引用格式:董文钱,董良,向琳,等.基于贝叶斯时空模型的城管事件时空变化分析[J].地球信息科学学报,2020,22(5):1073-1082. [Dong W Q, Dong L, Xiang L, et al. Spatiotemporal variability of urban management events based on the Bayesian spatiotemporal model[J]. Journal of Geo-information Science, 2020,22(5):1073-1082.] **DOI:**10.12082/dqxxkx.2020.190413

基于贝叶斯时空模型的城管事件时空变化分析

董文钱[1],董　良[2,3],向　琳[1*],陶海军[1],赵传虎[4],曲寒冰[2,3]

1. 中国计量大学信息工程学院,杭州 310018; 2. 北京市科学技术研究院,北京 100089; 3. 北京市新技术应用研究所,北京 100094; 4. 河北工业大学人工智能与数据科学学院,天津 300401

Spatiotemporal Variability of Urban Management Events based on the Bayesian Spatiotemporal Model

DONG Wenqian[1], DONG Liang[2,3], XIANG Lin[1*], TAO Haijun[1], ZHAO Chuanhu[4], QU Hanbing[2,3]

1. College of Information Engineering, China Jiliang University, Hangzhou 310018, China; 2. Beijing Academy of Science and Technology, Beijing 100089, China; 3. Beijing Institute of New Technology Applications, Beijing 100094, China; 4. School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China

Abstract: Urban management events are regional and periodic. The spatiotemporal laws and potential impact factors implied in urban management events are vital for improving urban management. However, research on the temporal and spatial changes of urban management events and influencing factors are rare. In this paper, by using the Bayesian space-time model, we modeled and analyzed the temporal and spatial evolution characteristics of three types of city management events-street order, urban environment, and publicity advertising-in the P district of H city, Northwest China, and explored the impact of urban management events as well as the underlying impact factors. We found that: (1) There were spatial differences in the relative risk distribution of the three types of urban management events. The street order type was concentrated in the residential and commercial areas of the city, while the urban environment type was concentrated in the residential areas of the city. The advertising type was mainly concentrated in the commercial areas of the city. The spatial risk posterior probability estimate indicated that the above two regions are hotspots of urban management events. (2) The relative risks of urban management events were more prominent on Tuesdays, Fridays, and Saturdays, but there was no obvious monotony in general trends. Meanwhile, the hourly trends had irregular fluctuation, everyday from 8 to 10 and from 14 to 15, it was a period of the high incidence of urban management events, and its relative risk was much higher than other periods. (3) For different built environments, the potential impacts of these factors were quite different. The relative risk of urban management events was significantly associated with restaurants, transportation, and living services, all positively correlated. (4) The relative risk of urban management events presented obvious spatial and temporal heterogeneity. and it is reasonable and necessary to consider the impact of spatial and temporal effects when analyzing urban management events data. Our findings are meaningful for

收稿日期:2019-07-30;**修回日期:**2019-11-20.

基金项目:国家自然科学基金重大研究计划重点支持项目(NSF91746207);科技部国家重点研发计划项目(2018YFF0301000、2018YFC0809700、2018YFC0704800)。[**Foundation items:** National Natural Science Foundation of China, No.91746207; National Key Research and Development Program of China, No.0301000, 0809700, 0704800.]

作者简介:董文钱(1995—),男,浙江温州人,硕士生,主要从事时空数据挖掘研究。E-mail: wqdong.chn@live.com

***通讯作者:**向　琳(1975—),女,浙江杭州人,副教授,主要从事人工智能应用研究。E-mail: xianglin@cjlu.edu.cn

relavant government departments to make effective policies to control and reduce the relative risk of urban management events, especially for the study area.

Key words: urban management events; Bayesian spatiotemporal model; structured effect; unstructured effect; relative risk; spatiotemporal variability; hot/cold spots; INLA

***Corresponding author:** XIANG Lin, E-mail: xianglin@cjlu.edu.cn

摘要：城市管理事件具有显著的地域特征和周期性特征，挖掘城管事件中隐含的时空规律和潜在影响因子对城市管理具有重要指导意义。然而，目前针对城管事件时空变化及影响因素驱动效应的研究仍较为少见。本文基于贝叶斯时空模型，针对西北某地H市P区的街面秩序类、市容环境类以及宣传广告类城管事件在一周之中的时空演变特征进行建模分析，并探究了影响城管事件案发风险的潜在影响因子。研究发现：① 在空间上，3类城管事件的相对风险分布存在差异，街面秩序类事件集中在城市的居住功能区和商业功能区，而市容环境类集中在城市的居住功能区，宣传广告类主要集中在城市的商业功能区，空间风险后验概率估计表明，以上2个区域是城管事件的热点区域。② 在时间上，每逢周二、周五以及周六，城管事件的相对风险较为突出，但总体上没有明显的单调性。每天的8—10时和14—15时是城管事件高发的时段，其相对风险远高于其他时段。③ 城市建成环境对城管事件的潜在影响存在差异。研究区域内餐饮、交通、生活服务等城市基础设施与城管事件的关联最为紧密，且都表现为正相关。④ 城管事件的案发风险呈现出明显的时空异质性，渐进性的建模过程表明在分析城管事件数据时考虑空间与时间效应的影响是合理且有必要的。研究表明，贝叶斯时空模型的分析结果满足城管部门日常工作中关于城管事件智能分析与重点管控的需求。

