Science Bulletin 65 (2020) 1297–1305

Contents lists available at ScienceDirect

Science Bulletin

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

Science Bulletin
www.scibull.com

Article

Using the contact network model and Metropolis-Hastings sampling to reconstruct the COVID-19 spread on the "Diamond Princess"

Feng Liu [a], Xin Li [b,c,*], Gaofeng Zhu [d]

[a] Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China
[b] National Tibetan Plateau Data Center, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China
[c] Center for Excellence in Tibetan Plateau Earth Sciences, Chinese Academy of Sciences, Beijing 100101, China
[d] Key Laboratory of Western China's Environmental Systems (Ministry of Education), Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 4 April 2020
Received in revised form 25 April 2020
Accepted 26 April 2020
Available online 5 May 2020

Keywords:
Contact network model
Small-world
Chain-binomial model
Airborne spread
Transmissibility
The basic reproductive number R_0

ABSTRACT

Traditional compartmental models such as SIR (susceptible, infected, recovered) assume that the epidemic transmits in a homogeneous population, but the real contact patterns in epidemics are heterogeneous. Employing a more realistic model that considers heterogeneous contact is consequently necessary. Here, we use a contact network to reconstruct unprotected, protected contact, and airborne spread to simulate the two-stages outbreak of COVID-19 (coronavirus disease 2019) on the "Diamond Princess" cruise ship. We employ Bayesian inference and Metropolis-Hastings sampling to estimate the model parameters and quantify the uncertainties by the ensemble simulation technique. During the early epidemic with intensive social contacts, the results reveal that the average transmissibility t was 0.026 and the basic reproductive number R_0 was 6.94, triple that in the WHO report, indicating that all people would be infected in one month. The t and R_0 decreased to 0.0007 and 0.2 when quarantine was implemented. The reconstruction suggests that diluting the airborne virus concentration in closed settings is useful in addition to isolation, and high-risk susceptible should follow rigorous prevention measures in case exposed. This study can provide useful implications for control and prevention measures for the other cruise ships and closed settings.

1. Introduction

The outbreak of COVID-19 on the "Diamond Princess" cruise ship (hereinafter referred to as Diamond Princess), which was one of the serious cases in the early outbreak, attracted extensive worldwide attention. This case provides ideal samples with little interference to study coronavirus epidemiology, considering that the infected origin is known, the number of susceptible individuals is constant and the prevention and control work are timely deployed after the first case is confirmed. A modeling analysis of the spread on the "Diamond Princess" identifies a more reasonable spread of the disease and further provides implications for global, national and regional control and prevention measures.

Mathematical modeling has long been an epidemiological tool [1]. Typical methods include regression [2] and compartmental models such as SIR (susceptible, infected and recovered) [3] and its variants are frequently used to model disease spread, such as SARS (Severe Acute Respiratory Syndrome) [4–6] and COVID-19 [7–12]. However, these models assume that epidemics transmit in a homogeneous population, i.e., new infectees are from the group of susceptible individuals who can be infected with the same transmission probability. This assumption is far from the social contact patterns, especially when population structure is a spatial–temporal variant with quarantine and socioeconomic activities [13]. Therefore, a more reasonable epidemic model considers social contact patterns.

The advent of complex network theory [14,15] provides a method for studying heterogenous epidemiological dynamics [16]. The contact network model extends SIR-type models to fit more complex epidemic situations. It can respond to prevention and control work, such as a decline in the number of nodes indicates the rigorous quarantine of close contacts and the infector is isolated, and the restrictions on movement and public gatherings lead to less contact frequency (reduced transmissibility) between individuals.

Another potential contribution of the contact network model is that it provides an alternative statistic, the average transmissibility, for the basic reproductive number R_0 to determine the severity of an epidemic [17]. R_0 varies with the population and presents

* Corresponding author.
E-mail address: xinli@itpcas.ac.cn (X. Li).

https://doi.org/10.1016/j.scib.2020.04.043
2095-9273/

substantial heterogeneity [18]; for example, the R_0 of COVID-19 ranges from 6.47 [8] to 3.6–4.0 [9] and 2–2.5 [19]. To more efficiently estimate the transmission rate, it is urgent to define new statistics, such as a data-driven rate that is easy to obtain [20] and a time-dependent rate [21]. The average transmissibility, however, denotes the mean probability of disease transmission among individuals and is free of the population; therefore, it is theoretically robust to benefit the public health strategy decision making.

However, the contact network model is still limited and fails to meet the scenario without contact, such as the spread in the presence of dense virus-laden aerosols. Ong et al. [22] found that COVID-19 can deposit on air exhaust outlets, which makes airborne spread possible. Therefore, a new mathematical model for no-contact between individuals is required.

In this study, we develop contact network and no-contact models to keep track of the disease spread in heterogeneous populations and contact patterns, such as unprotected and protected contact and airborne spread. We expect that epidemic reconstruction can provide useful implications for implementing control and prevention measures in larger regions.

2. Materials and methods

2.1. Basic facts and two stages of spread and prevention

The Diamond Princess set sail from Yokohama on 20 January (hereinafter omit 2020), via Hong Kong (25 January), and returned to Yokohama on 3 February. At this time, the ship had 3711 members including 2666 passengers and 1045 crew, and it is not clear whether and how many people went ashore during the stopover.

As shown in Fig. 1, a passenger from Hong Kong was on board on 20 January and disembarked on 25 January, which is during his symptomatic period, and he was confirmed as a case of COVID-19 on 1 February [23]. On 5 February, all the passengers and crew were asked to self-isolate in their cabins until 19 February.

As of 26 February, a cumulative total of 705 cases were reported (another case was reported on 2 March and 712 cases were reported as of press time). The daily report of confirmed cases from 5 February to 19 February is shown in Fig. 1.

The epidemic is undertaken in two stages (see the bottom of Fig. 1) according to the quarantine and the major routes of virus transmissions. Stage 1 with unprotected contact (from 20 January to 4 February). Full transmission of COVID-19 occurred. We assume that the major driver of transmission was droplets and fomites due to the close unprotected contact among the population. This situation agrees with the contact network model. We assume that C_1 crews and P_1 passengers were infected, and the corresponding infection rates for passengers and crews are α_1 and α_2, respectively.

