

基于贝叶斯模型的中国房地产投资发展研究

牛佳安*, 丁咏梅

武汉科技大学理学院, 湖北 武汉

收稿日期: 2024年1月14日; 录用日期: 2024年2月4日; 发布日期: 2024年4月19日

摘要

房地产投资的发展程度会受到很多因素的影响, 而且这些因素并不是单独存在, 是相互组成一个复杂的系统共同影响着中国大陆各地区的房地产投资发展在时空上演进、变化, 本文运用BYM、FBM两种贝叶斯时空模型研究分析房地产投资的发展程度, 分析各个影响因子的影响程度, 最后根据DIC值评估两种模型的拟合优劣。除此之外, 本文运用FBM模型研究了2012~2021年中国大陆地区各个省、自治区、直辖市房地产投资的总体发展热度以及局部发展趋势。

关键词

BYM模型, FBM模型, 房地产投资发展程度, 房地产投资发展热度, 房地产投资发展趋势

Research on the Development of Real Estate Investment in China Based on Bayesian Models

Jia'an Niu*, Yongmei Ding

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

Received: Jan. 14th, 2024; accepted: Feb. 4th, 2024; published: Apr. 19th, 2024

Abstract

The development degree of real estate investment will be affected by many factors, and these factors do not exist alone. They form a complex system and jointly affect the evolution and change of real estate investment development in Chinese Mainland in time and space. This paper uses BYM and FBM Bayesian space-time models to study and analyze the development degree of real estate

*通讯作者。

investment and the impact degree of each influencing factor, Finally, evaluate the fitting quality of the two models based on the DIC value. In addition, this paper uses the FBM model to study the overall development heat and local development trend of real estate investment in Chinese Mainland from 2012 to 2021.

Keywords

BYM Model, FBM Model, Degree of Real Estate Investment Development, Heat of Real Estate Investment Development, Trend of Real Estate Investment Development

Copyright © 2024 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

我国房地产投资快速增长, 2022 年房地产开发投资占全社会固定资产投资的比重为 22.93% [1]。影响房地产投资的潜在因素有很多, 主要包括基础设施环境因素、经济环境因素、政策环境因素等方面, 完备的基础设施能带动人口的流入和地区贸易的增长, 对房地产市场的发展壮大具有强大的推动作用[2]。韦兰宁、唐晓莲基于面板向量自回归模型, 发现房地产投资与就业量变化之间的互相影响在相应的周期内非常显著[3]。Weinberg 等指出, 消费水平对于房地产投资也存在影响[4]。乔林等在研究国内不同发展层次城市房价的影响因素时发现, 人均收入对房地产影响达到 43.16% [5]。然而, 上述文献仅研究单个影响因子, 没有考量尽可能多的影响因子, 为此本文引入两种贝叶斯时空模型, 对多个影响因子加以评估, 并对它们组成的复杂整体进行系统分析。

贝叶斯方法是将参数都视为随机变量并且服从于一定形式的分布, 数据可以是离散的, 可以服从于泊松分布、伯努利分布等。通过层次结构, 贝叶斯方法可以将时间信息, 空间信息和时间交互信息都纳入到一个模型中, 并且通过 MCMC 模拟估计出来[6]。这里我们引入 BYM 和 FBM 两种模型, BYM (Besag-York-Mollie)模型由 Besag 等人提出[7], FBM (Familiar Bayesian Spatio-Temporal)模型是 G. Li 等提出的一种模型[8]。在贝叶斯时空模型中, 常添加空间随机效应参数, 空间随机效用可以从邻近地区借力以稳定特定地区的风险, 克服由于数据较少导致的不稳定估计[9] [10]。同时, 空间随机效应一般用来替代无法解释的空间结构, 并且当数据存在过散分布和空间自相关时, 空间随机效应模型的表现也较良好[11]。并且, 在贝叶斯模型中加入时间项, 可以解释样本的时间变化特点, Yu 等[12]将时间项引入贝叶斯空间模型中, 研究了酒水供应量减少后对于暴力事件发生率的影响。Bernardinelli 等[13]将贝叶斯层次模型与时空交互模型结合起来, 既包含了地区随机效应, 也包含了平均趋势项和地区趋势项, 使得模型可以评估时空相互影响。贝叶斯层次时空模型在社会学等领域中应用广泛[14] [15] [16], Law 等[17]采用了这个模型, 研究了加拿大的随时间变化的小区域层级犯罪模式。Junming Li 等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[18]。韩秀兰、李俊明构建了贝叶斯时空层次非线性模型和贝叶斯分域时空回归模型研究了 2002~2015 年中国 35 个主要城市住宅商品房价格时空演变规律和影响机制, 并考虑年度人均 GDP、第三产业值、人口密度等影响因素[19], 上述文献虽然考量了多个影响因子, 但缺乏对影响因子的筛选, 没有验证分析影响因子之间可能存在的多重共线性, 并且对于研究区域整体的发展热度和局部的发展趋势缺乏研究。因而根据这些前人的研究成果, 本文引入第三产业产值、客运总量、就业人数、人均 GDP、小初高学校数量、医疗床位利用情况以及各地区的