关键词：城管事件；贝叶斯时空模型；结构效应；非结构效应；相对风险；时空变化；冷热点；INLA

1 引言

数字化城市管理是现代城市管理领域的一次突破性整合与创新。在数字化城管系统中，城管网格巡查员与中心控制人员分工协作，网格巡查员在责任片区内巡查和记录城管事件发生的时间、地点、事件类型、处置详情，并将信息发送至数字城管控制中心，由控制中心对事件做进一步的分发，由相关责任单位派员进行事件的后续处理。

然而，随着我国城市现代化建设步伐的加快，城市中各种社会要素的不断聚集和碰撞，城市管理问题的复杂性也随之提高。偏向于数据采集与传递的数字城管显得缺乏积极性和预见性，处于一种被动反应的状态，破解这一难题需要数字城管逐步向智慧城管转型，化被动为主动。与数字城管相比，智慧城管更多的关注城市管理问题中信息的分析、知识与规律的发现。需要指出的是，这种转型不是将数字城管推倒重来，而是深化和发展数字城管，在城市系统中，数字是基础，智慧是提升[1]。在这个背景下，研究数字化城管系统中累积的城市管理事件数据，用数据支持决策对城管部门的管理工作具有指导意义。

城管事件的产生及其特定时空分布模式的形成是社会、经济、环境等多种因素共同作用的结果，背后隐藏着一定的时空规律，探析其中的时空分布模式与演变趋势，研究城管事件与周边城市建成环境的关系，能够促进城市管理资源的合理分配，改善城管部门的工作。已有的城管事件研究多利用事件点的空间坐标，根据点簇的距离或密度来分析城管事件的空间热点分布和聚集尺度等特征，提供了一种直观的可视化分析手段。如佘冰等[2]使用核密度估计与L函数分析了武汉市江汉区城管事件的空间聚集分布模式。靖常峰等[3]使用基于欧氏距离的近邻指数为权重分析了北京市西城区城管事件的空间分布模式，并使用K函数度量了其空间聚集尺度。同时，也有研究使用基于密度的DBSCAN聚类方法对不同类型的城管事件进行聚簇分析[4]，并根据聚簇的空间分布探究了不同聚簇之间的相关性。但以上这些方法本质上属于探索性空间数据分析的范畴，在对隐藏在数据中的关系、模式和趋势进行探索时，多只考虑了城管事件本身的空间属性，而忽略了事件的时间属性、事件所处的外部环境等信息。因此已有的研究多侧重于城管事件空间关系的分析，没有综合分析城管事件数据中的时空关系以及城管事件的外部影响因素，存在一定的不足与局限性。

贝叶斯时空模型是一种为分析时空数据而建立的时空统计模型，在进行统计推断时除了使用总体信息与样本信息还可以利用先验信息，实现时间、空间以及相关影响因素的同时分析。时空模型

通常是分层的，数据嵌套在模型的分层结构之中，时空上邻近的数据隐含着相似的时空关系，而贝叶斯方法非常适合构建分层模型，并且可以有效的估计复杂分层模型的随机参数[5-7]。这种分层结构赋予了贝叶斯时空模型灵活的扩展性，使其能够充分挖掘数据中隐含的时间、空间信息，满足多层次的分析需求[8-9]，因而在社会学、生态学等多个领域得到了广泛的应用。空间流行病学中也有基于贝叶斯时空模型的疾病制图（Disease Mapping）研究[10-12]。

目前使用贝叶斯时空模型对城管事件数据进行时空分析的研究还较为少见。本文以西北某市数字城管系统中的城管事件数据为研究对象，建立一个贝叶斯时空模型。将城管事件视为研究区域内的一种风险事件，研究其相对风险（Relative Risk, RR）的时空分布情况、变化趋势以及研究区域内的城市建成环境对事件相对风险的影响情况，从中探讨贝叶斯时空模型在城管事件数据分析中的应用价值，为城管部门制定区域化管控措施提供辅助支持。

