

贝叶斯时空模型在新冠风险评估中的应用

牛佳安*, 余佳雄

武汉科技大学理学院, 湖北 武汉

收稿日期: 2023年4月24日; 录用日期: 2023年5月19日; 发布日期: 2023年5月25日

摘要

基于2020~2021年河南省18个地级市新冠肺炎确诊病例数据, 采用BYM (Besag-York-Mollie)、FBM (Familiar Bayesian Spatio-temporal)模型来分析新冠肺炎发病风险的时空分布特征, 并探讨社会经济、生态环境等因素对疾病风险的影响效应, 同时就应用结果对比两种模型的优劣。方法: 收集2020~2021年河南新冠肺炎确诊数据, 采用相关性分析排除部分相关性不大的影响因子, 并运用因子分析降维诸指标, 而后分别代入两模型进行分析、讨论。结果: 1) 2020~2021年河南全省的疫情风险程度是呈现下降趋势的, 高风险地区主要集中于豫北地区。2) 因子分析共构建基本经济因子、社会因子、温度因子、降水因子以及空气质量因子五个综合指标。3) 通过模型运行后的结果和DIC值可以得出FBM模型总体要优于BYM模型。

关键词

新冠疫情, 疾病风险评估, BYM模型, FBM模型

Application of Bayesian Spatio-Temporal Model in the Risk Assessment of COVID-19

Jia'an Niu*, Jiaxiong Yu

College of Science, Wuhan University of Science and Technology, Wuhan Hubei

Received: Apr. 24th, 2023; accepted: May 19th, 2023; published: May 25th, 2023

Abstract

Based on the data of confirmed cases of COVID-19 in 18 prefecture-level cities in Henan Province from 2020~2021, BYM (Besag-York-Mollie) and FBM (Familiar Bayesian Spatio-temporal) models were used to analyze the spatial and temporal distribution characteristics of the risk of COVID-19,

*通讯作者。

and explore the influence of socioeconomic and ecological factors. The results of this study were also used to compare the advantages and disadvantages of the two models. Methods: The data of confirmed cases of COVID-19 in Henan from 2020 to 2021 were collected, correlation analysis was used to exclude some influencing factors with little correlation, and factor analysis was applied to downscale the indicators, which were then analyzed and discussed by substituting into the two models. Results: 1) The risk level of the epidemic in Henan province from 2020 to 2021 showed a decreasing trend, and the high-risk areas were mainly concentrated in the northern part of Henan. 2) The factor analysis constructed five comprehensive indicators: basic economic factor, social factor, temperature factor, precipitation factor and air quality factor. 3) The results and DIC values after the model run can be concluded that the FBM model is better than the BYM model in general.

Keywords

COVID-19, Disease Risk Assessment, BYM Model, FBM Model

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

新冠肺炎大流行历经三年，其中较为严重的是前两年的初始毒株和 Delta 大流行，通过贝叶斯时空模型可以良好地探索新冠肺炎在时间和空间维度上的发展趋势。对疾病的时空数据进行探索分析，这有助于我们知晓某种流行病的时空演变特征[1]，而对 2020~2021 新冠大流行的研究也是这个范畴。

BYM (Besag-York-Mollie)模型、FBM (Familiar Bayesian Spatio-temporal)模型是两种经典的贝叶斯时空统计模型，BYM 模型是当前使用较多的、被充分验证过的贝叶斯时空统计模型，由 Besag 等人提出[2]，而 FBM 模型是 G. Li 等以 BYM 为根基而发展出的一种模型[3]，本文的后续讨论采用这两种模型。这两种模型目前已经被广泛运用于分析时空演变的问题当中，包括疾病、犯罪等问题的风险评估，例如，于国伟等利用 BYM 模型分析了甘肃省武威市凉州区胃癌死亡的时空分布及演变[4]，潘蕾等采用 BYM 模型分析了北京市 2005~2009 各区的结核病在气温、人口等影响因子的作用下的时空演变特征[5]，张伟文等运用 BYM 模型分析了新疆涂阳地区的肺结核发病的时空演变[6]，Daiane Leite da Roza 等以 BYM 模型分析了巴西东南部 São Paulo 州 Ribeirão Preto 结核病发病率的时空格局及其与社会脆弱性的关系[7]，Daqian Liu 等用 BYM 模型，加入相应的分析参数，基于贝叶斯时空模型研究了长春市犯罪时空格局[8]，上述研究的问题都有两个特点，那就是研究的对象都具有时间性和空间性两个性质，对于本文的研究很有借鉴价值，因而我们可以尝试将 BYM 模型应用于新冠肺炎的研究。FBM 模型起初由 G. Li 等提出并运用于英国彼得区的入室盗窃犯罪风险评估，并分析了该区域内的热点、冷点区域及其发展趋势[9]，Danyan Liang 等采用 FBM 模型讨论了 2010~2015 年内蒙古地区的布鲁氏菌疫情的时空发展演变[10]，Li Wang 等运用 FBM 模型分析了社会经济发展不平衡地区的结核病的时空演变[11]，Ram K. Raghavan 等采用 FBM 模型论证了美国落基山脉特有的斑点热疾病的时空分布和演变，并评估其风险，同时探讨了相关的驱动因素[12]，Junming Li 等使用该方法分析论证了中国大陆地区卫生支出的时空演变并探讨了驱动因素以及驱动因素对演变的影响程度[13]，Xiangxue Zhang 等采用该方法分析了 2000~2018 年中国 PM2.5 与气象和社会经济因素的时空异质性[14]，Wei Wang 等运用 FBM 模型探索了 2006~2020 年中国大陆 2844