We divide stage 2 (from 5 February to 19 February) into two distinct scenarios:

(1) Stage 2a with protected contact for crew. The contact between crew was restricted because they should maintain ship operations, and provide limited services with personal protections. For the same reason, we assume that the transmission between crew and passengers can be ignored. A contact network model considering quarantine also works in this scenario. The daily removed infected crew is supposed to be the number of daily confirmed cases multiplied by α_1, and we assume that C_{2a} crew were extra infected and R_{2a} infected crew were left on 19 February.

(2) Stage 2b with the airborne spread for passengers. We assume there was no physical contact between the isolated passengers, and the closed space and air recirculation throughout the cabins facilitated the airborne spread of the virus. We develop a differential equation epidemic model to study this scenario. We assume P_{2b} passengers were extra

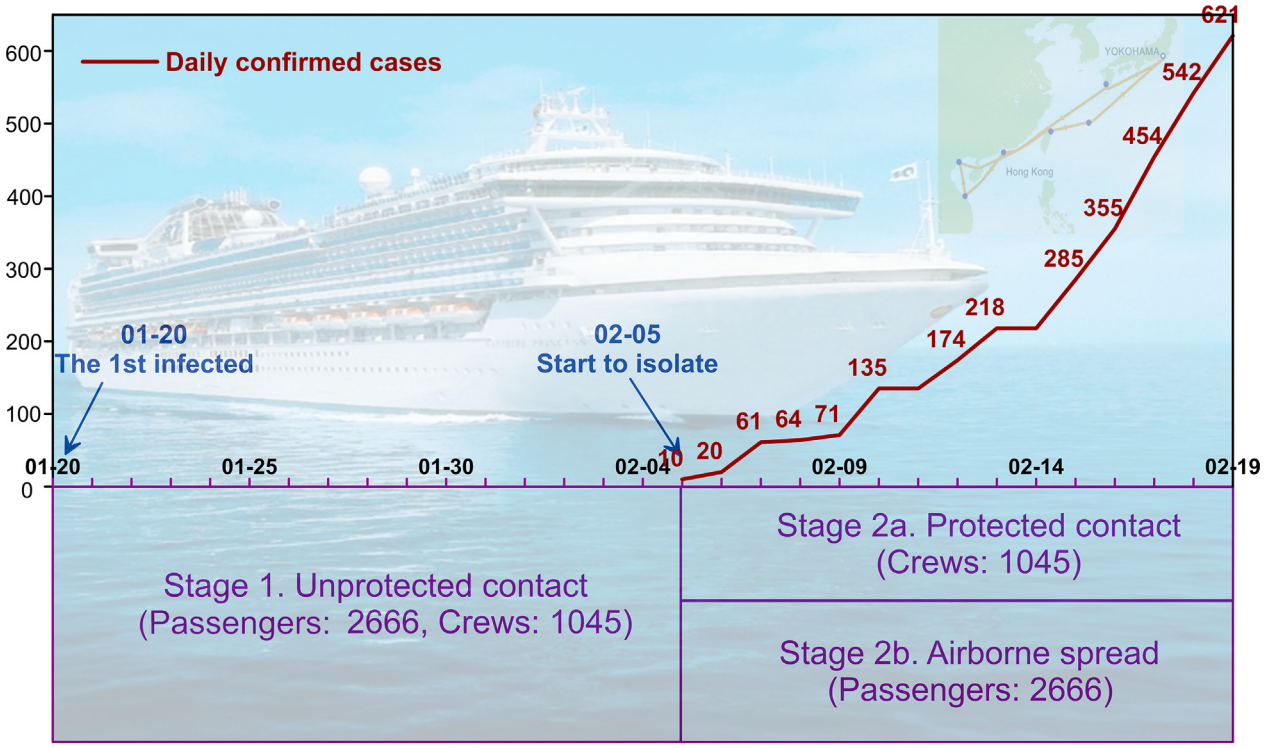

Fig. 1. The daily confirmed cases and two stages (1, 2a and 2b) in the "Diamond Princess" epidemic.

infected in this scenario. The number of daily removed infected passengers is supposed to be the number of daily confirmed cases multiplied by α_2, and R_{2b} infected passengers remained on 19 February.

The aim of this study is to model the epidemic dynamics in these stages and then reconstruct the epidemic curve on the Diamond Princess realistically. We determine the unknown figures including P_1, C_1 and the models in the following text.

2.2. Data analysis

We first estimated the unknown numbers presented in the above section.

It was assumed that all 705 cases were infected on Diamond Princess before 19 February. An investigation from the Ministry of Health, Labor and Welfare of Japan reported that 65 crews and 466 passengers were infected on 18 February [24]. The corresponding infection rates for crews and passengers were α_1 =12.24% and α_2 =87.76%, respectively. Assuming these rates are constant, the infected passengers and crew were 619 (sum of P_1 and P_{2b}) and 86 (sum of C_1 and C_{2a}) until 19 February, respectively.

Since there are no available data to directly calculate the details of the infected passengers and crews, we have to refer to the indirect information of the SARS outbreak in Amoy Gardens of Hong Kong, cases of which were 331 during the incubation and infectious period [25]. This epidemic typically provided a confirmation of airborne spread of coronavirus [26]; we consequently assume that all the cases were infected through viral aerosol, and based on the population census from Hong Kong [27], the corresponding infection rate was 3.01%.

Assuming the airborne infection rate of COVID-19 is similar to SARS, then P_{2b} = 80 leaving out the infected in stage 1. We now have P_1 = 539 and the corresponding infection rate for passengers through unprotected contact 87.08%. If the ratio of infected crew is the same as passengers, then all numbers are clear (see Table 1).

2.3. Modeling strategy

We designed a combination of epidemic models to understand the two distinct stages.