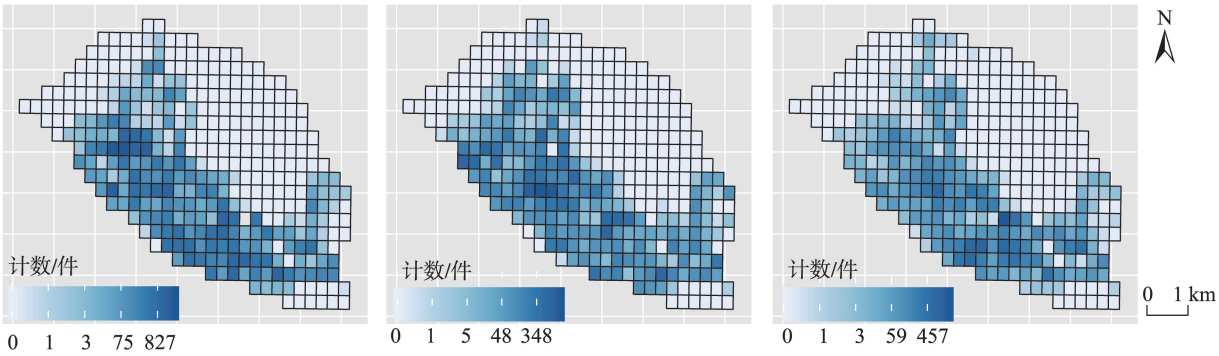
2　研究区概况与数据来源

2.1　研究区域与城管数据

研究区域位于西北某地H市下属的市辖区P区。P区是H市的核心工业区，地属平原，南面环山，北面黄河绕城而过，城市布局向东西两面扩展，下辖5镇1乡7个街道，总人口37万，其中城市人口28万。全区产业结构以重化工为中心，生产要素聚集紧密，城市管理问题较为突出。H市自2013年开始在市辖区内推行城市网格化管理和精细化服务，建立了数字化城管系统。本文选取P区数字化城管子系统累积的历史数据中占比较大的街面秩序类、市容环境类以及宣传广告类事件进行研究。街面秩序类包含的事件内容主要有乱堆物堆料、占道经营、流动摊贩、机动车乱停放等；市容环境类包含的事件内容主要有暴露垃圾、道路不洁、绿化不洁、积存垃圾渣土等；宣传广告类包含的事件内容主要有散发小广告、违章悬挂物品、占道广告牌等。除了事件类别、事件内容等业务字段外，城管事件数据还带有空间地理坐标和时间戳信息。

P区的总面积为358 km^2，城区建成面积为40 km^2，而城管事件数据的采集实际上主要在城区面积内，范围相对有限，因此其在街道或乡镇一级上的行政区划不适合作为城管事件的基本分析单元。数据预处理过程中，首先根据三类事件数据的空间分布确定P区主城区内的7个街道为具体的研究区域，并结合研究区域的实际情况，进一步确定模型的基本计算单元是分辨率为275 m×335 m的网格，整个研究区域可以划分为370个单元。城管事件数据的时间戳跨度为2016年3月11日至2016年7月18日，前后间隔为130 d，为了在更细的粒度上分析城管事件相对风险的变化规律，将城管事件的时间跨度设置为“周”，时间分辨率为“小时”。

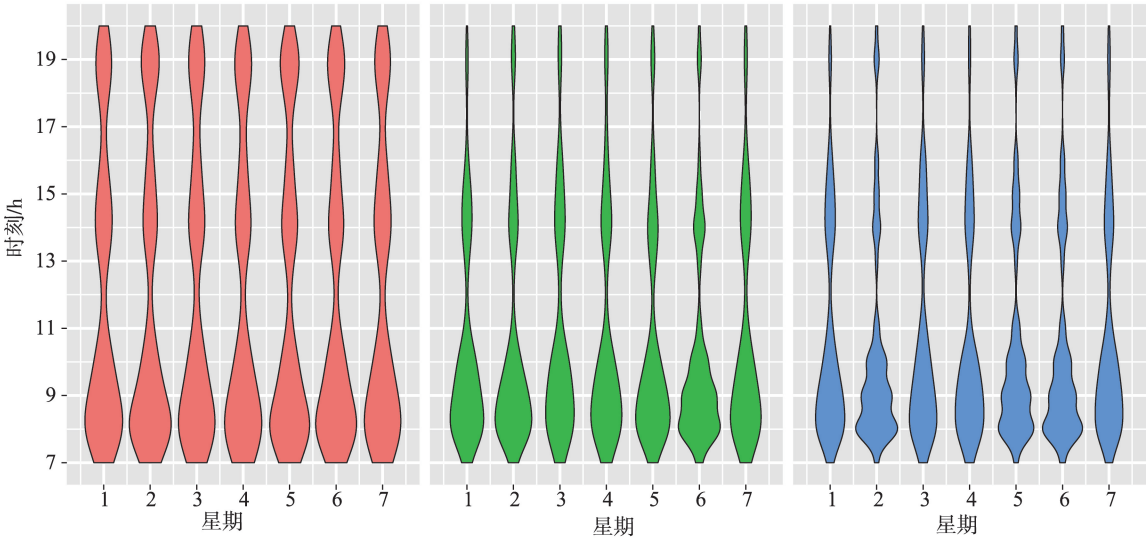
数据预处理过程中，根据上述的时空尺度，以网格为底图，与事件数据进行匹配，获得不同时段网格上的事件计数。图1展示了全部事件数据在网格上的计数结果，3类事件所覆盖的范围大致上是相近的，但是高发频发的区域有所不同，探析其相对风险的空间分布变化是本文的研究目的之一。城管事件在时间上的分布如图2所示，3类事件在时间上的模式非常相近。从横轴上看，星期一到星



(a) 街面秩序事件　(b) 市容环境事件　(c) 宣传广告事件

图1　城管事件在网格上的计数

Fig. 1　Spatial distribution of urban management event number in the study area



(a) 街面秩序 (b) 市容环境 (c) 宣传广告

注:横轴上每个小提琴图的宽窄表示当天数据量的多少,纵轴上小提琴的起伏表示数据分布在时间段上的密度。

图2 城管事件在一周内的时间分布趋势

Fig. 2 Temporal trend of urban management event number in a week

期日每天的事件数量都维持在一个相近的水平,没有特别明显的单调增加或减少的趋势,但可以注意到,星期二、星期五以及星期六对应的小提琴图形宽度稍显突出,说明每逢周二、周五与周六,城管事件的案发数量会有所增长。从纵轴上看,街面秩序类事件从早到晚有3个波峰,而市容环境类事件和宣传广告类事件集中在早晨的时间段内。本文的研究目的之一是探析事件在一周之中、早晚14 h之内相对风险的变化规律。