城市人口密度作为影响因子, 并运用 BYM 和 FBM 建立贝叶斯时空模型。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[20]。

2. 理论分析

2.1. 数据相关性及因子分析

对 2012~2021 年中国大陆各地区的各影响因子、房地产投资的数据取平均值, 而后对各潜在的影响因子同房地产投资额做相关性分析, 这里采用简单相关性分析的方法, 得出结果后对相关性系数进行检验, 检验结果 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。由于在实际问题当中可能出现多重共线性的问题, 当有多重共线性的情况发生时, 参数估计的结果不再具有有效性, 而本文研究中探究多个影响因子对房地产投资的影响, 所以在总体分析前需要消除共线性。我们通常需要通过 VIF 检验来排除掉某些有多重共线性的变量。VIF, Variance Inflation Factor, 即方差膨胀因子。VIF 指的是解释变量之间存在多重共线性时的方差与不存在多重共线性时的方差之比, 可以反映多重共线性导致的方差的增加程度。这里 VIF 越大, 显示共线性越严重。经验判断方法表明: 当 $0 < VIF < 10$, 不存在多重共线性; 当 $10 \leq VIF < 100$, 存在较强的多重共线性; 当 $VIF \geq 100$, 存在严重多重共线性。这里我们将 $VIF > 10$ 的影响因子排除。

完成复相关性分析后, 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。归一化的取值范围是 0~1。

在研究问题时, 尽可能多的收集资料便于对问题有充分了解, 实际数据建模中, 有些变量不一定可以真正发挥作用, 还可能加大计算工作量, 所以要进行因子分析, 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。

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(MarkovChain 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。由于各省的房地产投资值差异较大, 本文采用泊松对数正态模型, 在连接函数中加入描述数据离散性的误差项, 容许数据过度离散的现象。

2.2.1. BYM 模型的构建

设区域 i 在 j 时间段内房地产投资值为 Y_{ij} , 贝叶斯时空模型建模如下:

$$Y_{ij} \sim \text{Poisson}(e_{ij}\theta_{ij}) \quad (1)$$

$$\theta_{ij} = \exp(\alpha_0 + \mu_i + g_j + Psi_{ij} + v_i + \alpha_1 * time1_j + d_i * time1_j + \sum \beta_k * x_{kij}) \quad (2)$$

其中 e_{ij} 是区域 i 在 j 时间段内的预期投资值。这里, 我们采用时间序列模型——ARIMA 模型, 以 2002~2011 年的房地产投资真实值为基准, 预估 2012~2021 年的房地产投资值。ARIMA 模型的全称叫做自回归移动平均模型, ARIMA 模型有三个参数: p, d, q 。其中, p 代表预测模型中采用的时序数据本身的滞后数; d 代表时序数据需要进行几阶差分, 才是稳定的; q 代表预测模型中采用的预测误差的滞后数。在 p, d, q 已知的情况下, ARIMA 模型的数学形式如下:

$$\hat{y}_i = \mu + \phi_1 * y_{i-1} + \dots + \phi_p * y_{i-p} + \theta_1 * e_{i-1} + \dots + \theta_q * e_{i-q} \quad (3)$$

θ_{ij} 是区域 i 在第 j 年的综合投资影响系数。 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:

$$\mu_i | \mu_{j, j \neq i} \sim N \left(\sum_{j=1}^{n_i} w_{ij} \mu_j, \sigma_\mu^2 / n_i \right) \quad (4)$$

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, 这里我们探寻的是中国大陆 31 个省、自治区和直辖市的空间同质性, 所以, 若两省级单位相邻则令 $w_{ij} = 1$, 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:

$$v_i \sim N(0, \sigma_v^2) \quad (5)$$

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:

$$\theta_{ij} = \exp(\alpha_0 + s_i + b_0 t^* + v_j + b_{it} t^* + \varepsilon_{ij} + \sum \beta_k x_{kij}) \quad (6)$$

$$s_i = \exp(\alpha_1 + \mu_i + v_i) \quad (7)$$

$$t^* = t - t_{mean} \quad (8)$$

这里, α_0 表示截距, s_i 表示房地产投资的空间分布, $(b_0 t^* + v_j)$ 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。 $b_{it} t^*$ 代表各区域各自的时间变化趋势; 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。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

