

# A Comparison of Modelling the Spatio-temporal Pattern of Disease: A Case Study of Schistosomiasis in Anhui Province, China

**Qing Su**

Fudan University School of Public Health

**Robert Bergquist**

Ingerod Brastad Sweden

**Yongwen Ke**

Schistosomiasis station of prevention and control in Guichi district, Anhui, China

**Jianjun Dai**

Schistosomiasis station of prevention and control in Guichi district, Anhui, Province

**Zonggui He**

Schistosomiasis Station of Prevention and Control in Guichi district, Anhui, China

**Fenghua Gao**

Anhui Provincial Institute of Parasitic diseases, Hefei ra

**Zhijie Zhang**

Fudan University School of Public Health

**Yi Hu** (✉ [huyi@fudan.edu.cn](mailto:huyi@fudan.edu.cn))

Laboratory for Spatial Analysis and Modeling, School of Public Health, Fudan University, Shanghai 200032, China

---

## Research

**Keywords:** Schistosomiasis, Spatio-temporal, Dynamic, China

**Posted Date:** January 13th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-142625/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**A comparison of modelling the spatio-temporal pattern of disease: a case study  
of schistosomiasis in Anhui Province, China**

Qing Su<sup>1,2,3</sup>, Robert Bergquist<sup>4</sup>, Yongwen Ke<sup>5</sup>, Jianjun Dai<sup>5</sup>, Zonggui He<sup>5</sup>, Fenghua  
Gao<sup>6</sup>, Zhijie Zhang<sup>1,2,3\*</sup>, Yi Hu<sup>1,2,3\*</sup>

<sup>1</sup> Department of Epidemiology and Biostatistics, School of Public Health, Fudan  
University, Shanghai 200032, China

<sup>2</sup> Key Laboratory of Public Health Safety, Ministry of Education, Shanghai 200032,  
China

<sup>3</sup> Laboratory for Spatial Analysis and Modeling, School of Public Health, Fudan  
University, Shanghai 200032, China

<sup>4</sup> Ingerod, Brastad, Sweden

<sup>5</sup> Schistosomiasis Station of Prevention and Control in Guichi Distirct, Anhui  
Province, China

<sup>6</sup> Anhui Provincial Institute of Parasitic Diseases, Hefei, China;

\* Corresponding author: Yi Hu ([huyi@fudan.edu.cn](mailto:huyi@fudan.edu.cn)), Zhijie Zhang  
([epistat@gmail.com](mailto:epistat@gmail.com))

## Abstract

**Background:** Various methods have been proposed in modelling spatio-temporal pattern of diseases in recent years. The construction of the spatio-temporal models can be either descriptive methods or dynamic. While the former mainly explores correlations by functions, the latter also depicts the process of transmission quantitatively. In this study, we aim to evaluate the differences in model fitting between a descriptive, spatio-temporal model and dynamic spatio-temporal model of schistosomiasis transmission in Guichi, Anhui Province, China.

**Methods:** The parasitological and environmental data at the village level from 1991 to 2014 were obtained by cross-sectional survey. The space-time changes of schistosomiasis risk were explored by two different spatio-temporal models, the Fixed Rank Kriging (FRK) model and the Integral-Difference Equation (IDE) model, and the performance of the two models in fitting schistosomiasis risk were compared.

**Results:** In both models, the average daily precipitation and the normalized difference vegetation index (NDVI) are significantly positively associated with schistosomiasis prevalence while distance to water bodies, the hours of daylight and day land surface temperature ( $LST_{day}$ ) are significantly negatively associated. The overall root mean squared prediction error (MSPE) of the Integral-Difference Equation (IDE) and the FRK model were  $0.35e-02$  and  $0.63e-02$ , respectively, and the correlation between the predicted and observed values of the IDE model (0.71) ( $p < 0.01$ ) was larger than the

FRK one (0.51) ( $p=0.02$ ). Our results also showed that the prediction error of the IDE model is lower than that of FRK model.

**Conclusions:** In capturing schistosomiasis risk combined with environment factors, the IDE model fits better than the descriptive FRK one. As a dynamic spatio-temporal method, the IDE approach models the process of disease transmission quantitatively and therefore provides more accurate predictions, which is of great importance with reference to schistosomiasis and other diseases.

**Key words:** Schistosomiasis; Spatio-temporal; Dynamic; China.

## Introduction

With the progress of computer technology and geographic information systems (GIS), the popularity of geographic modelling contributes to the assessment of the risk patterns of diseases in a growing number of studies. In addition, full awareness of the spatio-temporal dynamic process of infectious diseases strongly influences prevention and control in general[1]. In principle, spatio-temporal models are either constructed by a descriptive approach or a dynamic one[2]. The former kind uses mean function and covariance to analyse the temporal and spatial variations of the variables of interest[3, 4], while the latter describes the dynamic evolution quantitatively under different space-time states, which can easily be simulated and explained[2]. Descriptive methods are widely used in spatio-temporal modelling studies[5, 6, 7, 8, 9]

to capture the statistical correlation of space-time phenomena. Examples include Kriging[10, 11, 12], Hierarchical Bayesian[13, 14, 15] and regression models[16, 17]. Dynamic methods, on the other hand, have been applied in various fields, such as environmental science[18], public health[19], etc., to explore the basic process of spatio-temporal transmission dynamics. However, there is a paucity of comparative studies on the performance of the descriptive spatio-temporal models and dynamic spatio-temporal models in practical applications.

Generally, when exploring the spatio-temporal variations of diseases, we are likely to select descriptive, spatio-temporal models in the absence of kinetic knowledge. Its main idea is to specify the dependence structure in the random process[2], which fails to capture the evolution of the dependent random process through time quantitatively. By describing the mechanical process via scientific knowledge and probability distributions, the dynamic space-time model quantifies the evolution of current and future spatial fields from past fields[2], which is more suitable for prediction in the complex dynamic process[18]. In addition, there is evidence that dynamic models are well suited for modelling dynamic processes[20, 21, 22] with high predictive accuracy. These models quantify the dependencies of the underlying dynamic processes of disease transmission. The process of disease transmission may be influenced by many unknown and uncertain factors, especially in complex health environment. Elements contributing to this complexity include human behaviour, pathogen evolution, environmental change and socio-economic status, etc[23, 24, 25, 26]. Hence, dynamic

spatio-temporal models are considered to have the potential to explore the apparently intractable complexities of disease dynamics[24].

In our study, we employed a descriptive method (Fixed Rank Kriging - FRK)[2] and a dynamic method (Integral-Difference Equation - IDE)[2] to model the spatio-temporal pattern of risk due to schistosomiasis japonica in Guichi, Anhui Province, China. The results of the models were compared, which could provide better options for spatio-temporal modelling in further studies.

