Artificial Intelligence Technologies and Applications
C. Chen (Ed.)

© 2024 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/FAIA231402

Analysis of the Spatiotemporal Variations and Influencing Factors of COVID-19 in Various Cities in East China: Bayesian Spatio-Temporal Model

Wang MAN, Junhao SUN[1], Xianqiang WEI
Xiamen University of Technology,
Xiamen, Fujian, 361024, China

Abstract. Since 2019, COVID-19 has spread worldwide in a pandemic and caused enormous losses. Even with numerous studies focusing on the spatiotemporal analysis of COVID-19, a noticeable research gap exists as Bayesian spatiotemporal models have not been extensively utilized to analyze the localized spatial patterns of COVID-19 in China. The purpose of this study is to analyze the temporal and spatial patterns of COVID-19 in East China while delving into the impact of socioeconomic factors on the incidence of the virus. Employing a Bayesian spatial-temporal hierarchical model, this paper thoroughly examines the case data spanning 24 months from various cities in East China. The results show that location and time have significant effects, and the two interact. From a spatial perspective, the cities with the highest relative risk are Shanghai and Xiamen. In terms of time, the relative risk of onset is highest in January and February 2020, and remains stable in other time periods. It shows that since 2020, the prevention and control measures to reduce population mobility and home isolation in China are effective.

Keywords. COVID-19, Bayesian spatio-temporal model, East China, relative risk

1.Introduction

In December 2019, China emerged as the focal point of an outbreak of unidentified pneumonia, which aroused great concern internationally. Most patients infected with COVID-19 will have different degrees of respiratory symptoms, namely fever, cough, shortness of breath, etc, but for patients with some underlying diseases, severe illness or even death may occur. The coronavirus is mainly transmitted through respiratory droplets and direct contact[1]. Because of its extended incubation period and high infectivity, the COVID-19 spread rapidly in China, and many cities were closed, triggering a public health crisis. The Chinese government promptly implemented measures to mitigate population movement in order to supress the outbreak of COVID-19.

Exploring the spatiotemporal patterns of COVID-19 is crucial for epidemiologists. Numerous studies have highlighted the potential of GIS and spatial statistics for monitoring the epidemic. For example, Peng et al. employed global autocorrelation

[1]Corresponding author: Junhao SUN, Xiamen University of Technology, Xiamen, Fujian, 361024, China; Email: 595394186@qq.com

analysis to elucidate the spatial pattern of COVID-19. Furthermore, they investigated the association between COVID-19 and diverse factors [2]. Ever since the onset of the COVID-19 pandemic, multiple studies have shown that various influencing factors have different effects on the spread and spatiotemporal distribution of COVID-19. Wang et al. studied the spatiotemporal characteristics of COVID-19 by collecting data on confirmed cases of COVID-19 and meteorological data from various provinces in China, and using techniques such as spatial autocorrelation, hotspots, and spatio-temporal scanning[3]. Simultaneously, the researchers utilized multiple linear regression analysis to examine how influencing factors relate to the confirmed cases of COVID-19. In addition to traditional regression models and clustering methods, many scholars conduct spatiotemporal analysis using Bayesian spatiotemporal models. Alfred Ngwira et al. utilized the Bayesian spatiotemporal model to fit the monthly count of confirmed cases, while considering socio-demographic factors, and explored the spatio-temporal distribution of COVID-19. The study found that the risk of incidence in cities was higher than that in rural areas. Across time, the risk of incidence has shown a pattern of initial rise followed by subsequent decline in most regions[4].

Currently, the research methods employed to study the COVID-19 epidemic still largely rely on traditional regression models and clustering, while overlooking the impact of space-time factors. At the same time, due to limited data, most studies using Bayesian spatio-temporal models did not take time and other geographical factors into consideration. This paper will utilize a Bayesian spatio-temporal model to investigate the temporal and spatial distribution of COVID-19 in East China, as well as investigate the relationship between socioeconomic factors and the incidence rate. At the same time, it provides reference for prevention and control in other countries and regions.