2.3.1. Small-world network-based chain-binomial model for stage 1. For the typical properties of economic and social operations, the epidemiologically relevant contact network regards one node as an individual, the degree of a node K as the possible contacts of an individual, and transmissibility T as the average probability that the disease transmits between two individuals. A brief depiction of these three components is shown in Fig. 2a. The degree for each node is 4 and if the transmissibility is defined as 0.3, then in the next time interval (Fig. 2b, t_2), at least one neighbor of the white infected node will probably be infected.

Small-world [14] is a typical realistic network between regular and disordered networks, with the nodes highly clustered with their neighbors and sparsely connected with the nodes far from them. In Fig. 2, nodes meet with their 4 neighbors at time t_2, and then disconnect some of their neighbors and link with farther nodes (dashed red edges) with probability p at the following time t_3. The small-world network corresponds to the unprotected contact pattern in stage 1, in which passengers can communicate with their companions and stewards, and meet new people in public activities; alternatively, the crew works with other colleagues and provides daily services for special passengers and occasionally responds to inquiries from other passengers. This contact pattern implies that the connectivity among population on the ship is high local clustering, but a fraction of the relationship is randomly changed.

The chain-binomial model [28] is a classic network models. The chain-binomial model stores the infected time for each individual and simulates the transmission along each edge by the Bernoulli trial. Individuals are recovered if their infected times are larger than the infected period. A brief flow diagram of chain-binomial model is shown in Fig. 3.

The basic reproductive number R_0 in chain-binomial model is calculated [17] by

$$R_0 = \frac{T}{T_C} = \left[1-(1-t)^{ip}\right]\left(\frac{\langle k\rangle}{\langle k^2\rangle - \langle k\rangle}\right)^{-1}. \quad (1)$$

Here $T = \left[1-(1-t)^{ip}\right]$ and $T_C = \frac{\langle k\rangle}{\langle k^2\rangle - \langle k\rangle}$ denote the probability of transmission and critical transmissibility between infected and susceptible individuals, where ip is infected period, t is the average transmissibility for each contact, and $\langle k\rangle$ and $\langle k^2\rangle$ present the first and second moments of degree K. T_C is regarded as an epidemic threshold, i.e., $T > T_C$ implies $R_0 > 1$ and outbreak will continuously expand; otherwise, the disease only spreads through a finite fraction of the population. Chain-binomial model is applied to simulate the transmission dynamics of the small-world networks in stage 1. The detailed parameters of the small-world network-based chain-binomial model are listed in Table 2.

2.3.2. Contact network epidemic model for stage 2a. Fig. 2d presents the protected contact pattern, where some of the infected nodes are identified and removed from the network. In Fig. 3 (dashed box), we further improve the regular procedure of chain-binomial model for this scenario, where the isolated nodes and their corresponding edges should be removed from the infected list at the beginning of each loop.

2.3.3. No-contact susceptible and infected model (NCSI) for stage 2b. NCSI is developed to model the epidemic in which there is no contact between individuals and only airborne spread occurs. Assume that all people are in closed spaces with the presence of high-level airborne COVID-19,

$$\begin{cases} \mathrm{d}I/\mathrm{d}t = \delta\sum_{i=0}^{m-1} I_i S, \\ \mathrm{d}S/\mathrm{d}t = -\delta\sum_{i=0}^{m-1} I_i S, \end{cases} \quad (2)$$

where S and I denote the susceptible and infected individuals, I_i is the total number of infected individuals at the ith day before this day, m is the number of days that COVID-19 is viable in the air, and δ is a coefficient bonding the valid infected $\sum_{i=0}^{m-1} I_i$ with the virus concentration.

Table 1
Estimated numbers for modeling virus transmission.

	Unprotected contact	Protected contact	Airborne spread
Days of the epidemic	16 (stage 1)	15 (stage 2a)	15 (stage 2b)
Population	3711	1045	2666
Number of initial infected	1	75 (C_1)	539 (P_1)
Total number of infected individuals	614 ($P_1 + C_1$)	11 (C_{2a})	80 (P_{2b})
Final number of infected individuals staying on ship		10 (R_{2a})	74 (R_{2b})

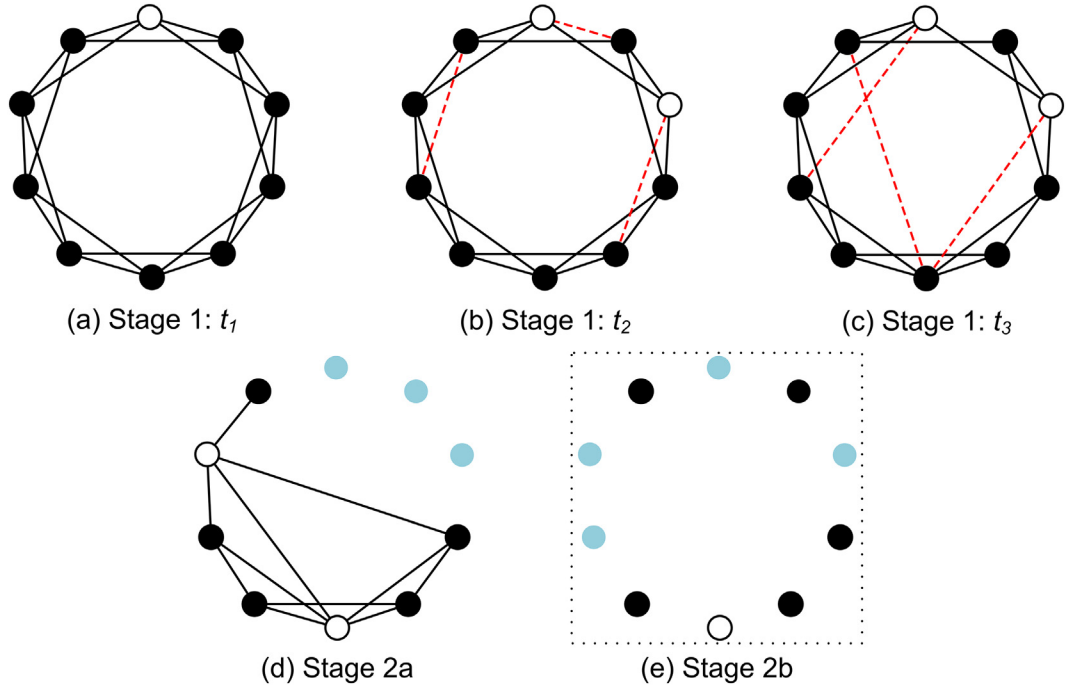

Fig. 2. Brief depiction of the contact network, where white, black and aquamarine nodes present infected, susceptible and isolated individuals, respectively. The dashed red edges represent the changed edges.