2012~2021 年中国大陆地区房地产投资值数据来源于 2012~2021 年中国统计年鉴 (<http://www.stats.gov.cn/sj/ndsj/>), 诸影响因子来源于 2012~2021 年中国统计年鉴。共引入人均 GDP、第三产业产值、高中学校数量、初中学校数量、小学学校数量、人均工资、人口密度、各省客运总量、各省就业人数、各省居民消费水平、医疗床位利用情况共计 11 个影响因子。中国基础地理信息数据来源于 2019 中国地图 - 审图号 GS(2019)1822 号。

3.2. 相关性及因子分析

将房地产投资值和各感兴趣的影响因子进行两两相关性分析, 结果见表 1:

Table 1. Correlation analysis between real estate investment and various influencing factors

表 1. 房地产投资同各影响因子的相关性分析

影响因子	房地产投资
人口密度	0.40
小学数量	0.39
高中数量	0.77
第三产业	0.94
消费水平	0.38
人均 GDP	0.44
人均工资	0.37
医疗床位情况	0.49
就业人数	0.96
初中数量	0.61
客运总量	0.86

相关性分析完成后, 我们对结果进行检测, 检测结果见表 2:

Table 2. Detection of correlation analysis results between real estate investment value and various influencing factor indicators

表 2. 房地产投资值与各影响因子指标相关性分析结果检测

影响因子	P 值
人均 GDP	0.01
人均工资	0.01
医疗床位情况	0.01
就业人数	0
初中数量	0
客运总量	0
人口密度	0
小学数量	0.03
高中数量	0.01
第三产业	0.04
消费水平	0.02

通过上述分析, 我们可以看到相关性分析的结果为各影响因子对房地产投资的影响值均大于 0, 而且相关性分析的检测结果显示, 检测结果 P 值均小于 0.05, 因而均具有统计学意义。为消除多重共线性, 进行复相关性分析及方差膨胀性因子的获取, 结果见表 3。

Table 3. Calculation results of variance inflation factor (VIF) for each influencing factor
表 3. 各影响因子的方差膨胀因子(VIF)计算结果

影响因子	VIF
人均 GDP	4.2
人均工资	3.5
医疗床位情况	1.5
就业人数	5.0
初中数量	5.3
客运总量	7.2
人口密度	1.3
小学数量	3.3
高中数量	6.1
第三产业	4.9
消费水平	2.5

通过表 3 我们可以发现, 各个影响因子的 VIF 值均大于 0 小于 10, 所以各影响因子之间不存在多重共线性。接下来, 在完成数据的归一化后, 我们对各影响因子进行因子分析, 降维处理, Bartlett 球度检验和 KMO 检验结果见表 4:

Table 4. Bartlett and KMO tests
表 4. 巴特利特和 KMO 检验

KMO 取样适切性量数		0.746
巴特利特球形度检验	近似卡方	401.305
	自由度	55
	显著性	<0.001

由表 4 可知, KMO 值大于 0.7, 且巴特利特球形度检测的显著性是小于 0.05 的, 所以适合进行因子分析。因子分析成分矩阵碎石图和空间组件图分别见表 5:

Table 5. Component matrix after rotation
表 5. 旋转后的成分矩阵

成分	1	2	3
高中数量	0.941	-0.074	0.208
客运总量	0.935	0.115	-0.003
初中数量	0.895	-0.290	0.243
就业人数	0.891	0.333	-0.101
第三产业	0.796	0.538	-0.098
小学数量	0.703	-0.443	0.286
医疗床位情况	0.499	0.298	0.482

续表

人均 GDP	0.051	0.974	-0.036
消费水平	0.016	0.970	0.086
人均工资	0.031	0.860	-0.264
人口密度	0.048	-0.191	0.890

通过表 4 我们可以将十一个影响因子分成三类, 其中高中学校数量、初中学校数量、小学学校数量、客运总量、第三产业产值以及医疗床位利用情况为一类, 重新命名为社会因子; 人口密度单独为一类, 即人口密度因子; 人均 GDP、人均工资和消费水平为一类, 重命名为经济因子。接下来, 我们将以社会因子、经济因子、人口密度因子为影响因子, 建立贝叶斯时空模型。

3.3. ARIMA 模型预测

采集 2002~2011 年中国大陆各省、自治区和直辖市的房地产投资值, 运用 ARIMA 模型预测 2012~2021 年中国大陆各省、自治区和直辖市的房地产投资值, 作为本文研究的预估值, 分别整理 31 个省、自治区和直辖市 2002~2011 年度的房地产投资数据, 选择合适的 p 、 d 、 q 值, 对中国大陆地区各个省、直辖市、自治区以 2002~2011 年数据为基准, 进行时间序列的预测, 结果见表 6。

