmed cases, and the geographic scope and number of outbreaks spreading expand significantly (Fig. 12). In terms of spatial distribution, the number of confirmed cases continued to expand in eastern China, where there are more cases, while in northwestern China, where there are fewer confirmed cases, the coverage of confirmed cases remained relatively stable. The epidemic spread from eastern and southern China to most parts of China, with relatively more spread in Guangdong Province and its neighboring areas, as well as in some economically developed areas, and less spread in western China and remote areas of northern China. The spatial distribution of the epidemic not only reflects the influence of population mobility and geographic factors, but also may be related to economic development and accessibility of public health facilities in different regions.

We use a downscaling approach to decompose epidemic data to a resolution of $1^\circ \times 1^\circ$, covering the entire country. Since the development trend of the epidemic from March to May is relatively stable and the growth of positive cases is slow, we choose May 1st for prediction, with a duration of 30 days. As shown in Fig. 13, the system has reliable predictions of various local epidemics in China. The actual development trend of epidemics closely aligns with our projections, and final predictions correspond well with actual reported data. It can be found that in the evolution trend of the cumulative number of confirmed cases, the model is relatively consistent with actual development, and they are highly overlapped in the first 25 days. By the 30th day, predicted results are more than the actual number of positive cases. This indicates that when data are

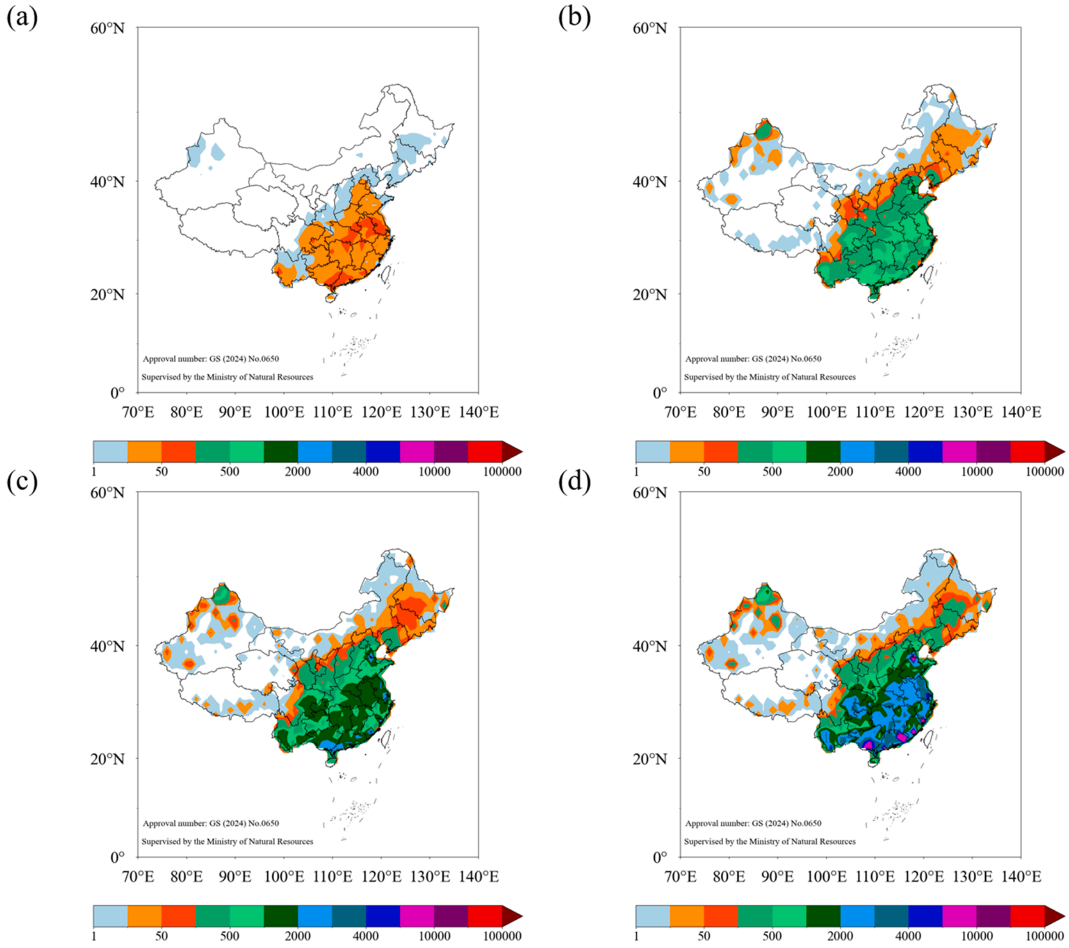

Fig. 14. Spatial distribution prediction of cumulative number of positive individuals on May 1st (a), May 10th (b), May 20th (c) and May 31st (d).