## **Materials and methods**

### **Study area**

Guichi, a district situated in the eastern part of Anhui Province, China, covers an area of about 2,500 km<sup>2</sup>, with a population of approximately 0.67 million (2018) (<http://tjj.ah.gov.cn/oldfiles/tjj/tjjweb/tjnj/2019/cn.html>). It's located in the middle and lower reaches of the Yangtze River close to Qiupu, another major river that runs across the region (Figure 1). The presence of the two rivers and the humid subtropical monsoon climate provide a favourable environment, on the one hand for the survival and breeding of the schistosomiasis intermediate snail host[27] and, on the other, for the sustained parasite life cycle that involves humans and other mammals as definitive hosts.

*Figure 1 about here*

### **Parasitological data**

The prevalence of *Schistosoma japonicum* infection from 1991 to 2014 were collected by cross-sectional surveys each year, which were through the organization of Health professionals in Anhui Institute of Parasitic Diseases (AIPD). The data were obtained by annual village-based area surveys including residents aged 5 to 65 years based on a two-pronged diagnostic approach: First, all participants were screened by the Indirect Hemagglutination Assay (IHA)[28] test for anti-schistosomiasis antibodies and then the Kato-Katz technique[29] was applied using three thick smears from stool samples of all seropositive individuals. The number of sample villages ranging from 60 to 132 in our study, and we included those with more than 100 participants .

### **Environmental data**

The environmental factors studied included hours of daylight, distance to water bodies, average daily precipitation, the normalized difference vegetation index (NDVI), and land surface temperature at daytime (LST<sub>day</sub>). Data on hours of daylight and average daily precipitation were acquired from the China Meteorological Data Sharing Service System (<http://cdc.cma.gov.cn/home.do>). Inverse distance weighting (IDW) interpolation was applied to analyse the meteorological data available at 610 weather stations nationwide by ArcGIS 10.4 (ESRI, Redlands, CA, USA), then selecting the data within the study area. NDVI and LST<sub>day</sub> were downloaded from the Level 1 and Atmosphere Archive and Distribution System (<http://ladsweb.nascom.nasa.gov>) with all 8-day global 1-km products for LST<sub>day</sub> and monthly global 1-km products for NDVI. Water-body data were extracted from the World Wildlife Fund's Conservation

Science Data Sets (<http://worldwildlife.org>). Euclidean distance from each village to water bodies were calculated using ArcGIS 10.4.

### **Statistical analysis**

A Box-Cox transformation[30] of the crude prevalence was performed to apply the Gaussian model before implementing the spatio-temporal model. For covariate selection, univariate analysis was used for preliminary screening with the exclusion criteria of  $p>0.1$ . A multicollinearity analysis was then conducted among the remaining variables. Correlation coefficients $>0.6$  were considered collinearity, and we eliminated meaningless variables by the backward stepwise regression method. Finally, we tested the spatial autocorrelation of the residual of the multiple regression (non-spatial model) via semi-variogram[31]. Lack of spatial autocorrelation with the residuals would indicate spatiotemporal variability of the variables in the non-spatial model; otherwise, spatio-temporal models considering spatial autocorrelation were employed.

### ***The FRK model***

The FRK model was used to explore the spatio-temporal pattern of schistosomiasis and to evaluate the influence of environmental factors. We assumed a two-stage model based on the idea of hierarchical model (HM), which was specified as a data model with the conditional probability distribution given by a latent process and some parameters, and a process model that described the process and its uncertainties given



by some parameters.

1. The data model: let  $Z(s; t)$  be the observed value of schistosomiasis in village  $s$  at year  $t$ , which considered an additive measurement error. The data model used was expressed as:

$$Z(s; t) = Y(s; t) + \varepsilon(s; t) \quad (1)$$

where  $Y(s; t)$  is the latent process, and  $\varepsilon(s; t)$  is the observation errors independent of  $Y(s; t)$ , which is assumed to be the independent identically distributed (iid) mean-zero measurement error; and

2. the process model was expressed by the equation:

$$Y(s; t) = x(s; t)' \beta + \eta(s; t) \quad (2)$$

where  $x(s; t)$  is a fixed term indicating the covariant environmental factors, and  $\eta(s; t)$  denotes a mean-zero random process that is dependent on space-time. Random effects can be divided into two parts, as follows:

$$\eta(s; t) = \sum_{k=0}^n \phi_i(s; t) \alpha_i + v(s; t) \quad (3)$$

where  $\{\phi_i(s; t), i = 1, \dots, \alpha_i\}$ , corresponding to location  $(s; t)$ , specifies a spatiotemporal basis function,  $\{\alpha_i\}$ , which are random effects explaining variations in schistosomiasis risk, and  $v(s; t)$  represents the small-scale spatiotemporal random effects not captured by the basis function.

### *The IDE model*

To capture the spatiotemporal dynamic evolution of schistosomiasis in the villages, we also constructed a dynamic model based on HM, as follows:

1. The data model:

$$Z_t = b_t + H_t Y_t + \varepsilon_t, \quad t = 1, \dots, T \quad (4)$$

where  $Z_t$  is the data vector, including arbitrary position at time  $t$ ;  $b_t$  an additive offset term that describes the main trend of  $Z_t$  corresponding to non-dynamic spatio-temporal structure;  $H_t$  the observation mapping matrix which relates the process to the observations gained by mapping;  $Y_t$  the latent dynamic spatio-temporal process; and  $\varepsilon_t$  an additive error process that is time varying (but statically dependent in time, assumed to be a mean-zero, Gaussian process); and

2. the process model, an approach represented by a first-order, integral-difference equation in a continuous spatial field. It is generally assumed that the process satisfies a Markovian spatio-temporal process, which indicates that its value at the current, given position is composed of two parts: a weighted combination of processes in the entire space domain at the previous time, and an additive, Gaussian and spatially coherent "innovation" expressed as:

$$Y_t(s) = \int_{D_s} m(s, x; \theta_p) Y_{t-1}(x) dx + \eta_t(s), \quad s, x \in D_s, \quad t = 1, 2, \dots, \quad (5)$$

where  $m(s, x; \theta_p)$  is a transition kernel, depending on the parameters  $\theta_p$ ; and  $\eta_t(s)$  a mean-zero Gaussian spatial process independent of  $Y_{t-1}(\cdot)$  that changes over time but is statistically independent. In a one-dimensional spatial domain,  $m(s, x; \theta_p)$  is assumed to be a Gaussian-shape kernel as a function of  $x$  relative to the location  $s$ :

$$m(s, x; \theta_p) = \theta_{p,1}(s) \exp\left(-\frac{1}{\theta_{p,2}(s)} [(x_1 - \theta_{p,3}(s) - s_1)^2 + (x_2 - \theta_{p,4}(s) - s_2)^2]\right) \quad (6)$$

where the coefficients of diffusion are  $\theta_{p,1}$  (the spatially varying amplitude) and  $\theta_{p,2}$  (the spatially varying kernel aperture); and the mean (shift) parameters  $\theta_{p,3}$  and  $\theta_{p,4}$  correspond to a shift of kernel relative to location  $s$ . Here we specified a spatially invariant kernel. All the parameters were estimated using the expectation-maximization (EM) algorithm and the inference for the latent process  $Y_t$  was implemented using Kalman filtering and smoothing[32]. The parametric significance ( $p$ -value) was calculated using a permutation approach with 99 random simulations of the observed prevalence.