Table 6. Estimated real estate investment values from 2012 to 2021

表 6. 2012~2021 房地产投资预估值

省市	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
北京	3252.9	3459.5	3666.2	3872.8	4079.5	4286.1	4492.8	4699.4	4906.1	5112.7
天津	1360.8	1547.5	1746.5	1957.9	2181.6	2417.6	2666	2925.7	3199.7	3485.1
河北	3238.3	3503.8	3769.4	4034.9	4300.5	4566	4831.6	5097.1	5362.7	5628.2
山西	1030.5	1065.2	1094.1	1117.2	1133.5	1134.5	1146	1151.4	1210.4	1251.6
内蒙	993.3	1061.9	1120.3	1168.7	1207.1	1235.3	1253.4	1261.5	1259.5	1247.4
辽宁	5939.2	6660.9	7382.6	8104.3	8826	9547.7	10269.4	10991.1	11712.8	12434.5
吉林	1262.1	1295.6	1327.1	1356.5	1383.8	1409.1	1432.3	1453.5	1472.7	1489.7
黑龙江	1491.8	1557.7	1627	1699.7	1775.6	1855.3	1938.1	2024.3	2113.9	2206.8
上海	2560.1	2733.9	2907.6	3081.4	3255.2	3429	3602.7	3776.5	3950.3	4124.1
江苏	6912.3	7554	7994	8351.7	8660.7	8936.8	9188.7	9421.8	9639.8	9845.3
浙江	5238.7	5401.9	5596.9	5823.8	6082.7	6373.4	6696	7050.4	7436.8	7855
安徽	3134.3	3576	3922.3	4573.4	4729.1	4889.5	5054.6	5225.5	6398.9	6578.1
福建	3463.5	3878	4153.1	4574.7	4595.4	4735.4	4890.2	5244.4	5867.1	6293.1
江西	1104	1222	1299	1360.8	1413.8	1461.1	1504.1	1543.9	1581	1615.9
山东	5578.4	6134.6	6517.6	6828.9	7098.2	7338.8	7558.4	7761.4	7951.7	8130.9
河南	3405.3	3978.8	4599.9	5268.5	5984.6	6748.3	7559.6	8418.4	9324.7	9878.6
湖北	3007.3	3608	4288.4	5048.5	5888.4	6808	7807.3	8886.4	9283.7	10045.2
湖南	2154.9	2294.8	2434.7	2714.5	2854.4	2994.3	3134.2	3274.1	4414	4825.5

续表

广东	5580.5	6087	7593.5	8099.9	8606.4	9112.9	9619.4	10125.9	11632.4	12138.9
广西	1000.9	1166	1348.4	1547.9	1764.7	1998.7	2249.9	2518.3	2803.9	3106.7
海南	318.9	389.5	467.4	552.6	645.1	744.9	851.9	966.2	1087.8	1216.6
重庆	3017.2	3328.1	3551.2	3734.8	3894.5	4037.7	4168.8	4290.3	4404.1	4511.5
四川	3165.9	3375.2	3584.5	3793.8	4003	4512.3	5021.6	5630.9	6250.2	7049.5
贵州	1055.7	1316.1	1605.7	1924.7	2273	2650.6	3057.5	3493.7	3959.3	4454.2
云南	1765.6	2002.7	2579.7	3096.5	3496.6	3653.1	4035.1	4649.6	5286	5962.2
西藏	12.5	14.5	16.5	18.5	21	23	25	26.5	29	31
陕西	2166	2419.3	1659.7	2894.3	3125.6	3354.9	3582.7	3809.5	4035.6	4261.1
甘肃	548.6	570.5	591.3	611.2	630.2	648.1	665.1	681.1	696.1	710.2
青海	201.5	225.8	252.4	281.3	312.5	346	381.7	419.7	460	502.5
宁夏	515.9	606.7	700.3	796.7	895.9	997.9	1102.7	1210.3	1320.7	1433.9
新疆	692.7	758.4	819.4	878.3	936	993	1049.6	1105.9	1162	1218

3.4. 贝叶斯时空模型建模

3.4.1. BYM 建模

假设中国大陆地区各省、自治区和直辖市 i 第 j 年的房地产投资发展程度为 dl_{ij} , 社会因子为 x_{1ij} , 经济因子为 x_{2ij} , 人口密度为 x_{3ij} , 各影响因子的协变量系数为 β_1 、 β_2 、 β_3 。那么构建的 BYM 数学模型如下:

$$Y_{ij} \sim \text{Poisson}(e_{ij} dl_{ij})$$

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

在 WinBUGS1.4 中进行大约 250,000 次迭代, 模型基本达到了收敛状态。模型达到收敛后, 我们可以得出中国大陆 31 省、自治区和直辖市 2012 年到 2021 年的房地产投资发展程度的后验均值(见表 7), 同时可以得出五个影响因子对房地产投资的影响评估(见表 8)。

Table 7. The development of real estate investment in different regions of Chinese Mainland from 2012 to 2021 under the BYM model