Table 2
The major estimated parameters of small-world network-based chain-binomial model and NCSI for computational implementation of MH sampling.

Parameter	Meaning	Sampling ranges		
		Ranges		References
K	Average degree for each node in the network	16–34 (reduced by half in simulations)		Average daily contacts of the first cases is around 25 people from a report
p	Probability of rewiring each edge	0.01–0.5		Small-world networks are regular ($p = 0$) or disordered ($p = 1$) [14]; assume the contact pattern is close to a regular graph
t	Probability of transmission for each contact	Stage 1	0.001–0.2	Trial[a)]
		Stage 2a	0.0001–0.2	t is very small in protected contact; trial [a)]
ip	Infected period	15, 31		The days of stage 2 and stage 1 plus 2 coincide with the official infected period [19]
δ	Coefficient bonding the valid infected and virus concentration	Stage 2b	10^{-8}–10^{-4}	Trial[a)]
m	Days that virus is viable in the air	Stage 2b	3	Virus' stability and identified RNA can last for 72 h [29] and 17 d [30]

[a)] Trial test means the result is from the explorative simulation by using the same data. A simulation will be markedly different from the truth if the parameter is beyond this range.

For the quarantine, we have

$$\begin{cases} \mathrm{d}I/\mathrm{d}t = \delta\sum_{i=0}^{m-1} I_i S - O, \\ \mathrm{d}S/\mathrm{d}t = -\delta\sum_{i=0}^{m-1} I_i S, \end{cases} \tag{3}$$

where O is the number of isolated infected.

2.4. Parameter estimation and quantification of uncertainties

We use Bayesian inference with Monte Carlo Markov Chain (MCMC) sampling algorithms such as Metropolis-Hastings (MH) [31] to discover the stationary parameter estimation for the above epidemic models in a multidimensional probability space. This technique has been widely applied in various fields including ecological modeling [32,33], hydrology [34] and geology [35].

Assume that Ω is the available observation set where the corresponding parameters are unknown; then, the posterior distribution of parameters θ conditioned on Ω is $P(\theta|\Omega) \propto P(\Omega|\theta)P(\theta)$, where $P(\theta)$ is the prior of parameters, and $P(\Omega|\theta)$ is the likelihood function. Generally to directly obtain a full posterior probability distribution $P(\Omega|\theta)$ is intractable, an alternative is to generate approximate posterior samples according to a proposal ordinary distribution (for example, uniform or normal distribution) by MCMC algorithms such as MH. By comparing the differences between the observations and the outputs of the model that receives possible samples of parameters θ, MH algorithm determines which sample can be accepted as a candidate for a parameter.

The detailed procedure of parameter estimation using MH algorithm is described below. The total number of confirmed cases is regarded as the observation and further compared with the model outputs conditioned on the candidate sample of parameters. Accordingly, MH algorithm calculates an acceptance rate to determine which sample can be accepted. We implement an initial test 10^4 times by using a uniform proposal distribution and produce

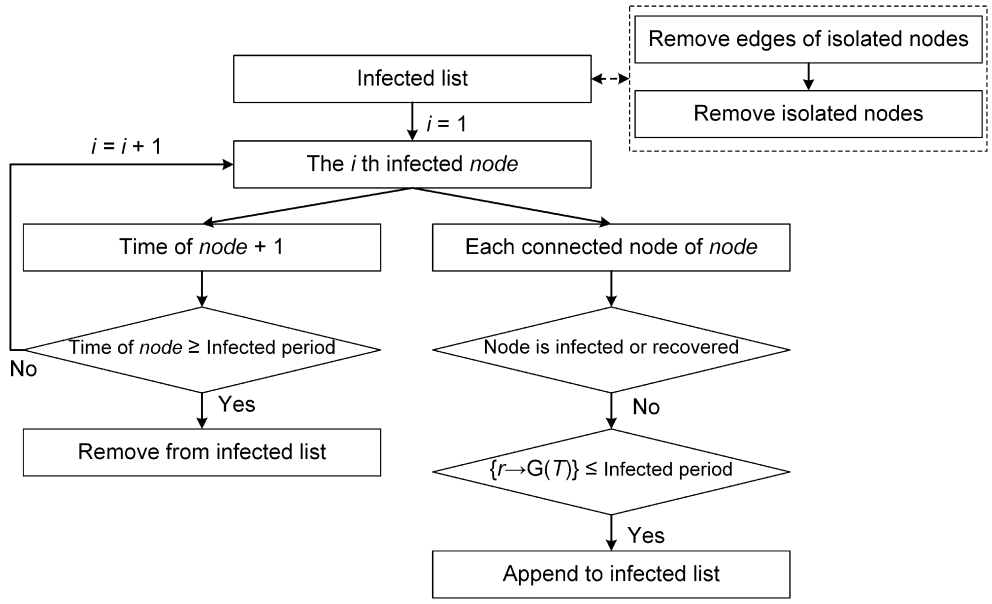

Fig. 3. The flow diagram of chain-binomial model, where $G(t)$ obeys the geometric distribution with the probability of success t and the maximum times that the Bernoulli trial can be implemented, r is a random sampling of $G(t)$.