2.Method

2.1. Study Area

The research area encompasses the eastern region of mainland China, commonly referred to as "East China," and is situated in the eastern part of the China. The East China region comprises a total of 77 prefecture-level administrative units.

2.2. Data Sources

The COVID-19 case data from January 2020 to December 2021 was obtained from the National Health Commission of China. Data was collected for 77 cities encompassing three types of variables: demographic variables, economic variables and medical variables. Among them, economic variables include GDP and per capita GDP(GDPPC), and medical variables include the number of and hospitals(NOCHI). These three kinds of data are all from China Statistical Yearbook.

2.3. Data Analysis

In this paper, Bayesian hierarchical spatio-temporal model is used to study the long-term spatio-temporal effect of COVID-19, and the framework used is the time-space

inseparable framework[5, 6]. The assessment entails investigating the correlation between incidence risk and the variables.

Firstly, the data model equation(1) is formulated to depict the spatial distribution of COVID-19.The Bayesian spatio-temporal model assumes that the COVID-19 number reported by the i city obeys Possion distribution(y_{it}):

$$y_{it} \sim Possion(\mu_{it}) \tag{1}$$

$$\mu_{it} = n_i \cdot \theta_{it} \tag{2}$$

As in the data model equation(2),Where μ_{it} is the product of n_i and θ_{it}. μ_{it} represents the number of cases in COVID-19. Assume that the total population of a region i is time-varying. n_i represents the population of each city. θ_{it} represents the potential unknown incidence of COVID-19 at t region at i months.

$$\log(\theta_{it}) = \alpha + \sum_{k=1}^{p} \beta_k \chi_{ki} + S + U + v_t + \delta_{it} \tag{3}$$

As in equation(3),the formula is a Bayesian spatio-temporal effect estimation model. Where α stands for intercept term, which describes the difference between local area and the whole. β_k represents the regression coefficient of GDP and medical variables respectively. In this model, S is an ICAR model[7] with spatial structure defined by spatial weight matrix W . U stands for spatial unstructured effect and obeys normal distribution. v_t stands for time random effect, and the model specifies a first-order random walk model RW for specific parameters at time points. δ_{it} stands for space-time interaction effect, which all spatio-temporal interaction parameters are independent of each other. But each δ follows the same normal distribution $N(0,\sigma_\delta^2)$ .In the parametric model, Assign the fuzzy uniform prior distribution $U(0.0001,10)$ to the standard deviation σ_δ .

After that, the model operation is fitted by Markov Monte Carlo (MCMC) method, the posterior distribution of model parameters is obtained by WINBUGS software, Finally, the convergence is evaluated by checking Gelman-Rubin statistics[8].

3. Results

3.1. Descriptive Analysis

Figure 1 shows the spatial distribution of the cumulative incidence rate of municipal COVID-19 in East China. The areas with higher incidence rate are mainly distributed in Shanghai, Xiamen, Nanjing, Putian, Fuzhou, Suzhou and Nanchang.