个县心血管死亡率的时空趋势和生态决定因素[15], 同样通过上述文献的论述, 我们可以看出凡是研究对象具有时间和空间两个维度的属性, 分析其时空发展的趋势可以采用 FBM 模型, 新冠原始毒株和 Delta 的大流行在空间上, 新冠流行于全世界, 时间上跨越三年, 因而将 FBM 模型应用于新冠肺炎的风险评估, 分析其时空发展趋势也是符合条件的。那么两种模型在分析评估 COVID-19 问题当中的应用结果如何? 通过上述前人的研究我们发现贝叶斯时空模型被较少地应用于新冠大流行的研究中, 同时驱动一个地区新冠流行的影响因子对新冠流行风险的影响程度鲜有系统地评估; 而且 BYM 与 FBM 模型的优劣对比也鲜有提及。因此, 本文以 2020~2021 年河南省地区的新冠肺炎疫情作为研究对象, 探讨两种模型在该流行病学研究中的应用情况和模型对比的结果, 同时探索研究该流行病的驱动因素及其对疾病风险的影响。这里我们采用离差信息准则(Deviance Information Criterion, DIC)来比较两种模型之间的拟合程度, 评价模型在应用中的优劣, 一般 DIC 值越小模型拟合情况越佳[16]。

2. 理论分析

2.1. 数据相关性及因子分析

首先, 对 2020~2021 年河南省 18 个地级市的各影响因子、确诊人数的数据取平均值, 而后对各潜在的影响因子同确诊人数做相关性分析, 这里采用皮尔逊相关性分析的方法, 公式如下:

$$r_{ij} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}} \quad (1)$$

得出结果后对相关性系数进行检验, 得出结果 P 值小于 0.05 具有统计学意义, 大于 0.05 不具有统计学意义, 予以舍弃, 从而筛选出合适的影响因子。

完成相关性分析后, 对筛选完成的影响因子进行因子分析。由于各个影响因子在单位及量纲上都不一致, 因而需要对各影响因子进行归一化, 从而方便后续的建模分析。归一化的方法公式如下:

$$\text{Normalization-Index} = \frac{X_{kij} - \text{Min}}{\text{Max} - \text{Min}} \quad (2)$$

其中, X_{kij} 表示区域 i 第 j 年第 k 个变量的原始数据, Min 表示第 k 个变量的最小值, Max 表示第 k 个变量的最大值, Normalization-Index 的取值范围是 0~1。

为使各影响因子间消除共线性问题, 这里采用因子分析对指标间进行降维, 从而对各影响因子进行新的分类重组。运用 Bartlett 球度检验和 KMO 检验, 检查变量是否独立或相关, Bartlett 球度检验统计量设为 0.05, KMO 取值介于 0~1 之间, 取值越接近 1, 则变量间的相关程度越强, KMO 统计量大于 0.7 以上时, 因子分析效果较好。因子分析的数学模型表达如下:

$$X = \begin{cases} X_1 = a_{11}F_1 + a_{12}F_2 + \cdots + a_{1m}F_m + \varepsilon_1 \\ \vdots \\ X_n = a_{n1}F_1 + a_{n2}F_2 + \cdots + a_{nm}F_m + \varepsilon_n \end{cases} = AF + \varepsilon \quad (3)$$

2.2. 贝叶斯时空模型的构建

BYM 和 FBM 模型均是利用先验分布对模型中所有的未知参数进行描述, 随后进行贝叶斯估计获得贝叶斯后验分布, 并利用马尔科夫蒙特卡洛方法(Markov Chain Monte Carlo methods, MCMC)进行后验分布的计算, 最终获得未知参数的估计值[17]。

2.2.1. BYM 模型的构建

设区域 i 在 j 时间段内感染病例数为 Y_{ij} , 当感染人数远小于当地标准人口数时, 可以认为 Y_{ij} 服从

Poisson 分布。

$$Y_{ij} | p_{ij} \sim \text{Poisson}(e_{ij}\theta_{ij}) \quad (4)$$

$$e_{ij} = n_{ij} * \left(\sum_{i=1}^n Y_{ij} / \sum_{i=1}^n n_{ij} \right) \quad (5)$$

$$\theta_{ij} = \exp(\alpha_0 + \mu_i + g_j + Psi_{ij} + v_i + \alpha_1 * time1_j + d_i * time1_j + \sum \beta_k * x_{kij}) \quad (6)$$

其中 p_{ij} 表示区域 i 在 j 时间段内的感染率, e_{ij} 是区域 i 在 j 时间段内的预期感染人数, n_{ij} 为区域 i 在 j 时间段内的标准人口数。 θ_{ij} 是区域 i 在第 j 年的相对风险度。 α_0 表示截距, μ_i 表示空间同质性, v_i 表示空间异质性, g_j 表示自回归时间效应, α_1 表示时间效应的系数, $time1_j$ 表示第 j 年的时间效应, d_i 表示距离效应, Psi_{ij} 表示时空交互效应, β_k 表示相应影响因子的协变量, x_{kij} 表示区域 i 第 j 年的第 k 个影响因子。