表 7. BYM 模型下 2012~2021 年中国大陆各地区房地产投资发展程度

省市	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
北京	0.97	0.93	0.95	0.98	1.30	0.92	1.04	1.03	0.93	0.90
天津	0.99	1.01	0.82	0.88	0.84	0.89	0.84	1.03	0.96	1.55
河北	2.76	0.83	1.03	1.39	1.01	0.72	0.85	1.02	0.94	0.83
山西	0.88	1.01	0.96	1.42	1.23	1.39	0.97	0.96	1.03	1.03
内蒙	0.96	1.15	1.10	0.95	0.96	0.89	0.97	0.91	1.15	1.07
辽宁	1.39	3.06	0.91	1.14	1.47	1.24	0.83	0.93	1.27	1.10
吉林	0.92	1.09	1.01	0.97	1.08	1.25	1.25	0.72	0.78	0.81

续表

黑龙江	1.10	1.03	1.30	1.11	1.10	1.02	0.89	0.95	0.93	1.18
上海	1.01	1.37	3.05	1.02	1.22	1.36	1.11	2.89	1.46	1.22
江苏	1.21	0.94	1.24	1.08	0.96	1.06	1.34	0.92	0.44	0.68
浙江	0.58	1.13	0.98	1.22	0.97	0.98	1.12	0.86	0.91	0.84
安徽	0.96	1.05	1.24	3.07	1.01	1.27	1.15	0.86	2.36	0.86
福建	1.25	1.19	0.80	1.14	0.98	1.05	1.09	1.41	0.94	0.24
江西	0.73	0.49	1.14	1.03	1.23	0.97	0.99	1.25	0.89	1.03
山东	0.73	1.04	1.20	1.36	2.76	0.96	1.32	0.95	0.77	2.15
河南	0.88	1.35	1.27	0.82	0.99	0.86	0.92	1.06	1.03	0.72
湖北	0.24	0.64	0.44	1.12	1.08	1.29	1.15	1.01	1.38	0.90
湖南	1.05	0.67	1.14	1.33	1.34	2.75	0.99	1.14	0.83	0.76
广东	1.61	0.93	1.45	1.18	0.66	1.04	0.86	0.91	0.93	1.20
广西	0.70	0.25	0.82	0.49	1.12	1.20	1.49	1.18	1.01	1.45
海南	1.00	0.93	0.60	1.26	1.50	1.34	2.01	1.02	1.13	0.77
重庆	0.81	3.35	0.99	1.67	0.92	0.41	0.98	0.82	0.93	0.85
四川	1.44	0.82	0.26	0.90	0.48	1.12	1.28	1.52	1.28	1.08
贵州	1.50	1.11	0.89	0.58	1.36	1.57	1.52	1.38	1.04	1.17
云南	0.86	0.89	4.50	1.03	1.84	0.97	0.34	0.97	0.80	0.82
西藏	0.86	1.51	0.93	0.26	0.99	0.47	1.19	1.37	1.54	1.10
陕西	1.03	1.50	1.19	0.83	0.53	1.11	1.49	1.37	1.23	0.99
甘肃	1.17	0.86	0.85	5.31	1.09	1.94	0.92	0.33	1.08	0.81
青海	0.80	0.89	1.55	0.99	0.23	1.03	0.43	1.22	1.37	1.58
宁夏	1.10	0.98	1.56	1.21	0.80	0.61	1.13	1.44	1.20	1.13
新疆	0.97	1.11	0.76	0.72	4.26	1.04	2.14	0.88	0.33	1.23

Table 8. Impact assessment of various influencing factors under the BYM model**表 8.** BYM 模型下各影响因子的影响评估

协变量系数	结果
beta1	0.04
beta2	0.19
beta3	0.12

从表 8 的结果来看, 三种影响因子对房地产投资发展程度都是存在着正相关的影响, 经济因子对房地产投资的影响程度最大, 社会因子的影响程度较小。

3.4.2. FBM 建模

参照 BYM 模型, 构建 FBM 数学模型如下:

$$Y_{ij} \sim \text{Poisson}(e_{ij} dl_{ij})$$

$$dl_{ij} = \exp(\alpha_0 + s_i + b_0 t^* + v_j + b_1 t^* + \varepsilon_{ij} + \beta_1 * x_{1ij} + \beta_2 * x_{2ij} + \beta_3 * x_{3ij})$$

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

在 WinBUGS1.4 中迭代 250,000 次, 模型基本达到收敛。投资发展程度见表 9, 各影响因子的评估结果见表 10。

Table 9. Development degree of real estate investment in different regions of Chinese Mainland from 2012 to 2021 under the FBM model