## Model validation

The accuracy of the models determines the quality of the prediction results. A 10-fold cross-validation was implemented to compare the performance of the FRK and IDE models. As validation data, 10% of the sample was randomly selected, with the remaining 90% used as training set. The root mean square error (MSPE) is given by

the equation:

$$MSPE = \frac{1}{Tm} \sum_{j=1}^T \sum_i^m \{Z_v(s_i; t_j) - \widehat{Z}_v(s_i; t_j)\}^2 \quad (7)$$

where  $\{Z_v(s_i; t_j)\}$  are observations of 10% spatio-temporal validation samples; and  $\{\widehat{Z}_v(s_i; t_j)\}$  predictions based on the rest of the observations.

We also investigated the correlation between predicted and observed values and producing the correlation scatter plots of the two fitting models. All statistical analyses were performed using R3.6.3 software (R Development Core Team, Vienna, Austria - <http://cran.r-project.org>). The study area was divided at the 1-km resolution level to obtain the above environmental variables, and then to make predictions.

## Results

The median annual prevalence of schistosomiasis decreased from 0.037 in 1991 to 0.000 in 1998; however, a recurrent sharp upward move started in 2005 reaching a second lower top at 0.023 in 2007 after which a gradual decline followed over the next few years (Figure 2). The interquartile range (IQR) was 0.008-0.033 in 1991, decreased to 0 by 1998, increased again reaching 0.011-0.028 in 2007 to finally shrink back to 0-0.003 in 2014.

### *Figures 2 and 3 about here*

Figure 3 shows that the residual semi-variogram of the regression model increases gradually with the increase of time and distance indicating a spatial autocorrelation of

the residuals, with the spatio-temporal model outperforming the non-spatial one. The estimates for all environmental factors are presented in Table 1. In both models, FRK and IDE, the five covariates were all statistically significant, with the average daily precipitation ( $p < 0.01$ ) and NDVI ( $p \leq 0.02$ ) positively associated with schistosomiasis risk, and distance to water bodies ( $p \leq 0.02$ ), hours of daylight ( $p < 0.01$ ) and LST<sub>day</sub> ( $p < 0.01$ ) negatively associated.

***Table 1 about here***

The results of cross validation of the two models are shown in Figure 4. The overall MSPE of the IDE model and the FRK model were 0.35e-02 and 0.63e-02, respectively. In addition, the MSPEs of the IDE model was lower than that of FRK model in most years. Pearson correlation analysis showed that the correlation between the predicted values and the observed values of the IDE model ( $r = 0.71$ ) ( $p < 0.01$ ) was larger than that of the FRK model ( $r = 0.51$ ) ( $p = 0.02$ ).

Predictions of annual prevalence by the IDE model are presented in Figure 5. From 1991 to 2001, the disease indicated a generally decreasing trend; however, with an occasional relatively high rise of prevalence ( $> 0.14$ , shown as strong shades of yellow and red in the figure) but remained relatively stable state from 2006 to 2009 to finally gradually decline (shown as large blue patches in the figure). The FRK model showed a similar pattern except the higher values (more yellow parts). Figure 6 presents the

corresponding standard errors of the predictions. The uncertainty of the IDE model (right) was higher at places far from the sample villages during our study period. Note that the standard errors after 2005 were slightly lower than previous years. The FRK model in Figure 6 (left) indicates that the standard errors were higher than that shown by the IDE model.

## Discussion

This study compared the performance of two spatio-temporal models, FRK and IDE, by fitting to the prevalence of schistosomiasis in Guichi, China. We found that the dynamic approach is more accurate in terms of data fitting when exploring the evolution pattern of schistosomiasis. The IDE model had smaller MSPEs and a higher correlation between predictions and observations. Thus, the spatio-temporal patterns of disease dynamics are better captured by the IDE model, which should help to identify accurate areas for future schistosomiasis control and surveillance.

Our study shows that both models investigated fit better than non-spatial models, with the IDE model considered optimal. The process of disease transmission can be affected by many uncertainties, such as pathogen variation and evolution, climate change, variation of intermediate and definitive hosts and not the least active control activities, as coordinated interventions across times and locations would influence the patterns of disease transmission strongly[11]. Therefore, quantitative analysis is crucial to explore the complex mechanism of space-time dynamics[33]. In our study,

the dynamic model predicts the disease more accurately as it is committed to quantify the dynamic changes of disease risk infection by kinetic equations, while the descriptive model is only based on functions that describe the distribution of the disease. This gives dynamic models advantages in predicting the likelihood of disease reoccurrence, particularly when the mechanisms governing its spreads are understood so that effective measures can be instituted at the right time[19]. In assessing the association of schistosomiasis infection with covariates, the results of the two spatial models were similar (Table1). The environmental variables here have been investigated in previous studies[8, 9, 10, 11] and are interpretable with respect to the epidemiology of schistosomiasis and snail biology.

The prevalence of schistosomiasis showed different distribution patterns during the study period. Figure 5 shows that the high-density incidence areas are mainly located near the rivers (parts of the central and northern regions), indicating that these regions should remain foci for active schistosomiasis control[34]. The goal of schistosomiasis control is to prevent infection by interrupting the life cycle of the parasites, e.g. by eliminating the intermediate snail hosts[35]. It is clear that the identification of high-risk areas for schistosomiasis is conducive to the implementation of effective integrated control measures, such as environmental modification, snail elimination and health education[36, 37]. Indeed, the national schistosomiasis control strategies (interrupting transmission, controlling the source of infection, chemotherapy of the definitive hosts, etc.) implemented at different stages in Guichi[38] have effectively

controlled the infection of schistosomiasis and kept it at a stable low level after 2005.

The limitations of this study also need to be elucidated. First, we considered only environmental factors, and some social factors associated with schistosomiasis should be considered, such as medical conditions and the general economy as expressed by the gross domestic product (GDP). Hence, more comprehensive data collection for research should be considered in the future. Second, since it is affected by many uncertain factors (human behaviour, pathogen variation, environmental change, etc.), the process of disease transmission is extremely complex and changing, which may not be adequately captured by linear models. Non-linear models will therefore be needed for the exploration of the more dynamic disease patterns in future studies.

## **Conclusions**

Our study constructed two spatio-temporal models (the FRK model and the IDE model) to compare the differences in capturing changing pattern of schistosomiasis risk. The former used a descriptive approach and the latter a dynamic one. The results show that the IDE model outperforms the FRK model. Dynamic models should be given more consideration for spatio-temporal modelling of disease patterns in the future.

## **Ethics approval and consent to participate**

The Ethics Committee of Fudan University approved the oral consent and other



aspects of the investigation (IRB# 2017-09-06.35). Written informed consent was also obtained from adult participants (18 years old). Parental consent was sought for participants ages 5 to 17. To protect patient privacy, all personal information about participants was anonymized by removing personal identifiers such as name, parent name, home address, and phone number.