2.2 POI数据

城市兴趣点(Point of Interest, POI)数据具备丰富的语义特征和时空动态关联特性,是分析自然地理环境、感知人类社会活动规律的重要资源[13-14]。本文使用研究区域内的POI数据作为城管事件相对风险的隐含影响变量,以研究城市建成环境对城管事件相对风险的影响情况。POI数据来自高德地图,数据经过筛选共有8个类别,同样以网格为底图与POI数据进行匹配,得到不同类别POI在网格上的计数作为密度协变量,并进行对数转换,以减少异方差性对模型估计的潜在影响。考虑到少数几个月内,POI信息的变动不会很大,故可以忽略这方面的影响将其加入计算。POI数据的空间分布如图3所示,描述信息和简要统计如表1所示。餐饮服务类、生活服务类以及医疗保健类POI在城区中心聚集,围绕在居民住宅类POI周边,这说明居民住宅区和生活配套设施主要集中在城区的中心部分。购物服务类POI的分布则表明居民住宅区和商业功能区有部分重叠。

3 研究方法

3.1 贝叶斯时空模型的基本思想

在贝叶斯统计思想中,总体中任何一个未知参数都可以看做一个随机变量,并用一个概率分布来表示,这个概率分布就是关于未知参数的先验信息,称为先验分布。进行贝叶斯统计推断时,除了利用总体信息与样本信息之外,还要结合先验信息。

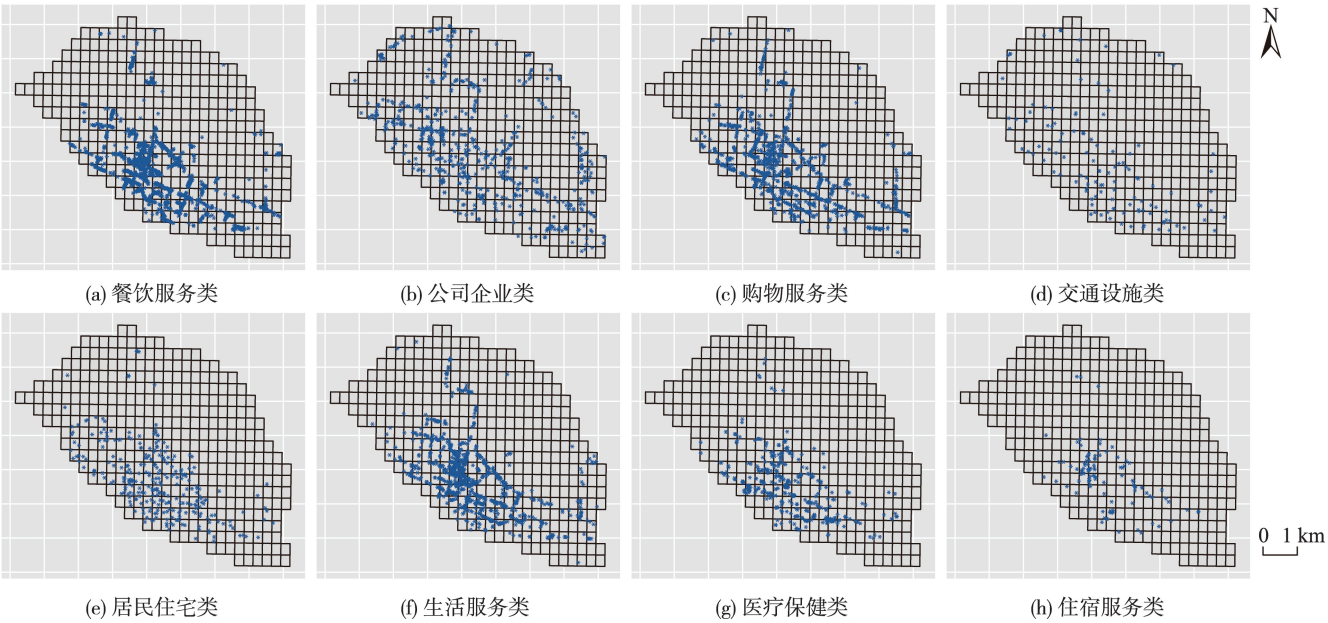
3.2 贝叶斯时空模型的数学形式

假定网格 i 在一周中第 j 天第 k 小时的城管事件数量为 Y_{ijk} 服从泊松分布,在此基础上,贝叶斯时空模型可以分解为协变量项、空间效应项和时间效应项,即:

$$Y_{ijk} \sim Pois\,son(\lambda_{ijk}) \tag{1}$$

$$\log(\lambda_{ijk}) = b_0 + \sum_p \beta_p X_{pijk} + \mu_i + v_i + \gamma_j + \varphi_j + \tau_k + \log(E_{ij}) \tag{2}$$