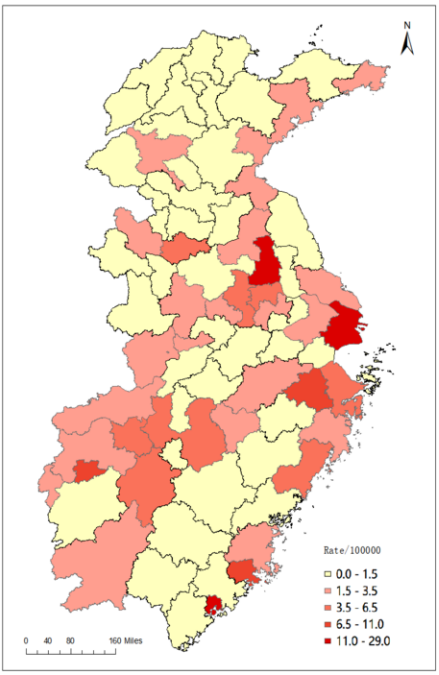

Figure 1. Incidence rate in prefecture-level cities

3.2. Temporal Relative Risk

This paper studies the change trend of relative risk of onset time in COVID-19 from January 2020 to December 2021. Figure 2 shows the posterior mean $(RR = \exp(\nu_t))$ [9]and 95% uncertainty band of relative risk in time and space during the study period. From the first month to the second month, there was an upward trend, which reached the peak in the second month, and then the relative risk decreased rapidly and remained stable until December 2021.

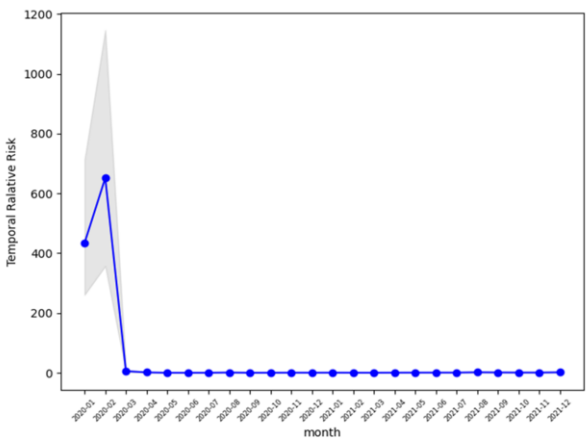

Figure 2. Temporal Relative Risk

3.3. Spatial Relative Risk

The figure 3 shows the spatial relative risk in East China ($RR = \exp(S_i + U_i)$) [9]. Among them, Shanghai, Xiamen, Qingdao, Nanjing and Fuzhou are the cities with relatively high spatial risks, and the city with the highest risk is Shanghai, with a relative risk value of 10.82. Other cities have relatively low space risks.

3.4. Spatial-temporal Relative Risk Trend

The relative risk of temporal and spatial trend ($RR = \exp(\delta)$) [9] indicates the risk of COVID-19 incidence in the whole region and the whole time period compared with the region i and time t. From the figure 4 and figure 5, we can see the change trend of relative risk of two adjacent areas with time. The temporal trend of the incidence risk in two neighboring regions is regarded as stochastic, and it is unrelated to the spatial characteristics of these regions.