### **Consent for publication**

Not applicable

### **Availability of data and materials**

The datasets used and analysed during the current study are available from the corresponding author on reasonable request and in the website repository (<http://cdc.cma.gov.cn/home.do>, <http://ladsweb.nascom.nasa.gov>, <http://worldwildlife.org>).

### **Competing Interests**

We declare that we have no conflicts of interest.

### **Funding**

This research was supported by the National Natural Science Foundation of China (81773487 to YH, 81673239 to ZZ, 81973102 to ZZ) and Public Health Talents Training Program of Shanghai Municipality (GWV-10.2-XD21 to ZZ).

355

356 Authors' contributions

357 QS analysed and interpreted the data, QS RB YH ZZ wrote the paper, YK JD and ZH  
358 contributed reagents/materials/analysis tools.

359

## 360 **Acknowledgements**

361 We thank the staff at the Anti-Schistosomiasis Station of Guichi for their kind  
362 collaboration and the reviewers for their comments and suggestions.

363

364

365

366

## References

1. Chowell G, Rothenberg R: Spatial infectious disease epidemiology: on the cusp. *BioMed Central*; 2018.
2. Wikle CK, Zammit-Mangion A, Cressie N. *Spatio-temporal Statistics with R*. CRC Press, Boca Raton, FL; 2019: Chapter 5.
3. Gneiting T, Genton MG, Guttorp P. Geostatistical space-time models, stationarity, separability, and full symmetry. *Monographs On Statistics and Applied Probability*. 2006;107:151.
4. Montero J-M, Fernández-Avilés G, Mateu J. Spatial and spatio-temporal geostatistical modeling and kriging. vol. 998: John Wiley & Sons; 2015.
5. Diggle PJ, Tawn JA, Moyeed RA. Model -based geostatistics. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*. 1998;47 3:299-350.
6. Hay S, Graham A, Rogers D: *Global Mapping of Infectious Diseases: Methods, Examples and Emerging Applications*. 2008.
7. Zhang Z, Carpenter T, Lynn H, Chen Y, Bivand R, Clark A, et al. Location of active transmission sites of *Schistosoma japonicum* in lake and marshland regions in China. *Parasitology*. 2009;136 7:737.
8. Hu Y, Xiong C-L, Zhang Z-J, Bergquist R, Wang Z-L, Gao J, et al. Comparison of data-fitting models for schistosomiasis: a case study in Xingzi, China. *Geospatial health*. 2013:125-32.
9. Ekpo UF, Hu`rlimann E, Schur N, Oluwole AS, Abe EM, Mafe MA, et al. Mapping and prediction of schistosomiasis in Nigeria using compiled survey data and Bayesian geospatial modelling. *Geospatial health: health applications in geospatial science*. 2013;7:355-66.
10. Hu Y, Gao J, Chi M, Luo C, Lynn H, Sun L, et al. Spatio-temporal patterns of schistosomiasis japonica in lake and marshland areas in China: the effect of snail habitats. *The American journal of tropical medicine and hygiene*. 2014;91 3:547-54.
11. Hu Y, Li R, Bergquist R, Lynn H, Gao F, Wang Q, et al. Spatio-temporal transmission and environmental determinants of schistosomiasis japonica in Anhui Province, China. *PLoS Negl Trop Dis*. 2015;9 2:e0003470.
12. Gething PW, Noor AM, Gikandi PW, Ogara EA, Hay SI, Nixon MS, et al. Improving imperfect data from health management information systems in Africa using space-time geostatistics. *PLoS Med*. 2006;3 6:e271; doi: 10.1371/journal.pmed.0030271.
13. Hu Y, Ward MP, Xia C, Li R, Sun L, Lynn H, et al. Monitoring schistosomiasis risk in East China over space and time using a Bayesian hierarchical modeling approach. *Scientific reports*. 2016;6 1:1-9.
14. Aswi A, Cramb SM, Moraga P, Mengersen K. Bayesian spatial and spatio-temporal approaches to modelling dengue fever: a systematic review. *Epidemiol Infect*. 2018;147:1-14; doi: 10.1017/s0950268818002807.
15. Zhang B, Cressie N. Bayesian Inference of Spatio-Temporal Changes of Arctic Sea Ice. *Bayesian Analysis*. 2020.

- 410 16. Zhao Y, Ge L, Zhou Y, Sun Z, Zheng E, Wang X, et al. A new Seasonal Difference  
411 Space-Time Autoregressive Integrated Moving Average (SD-STARIMA) model and  
412 spatiotemporal trend prediction analysis for Hemorrhagic Fever with Renal  
413 Syndrome (HFRS). *PloS one*. 2018;13 11:e0207518.
- 414 17. Shim J, Hwang C. Kernel-based geographically and temporally weighted  
415 autoregressive model for house price estimation. *PLoS One*. 2018;13 10:e0205063;  
416 doi: 10.1371/journal.pone.0205063.
- 417 18. Zammit-Mangion A, Wikle CK. Deep integro-difference equation models for  
418 spatio-temporal forecasting. *Spatial Statistics*. 2020;100408.
- 419 19. Woolhouse M. How to make predictions about future infectious disease risks.  
420 *Philos Trans R Soc Lond B Biol Sci*. 2011;366 1573:2045-54; doi:  
421 10.1098/rstb.2010.0387.
- 422 20. Calder CA, Berrett C, Shi T, Xiao N, Munroe DK. Modeling space - time dynamics  
423 of aerosols using satellite data and atmospheric transport model output. *Journal*  
424 *of agricultural, biological, and environmental statistics*. 2011;16 4:495-512.
- 425 21. Kot M, Schaffer WM. Discrete-time growth-dispersal models. *Mathematical*  
426 *Biosciences*. 1986;80 1:109-36.
- 427 22. Xu K, Wikle CK, Fox NI. A kernel-based spatio-temporal dynamical model for  
428 nowcasting weather radar reflectivities. *Journal of the American statistical*  
429 *Association*. 2005;100 472:1133-44.
- 430 23. Lessler J, Azman AS, Grabowski MK, Salje H, Rodriguez-Barraquer I. Trends in  
431 the Mechanistic and Dynamic Modeling of Infectious Diseases. *Curr Epidemiol Rep*.  
432 2016;3 3:212-22; doi: 10.1007/s40471-016-0078-4.
- 433 24. Heesterbeek H, Anderson RM, Andreasen V, Bansal S, De Angelis D, Dye C, et al.  
434 Modeling infectious disease dynamics in the complex landscape of global health.  
435 *Science*. 2015;347 6227:aaa4339; doi: 10.1126/science.aaa4339.  
436 <http://science.sciencemag.org/content/347/6227/aaa4339.abstract>.
- 437 25. Dye C. After 2015: infectious diseases in a new era of health and development.  
438 *Philos Trans R Soc Lond B Biol Sci*. 2014;369 1645:20130426; doi:  
439 10.1098/rstb.2013.0426.
- 440 26. Bloom DE, Cadarette D. Infectious Disease Threats in the Twenty-First Century:  
441 Strengthening the Global Response. *Front Immunol*. 2019;10:549; doi:  
442 10.3389/fimmu.2019.00549.
- 443 27. Zhang Z, Carpenter TE, Chen Y, Clark AB, Lynn HS, Peng W, et al. Identifying  
444 high-risk regions for schistosomiasis in Guichi, China: a spatial analysis. *Acta*  
445 *Trop*. 2008;107 3:217-23; doi: 10.1016/j.actatropica.2008.04.027.
- 446 28. Zhao WX. Practical schistosomiasis japonica (1st edition). People's Health Press,  
447 Beijing. 1997:97-115.
- 448 29. Yu JM, de Vlas SJ, Jiang QW, Gryseels B. Comparison of the Kato-Katz technique,  
449 hatching test and indirect hemagglutination assay (IHA) for the diagnosis of  
450 *Schistosoma japonicum* infection in China. *Parasitol Int*. 2007;56 1:45-9; doi:  
451 10.1016/j.parint.2006.11.002.
- 452 30. Box GE, Cox DR. An analysis of transformations. *Journal of the Royal Statistical*  
453 *Society: Series B (Methodological)*. 1964;26 2:211-43.