表 9. FBM 模型下 2012~2021 年中国大陆各地区房地产投资发展程度

省市	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
北京	0.97	0.93	0.95	0.98	1.30	0.92	1.04	1.03	0.93	0.90
天津	0.99	1.01	0.82	0.88	0.84	0.89	0.84	1.03	0.96	1.55
河北	2.76	0.83	1.03	1.39	1.01	0.73	0.85	1.02	0.94	0.83
山西	0.88	1.01	0.96	1.42	1.23	1.39	0.97	0.96	1.03	1.03
内蒙	0.96	1.15	1.10	0.96	0.96	0.89	0.97	0.91	1.15	1.07
辽宁	1.38	3.06	0.91	1.14	1.47	1.24	0.81	0.93	1.27	1.10
吉林	0.92	1.09	1.01	0.97	1.08	1.25	1.25	0.72	0.78	0.81
黑龙江	1.10	1.03	1.30	1.11	1.10	1.02	0.89	0.95	0.93	1.18
上海	1.01	1.36	3.05	1.02	1.22	1.36	1.10	2.91	1.46	1.22
江苏	1.22	0.94	1.24	1.08	0.96	1.06	1.34	0.93	0.44	0.68
浙江	0.59	1.13	0.98	1.22	0.97	0.98	1.12	0.86	0.91	0.84
安徽	0.96	1.05	1.23	3.07	1.01	1.27	1.15	0.86	2.46	0.86
福建	1.25	1.19	0.80	1.14	0.98	1.05	1.09	1.41	0.94	0.24
江西	0.74	0.49	1.14	1.03	1.23	0.97	0.99	1.25	0.89	1.03
山东	0.73	1.04	1.20	1.36	2.76	0.96	1.32	0.95	0.77	2.13
河南	0.88	1.35	1.27	0.81	0.99	0.86	0.92	1.06	1.03	0.72
湖北	0.24	0.65	0.44	1.13	1.08	1.29	1.15	1.01	1.38	0.90
湖南	1.05	0.67	1.14	1.33	1.34	2.75	0.99	1.14	0.83	0.76
广东	1.64	0.92	1.45	1.18	0.66	1.05	0.86	0.91	0.93	1.20
广西	0.71	0.25	0.82	0.49	1.12	1.20	1.49	1.18	1.01	1.44
海南	1.00	0.93	0.60	1.26	1.50	1.34	2.01	1.02	1.14	0.77
重庆	0.81	3.43	0.99	1.67	0.92	0.41	0.98	0.82	0.93	0.85
四川	1.44	0.83	0.26	0.91	0.48	1.12	1.27	1.52	1.28	1.08
贵州	1.45	1.11	0.89	0.58	1.36	1.57	1.51	1.38	1.04	1.17
云南	0.86	0.89	4.56	1.02	1.84	0.97	0.34	0.97	0.80	0.82
西藏	0.86	1.51	0.93	0.26	0.99	0.47	1.19	1.37	1.53	1.10
陕西	1.03	1.50	1.19	0.83	0.53	1.11	1.49	1.37	1.23	0.99

续表

甘肃	1.17	0.86	0.85	5.39	1.09	1.94	0.92	0.33	1.09	0.81
青海	0.80	0.89	1.55	0.99	0.23	1.03	0.43	1.22	1.37	1.58
宁夏	1.10	0.98	1.56	1.21	0.80	0.61	1.13	1.44	1.20	1.13
新疆	0.97	1.11	0.76	0.72	4.32	1.04	2.14	0.88	0.33	1.23

Table 10. Impact assessment of various influencing factors under the FBM model**表 10.** FBM 模型下各影响因子的影响评估

协变量系数	结果
beta1	0.01
beta2	0.08
beta3	0.05

从表 10 的评估结果来看, 结论与 BYM 相似。根据 DIC 值, 我们可以对 BYM 和 FBM 模型进行一个对比, 已知 DIC 值越小, 模型拟合效果越好。DIC 值见表 11。

Table 11. Comparison of fit levels between two models**表 11.** 两种模型的拟合程度比较

模型	Dbar	pD	DIC
BYM	3288.970	307.846	3596.820
FBM	3286.920	265.841	3552.760

注: DIC = Dbar + pD。

通过表 11, 我们可以较为容易的得出结论: FBM 模型拟合程度要优于 BYM。

显然, 房地产投资发展程度大于 1, 则说明该地区的房地产投资额超过预期, 反之则是低于预期。在社会因子、经济因子、人口密度因子的综合影响下, 中国大陆地区各省、自治区、直辖市的房地产投资与预期值或多或少的发生了偏离。从表 11 结果来看, 中国大陆地区绝大多数省、自治区、直辖市在 2019 年后的这 2 年中房地产投资额都低于预期, 这也契合了新冠肺炎疫情爆发的三年以来, 房地产市场的不景气。