空间同质性 μ_i 反映空间相关性, 其先验信息一般通过条件自回归模型(Conditional Autoregressive Process, CAR)建模表示。模型公式如下:

$$\mu_i | \mu_{j, j \neq i} \sim N\left(\sum_{j=1}^{n_i} w_{ij} \mu_j, \sigma_\mu^2 / n_i\right) \quad (7)$$

其中, $\mu_i | \mu_{j, j \neq i}$ 表示区域 i 和区域 j 之间的空间效应; σ_μ^2 / n_i 代表区域 i 受到相邻 m 个区域空间效应影响的方差平均值, $\sum_{j=1}^{n_i} w_{ij} \mu_j$ 代表区域 i 受到相邻 m 个单位空间效应影响的平均值, 这里我们探寻的是河南省 18 个地级市的空间同质性, 所以, 若两地市相邻则令 $w_{ij} = 1$, 反之则等于 0。空间异质性由空间非结构效应反映, 一般通过正态分布建模, 即:

$$v_i \sim N(0, \sigma_v^2) \quad (8)$$

时间效应、时空交互效应一般也通过正态分布建模, 而距离效应的建模方法与空间同质性的建模方法相同。这里所有的参数的超先验分布均采用标准差为(0, 10)的均匀分布, 对于截距 α_0 的先验分布, 这里采用无信息先验分布。

2.2.2. FBM 模型的构建

FBM 的数学形式如下:

$$\theta_{ij} = \exp(\alpha_0 + s_i + b_0 t^* + v_j + b_{li} t^* + \varepsilon_{ij} + \sum \beta_k x_{kij}) \quad (9)$$

$$s_i = \exp(\alpha_1 + \mu_i + v_i) \quad (10)$$

$$t^* = t - t_{mean} \quad (11)$$

这里, α_0 表示截距, s_i 表示疾病风险的空间分布, 这部分可以用 BYM 模型来表示。 $(b_0 t^* + v_j)$ 代表在各区域里共同的时间变化趋势, $b_0 t^*$ ($t^* = t - t_{mean}$) 代表共同的时间趋势中线性组成部分, v_j 通过正态分布建模代表时间效应中的随机变化部分。 $b_{li} t^*$ 代表各区域各自的独立的时间变化趋势; ε_{ij} 代表不包含在模型中但是对观测数据产生影响的随机误差项。 β_k 表示相应影响因子的协变量, x_{kij} 表示区域 i 第 j 年的第 k 个影响因子。各参数采用正态分布建模, 各参数的超先验分布采用方差(0, 10)的均匀分布; μ_i 和 v_i 的先验分布与 BYM 模型相同; b_0 、 b_{li} 的取值通过计算河南省整体确诊人数变化率和各地级市确诊人数变化率获得。本文中的两种模型均在 WinBUGS1.4 中编译运行。

3. 实证分析

3.1. 数据来源

2020~2021 年河南省新冠肺炎确诊人数数据来源于河南省疾病预防控制中心, 诸影响因子来源于 2020 年和 2021 年河南省统计年鉴。共引入 GDP、农林牧渔产值、工业产值、建筑业产值、批发零售产值、交通运输产值、住宿餐饮产值、金融产值、房地产产值、其他服务业产值、居民人均家庭收入、第一产业单位数、第二产业单位数、第三产业单位数、年平均空气质量指数(AQI)、年平均气温、年平均降水、医疗机构、医疗床位数、医疗人员数共计 20 个影响因子。河南省基础地理信息数据来源于天地图。

3.2. 相关性及因子分析

将确诊人数和各感兴趣的影响因子进行两两相关性分析, 分析结果如图 1:

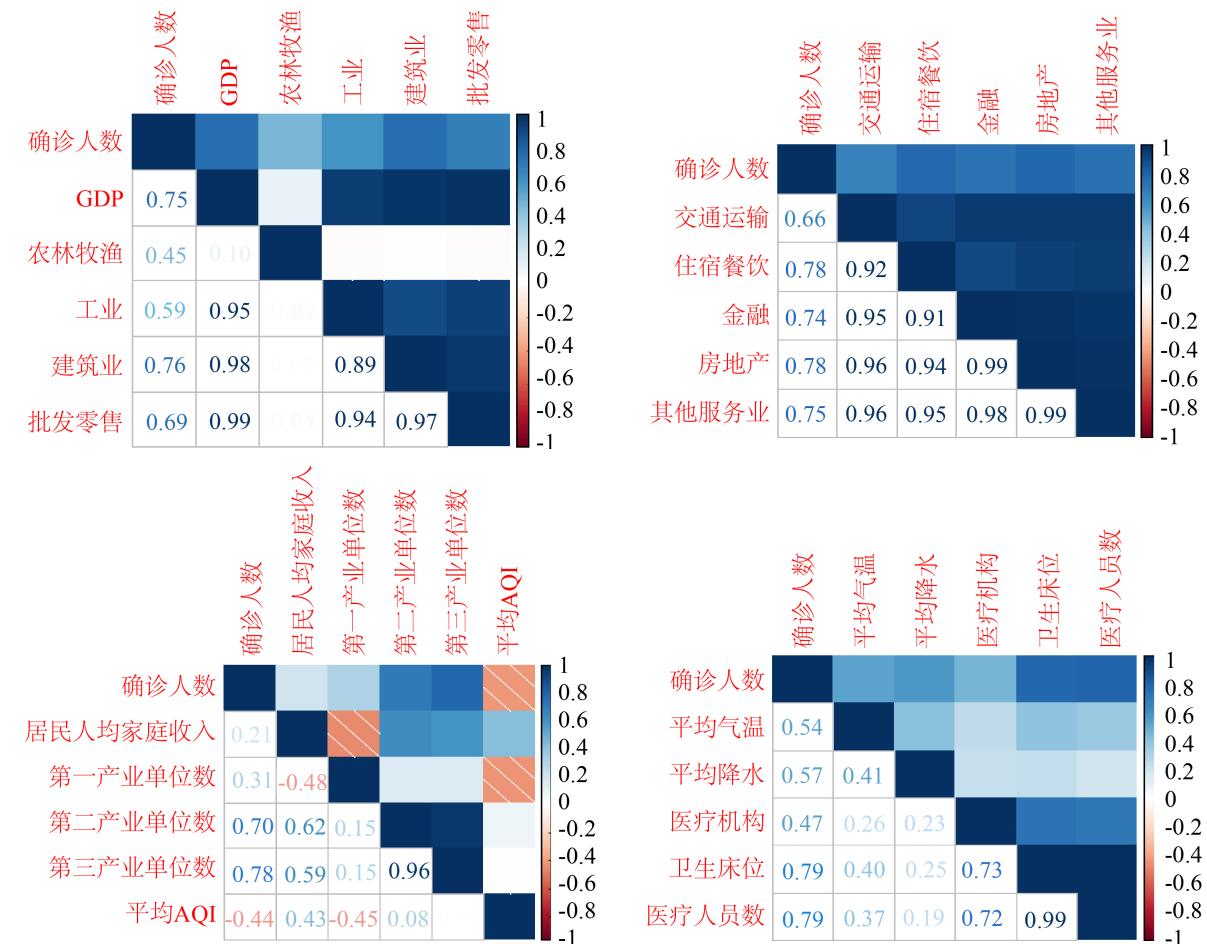


Figure 1. Correlation analysis of the number of diagnosed patients with each influence factor
图 1. 确诊人数同各影响因子的相关性分析

相关性分析完成后, 我们对结果进行检测, 检测结果见表 1。

通过上述分析, 我们可以发现 GDP、工业、建筑业、批发零售、交通运输、金融、其他服务业、第二产业单位数、医疗机构的 P 值大于 0.05, 不具备统计学意义, 故而舍弃上述影响因子。筛选后的影响

Table 1. Results of correlation analysis between the number of confirmed diagnoses and each impact factor indicator detection

表1. 确诊人数与各影响因子指标相关性分析结果检测

影响因子	P 值
GDP	0.12
农林牧渔	0.00
工业	0.81
建筑业	0.11
批发零售	0.55
交通运输	0.73
住宿餐饮	0.04
金融	0.20
房地产	0.03
其他服务业	0.14
居民人均家庭收入	0.01
第一产业单位数	0.03
第二产业单位数	0.39
第三产业单位数	0.03
平均空气质量指数(AQI)	0.00
平均气温	0.03
平均降水	0.00
医疗机构	0.06
医疗床位	0.02
医疗人员数	0.02

因子, 显然平均 AQI、平均气温、平均降水属于气象因子, 其余属于社会经济因子, 分成两组进行因子分析, 分析结果如下:

Table 2. Test results of socio-economic and meteorological factors

表2. 社会经济因子和气象因子的检验结果

KMO 和巴特利特检验	社会经济因子	气象因子
KMO 取样适切性量数	0.741	0.601
近似卡方	249.495	4.226
自由度	28	3
显著性	0.000	0.238

从表2的分析结果看, 社会经济因子的KMO值大于0.7, 且显著性小于0.001, 适合进行因子分析, 而气象因子的KMO值小于0.7, 且显著性大于0.001, 故而不适合进行因子分析。所以, 对社会经济因子进行因子分析, 对其进一步降维, 而气象环境因子中的平均气温、平均AQI、平均降水我们则将其设定为

三个独立变量代入后面的贝叶斯统计建模中。社会经济因子的降维如图2、表3:

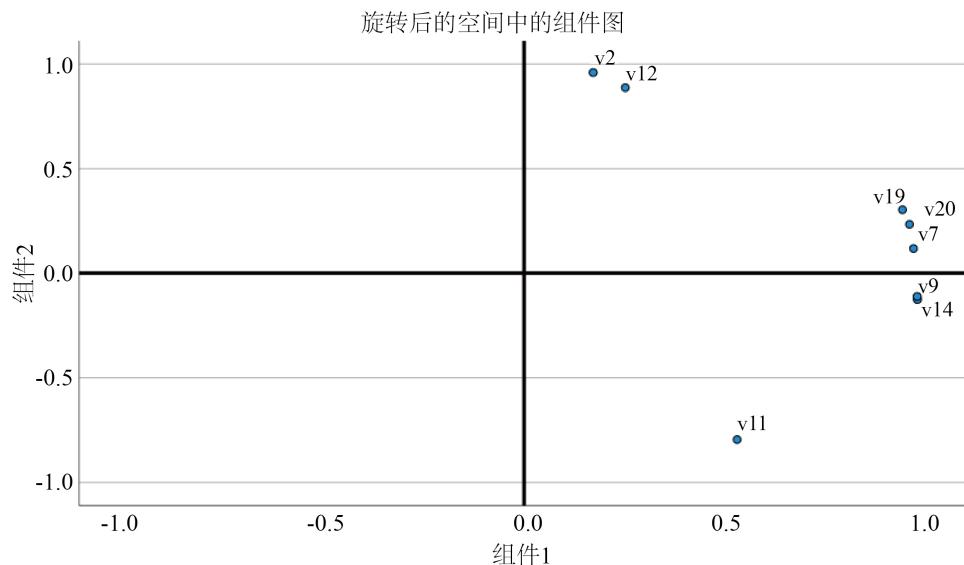


Figure 2. Spatial component diagram after rotation
图 2. 旋转后的空间组件图

Table 3. Component score coefficient matrix

表3. 成分得分系数矩阵

成分得分系数矩阵	成分 1	成分 2
V2	0.007	0.379
V7	0.190	0.019
V9	0.199	-0.078
V11	0.129	-0.333
V12	0.025	0.347
V14	0.199	-0.072
V19	0.180	0.094
V20	0.185	0.066