31. Cressie NAC. statistics for spatial data. 1993; doi: 10.1002/9781119115151.
32. Kalman RE. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering. 1960;82 1:35-45; doi: 10.1115/1.3662552. <https://doi.org/10.1115/1.3662552>.
33. Camara B. Dynamics omplexity of diffusive predator-prey models and applications. 2009.
34. Zhang ZJ, Carpenter TE, Lynn HS, Chen Y, Bivand R, Clark AB, et al. Location of active transmission sites of Schistosoma japonicum in lake and marshland regions in China. Parasitology. 2009;136 7:737-46; doi: 10.1017/S0031182009005885. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-67651146576&doi=10.1017%2fS0031182009005885&partnerID=40&md5=ff3056c0ff25f7e02d0f55f718459627>.
35. Seto EY, Remais JV, Carlton EJ, Wang S, Liang S, Brindley PJ, et al. Toward sustainable and comprehensive control of schistosomiasis in China: lessons from Sichuan. PLoS Negl Trop Dis. 2011;5 10:e1372; doi: 10.1371/journal.pntd.0001372.
36. Li SZ, Luz A, Wang XH, Xu LL, Wang Q, Qian YJ, et al. Schistosomiasis in China: acute infections during 2005-2008. Chin Med J (Engl). 2009;122 9:1009-14.
37. Zhou YB, Liang S, Chen GX, Rea C, He ZG, Zhang ZJ, et al. An integrated strategy for transmission control of Schistosoma japonicum in a marshland area of China: findings from a five-year longitudinal survey and mathematical modeling. Am J Trop Med Hyg. 2011;85 1:83-8; doi: 10.4269/ajtmh.2011.10-0574.
38. Hu Y, Li S, Xia C, Chen Y, Lynn H, Zhang T, et al. Assessment of the national schistosomiasis control program in a typical region along the Yangtze River, China. Int J Parasitol. 2017;47 1:21-9; doi: 10.1016/j.ijpara.2016.09.003.

**Figure legend**

**Figure 1. Endemic area of schistosomiasis in Guichi, Anhui Province, China.**

Guichi in Anhui province situates in the middle-lower reaches of the Yangtze River in eastern China, and the Qiupu River is in the study area.

**Figure 2. Box plot for the observed prevalence of schistosomiasis for sample**

**villages in Guichi, Anhui, China, from 1991 to 2014.** The blue line represents the average annual prevalence, and the red point is median annual prevalence, the boxes denote minimum, median, maximum, and interquartile ranges.

**Figure 3. Empirical spatio-temporal semi-variogram of the residuals in regression model.**

**Figure 4. MSPE of the FRK predictions (red), the IDE predictions (blue) and the non-spatial model(yellow) as a function of time.**

**Figure 5. Plots for the annual prevalence of schistosomiasis predicted from 1991**

**to 2014 in Guichi, Anhui Province, China.** The left is the forecast graph generated by FRK model, and the right is by IDE model.

**Figure 6. Plots for the annual standard error of predicted prevalence of**

**schistosomiasis in Guichi, Anhui Province, China, from 1991 to 2014.** The left is  
the forecast graph generated by FRK model, and the right is by IDE model.