每年的房地产投资发展程度评估只能体现中国大陆各地区当年的房地产投资发展状况以及空间分布, 对于各地区整体在 2012~2021 这 10 年的整体发展热度以及局部发展趋势无法作出分析。因而, 我们可以根据各省级行政区的空间相对大小分为热点、温点和冷点区域。区域分类的准则基于 Richardson 等提出的分类准则[21]: 如果 i 地区的房地产投资空间相对大小大于 1 的后验概率 $P(\exp(s_i) > 1 | data)$ 大于 0.8, 则属于热点区域; 若小于 0.2 则属于冷点区域; 若介于 0.2 到 0.8 之间, 则属于温点区域。此外, 也可以按照这个分类准则, 将各地区的局部变化趋势分为快速增长区域、稳定增长区域和增长放缓区域, 即若 i 地区的局部变化趋势大于 0 的后验概率 $P(b_i > 0 | data)$ 大于 0.8 则属于快速增长区域, 小于 0.2 则属于增长放缓区域, 介于 0.2 到 0.8 之间则属于稳定增长区域。运用 FBM 模型分析, 结果分别如表 12、表 13 所示:

根据表 12 和表 13, 从总体发展热度上看, 中国大陆地区各个省、自治区和直辖市的房地产投资发展热度都在大于 0.2, 小于 0.8 这个区间范围内, 因而 2012~2021 年这十年的总体房地产投资发展热度各

地区均属于温点地区, 发展平稳。从局部发展趋势上来看, 2012~2021 年这十年间各地区的局部发展趋势也在[0.2, 0.8]这个区间范围内, 因而各地区房地产投资的局部发展趋势也是处于稳定增长, 总量在稳定增长而增速在逐渐降低。

Table 12. The overall development heat of real estate investment in all regions of Chinese Mainland is greater than 1 posterior probability

表 12. 中国大陆各地区房地产投资总体发展热度大于 1 后验概率

省市	大于 1 的后验概率估计
北京	0.29
天津	0.27
河北	0.30
山西	0.31
内蒙	0.31
辽宁	0.38
吉林	0.30
黑龙江	0.33
上海	0.51
江苏	0.31
浙江	0.29
安徽	0.38
福建	0.28
江西	0.29
山东	0.38
河南	0.29
湖北	0.25
湖南	0.31
广东	0.31
广西	0.22
海南	0.33
重庆	0.29
四川	0.25
贵州	0.31
云南	0.26
西藏	0.24
陕西	0.31
甘肃	0.32
青海	0.25
宁夏	0.32
新疆	0.30

Table 13. Partial development trend of real estate investment in various regions of Chinese Mainland is greater than 0 posterior probability**表 13.** 中国大陆各地区房地产投资局部发展趋势大于 0 后验概率

省市	大于零的后验概率估计
北京	0.45
天津	0.43
河北	0.36
山西	0.53
内蒙	0.41
辽宁	0.54
吉林	0.35
黑龙江	0.48
上海	0.72
江苏	0.25
浙江	0.40
安徽	0.66
福建	0.23
江西	0.53
山东	0.69
河南	0.35
湖北	0.56
湖南	0.49
广东	0.47
广西	0.63
海南	0.59
重庆	0.30
四川	0.51
贵州	0.65
云南	0.32
西藏	0.49
陕西	0.51
甘肃	0.40
青海	0.37
宁夏	0.52
新疆	0.47

4. 结论

本文通过运用 BYM 以及 FBM 两种贝叶斯时空统计模型, 对 2012~2021 年中国大陆地区各省、自治

区、直辖市的房地产投资的发展程度作了评估, 并对社会因子、经济因子、人口密度因子三个综合影响因子对房地产投资的影响也作出了相应的评估, 其中经济因子对房地产投资的影响程度最大, 社会因子的影响程度较小。利用 FBM 模型, 对 31 个省、自治区、直辖市的房地产投资发展程度作出了相应的评估分析。可以看出房地产投资发展程度越大说明该地区的投资风险相对较小, 房地产实际投资额是超过我们的预期值的, 该地区的房地产发展良好, 因而适合将资金注入该地区; 如果房地产投资发展程度较小, 略微大于或者等于 1, 那么说明该地区房地产的投资相对饱和, 同我们的预期较为吻合; 如果房地产投资发展程度小于 1, 说明该地区存在一定的投资风险, 这里的房地产投资额是低于我们的预期值的。从总体发展热度上来看, 以 2012~2021 这 10 年的跨度综合评判, 中国大陆地区各个省、自治区、直辖市的房地产投资发展热度都是属于温点区域。同样以这 10 年为准, 各地区房地产投资局部发展趋势, 我们可以看出, 大陆各地区的局部发展趋势也是趋于稳定增长。从 2021 年 31 个省、自治区、直辖市的发展程度来看, 在总体发展冷热度和局部发展趋势都趋于稳定发展的情况的前提下, 天津、山西、内蒙古、辽宁、黑龙江、上海、江西、山东、广东、广西、四川、贵州、西藏、青海、宁夏、新疆的房地产投资发展程度较好(大于 1), 在未来的投资规划中可以重点考察上述省份。同时, 我们建议相关部门在作房地产投资发展规划时应着重考虑区域的人均 GDP、人均工资、居民消费水平以及当地的人口情况, 这些因素影响地区的房地产开发和投资。