式中:b_0是截距项,表示整个研究区域城管事件的平均相对风险;X表示对城管事件相对风险变化有影响的协变量,即网格上各类POI的密度值;β是对



(a) 餐饮服务类　(b) 公司企业类　(c) 购物服务类　(d) 交通设施类

(e) 居民住宅类　(f) 生活服务类　(g) 医疗保健类　(h) 住宿服务类

图3　POI数据的空间分布

Fig. 3　Spatial distribution of the POI data

表1　POI数据的简要信息

Tab. 1　Summary of the POI data

POI类别	POI内容	均值(标准差)
餐饮服务类	中餐厅、快餐厅、外国餐厅、休闲餐饮场所等	0.77(1.20)
交通设施类	公交车站、物流站等	0.20(0.37)
居民住宅类	住宅区	0.35(0.58)
购物服务类	便利店、超市、烟酒专卖店、购物相关场所等	0.71(1.07)
生活服务类	美容美发店、洗浴推拿场所、生活服务场所等	0.71(1.00)
住宿服务类	酒店、旅馆、招待所等	0.13(0.38)
医疗保健类	医院、社区诊所、卫生院、药店、药房等	0.37(0.69)
公司企业类	公司、企业、工厂等	0.68(0.78)

应的回归系数，实验中使用方差膨胀因子（Variance Inflation Factor, VIF）来检验协变量之间的多重共线性[15]；μ_i 表示空间结构效应，即各个网格所代表的区域之间的相关性，v_i 表示空间非结构效应，即区域本身的异质性，这两者构成了空间分析中最常用的BYM规范[16]；γ_j 表示跨度是周的时间结构效应，即一周的总体趋势，φ_j 是周的非结构效应，表示一周之中每天的时间异质性；τ_k 是小时项的时间非结构效应，表示一天之中小时的时间异质性。最后，$\log(E_{ij})$ 表示一个偏移项，E_{ij} 是事件数量的均值[17]。

模型中空间结构效应 μ_i 的先验分布假定是一个条件自回归过程（Conditional Autoregressive, CAR），即某个区域中事件发生的相对风险受到附近区域的影响，距离越近，这些空间区域呈现的相对风险就越相似，这种结构的空间特性称为空间自相关[18-19]。相反地，空间非结构效应 v_i 服从一个均值为零的正态分布，表示空间区域的彼此独立。时间效应也由2个部分组成，其中周的时间结构化效应假定为一阶随机游走（Random Walk，RW1），与条件自回归类似，随机游走考虑了时间上的相关性，某个时间点事件发生的相对风险受到相邻时间点的影响，而周的时间非结构项 φ_j 以及小时的时间非结构项 τ_k 均服从一个均值为零的正态分布，表示时间上的异质性[20-21]。

协变量对城管事件相对风险的影响可以通过 $RR=e^{\beta}$ 获得，RR 值高于1表明该变量是正相关风险因子，低于1表示是负相关风险因子，等于1表示不相关因子，以此来识别对城管事件相对风险具有显著影响的风险因子。同样的，还可以分别获得空间效应的总体相对风险估计值 $RR_i=e^{\mu_i+v_i}$，“周”的时间结构效应相对风险估计值 $RR_j=e^{\gamma_j}$、时间非结构效应相对风险估计值 $RR_j=e^{\varphi_j}$，“小时”的时间非结构效应相对风险估计值 $RR_k=e^{\tau_k}$ 等关心的随机效应结果。以空间效应为例，其 RR 值大于1表示该空间单位是高风险区域，RR 值小于1表示该空间单位是低风险区域，RR 值等于1表示平均水平[17]。

3.3 贝叶斯时空模型的计算与比较

贝叶斯时空模型包含3个层次，分别是数据分布、时空过程以及参数或超参数[22]。其中数据分布的似然函数假定为一个泊松分布；在时空过程上，通过CAR和RW等不同子过程的组合来考虑空间和时间上的随机效应；参数的非信息先验能使数据本身对后验分布的影响最大[23]，因此模型中未知方差参数的先验分布均指定为一个反伽马分布。

贝叶斯时空模型的渐进式建模过程如表2所示。首先仅考虑协变量项的组合，即研究区域内不同类别的POI密度对城管事件相对风险的影响，该过程产生模型M1。模型中不包含任何时间或空间效应，在这种情况下，M1近似于一个广义线性模型（Generalized Linear Model, GLM）。然后，考虑到仅有协变量项的组合可能忽略了数据中隐含的空间关系，因此有必要在模型中添加空间效应的影响，该过程产生了模型M2。基于类似的原因，再对M2进行扩展，添加时间效应组件产生模型M3。并通过偏差信息准则（Deviance Information Criterion, DIC）来比较不同模型的拟合优度，选取DIC值最小的模型为优选模型[24]。模型在R环境中使用积分嵌套拉普拉斯逼近（Integrated Nested Laplace Approximations, INLA）来进行贝叶斯推断和计算，相比于常用的马尔科夫链蒙特卡罗（MCMC）方法，INLA的一个主要优点是能在更短的时间内返回准确的参数估计值[25]。

表2 不同组件的候选模型

Tab. 2 Candidate models with different components

模型	模型结构
M1	$\log(\lambda_{ijk}) = b_0 + \sum_p \beta_p X_{pijk} + \log(E_{ij})$
M2	$\log(\lambda_{ijk}) = b_0 + \sum_p \beta_p X_{pijk} + \mu_i + v_i + \log(E_{ij})$
M3	$\log(\lambda_{ijk}) = b_0 + \sum_p \beta_p X_{pijk} + \mu_i + v_i + \gamma_j + \varphi_j + \tau_k + \log(E_{ij})$

4 结果与分析

4.1 模型选择

表3显示了协变量多重共线性分析的结果。所有变量的VIF值都小于10，协变量之间没有严重的多重共线性，因此可以将所有协变量都包含在模型中，分别为3类城管事件实施渐进式的建模。