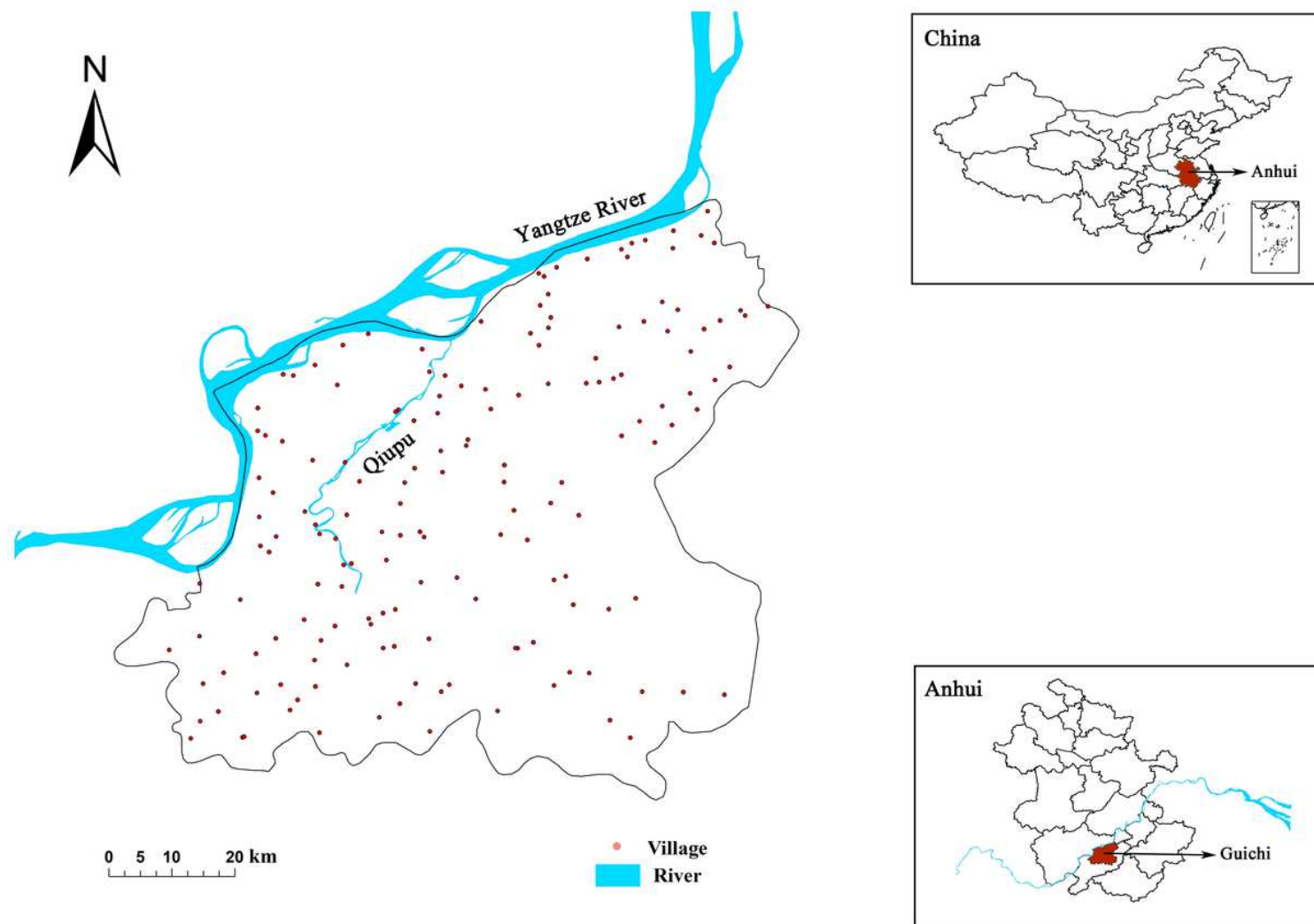
# Table

Table 1 Environmental factors of schistosomiasis japonica: the spatio-temporal FRK model and the spatio-temporal IDE model.

Parameters	spatio-temporal FRK				spatio-temporal IDE			
	Estimate	Standard error	<i>z</i>	<i>p</i>	Estimate	Standard error	<i>z</i>	<i>p</i>
Intercept	-2.22E-03	20.14	-24.41	<0.01	-2.22E-03	7.21	13.31	<0.01
Average daily precipitation	5.33E-04	2.32E-03	-3.30	<0.01	0.01	1.90E-06	-6.32	<0.01
Distance to water bodies	-0.68	3.03E-01	-2.25	0.02	-0.24	2.49E-05	5.35	<0.01
Hours of daylight	-6.34E-05	2.03E-05	-3.13	<0.01	-6.74E-06	3.59E-03	3.46	<0.01
NDVI	2.59E-05	1.90E-05	3.37	<0.01	2.59E-05	5.32	2.47	0.02
LST at day	-6.34E-05	2.57E-05	3.63	<0.01	-1.06E-06	8.21	-3.31	<0.01