同时, 本文通过对比 BYM 模型和 FBM 模型的 DIC 值, 得出结论, FBM 模型在房地产投资评估的模型分析中是要优于 BYM 模型, 而且 FBM 模型能够更好地分析研究对象的时空演化, 因而在对房地产投资类的时空建模问题中可以引入 FBM 模型加以解决。

参考文献

- [1] 黄徐亮, 徐海东, 倪鹏飞. 带动还是挤出? 房地产投资对其他固定资产投资的影响研究[J]. 经济学家, 2023(6): 78-87.
- [2] 刘芸. 长三角重点城市房地产投资潜力评价[J]. 现代营销(下旬刊), 2023(9): 96-99.
- [3] 韦兰宁, 唐晓莲. 房地产投资与就业关系研究[J]. 合作经济与科技, 2022(16): 64-66.
- [4] Weinberg, D.H., Friedman, J.M. and Stephen, K.M. (1981) Intraurban Residential Mobility: The Role of Transactions Costs, Market Imperfections, and Household Disequilibrium. *Journal of Urban Economics*, 9, 332-348. [https://doi.org/10.1016/0094-1190\(81\)90031-0](https://doi.org/10.1016/0094-1190(81)90031-0)
- [5] 乔林, 孔淑红. 我国不同发展层次城市房价影响因素的差异分析及对策研究[J]. 特区经济, 2012(5): 210-212.
- [6] 黄俊宝. 基于贝叶斯层次时空模型的中国区域科技创新能力研究[D]: [硕士学位论文]. 深圳: 深圳大学经济学院, 2020.
- [7] 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>
- [8] 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, 17(34): 55-57.
- [9] Wang, W., Li, J., 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>
- [10] Congdon, P. (2000) Monitoring Suicide Mortality: A Bayesian Approach. *European Journal of Population/Revue Européenne de Démographie*, 16, 251-284. <https://doi.org/10.1023/A:1026587810551>
- [11] Haining, R., Law, J. and Griffith, D. (2009) Modelling Small Area Counts in the Presence of Overdispersion and Spatial Autocorrelation. *Computational Statistics & Data Analysis*, 53, 2923-2937. <https://doi.org/10.1016/j.csda.2008.08.014>
- [12] Yu, Q., Scribner, R., Carlin, B., et al. (2008) Multilevel Spatio-Temporal Dual Change-point Models for Relating Alcohol Outlet Destruction and Changes in Neighborhood Rates of Assaultive Violence. *Geospatial Health*, 2, 161-172. <https://doi.org/10.4081/gh.2008.240>

-
- [13] Bernardinelli, L., Clayton, D., Pascutto, C., *et al.* (1995) Bayesian Analysis of Space-Time Variation in Disease Risk. *Statistics in Medicine*, **14**, 2433-2443. <https://doi.org/10.1002/sim.4780142112>
- [14] Kleinschmidt, I., Sharp, B., Mueller, I., *et al.* (2002) Rise in Malaria Incidence Rates in South Africa: A Small-Area Spatial Analysis of Variation in Time Trends. *American Journal of Epidemiology*, **155**, 257-264. <https://doi.org/10.1093/aje/155.3.257>
- [15] Aguero-Valverde, J. and Jovanis, P.P. (2006) Spatial Analysis of Fatal and Injury Crashes in Pennsylvania. *Accident Analysis & Prevention*, **38**, 618-625. <https://doi.org/10.1016/j.aap.2005.12.006>
- [16] Osnas, E.E., Heisey, D.M., Rolley, R.E., *et al.* (2009) Spatial and Temporal Patterns of Chronic Wasting Disease: Fine-Scale Mapping of a Wildlife Epidemic in Wisconsin. *Ecological Applications*, **19**, 1311-1322. <https://doi.org/10.1890/08-0578.1>
- [17] Law, J., Quick, M. and Chan, P. (2014) Bayesian Spatio-Temporal Modeling for Analysing Local Patterns of Crime over Time at the Small-Area Level. *Journal of Quantitative Criminology*, **30**, 57-78. <https://doi.org/10.1007/s10940-013-9194-1>
- [18] Li, J. and Han, X. (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>
- [19] 韩秀兰, 李俊明. 中国城市房价时空特征与影响机制研究——基于贝叶斯分域时空模型的实证[J]. *数理统计与管理*, 2018, 37(5): 940-950.
- [20] 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.
- [21] Richardson, S., Thomson, A., Best, N., *et al.* (2004) Interpreting Posterior Relative Risk Estimates in Disease-Mapping Studies. *Environmental Health Perspectives*, **112**, 1016-1025. <https://doi.org/10.1289/ehp.6740>