这里V2代表农林牧渔业, V7代表住宿餐饮业, V9代表房地产, V11代表居民人均家庭收入, V12代表第一产业单位数, V14代表第三产业单位数, V19代表医疗床位数, V20代表医疗人员数。根据分析结果, 这八个影响因子可以分成两个部分, 第一部分包含农林牧渔业、居民人均家庭收入和第一产业单位数这三个影响因子, 将这一部分命名为基础经济因子; 其余为第二部分, 将其命名为社会因子。至此, 我们完成了数据的相关性分析和因子分析。

3.3. 贝叶斯时空模型建模

3.3.1. BYM 建模

假设河南省地级市*i*第*j*年的相对风险为 rr_{ij} , 基础经济因子为 x_{1ij} , 社会因子为 x_{2ij} , 平均AQI为 x_{3ij} , 平均气温为 x_{4ij} , 平均降水为 x_{5ij} , 各影响因子的协变量系数为 $\beta_1 \sim \beta_5$ 。那么构建的BYM数学模型如下:

$$Y_{ij} \mid p_{ij} \sim \text{Poisson}(e_{ij} rr_{ij})$$

$$rr_{ij} = \exp(\alpha_0 + \mu_i + g_j + Psi_{ij} + v_i + \alpha_1 * time1_j + d_i * time1_j + \beta_1 * x_{1ij} + \beta_2 * x_{2ij} + \beta_3 * x_{3ij} + \beta_4 * x_{4ij} + \beta_5 * x_{5ij})$$

在WinBUGS1.4中进行大约300000次迭代, 模型基本达到了收敛状态。

模型达到收敛后, 我们可以得出河南省18个地级市2020年和2021年的相对风险程度的平均数值(见表4), 同时可以得出五个影响因子对相对风险的影响评估(见表5)。

Table 4. Relative risk level of prefecture-level cities in Henan Province, 2020~2021 in BYM Model
表4. BYM模型下2020~2021年河南省各地级市相对风险程度

城市	2020 年	2021 年
郑州	1.19	0.34
洛阳	0.39	0.78
濮阳	0.77	0.89
安阳	0.70	0.25
新乡	0.59	0.44
鹤壁	0.71	0.89
焦作	0.70	1.04
三门峡	0.69	1.21
济源	1.53	3.30
开封	6.18	0.05
商丘	0.08	0.05
平顶山	0.05	0.02
许昌	0.07	0.02
漯河	0.39	0.51
周口	0.76	0.08
南阳	0.09	0.05
驻马店	0.85	0.04
信阳	0.19	0.05

Table 5. Impact assessment of each impact factor in BYM Model
表5. BYM模型下各影响因子的影响评估

影响因子	结果
基础经济因子(β_1)	0.91
社会因子(β_2)	1.81
平均 AQI(β_3)	0.46
平均气温(β_4)	1.27
平均降水(β_5)	0.49

根据表4的结果, 绘制河南省2020~2021年新冠肺炎疾病风险专题地图, 见图3、图4:

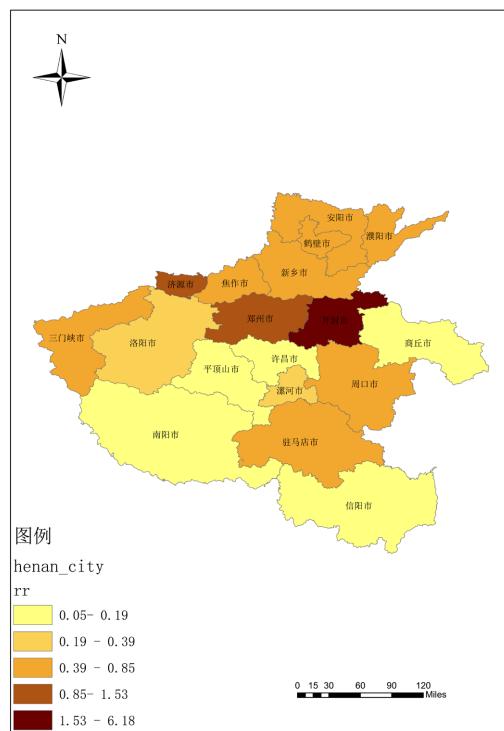


Figure 3. Risk assessment map of COVID-19 epidemic by prefecture-level cities in Henan Province in 2020 (BYM)

图3. 2020年河南省各地级市新冠疫情风险评估图(BYM)