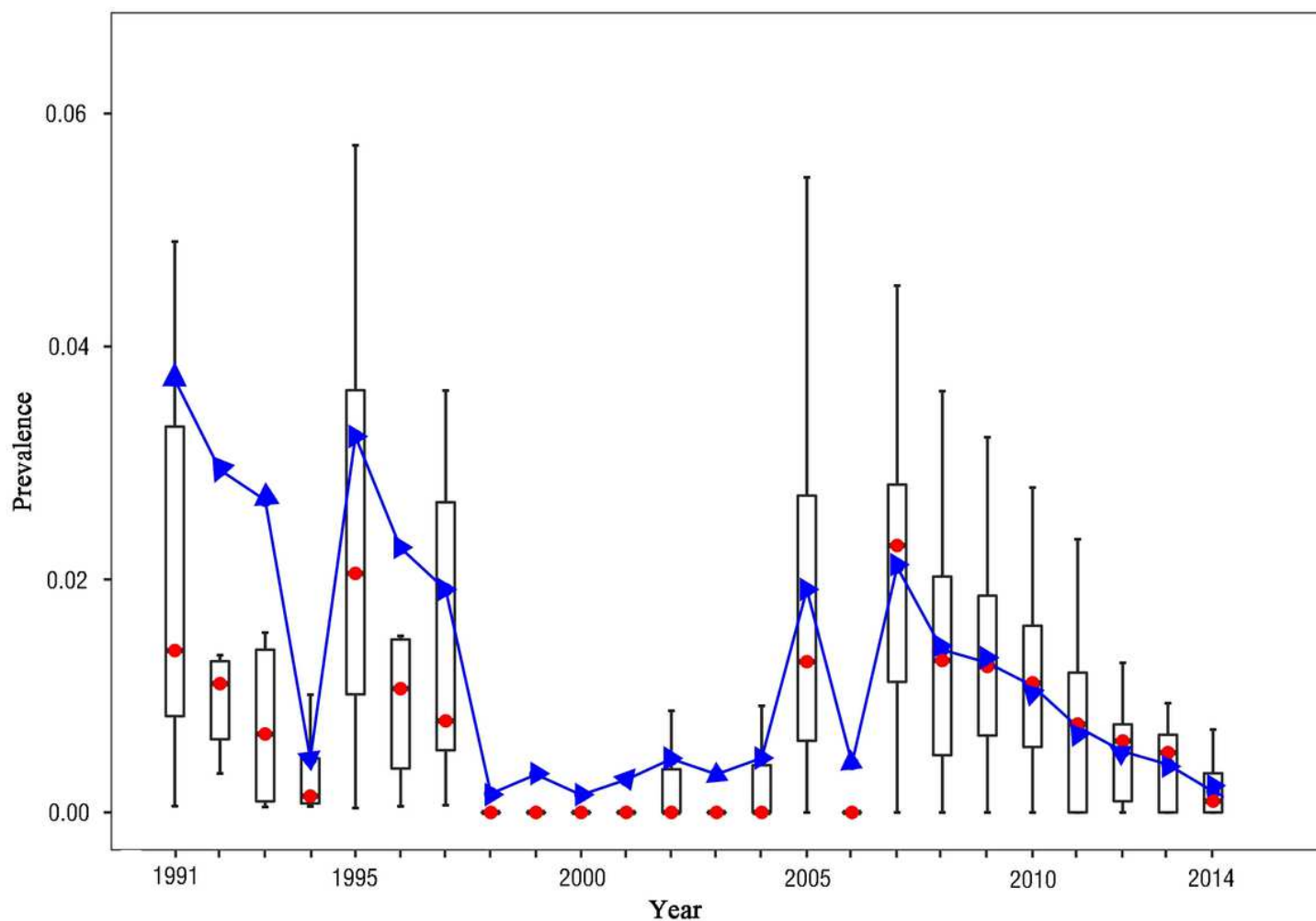


# Figures



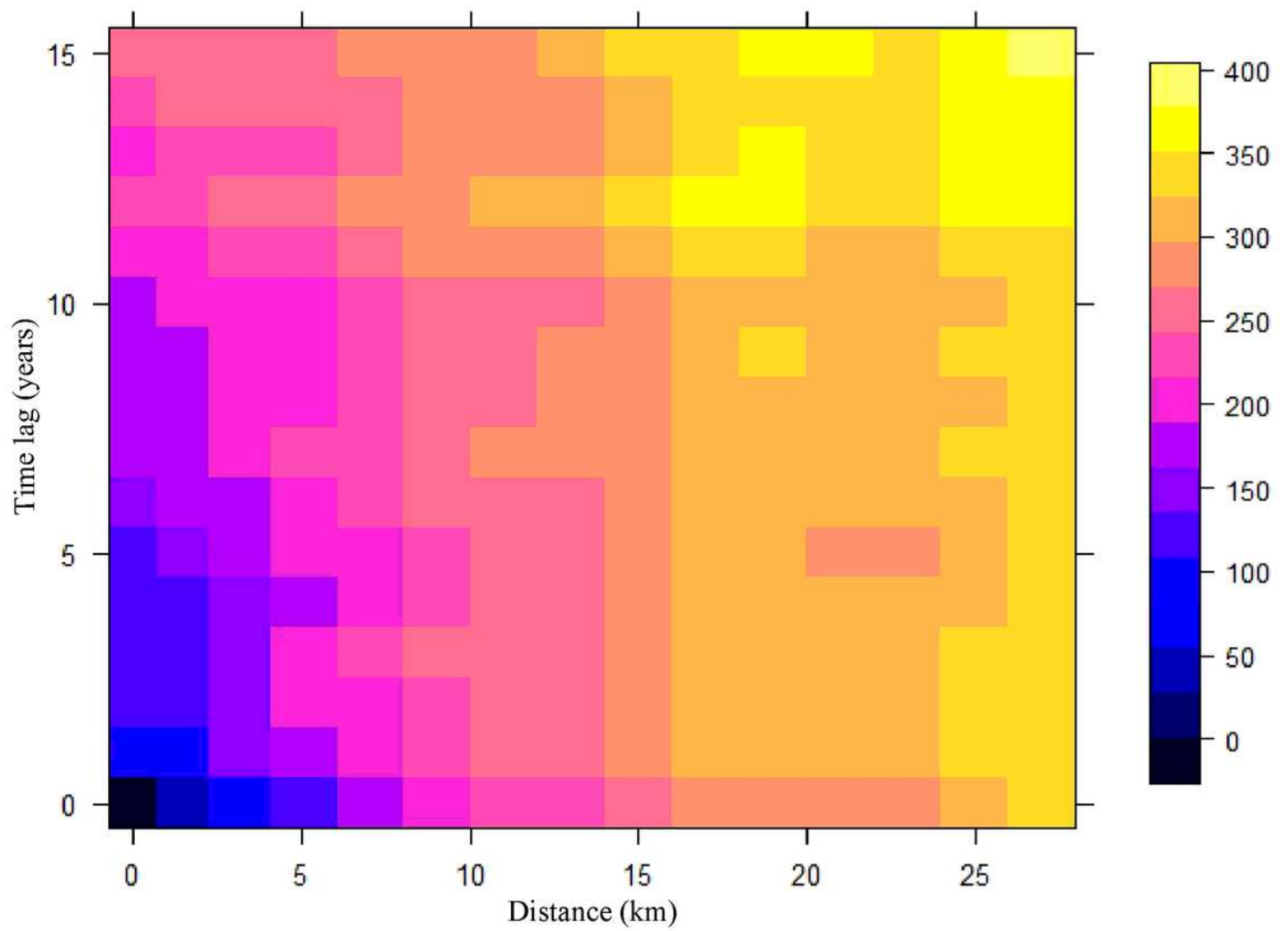
**Figure 1**

Endemic area of schistosomiasis in Guichi, Anhui Province, China. Guichi in Anhui province situates in the middle-lower reaches of the Yangtze River in eastern China, and the Qiupu River is in the study area. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.



**Figure 2**

Box plot for the observed prevalence of schistosomiasis for sample villages in Guichi, Anhui, China, from 1991 to 2014. The blue line represents the average annual prevalence, and the red point is median annual prevalence, the boxes denote minimum, median, maximum, and interquartile ranges.



**Figure 3**

Empirical spatio-temporal semi-variogram of the residuals in regression model.

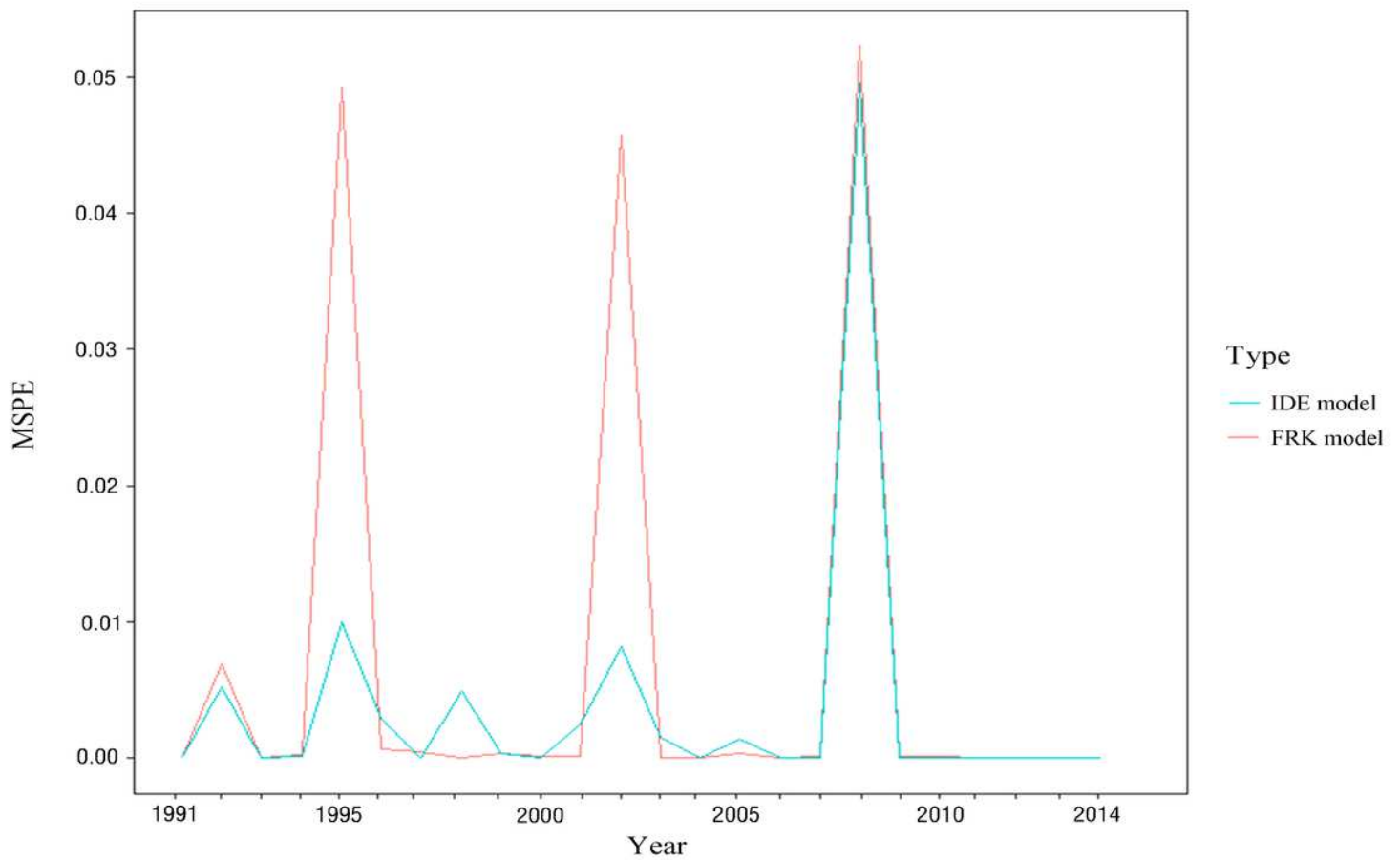


Figure 4

MSPE of the FRK predictions (red), the IDE predictions (blue) and the non-spatial model(yellow) as a function of time.