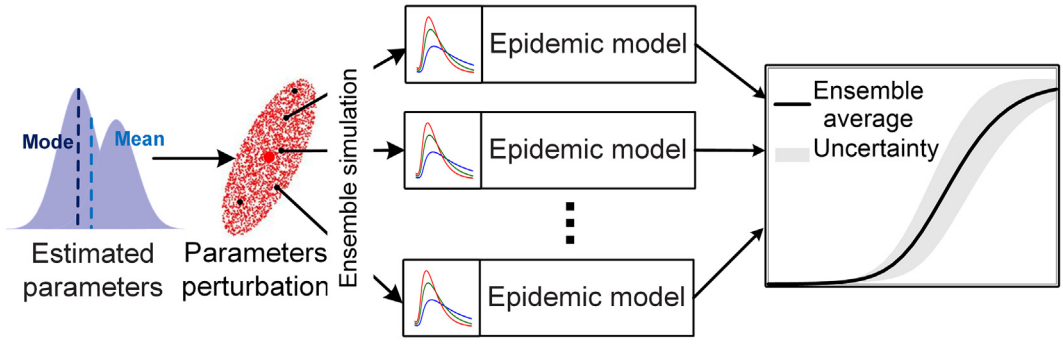

Fig. 4. Flowchart of parameter perturbation and ensemble simulation for epidemic models.

the corresponding covariance matrix cov of the parameter's samples. Generally the accepting rate in the initial test is too low to provide the desired samples. A test based on a normal proposal distribution such as $\theta\, N\left(\theta^{(k)}, cov\right)$ is consequently conducted to improve the accepting rate. The latter test generates 5×10^3 reasonable samples for estimating the posterior probability distribution of parameter θ.

Contact network epidemic simulations perform differently when using the same parameters and initial values. We therefore employ an ensemble simulation technique to make the simulations more reasonable and stable. We implement an ensemble of models that are driven in a parameter field, and produce the results by the ensemble averages. The ensemble simulation provides probabilistic solutions for stochastic nonlinear models and the uncertainty ranges of outputs. In this study, a 10% perturbation of the modes of the samples is introduced as the parameter field, and 100 epidemic models are created to fulfill the ensemble simulation (Fig. 4). Note that the configuration is arbitrary due to the lack of knowledge of ensemble simulation in epidemiology, which highlights the need for analyzing the sensitivity of parameters and initial values in epidemic models.

3. Results

We implement epidemic reconstruction based on a common software for data assimilation development (ComDA [36]). This software is used to fuse the available information into dynamics and produce more reliable predictions or simulations. Typical MCMC techniques such as MH and dynamic models in land-surface/hydrology are integrated into ComDA. The small-world network and chain-binomial model are also introduced by importing the open source library EpiFire [37].

3.1. Parameter estimation

Accordingly the possible statistics of the parameters of both small-world network-based chain-binomial model and NCSI are characterized by the samples, i.e., $\theta = \{K, p, T\}$ and $\theta = \{\delta\}$. All the major parameters and their sampling ranges are listed in Table 2, and the configurations and initial values for epidemic models are provided in Table 1.

The probability distributions of the parameter samples for each stage are shown in Fig. 5. The modes of samples are adopted as the parameters. Compared to the differential equation model such as