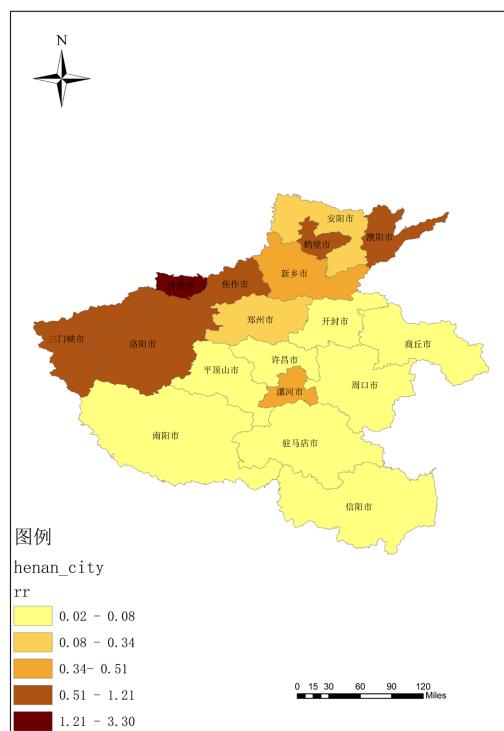


Figure 4. Risk assessment map of COVID-19 by prefecture-level cities in Henan Province in 2021 (BYM)

图4. 2021年河南省各地级市新冠疫情风险评估图(BYM)

从表5的结果来看,农林牧渔业、第一产业单位数等基础经济因子对新冠疫情的影响是正相关的,而空气质量指数即平均AQI对新冠疫情也存在着正相关的影响,且其中社会因子和平均气温对新冠疫情的影响较大,而从图3、图4的专题地图的表现来看,风险程度较高的主要集中于河南的中北部地区。河南的经济发达地区主要集中于这里,这里位于华北平原,农林牧渔业等基础经济因子也相较发达,同时,这里有着河南重要的工业基地,所以空气质量相对南部城市较差,人口也相对集中,所以疫情的风险集中于此。

3.3.2. FBM 建模

假设河南省地级市 i 第 j 年的相对风险为 rr_{ij} ,基础经济因子为 x_{1ij} ,社会因子为 x_{2ij} ,平均AQI为 x_{3ij} ,平均气温为 x_{4ij} ,平均降水为 x_{5ij} ,各影响因子的协变量系数为 $\beta_1 \sim \beta_5$ 。那么构建的FBM数学模型如下:

$$Y_{ij} | p_{ij} \sim \text{Poisson}(e_{ij} rr_{ij})$$

$$rr_{ij} = \exp(\alpha_0 + s_i + b_0 t^* + v_j + b_{1i} t^* + \varepsilon_{ij} + \beta_1 * x_{1ij} + \beta_2 * x_{2ij} + \beta_3 * x_{3ij} + \beta_4 * x_{4ij} + \beta_5 * x_{5ij})$$

$$s_i = \exp(\alpha_1 + \mu_i + v_i)$$

在WinBUGS1.4中迭代400,000次,初始值burn in设置为10,000,模型基本达到收敛,这里 $t^* = t - t_{mean}$, t 取1,2,则 $t_{mean} = 1.5$ 。

模型达到收敛后,我们可以得出河南省18个地级市2020年和2021年的相对风险程度的平均数值(见表6),同时可以得出五个影响因子对相对风险的影响评估(见表7)。根据表4的结果绘制河南省各地级市2020~2021年新冠疫情专题地图,见图5,图6,同时我们用WinBUGS1.4软件计算BYM和FBM两模型的DIC值,见表8。

Table 6. Relative risk level of prefecture-level cities in Henan Province, 2020~2021 in FBM Model
表6. FBM模型下2020~2021年河南省各地级市相对风险程度

城市	2020 年	2021 年
郑州	1.18	0.35
洛阳	0.37	0.78
濮阳	0.75	0.87
安阳	0.69	0.23
新乡	0.56	0.43
鹤壁	0.70	0.88
焦作	0.69	1.04
三门峡	0.68	1.21
济源	1.52	3.29
开封	6.17	0.04
商丘	0.07	0.09
平顶山	0.06	0.08
许昌	0.09	0.06
漯河	0.20	0.54
周口	0.82	0.05

Continued

南阳	0.12	0.09
驻马店	0.91	0.03
信阳	0.25	0.06

Table 7. Impact assessment of each impact factor in FBM Model
表7. FBM模型下各影响因子的影响评估

影响因子	结果
基础经济因子(β_1)	0.99
社会因子(β_2)	1.54
平均 AQI(β_3)	0.13
平均气温(β_4)	0.34
平均降水(β_5)	1.18