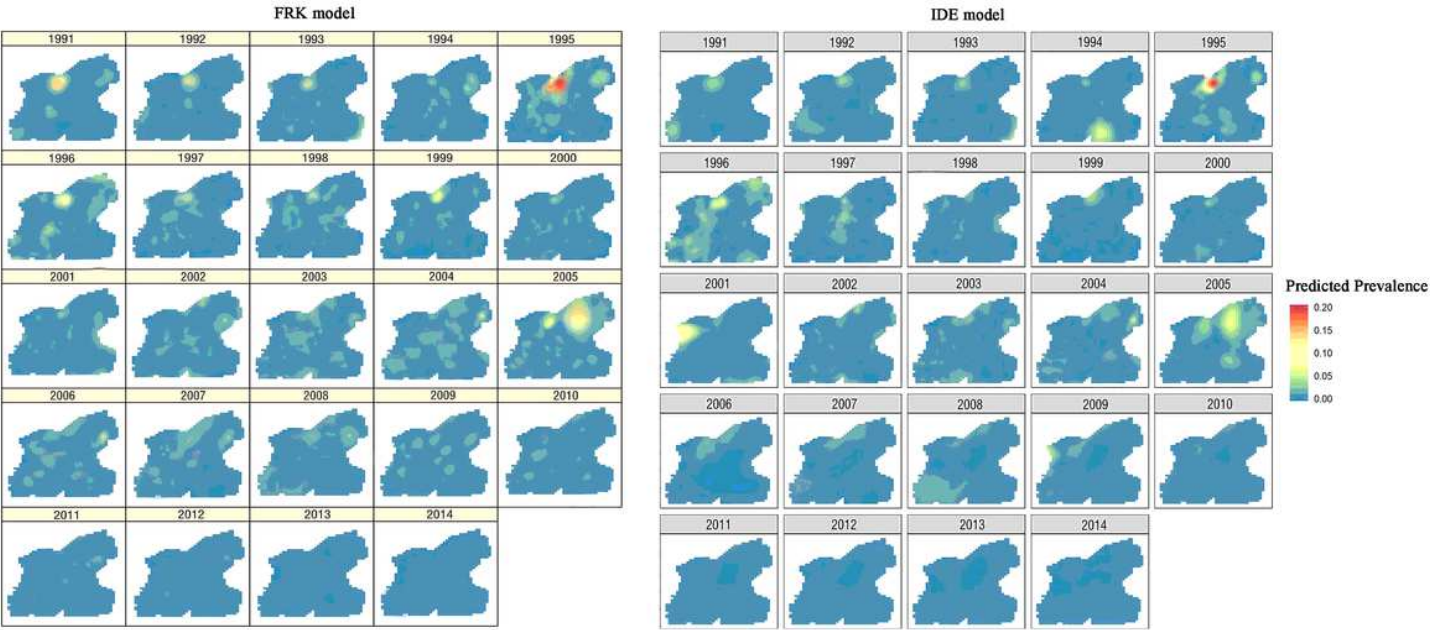
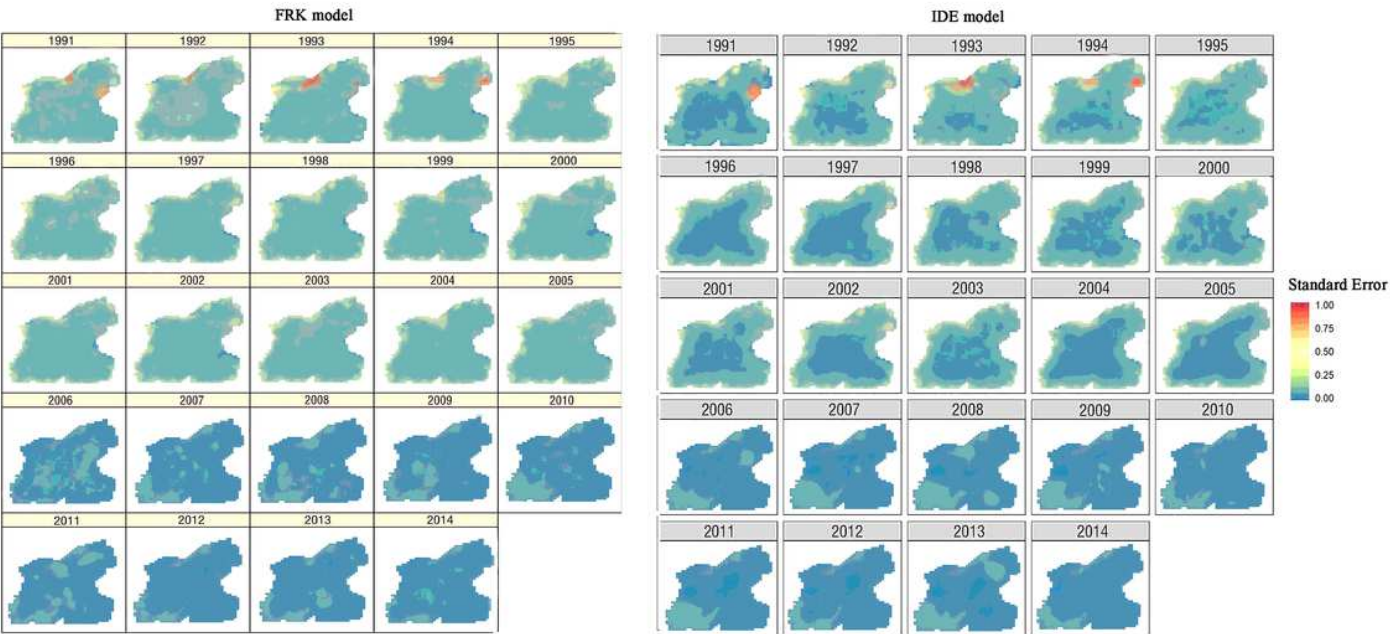


Figure 5

Plots for the annual prevalence of schistosomiasis predicted from 1991 to 2014 in Guichi, Anhui Province, China. The left is the forecast graph generated by FRK model, and the right is by IDE model.



**Figure 6**

Plots for the annual standard error of predicted prevalence of schistosomiasis in Guichi, Anhui Province, China, from 1991 to 2014. The left is the forecast graph generated by FRK model, and the right is by IDE model.