representative, the model can predict the trend of development and obtain good results. The general prediction accuracy of prediction results is more than 80 % (Fig. 15a), and the model can achieve relatively more accurate predictions in a short period.

In terms of spatial distribution, predicted results are shown in Fig. 14. We find that the general distribution area of the predicted results is almost consistent with the actual results, and local development is close to the actual situation. As of

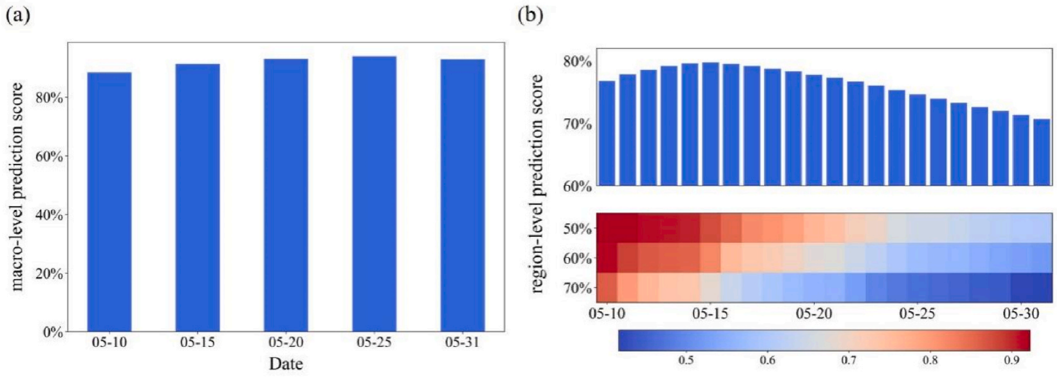

Fig. 15. Macro-level (a), region-level prediction score (b: upper) of cumulative number of positive individuals and daily region-level prediction score (b: bottom) of the MS calculated by cumulative cases on each grid set as greater than 50 %, 60 % and 70 %, respectively.

May 31st, the predicted score of BS reaches 70.7 % (Fig. 15b upper). When we set the MS of each grid is greater than 50 %, 60 % and 70 %, respectively. The daily region-level prediction scores all show good results in the first fifteen days (Fig. 15b bottom). It indicates that using diffusion models to predict real epidemics can provide comparatively more accurate epidemic analysis information for government departments, thus helping them to take targeted response measures in advance. Future work will specifically optimize for data fluctuations to enhance prediction accuracy.

9. Conclusions

We introduce diffusion terms into the second version of the GPEP-2, consider the interaction terms among grid points to the mathematical equation system, and extend the traditional prediction model to two dimensions. The new model not only considers the development of local epidemics, but also considers the outward spread of virus carriers, which is relatively more consistent with the actual development and spread of epidemics. We use mathematical methods to prove the well-posedness of the model, and conclude that the corresponding definite solution problem for each term of model is well-posedness. In addition, we conduct an ideal experiment on a regular grid to simulate the COVID-19 outbreak in Suzhou. The model achieves good results in predicting the cumulative number of confirmed cases, and the feasibility of the model in grid is verified. Meanwhile, we construct various parameterization schemes to simulate the spatial and temporal evolution of real epidemics, considering factors like the distribution of susceptible populations, medical resources, population density, and weather factors. We use the model to predict the evolution of the cumulative number of infected cases in Lanzhou and the cumulative number of positive cases of fever clinics in China, achieving reliable prediction results, and realize spatial visualization. This provides a research basis for high-resolution and accurate prediction, as well as early warning of future global epidemics. Through predicting the temporal evolution trend of epidemics in local region and the spatial development distribution across China, we can provide relatively more scientific references for the formulation and implementation of control measures.