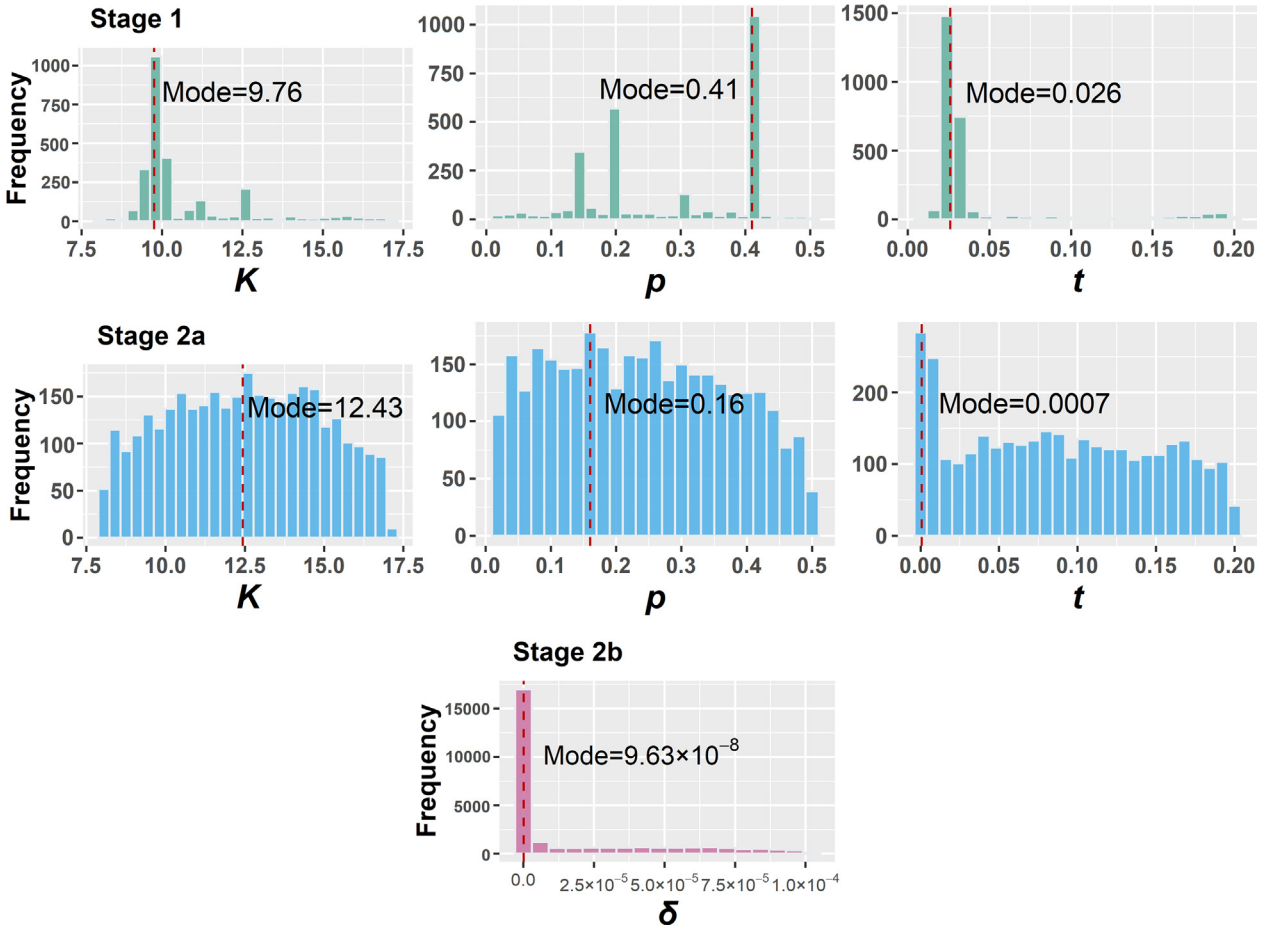

Fig. 5. Histograms of samples from the parameters' posterior distributions for epidemic models on the "Diamond Princess". The vertical dashed lines indicate the modes of samples, which can be regarded as the selected parameters, i.e., (9.76, 0.41, 0.026) for stage 1 (the first row), (12.43, 0.16, 0.0007) for stage 2a (the second row) and 9.63×10^{-8} for stage 2b (the third row).

NCSI, the contact network model is characterized by multiple parameters and strong randomness, which may result in samples being trapped in the suboptimal regions and bias estimation of the parameter. In the order of stage 2a, stage 1 and stage 2b, their population heterogeneities and model randomness gradually weaken, while estimated parameters show an increasing level of statistical significance. Apparently, the highly complex and random network leads to difficulty in parameter estimation in epidemic modeling.

3.2. Simulation results

The daily epidemic curves for each stage are shown in Fig. 6. Considering the strong randomness in parameter estimation, the simulations of unprotected and protected contact network models are inferior to that of airborne spread with the differential equation model, i.e., the ensemble averages of the network-based simulations slightly deviate from the true infected, while NCSI reconstructs the epidemic exactly. This is because, on the one hand, NCSI is not as sensitive as the contact network epidemic model to the parameter, resulting in a more stable model trajectory even when the parameter is perturbed. On the other hand, the estimated parameter for NCSI is very small ($\delta = 9.63 \times 10^{-8}$), implying that an additional 10% perturbation is insufficient to generate the obvious uncertainties.

Scenario-based modeling without any quarantine is also provided (Fig. 6a), which predicts that almost all the people on the ship will become infected in one month. Defining the infected period as 15 and 31, we found that compared with the epidemic curve and average transmissibility t, the basic reproductive number R_0 is more sensitive to the infected period, which is also confirmed in Eq. (1).

4. Discussion

The modeling strategy using parameter perturbation and ensemble simulation facilitates the understanding of strong random network epidemic models by quantifying their uncertainty. Uncertainty ranges imply the potential ceiling (the worst outcome) and floor (the best outcome) for the spread of the disease. Simulations are acceptable if the reference observation falls in the uncertainty ranges (for example, 614 cases in Fig. 6b).

The transmission dynamics of COVID-19 on the Diamond Princess are studied. However, the designed scenarios are simplified because COVID-19 can still transmit via air in stage 1 and 2a. We only consider the major driver because introducing extra transmission makes the simulations more complex, and provides limited knowledge.

A major uncertainty comes from the data, i.e., except for the population and total infected individuals, the available data are mainly deduced from direct or indirect information. Assume that the indirect data has limited reliabilities, considerable errors are consequently introduced into simulations. A promising method for eliminating this type of uncertainty is formulating inference