表3 协变量多重共线检验结果

Tab. 3 Results of multicollinearity evaluation

协变量(POI密度)	VIF	协变量(POI密度)	VIF
餐饮服务类	4.53	生活服务类	5.17
交通设施类	1.37	住宿服务类	2.52
居民住宅类	2.26	医疗保健类	3.21
购物服务类	3.88	公司企业类	1.41

最终各候选模型的DIC信息如表4所示，3类城管事件的候选模型中，充分考虑了时空效应的贝叶斯时空模型M3具有最低的DIC值，说明在模型中纳入空间效应与时间效应的影响是合理的，进一步的结果分析都基于模型M3。

表4 不同候选模型的DIC值

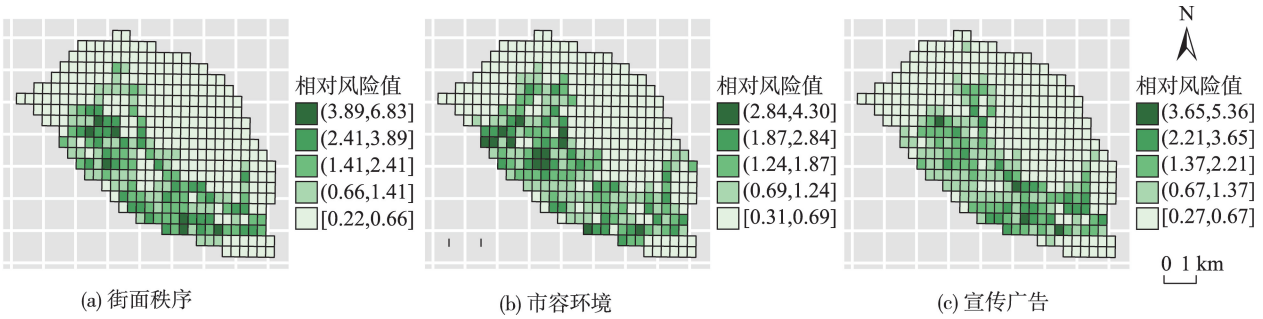
Tab. 4 DIC values for the various candidate models

事件类别	候选模型	DIC
街面秩序	M1	36 873.15
	M2	26 182.07
	M3	24 851.99
市容环境	M1	27 093.02
	M2	21 791.38
	M3	19 099.85
宣传广告	M1	29 935.99
	M2	23 364.75
	M3	20 375.71

4.2 城管事件相对风险的空间分布

空间效应的相对风险估计值 $RR_i = e^{\mu_i + v_i}$ 结果如图4所示，3类城管事件呈现出不同的空间聚集特征。这或许和不同区域隐含的城市功能不同有关，宣传广告类事件主要聚集在研究区域的东南方向，该区域分布着城市广场、商业城、美食城、建材市场等社会经济活动密集的商业建筑，是P区主要的商业中心；市容环境类事件主要聚集在P区的城区中心位置，该区域集中了主要的居民住宅小区和生活配套设施，是P区城市居住人口最为密集的区域。而街面秩序类事件则反映出城市居民的某种活动规律：从居住区前往商业区工作或者消费。

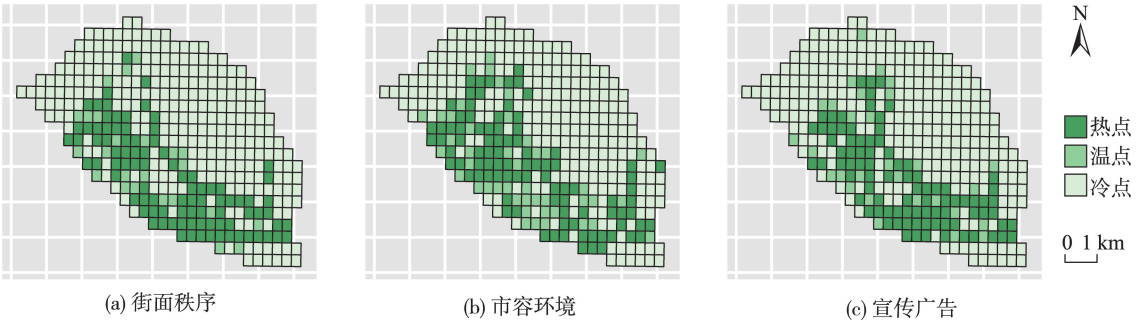
通过计算空间相对风险高于总体平均水平的后验概率值可以将区域划分为热点、温点和冷点：如果后验概率 $P(\exp(\mu_i + v_i) > 1 \mid Y) > 0.8$ 则归类为热点，如果后验概率小于0.2则归类为冷点，如果处于0.2与0.8之间则归类为温点[26]。其结果如图5所示，空间后验概率的估计表明，城市的居住功能区和商业功能区是3类城管事件高发频发的热点区域。且可以注意到，宣传广告类事件在居住聚集区也有明显的热点分布，这可能是由于居住区内的超市、便利店等许多商业点产生了违章悬挂广告牌、散发小广告等问题，而街面秩序类事件中包含的占道经营等问题，使得街面秩序类事件在这种情况下与宣传广告类事件的热点有部分重合现象。