Atmospheric sciences have been conducting numerical simulation predictions for many years, and it is one of the best subjects for numerical simulation predictions. The statistical-dynamic prediction method in atmospheric sciences provides an effective method to accurately invert coefficients. By inverting various parameters in epidemic model, we can conduct a prediction which is relatively closer to the development of real epidemics. Nevertheless, this method has not been used in public health previously. Therefore, the introduction of research methods from atmospheric sciences into the prediction of infectious disease modeling can provide substantial support for epidemic prevention.

10. Discussions

However, there are still limitations that need to be improved in our prediction system. Firstly, the model is sensitive to initial values and predicts poorly when the initial values are not accurate. Secondly, the data we currently obtain are mainly based on regional levels, and the accuracy of spatial downscaling of epidemic data should be improved to make refined predictions. In future research, we will add more covariates related to infectious diseases to improve the performance of the model and reduce the sensitivity of forecasts in the region to prior changes. On account of the differences in data observed at different spatial scales, we suggest establishing and maintaining a comprehensive and reliable database of classified infectious diseases. It would be invaluable to researchers and policymakers in addressing the global threat posed by viruses. There is no doubt that the application of statistical dynamic prediction methods in atmospheric sciences to parameter inversion in infectious disease model prediction is a major innovation in public health. However, the statistical process is limited by historical data. When historical data are scarce, parameters obtained may have significant deviations. Therefore, we will incorporate data assimilation methods in atmospheric sciences to process initial values and parameters in the future. Additionally, we need to combine artificial intelligence methods such as machine learning to obtain a set of stable coefficients that are close to the actual situation.

Currently, our prediction system incorporates three parameterization schemes, including control measures, vaccination and seasonal changes of infection rate, considering factors that affect the development and changes of infectious diseases. However, the development and changes of infectious diseases in the real world are complex. For example, most respiratory infectious diseases exhibit periodicity, with epidemic peaks appearing at regular intervals. In consequence, we need to add more parameterization schemes, such as periodic fluctuations of infectious diseases and large-scale population movements, to improve the universality and accuracy of the system for global prediction. Moreover, we establish a model at fine resolution, which can be coupled with various models, e.g., bioaerosol transport and climate system models. By coupling the model with climate system, we can better integrate infectious diseases with climate change, and provide relatively more accurate predictive warnings in the context of climate change.

CRediT authorship contribution statement

Jianping Huang: Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Wei Yan:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation. **Han Li:** Writing – original draft, Visualization, Methodology, Data curation. **Shujuan Hu:** Writing – review & editing, Methodology, Conceptualization. **Zihan Hao:** Writing – review & editing, Writing – original draft, Methodology, Data curation. **Licheng Li:**

Writing – review & editing, Writing – original draft, Methodology, Data curation. **Xinbo Lian:** Writing – review & editing, Writing – original draft, Methodology, Data curation. **Danfeng Wang:** Writing – review & editing, Writing – original draft, Data curation.

Funding

This work was jointly supported by the Major Project of Guangzhou Laboratory, Grant No. GZNL 2024A01004, National Key Research and Development Program of China (2023YFC3503400) and Gansu Province Intellectual Property Program (Oriented Organization) Project (22ZSCQD02).

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.

Acknowledgments

The authors acknowledge the Chinese Center for Disease Control and Prevention (CDC) and the Health Commission of Lanzhou City for providing the COVID-19 data. The authors acknowledge the Amp for providing Points of Interest (POI) data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.idm.2025.06.009.

References

Chen, T. M., Rui, J., Wang, Q. P., Zhao, Z. Y., Cui, J. A., & Yin, L. (2020). A mathematical model for simulating the phase-based transmissibility of a novel coronavirus. *Infectious diseases of poverty, 9*(1), 18–25.

Chowdhury, F. R., Ibrahim, Q. S. U., Bari, M. S., Alam, M. J., Dunachie, S. J., Rodriguez-Morales, A. J., & Patwary, M. I. (2018). The association between temperature, rainfall and humidity with common climate-sensitive infectious diseases in Bangladesh. *PLoS One, 13*(6), Article e0199579.