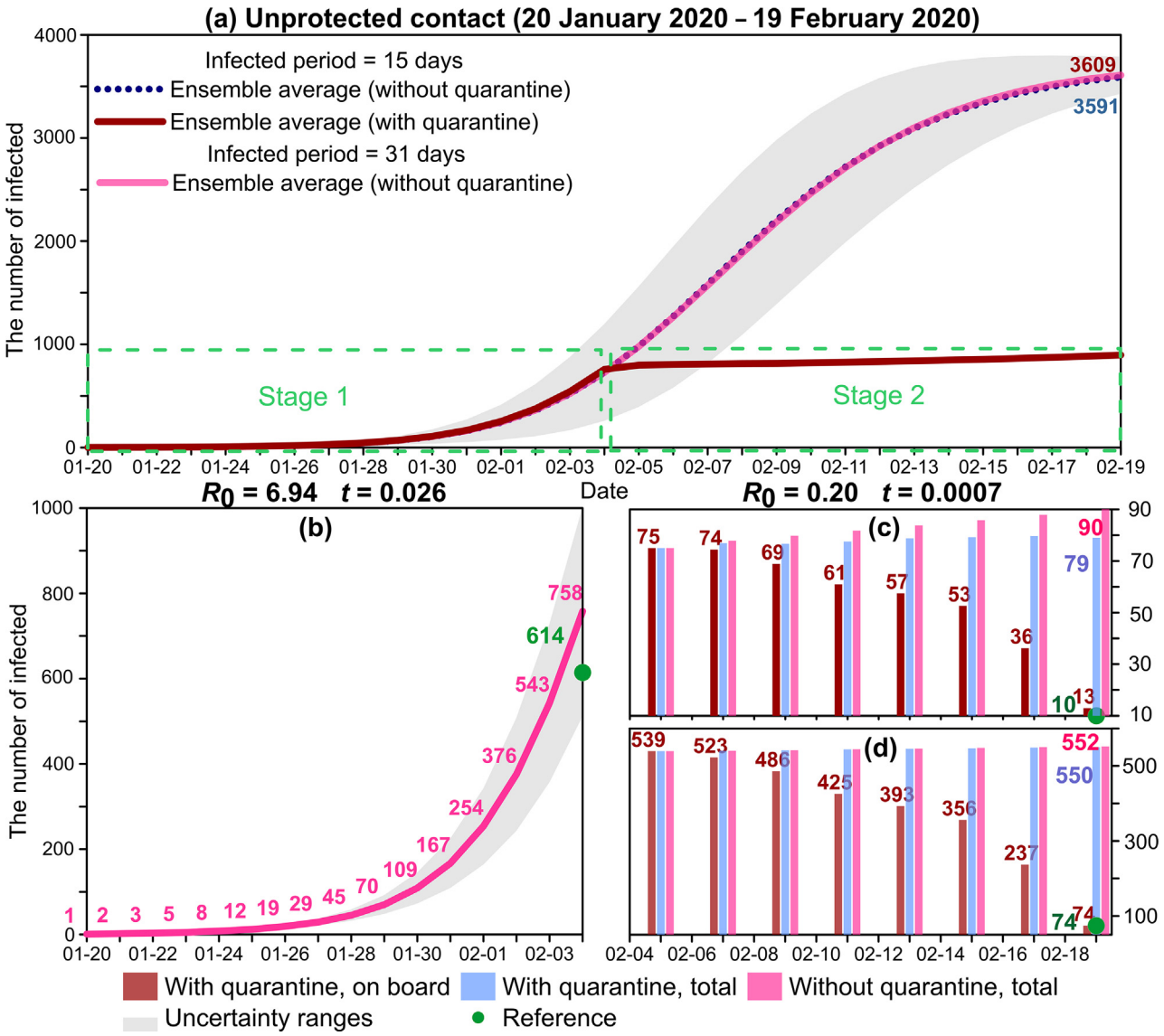

Fig. 6. The daily epidemic curves on the "Diamond Princess". The reference is the final number of infected staying on the ship. (a) Epidemic curves without any quarantine from 20 January to 19 February 2020; (b) stage 1; (c) stage 2a; (d) stage 2b.

processes and mapping indirect information into epidemic modeling, i.e., using data assimilation [38–41].

Quarantine is productive since R_0 dramatically drops from 6.94 for unprotected contact to 0.20 for protected contact. Eq. (1) implies that the average transmissibility t, the infected period ip, and the average degree K have notable impacts on R_0. By providing personal protective equipment and cutting down the close contact among individuals, both t and K are reduced and so is R_0. We further model the unprotected contact scenario for one month, and R_0 ranges from 6.94 to 11.20 when the infected period is from 15 to 31, while t remains at 0.026. This proves that the average transmissibility is more stable, and reasonable judgement can be derived by turning to t.

Based on the simulations and the related quantities, proper containment efforts can be deduced to address the various epidemic scenarios. Compared with protected contact, the airborne spread test is implemented in a larger population. However, its epidemic curves with or without quarantine are closer (data deviations of protected contact and airborne spread tests are 14.46% and 0.36%, respectively). We believe this difference is mainly attributed to incorrect control work, i.e., isolation works well in managing transmission in social contact scenarios; however, when a virus is transmitted via air, isolation may not be sufficient. Inspired by the NCSI Eq. (2), a better approach is to disperse the crowds and/or ventilate the closed spaces, which can dilute the virus concentration in air. It further provides the correct information that air exchanges are vital for the prevention of COVID-19 outbreaks in the closed special settings, such as hospitals, malls, and mass gathering activities.

The other containment efforts are for the intensive social contact scenario. Considering the unprotected and protected contact tests, both R_0 and t decrease when quarantine is implemented, which proves that personal protective materials such as coveralls, masks, and eye covers are useful for infection prevention. We further notice that R_0 in the unprotected contact test is much larger than 2.5, which was provided by the WHO report. This implies that if the intensive unprotected contact pattern is maintained for a couple of days, the epidemic curve may climb dramatically. As shown in the protected contact test, R_0 falls far below 2.5. However, this tiny R_0 does not represent the true basic reproduc-

tive number when using protective materials; we are inclined to assume that protective measures are not fully implemented and some crews were exposed. This highlights that high-risk populations such as medical staff or healthcare workers, should rigorously practice prevention measures during epidemics.

5. Summary

This study reconstructs the epidemic curve on the "Diamond Princess" cruise ship based on two stages (1, 2a and 2b), i.e., unprotected, protected contact and airborne spread of the virus in closed settings. The corresponding epidemic models are mainly based on contact networks, in which parameter estimation is performed with Bayesian inference and MH sampling. Considering that the daily confirmed case reports lag far behind the situation of disease spread, reconstruction is useful for understanding the true transmission dynamics. The results reveal that during early intensive social contacts, R_0 is 6.94, triple that reported by the WHO for China. If no quarantine is implemented, the population will be infected in one month. The decreasing R_0 indicates that following control and prevention measures manage the outbreak. Considering the ongoing outbreaks in many other cruise ships all over the world, this study provides a scientific reference and the practical prevention measures for the similar closed settings.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDA19070104), the National Natural Science Foundation of China (41801270), the 13th Five-year Informatization Plan of Chinese Academy of Sciences (XXH13505-06) and the Foundation for Excellent Youth Scholars of Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences (Y851D41). We thank Prof. Thomas J. Hladish at University of Florida for his help in using EpiFire.

Author contributions

Feng Liu contributed to the manuscript preparation, modelling strategy, data collection and experiment implementation. Xin Li contributed to the conceptualization, study management and manuscript revision. Gaofeng Zhu contributed to parameters estimation and manuscript revision. All authors contributed to experiment design, results analysis and discussion.

Appendix A. Supplementary materials

Supplementary materials to this article can be found online at https://doi.org/10.1016/j.scib.2020.04.043.