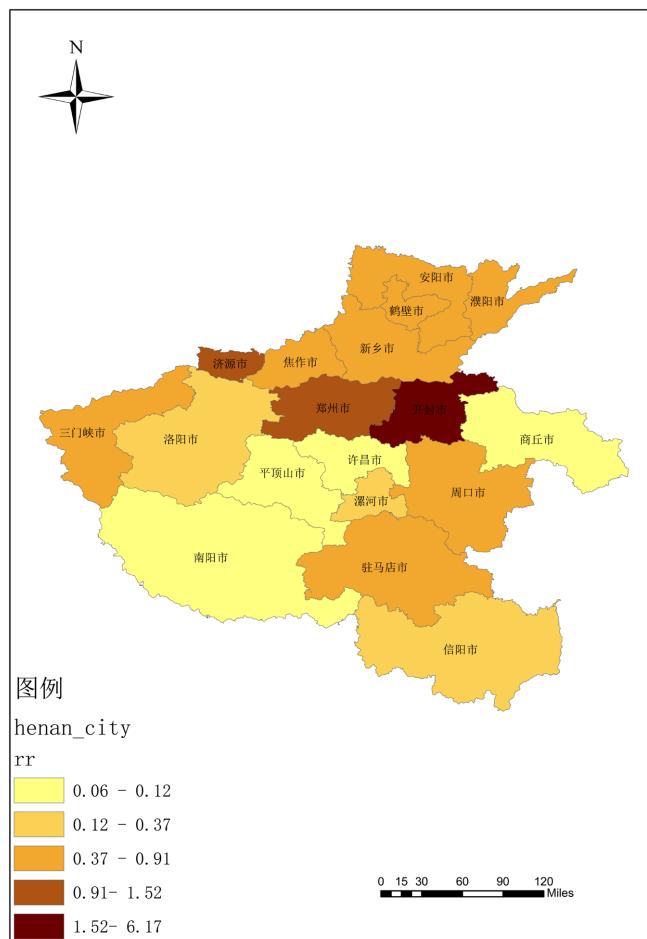


Figure 5. Risk assessment map of COVID-19 epidemic by prefecture-level cities in Henan Province in 2020 (FBM)

图5. 2020年河南省各地级市新冠疫情风险评估图(FBM)

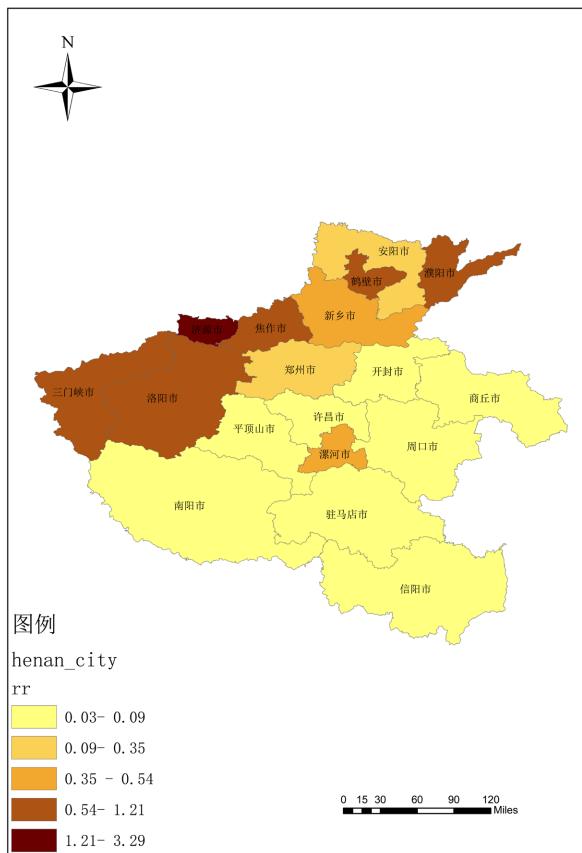


Figure 6. Risk assessment map of COVID-19 epidemic by prefecture-level cities in Henan Province in 2021 (FBM)

图6. 2021年河南省各地级市新冠疫情风险评估图(FBM)

Table 8. Comparison of the fit of the two models

表8. 两种模型的拟合程度比较

模型	Dbar	pD	DIC
BYM	175.788	36.983	212.771
FBM	167.482	30.028	197.509

注: DIC = Dbar + pD。

从最终的结果来看,基础经济因子、社会因子以及平均降水是对河南省新冠疫情的影响较大的影响因素,各影响因子都对相对风险存在正相关的影响,相对风险的程度都会随着各影响因子的指标的提升而提升。从两年的河南省各地级市的疫情风险评估专题地图也可以看出,河南省各地级市的新冠疫情相对风险程度呈现整体的下降趋势,在河南的北部地区新冠疫情的风险一直较为严重,其中济源市、三门峡市、洛阳市、焦作市、濮阳市、鹤壁市在2021年的疫情相对风险程度有所增加,郑州市略微减弱。而河南南部的城市,从2020年到2021年全部都成减弱趋势。

4. 结论

本文采用的BYM和FBM模型,通过2020~2021年河南省新冠疫情时空风险评估的应用,我们可以得出结论如下:

2020~2021年的新冠大流行的风险评估可以运用BYM和FBM来实现, 同时我们运用这两种模型还可以分析出相应的影响因子的影响系数。从两种模型最终的分析结果来看, 2020~2021年河南省新冠疫情较为严重的城市主要集中于河南的北方, 从时空发展的趋势来看, 从东北部向西北部转移。南部地区在大流行的第二年基本上风险系数都降低到了较低值。从两个模型中影响新冠大流行风险的影响因子评估结果来看, 基础经济因子和社会因子是影响程度较大的。从数据上看, 河南省的主要产业集中于以郑州为中心的城市群, 这里也是河南省人口稠密的地区, 所以可以说这里的风险较大是符合实际情况的。这表明BYM和FBM两种模型不仅可以良好地评估研究区域的风险程度, 而且可以良好地分析影响因子的对大流行风险的影响程度。

对于BYM模型和FBM模型的优劣对比, 通过表6的DIC值计算结果可以明显看出, FBM模型是优于BYM模型。同时, 通过对影响因子的评估结果分析我们可以看出, FBM模型的分析结果更加符合实际, 在2021年河南省因为发生了强降水导致了较大范围的洪灾, 洪灾过后疫情加重, 这说明降水对于新冠大流行的风险有着较大的影响, 而BYM模型的评估结果中, 平均降水的影响系数仅为0.49, 同时, 气温对于新冠大流行的影响已然被证明影响程度有限, 而BYM分析出的影响系数确是1.27, 对比FBM模型仅为0.34, 因而我们认为FBM模型对于新冠大流行风险的时空分析更为准确。