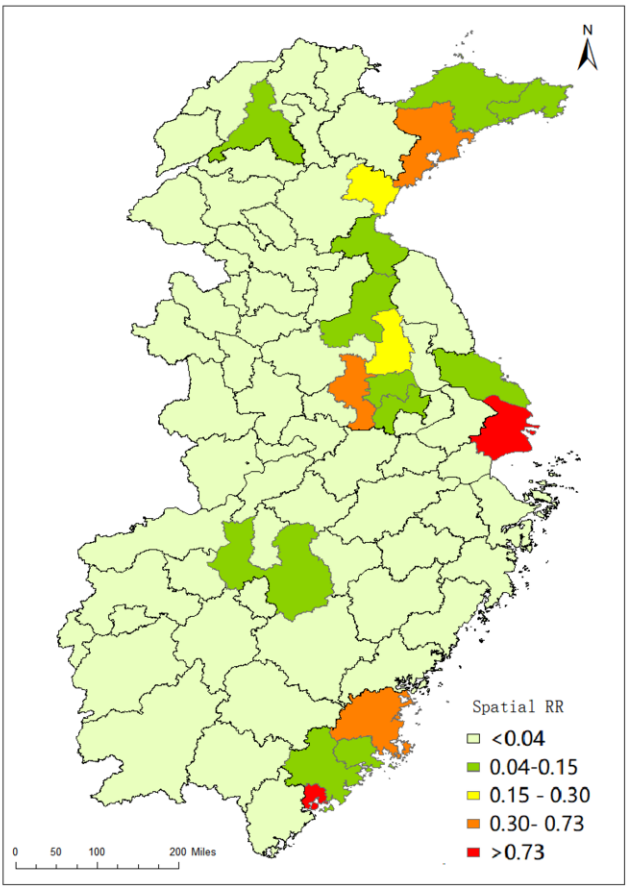

Figure 3. Spatial Relative Risk

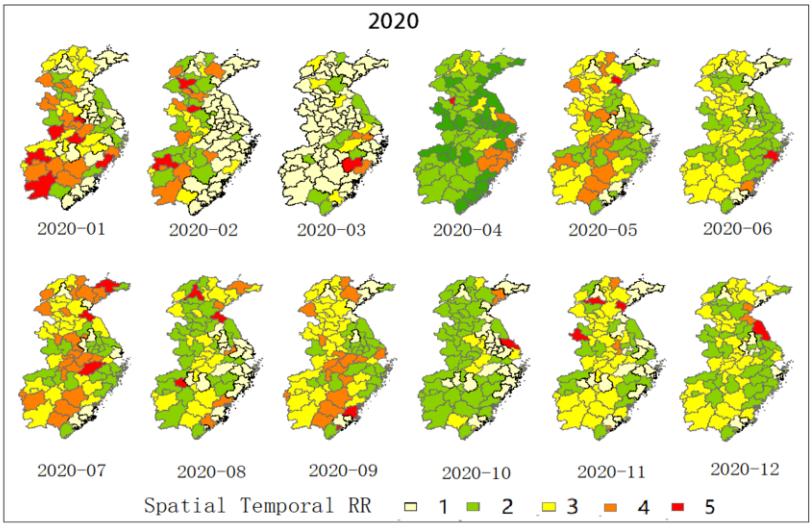

Figure 4.2020 Spatial-Temporal Relative Risk

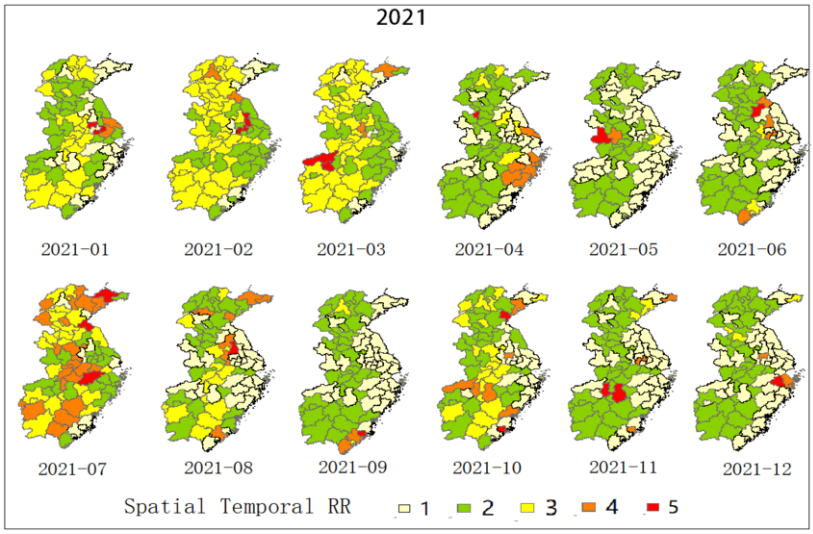

Figure 5.2021 Spatial-Temporal Relative Risk

3.5. Influencing Factors Analysis

The table 1 shows the relative risk of each factor ($RR = \exp(\beta)$) [9]. GDP and per capita GDP are negatively correlated with the overall incidence rate in East China.NOCHI is negatively correlated with the incidence rate of COVID-19.