Corti, D., Zhao, J., Pedotti, M., Simonelli, L., Agnihothram, S., Fett, C., … Lanzavecchia, A. (2015). Prophylactic and postexposure efficacy of a potent human monoclonal antibody against MERS coronavirus. *Proceedings of the National Academy of Sciences, 112*(33), 10473–10478.

Earn, D. J., Rohani, P., Bolker, B. M., & Grenfell, B. T. (2000). A simple model for complex dynamical transitions in epidemics. *Science, 287*(5453), 667–670.

Fitzgibbon, W. E., Morgan, J. J., Tang, B. Q., & Yin, H. M. (2021). Reaction-diffusion-advection systems with discontinuous diffusion and mass control. *SIAM Journal on Mathematical Analysis, 53*(6), 6771–6803.

Flynn, N. M., Hoeprich, P. D., Kawachi, M. M., Lee, K. K., Lawrence, R. M., Goldstein, E., … Wong, G. A. (1979). An unusual outbreak of windborne coccidioidomycosis. *New England Journal of Medicine, 301*(7), 358–361.

Garske, T., Cori, A., Ariyarajah, A., Blake, I. M., Dorigatti, I., Eckmanns, T., … Donnelly, C. A. (2017). Heterogeneities in the case fatality ratio in the West African Ebola outbreak 2013–2016. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1721), Article 20160308.

Heesterbeek, H., Anderson, R. M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., … Isaac Newton Institute IDD Collaboration. (2015). Modeling infectious disease dynamics in the complex landscape of global health. *Science, 347*(6227), Article aaa4339.

Huang, J., & Chou, J. (1990). Studies on the analogous rhythm phenomenon in coupled ocean-atmosphere system. *Science China Chemistry, 33*, 851–860.

Huang, Z., Huang, J., Gu, Q., Du, P., Liang, H., & Dong, Q. (2020). Optimal temperature zone for the dispersal of COVID-19. *Science of the Total Environment, 736*, Article 139487.

Huang, J., Liu, X., Zhang, L., Zhao, Y., Wang, D., Gao, J., … Liu, C. (2021). The oscillation-outbreaks characteristic of the COVID-19 pandemic. *National Science Review, 8*(8), Article nwab100.

Huang, J., & Wang, S. (1991). The monthly prediction experiments using a coupled analogy-dynamical model. *Acta Meteorologica Sinica, 5*, 8–15.

Huang, J., Zhang, L., Chen, B., Liu, X., Yan, W., Zhao, Y., … Wang, D. (2023a). Development of the second version of the global prediction system for epidemiological pandemic. *Fundamental Research*. https://doi.org/10.1016/j.fmre.2023.02.030

Huang, J., Zhang, L., Liu, X., Wei, Y., Liu, C., Lian, X., … Zhang, T. (2020). Global prediction system for COVID-19 pandemic. *Science bulletin, 65*(22), 1884.

Huang, J., Zhao, Y., Yan, W., Lian, X., Wang, R., Chen, B., & Chen, S. (2023b). Multi-source dynamic ensemble prediction of infectious disease and application in COVID-19 case. *Journal of Thoracic Disease, 15*(7), 4040.

Li, H., Huang, J., Lian, X., Zhao, Y., Yan, W., Zhang, L., & Li, L. (2023). Impact of human mobility on the epidemic spread during holidays. *Infectious Disease Modelling, 8*(4), 1108–1116.

Li, M. T., Sun, G. Q., Zhang, W. Y., & Jin, Z. (2017). Model-based evaluation of strategies to control brucellosis in China. *International Journal of Environmental Research and Public Health, 14*(3), 295.

Lian, X., Huang, J., Huang, R., Liu, C., Wang, L., & Zhang, T. (2020). Impact of city lockdown on the air quality of COVID-19-hit of wuhan city. *Science of the Total Environment, 742*, Article 140556.

Lian, X., Huang, J., Li, H., He, Y., Ouyang, Z., Fu, S., … Guan, X. (2023). Heat waves accelerate the spread of infectious diseases. *Environmental Research, 231*, Article 116090.