本文以河南省2020~2021年新冠疫情为例, 探索了可能影响新冠大流行风险的影响因子, 并探索了用BYM和FBM模型对新冠疫情的分析评估, 并对比两种模型的优劣。最终证明, 新冠疫情作为具有空间流行病特点的研究对象, 可以运用贝叶斯时空模型来分析处理, 同时在相同数据的前提下, FBM模型的分析结果要优于BYM模型。

参考文献

- [1] 刘大有, 陈慧灵, 齐红, 等. 时空数据挖掘研究进展[J]. 计算机研究与发展, 2013, 50(2): 225-239.
- [2] Goovaerts, P. and Gebreab, S. (2008) How Does Poisson Kriging Compare to the Popular BYM Model for Mapping Disease Risks? *International Journal of Health Geographics*, 7, Article No. 6. <https://doi.org/10.1186/1476-072X-7-6>
- [3] 李世元, 王学梅. 贝叶斯时空模型在空间流行病学中的研究进展[J]. 世界最新医学信息文摘, 2017(34): 55-57.
- [4] 廖伟斌, 孙建国, 于国伟, 等. 贝叶斯模型和时空扫描统计量在胃癌死亡空间分布特征中的应用[J]. 中华疾病控制杂志, 2016, 20(8): 847-850.
- [5] 潘蕾, 贺晓新, 罗艳侠, 等. 基于Bayes时空理论探讨北京市结核病演化特征及生态影响因素分析[J]. 中国卫生统计, 2013, 30(5): 658-664.
- [6] 张伟文, 曹明芹, 贺湘焱. 贝叶斯时空建模在新疆涂阳肺结核发病风险中的应用研究[D]: [硕士学位论文]. 乌鲁木齐: 新疆医科大学公共卫生学院, 2019.
- [7] da Roza, D.L., do Carmo Gullaci Guimarães Caccia-Bava, M. and Martinez, E.Z. (2012) Spatio-Temporal Patterns of Tuberculosis Incidence in Ribeirão Preto, State of São Paulo, Southeast Brazil, and Their Relationship with Social Vulnerability: A Bayesian Analysis. *Revista da Sociedade Brasileira de Medicina Tropical*, 45, 607-615. <https://doi.org/10.1590/S0037-86822012000500013>
- [8] Liu, D.Q., Song, W., Xiu, C.L. and Xu, J. (2021) Understanding the Spatiotemporal Pattern of Crimes in Changchun, China: A Bayesian Modeling Approach. *Sustainability*, 13, Article 10500. <https://doi.org/10.3390/su131910500>
- [9] Li, G., Haining, R., Richardson, S. and Best, N. (2014) Space-Time Variability in Burglary Risk: A Bayesian Spatio-Temporal Modelling Approach. *Spatial Statistics*, 9, 180-191. <https://doi.org/10.1016/j.spasta.2014.03.006>
- [10] Liang, D.Y., Liu, D., et al. (2021) Spatiotemporal Distribution of Human Brucellosis in Inner Mongolia, China, in 2010-2015, and Influencing Factors. *Scientific Reports*, 11, Article 24213. <https://doi.org/10.1038/s41598-021-03723-9>
- [11] Wang, L., Xu, C.D., et al. (2021) Spatio-Temporal Variation in Tuberculosis Incidence and Risk Factors for the Disease in a Region of Unbalanced Socio-Economic Development. *BMC Public Health*, 21, Article No. 1817. <https://doi.org/10.1186/s12889-021-11833-2>
- [12] Raghavan, R.K., Goodin, D.G., Neises, D., Anderson, G.A. and Ganta, R.R. (2016) Hierarchical Bayesian Spatio-Temporal Analysis of Climatic and Socio-Economic Determinants of Rocky Mountain Spotted Fever. *PLOS ONE*, 11, e0150180. <https://doi.org/10.1371/journal.pone.0150180>

-
- [13] Li, J.M. and Han, X.L. (2023) Spatiotemporal Evolution and Drivers of Total Health Expenditure across Mainland China in Recent Years. *International Journal of Environmental Research and Public Health*, **20**, Article 597. <https://doi.org/10.3390/ijerph20010597>
 - [14] Zhang, X.X. and Cheng, C.X. (2022) Temporal and Spatial Heterogeneity of PM_{2.5} Related to Meteorological and Socioeconomic Factors across China during 2000-2018. *International Journal of Environmental Research and Public Health*, **19**, Article 707. <https://doi.org/10.3390/ijerph19020707>
 - [15] Wang, W., Li, J.M., et al. (2022) Spatiotemporal Trends and Ecological Determinants of Cardiovascular Mortality among 2844 Counties in Mainland China, 2006-2020: A Bayesian Modeling Study of National Mortality Registries. *BMC Medicine*, **20**, Article No. 467. <https://doi.org/10.1186/s12916-022-02613-9>
 - [16] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and Van Der Linde, A. (2022) Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **4**, 583-639. <https://doi.org/10.1111/1467-9868.00353>
 - [17] 郑杨, 李晓松. 贝叶斯时空模型在疾病时空数据分析中的应用[J]. 中华预防医学杂志, 2010, 44(12): 1136-1139.