Table 1. Relative risk of influencing factors and 95%CI

Factors	Posterior Estimates of Risk (95% CI)
GDP	0.9986(0.9974-0.9999)
NOCHI	1.032(0.9708-1.095)
GDPPC	0.9995(0.9991-0.9998)

4.Conclusion

In this paper, we utilized a Bayesian spatiotemporal model to examine the spatiotemporal pattern of COVID-19 in a prefecture-level city located in East China, while also examining the impact of certain socio-economic factors on the epidemic's spread.According to the research results, from January 2020 to February 2020, Shanghai, Xiamen, Fuzhou, Nanjing and Qingdao were recognized as high-risk areas. This is because during the initial phase of the epidemic, China has taken measures to reduce population mobility and isolate people at home to control the domestic epidemic. However, with the worldwide popularity of COVID-19, a large number of foreign cases were imported into China. As a result, these cities with international airports are relatively higher risk.

From a temporal perspective, the peak incidence occurred in January and February 2020, mainly due to people's lack of preventive awareness and significant population movement. However, strict control measures by the Chinese government have stabilized the epidemic[10].By observing the result diagram of time-space interaction, we can find that the relative risk of the same city has changed in adjacent time periods. This is because when an epidemic occurred in every city, the local government quickly took emergency measures.

According to the analysis of relative risk value of covariate, GDP and per capita GDP are negatively correlated with the overall incidence rate in East China. This maybe due to the fact that some cities with high GDP are large in scale and the government has taken better preventions .The number of hospitals is negatively correlated with the incidence rate of COVID-19 may be caused by sending cases to places with good medical resources for treatment in some areas. The limitation of this paper is that other influencing factors are not taken into account, but this does not affect the validity of the results.

References

[1] C. C. f. D. Control and P. E. W. G. f. N. E. Response, "The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China," *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi,* vol. 41, no. 2, pp. 145-151, 2020. doi:10.3760/cma.j.issn.0254-6450.2020.02.003

[2] D. Peng *et al.*, "COVID-19 distributes socially in China: A Bayesian spatial analysis," *PLoS One,* vol. 17, no. 4, p. e0267001, 2022. https://doi.org/10.1371/journal.pone.0267001

[3] Q. Wang *et al.*, "Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors," *International Journal of Infectious Diseases,* vol. 105, pp. 675-685, 2021.https://doi.org/10.1186/s12879-021-05926-x

[4] A. Ngwira, F. Kumwenda, C. S. Munthali, and D. Nkolokosa, "Spatial temporal distribution of COVID-19 risk during the early phase of the pandemic in Malawi," *PeerJ,* vol. 9, no. 1, 2021. doi:10.7717/peerj.11003

[5] R. P. Haining and G. Li, *Regression Modelling Wih Spatial and Spatial-Temporal Data: A Bayesian Approach*. CRC Press, 2020. https://www.crcpress.com/go/ssbs

[6] L. Knorr-Held, "Bayesian modelling of inseparable space-time variation in disease risk," *Statistics in medicine,* vol. 19, no. 17-18, pp. 2555-2567, 2000.https://doi.org/10.1002/1097-0258(20000915/30)19:17/18<2555::AID-SIM587>3.0.co;2-%23

[7] J. Besag and J. Y. Mollié, "Bayesian image restoration, with two applications in spatial statistics," *Annals of the Institute of Statistical Mathematics,* 1991.

[8] N. Nazia, J. Law, and Z. A. Butt, "Identifying spatiotemporal patterns of COVID-19 transmissions and the drivers of the patterns in Toronto: a Bayesian hierarchical spatiotemporal modelling," *Scientific Reports,* vol. 12, no. 1, p. 9369, 2022. https://doi.org/10.1038/s41598-022-13403-x

[9] S. Bie, X. Hu, H. Zhang, K. Wang, and Z. Dou, "Influential factors and spatial–temporal distribution of tuberculosis in mainland China," *Scientific Reports,* vol. 11, no. 1, p. 6274, 2021. https://www.nature.com/articles/s41598-021-85781-7

[10] H. Lau *et al.*, "The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China," *Journal of travel medicine,* 2020.doi: 10.1093/jtm/taaa037