(a) 街面秩序　　(b) 市容环境　　(c) 宣传广告

图4　城管事件相对风险的空间分布

Fig. 4　Spatial distribution of urban management events' relative risk



(a) 街面秩序　　(b) 市容环境　　(c) 宣传广告

图5　城管事件相对风险的空间冷热点趋势

Fig. 5　Spatial hot/cold spots of urban management events

4.3 城管事件相对风险的时间趋势

时间维度上的相对风险趋势如图6和图7所示。一周的时间结构项的结果 $RR_j = e^{\gamma_j}$ 表明城管事件的相对风险在一周内没有显著的单调性，说明3类城管事件的数量在总体上接近一个均匀分布，但从非结构项的结果 $RR_j = e^{\varphi_j}$ 中可以发现周二、周五和周六的时间异质性较为突出，说明每逢周二、周五和周六，3类城管事件的风险趋势会相对增强。

本研究关注的另一个时间尺度，小时项的非结构化结果 $RR_k = e^{\tau_k}$ 给出了一天之中城管事件在各个小时时段上的异质特征。从图7中可见，街面秩序类事件在早午晚3个时段都有高发的现象，而市容环境类事件与宣传广告类事件在午后至晚上的时间段内的高发频发程度没有早晨剧烈。8—10时是城管事件案发相对风险最高的时段，午后的14—15时次之，晚间时间城管事件的风险也有小幅增长，并在19时左右达到高峰，说明P区城市居民的业余休闲与消费活动可能更加集中在这个时间段。分析的结果可以为城管部门更有针对性的管控工作提供参考。

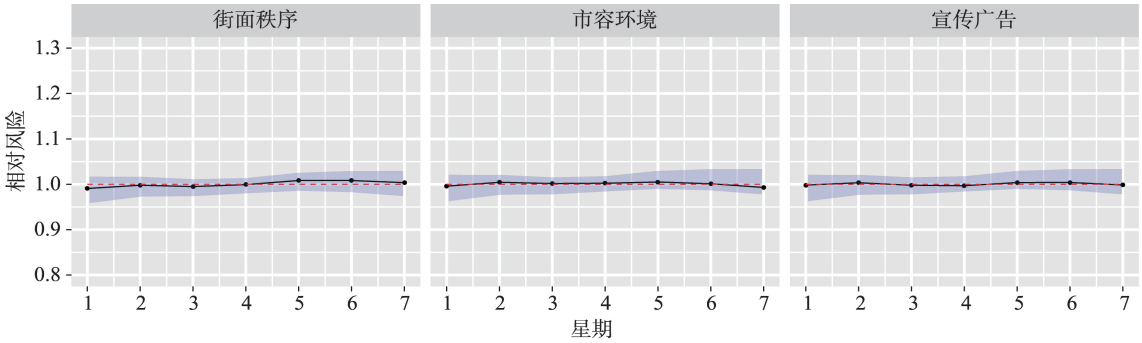
4.4 城管事件相对风险的影响因素分析

表5显示了各类POI密度协变量对3类城管事件相对风险影响的估计结果，包括后验估计均值以及置信区间。餐饮服务类、交通设施类、生活服务类以及住宿服务类POI的密度与城管事件的相对风险始终呈正相关，以餐饮服务类POI密度对街面秩序类事件的影响为例，当餐饮服务类的POI密度每增加一个标准单位，就会带来8.3%的相对风险的增加，同样的解读也适用于另外2类事件以及其他结果为显著的协变量。而其余结果为非显著的协变量在本研究中都没有统计学意义。

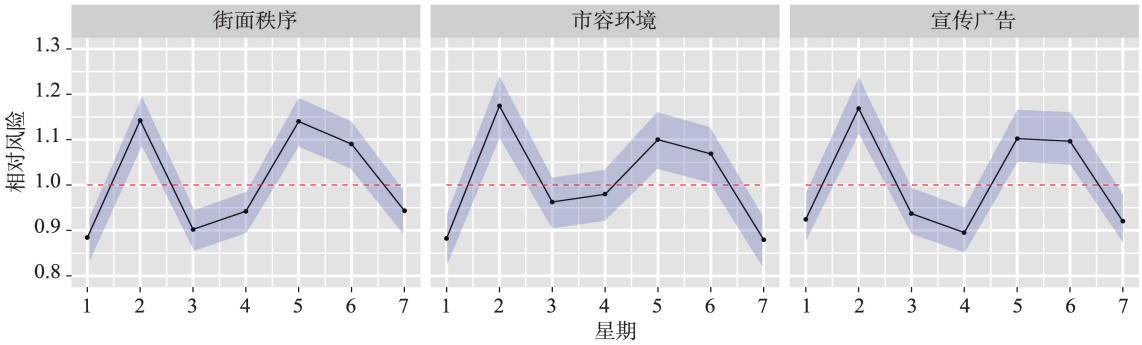
协变量的分析结果表明，城管事件与经济社会活动具有紧密的联系。以住宿服务类为例，本研究中住宿服务类POI对城管事件的相对风险影响平均约18%，而酒店、宾馆等住宿设施通常集中经济社会活动较为密集的商业区域，说明城市的经济社会活动在带来经济社会效益的同时，也会不可避免的产生相应的城市管理问题，这需要城管部门有重点的进行区域化管控，为城市的各项经济社会活动创造良好的秩序与环境。