References

[1] Dietz K, Heesterbeek JAP. Daniel Bernoulli's epidemiological model revisited. Math Biosci 2002;180:1–21.
[2] Gu C, Zhu J, Sun Y, et al. The inflection point about COVID-19 may have passed. Sci Bull 2020;65:865–7.
[3] Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc London A 1927;115:700–21.
[4] Lipsitch M, Cohen T, Cooper B, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science 2003;300:1966–70.
[5] Riley S, Fraser C, Donnelly CA, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. Science 2003;300:1961–6.
[6] Wang J, McMichael AJ, Meng B, et al. Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. Bull WHO 2006;84:965–8.
[7] Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science 2020. https://doi.org/10.1126/science.abb6105.
[8] Tang B, Wang X, Li Q, et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. J Clin Med 2020;9:462.
[9] Read JM, Bridgen JR, Cummings DA, et al. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. MedRxiv 2020. https://doi.org/10.1101/2020.01.23.20018549.
[10] Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 2020;395:689–97.
[11] Yang Z, Zeng Z, Wang K, et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. J Thorac Dis 2020;12:165–74.
[12] Zhao Z, Li X, Liu F, et al. Prediction of the COVID-19 spread in African countries and implications for prevention and controls: a case study in South Africa, Egypt, Algeria, Nigeria, Senegal and Kenya. Sci Total Environ 2020. https://doi.org/10.1016/j.scitotenv.2020.138959.
[13] Zhao W, Zhang J, Meadows M, et al. A systematic approach is needed to contain COVID-19 globally. Sci Bull 2020;65:876–8.
[14] Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature 1998;393:440–2.
[15] Barabási AL, Albert R. Emergence of scaling in random networks. Science 1999;286:509–12.
[16] Newman ME. Spread of epidemic disease on networks. Phys Rev E 2002;66:016128.
[17] Meyers L. Contact network epidemiology: bond percolation applied to infectious disease prediction and control. Bull Am Math Soc 2007;44:63–86.
[18] Cauchemez S, Nouvellet P, Cori A, et al. Unraveling the drivers of MERS-CoV transmission. Proc Natl Acad Sci USA 2016;113:9081–6.
[19] World Health Organization. Report of the WHO-China joint mission on coronavirus disease 2019 (COVID-19). World Health Organization. https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf. Accessed 4 April 2020.
[20] Huang NE, Qiao F. A data driven time-dependent transmission rate for tracking an epidemic: a case study of 2019-nCoV. Sci Bull 2020;65:425–7.
[21] Yue T, Fan Z, Fan B, et al. A new approach to modeling the fade-out threshold of Coronavirus disease. Sci Bull 2020;65:1225–7.
[22] Ong SWX, Tan YK, Chia PY, et al. Air, surface environmental, and personal protective equipment contamination by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from a symptomatic patient. JAMA 2020. https://doi.org/10.1001/jama.2020.3227.
[23] The Government of Hong Kong. CHP investigates additional case of novel coronavirus infection. https://www.info.gov.hk/gia/general/202002/01/P2020020100795.htm. Accessed 4 April 2020.
[24] National Institute of Infectious Diseases of Japan. Field briefing: diamond princess COVID-19 cases. National Institute of Infectious Diseases of Japan. https://www.niid.go.jp/niid/en/2019-ncov-e/9407-covid-dp-fe-01.html. Accessed 4 April 2020.
[25] Yu IT-S, Qiu H, Tse LA, et al. Severe acute respiratory syndrome beyond Amoy Gardens: completing the incomplete legacy. Clin Infect Dis 2014;58:683–6.
[26] The Government of Hong Kong. WHO Environmental Health Team Reports on Amoy Gardens. https://www.info.gov.hk/gia/general/200305/16/0516114.htm. Accessed 4 April 2020.
[27] The Government of Hong Kong. Hong Kong 2011 Population Census – Summary Results. https://www.statistics.gov.hk/pub/B11200552011XXXXB0100.pdf. Accessed 4 April 2020.
[28] Abbey H. An examination of the Reed-Frost theory of epidemics. Hum Biol 1952;24:201.
[29] van Doremalen N, Bushmaker T, Morris DH, et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. N Engl J Med 2020. https://doi.org/10.1056/NEJMc2004973.
[30] Moriarty LF. Public health responses to COVID-19 outbreaks on cruise ships—worldwide, February–March 2020. Morb Mortal Wkly Rep 2020;69:347–52.
[31] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970;57:97–109.
[32] Xu T, White L, Hui D, et al. Probabilistic inversion of a terrestrial ecosystem model: analysis of uncertainty in parameter estimation and model prediction. Glob Biogeochem Cycles 2006;20. https://doi.org/10.1029/2005GB002468.
[33] Zhu GF, Li X, Su YH, et al. Simultaneous parameterization of the two-source evapotranspiration model by Bayesian approach: application to spring maize in an arid region of northwest China. Geosci Model Dev 2014;7:1467–82.
[34] Zhu G, Li X, Ma J, et al. A new moving strategy for the sequential Monte Carlo approach in optimizing the hydrological model parameters. Adv Water Resour 2018;114:164–79.
[35] Hilley G, Young J. Deducing paleoearthquake timing and recurrence from paleoseismic data, Part I: evaluation of new Bayesian Markov-Chain Monte Carlo simulation methods applied to excavations with continuous peat growth. Bull Seismol Soc Am 2008;98:383–406.
[36] Liu F, Wang L, Li X, et al. ComDA: a common software for nonlinear and non-Gaussian land data assimilation. Environ Modell Softw 2020;127:104638.

[37] Hladish T, Melamud E, Barrera LA, et al. EpiFire: an open source C++ library and application for contact network epidemiology. BMC Bioinf 2012;13:76.
[38] Li X, Liu F, Fang M. Harmonizing models and observations: data assimilation for Earth System Science. Sci China Earth Sci 2020. https://doi.org/10.1007/s11430-019-9620-x.
[39] Shaman J, Karspeck A. Forecasting seasonal outbreaks of influenza. Proc Natl Acad Sci USA 2012;109:20425–30.
[40] Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science 2020. https://doi.org/10.1126/science.abb3221.
[41] Li X. Characterization, controlling and reduction of uncertainties in the modeling and observation of land-surface systems. Sci China Earth Sci 2014;57:80–7.

Feng Liu received his Ph.D. degree in Geographic Information System from University of Chinese Academy of Sciences. He is mainly engaged in interdisciplinary research work on mathematical modeling and land data assimilation, spatial data analysis and remote sensing.

Xin Li is currently a professor at the Institute of Tibetan Plateau Research, Chinese Academy of Sciences. His primary research interests include land data assimilation, the application of remote sensing and GIS in Hydrology and Cryosphere Science, and integrated watershed modeling. He is the lead scientist of the Watershed Allied Telemetry Experimental Research (WATER) and Heihe Watershed Allied Telemetry Experimental Research (HiWATER) Projects, which have been the most comprehensive remote sensing experiment conducted in China in recent years.