Liebig, J., Jansen, C., Paini, D., Gardner, L., & Jurdak, R. (2019). A global model for predicting the arrival of imported dengue infections. *PLoS One, 14*(12), Article e0225193.

Liu, H., Cai, J., Zhou, J., Xu, X., Ajelli, M., & Yu, H. (2024). Assessing the impact of interventions on the major omicron BA. 2 outbreak in spring 2022 in Shanghai. *Infectious Disease Modelling*. https://doi.org/10.1016/j.idm.2024.02.013

Liu, X., Huang, J., Li, C., Zhao, Y., Wang, D., Huang, Z., & Yang, K. (2021). The role of seasonality in the spread of COVID-19 pandemic. *Environmental Research, 195*, Article 110874.

Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., … Luo, B. (2020). Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Science of the Total Environment, 724*, Article 138226.

Nuzzo, J. B., Borio, L. L., & Gostin, L. O. (2022). The WHO declaration of monkeypox as a global public health emergency. *JAMA, 328*(7), 615–617.

Stensrud, D. J. (2009). *Parameterization schemes: Keys to understanding numerical weather prediction models*. Cambridge University Press.

Sun, G., Jin, Z., Liu, Q. X., & Li, L. (2007). Pattern formation in a spatial S–I model with non-linear incidence rates. *Journal of Statistical Mechanics: Theory and Experiment, 2007*(11), Article P11011.

Sun, G. Q., Jusup, M., Jin, Z., Wang, Y., & Wang, Z. (2016). Pattern transitions in spatial epidemics: Mechanisms and emergent properties. *Physics of Life Reviews, 19*, 43–73.

Tellier, R., Li, Y., Cowling, B. J., & Tang, J. W. (2019). Recognition of aerosol transmission of infectious agents: A commentary. *BMC Infectious Diseases, 19*, 1–9.

Tian, H., Liu, Y., Li, Y., Wu, C. H., Chen, B., Kraemer, M. U., … Dye, C. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science, 368*(6491), 638–642.

Wang, L., Chu, C., Yang, G., Hao, R., Li, Z., Cao, Z., … Song, H. (2014). Transmission characteristics of different students during a school outbreak of (H1N1) pdm09 influenza in China, 2009. *Scientific Reports, 4*(1), 5982.

Wang, J., Li, Z., Cheng, X., Hu, H., Liao, C., Li, P., … Chen, Z. (2020). Epidemiologic characteristics, transmission chain, and risk factors of severe infection of COVID-19 in Tianjin, a representative municipality city of China. *Frontiers in Public Health, 8*, 198.

Wilson, M. E. (2010). *Geography of infectious diseases, 1055*. Infectious diseases.

Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet, 395*(10225), 689–697.

Wu, X., Tian, H., Zhou, S., Chen, L., & Xu, B. (2014). Impact of global change on transmission of human infectious diseases. *Science China Earth Sciences, 57*, 189–203.

Xu, F., McCluskey, C. C., & Cressman, R. (2013). Spatial spread of an epidemic through public transportation systems with a hub. *Mathematical Biosciences, 246*(1), 164–175.

Zhang, L., Huang, J., Yan, W., Zhao, Y., Wang, D., & Chen, B. (2024). Global prediction for mpox epidemic. *Environmental Research, 243*, Article 117748.

Zhao, N., Liu, Y., Cao, G., Samson, E. L., & Zhang, J. (2017). Forecasting China's GDP at the pixel level using nighttime lights time series and population images. *GIScience and Remote Sensing, 54*(3), 407–425.

Zhu, Y. (2005). Ensemble forecast: A new approach to uncertainty and predictability. *Advances in Atmospheric Sciences, 22*, 781–788.

Zhu, P., & Guo, Y. (2021). The role of high-speed rail and air travel in the spread of COVID-19 in China. *Travel Medicine and Infectious Disease, 42*, Article 102097.

Zou, Z., Fairley, C. K., Shen, M., Scott, N., Xu, X., Li, Z., … Zhang, L. (2022). Critical timing and extent of public health interventions to control outbreaks dominated by SARS-CoV-2 variants in Australia: A mathematical modelling study. *International Journal of Infectious Diseases, 115*, 154–165.