5　结论与讨论

本文利用贝叶斯时空模型探讨了西北某地H市P区街面秩序、市容环境以及宣传广告3类城管



(a) 一周的时间结构变化趋势

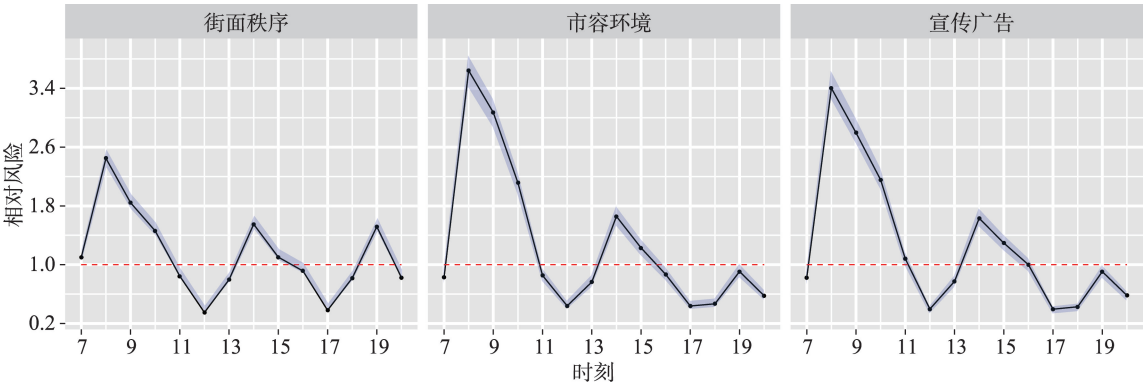


(b) 一周的时间非结构变化趋势

注:黑色实线:相对风险的后验均值;紫色色带:相对风险的95%置信区间;红色虚线:均值参考线。

图6 一周之中城管事件相对风险的时间变化趋势及其置信区间

Fig. 6 Temporal relative risk trend of urban management events during a week with the 95% CI



注:黑色实线:相对风险的后验均值;紫色色带:相对风险的95%置信区间;红色虚线:均值参考线。

图7 一天之中城管事件相对风险的时间变化趋势及其置信区间

Fig. 7 Temporal relative risk trend of urban management events during a day with the 95% CI

事件的时空分布及其与POI密度协变量隐含的关联。结果表明,3类城管事件的空间相对风险呈现出不同的分布,市容环境类事件主要在P区的城区中心汇集,宣传广告类事件主要在P区东南方向的商业区汇集,街面秩序类事件的分布则反映出城市居民在以上2个区域之间的某种活动规律,通过后验概率计算得到的冷热点分布则更加直观的展示了研究区域内空间相对风险的分布差异。

对城管事件相对风险在时间尺度上的分析结果表明,3类事件在时间上的模式是相近的,一周之中星期二、星期五以及星期六的城管事件案发数量较为突出,但总体上呈现出一个均匀分布的状态。8—10时是城管事件高发频发的时间段,中午时回落到平均水平以下,午后14—15时又开始呈现出活跃的状态,3类城管事件在傍晚时分至晚上20时之间仍有增长的趋势,说明P区的城市居民可能在这

表5 协变量后验估计值与置信区间

Tab. 5 Posterior mean and confidence intervals of potential covariates

协变量	街面秩序类	市容环境类	宣传广告类
截距项	0.586(0.549, 0.624)	0.654 (0.614, 0.697)	0.611 (0.573, 0.651)
餐饮服务类	1.083*(1.032, 1.136)	1.065*(1.017, 1.116)	1.086* (1.036, 1.138)
交通设施类	1.105*(1.047, 1.166)	1.056* (1.001, 1.113)	1.091* (1.035, 1.151)
居民住宅类	1.156*(1.017, 1.315)	1.073 (0.949, 1.213)	1.118 (0.990, 1.263)
购物服务类	1.055(0.963, 1.157)	1.039 (0.950, 1.136)	1.052 (0.962, 1.150)
生活服务类	1.084*(1.026, 1.145)	1.065* (1.009, 1.125)	1.081* (1.024, 1.141)
住宿服务类	1.158*(1.019, 1.315)	1.201* (1.070, 1.363)	1.205* (1.069, 1.359)
医疗保健类	1.070(0.956, 1.162)	1.001 (0.931, 1.093)	1.081 (0.997, 1.171)
公司企业类	0.984(0.928, 1.043)	0.995 (0.941, 1.052)	0.944 (0.921, 1.031)

注:协变量包括各个类别POI的网格密度;*代表POI协变量在95%置信区间的统计显著性。截距项表示事件相对风险的全局平均水平。

个时段开始了业余的休闲与消费活动,并在晚上19时左右达到高峰。

模型中还包括了一组POI密度协变量,用来分析城管事件与周边城市建成环境隐含的关联。协变量的后验估计结果表明,餐饮、交通、生活服务等城市基础设施与城管事件的关联最为紧密,且都表现为正相关。说明这些城市兴趣点潜在的城市功能吸引了最多经济社会活动,而城市的经济社会活动在带来效益的同时,也产生了相应的城市管理问题。这一分析结果可以作为城管部门进行区域化重点管控的依据与参考。

贝叶斯时空模型能够充分考虑时空数据中隐含的空间效应以及时间效应,得到良好的参数估计结果[27],因而在许多研究领域中得到了广泛的应用[28]。但目前关于贝叶斯时空模型在城管事件上的应用还较为少见,本研究通过渐进性的建模过程说明了在城管事件数据分析中纳入空间效应与时间效应的合理性,并基于优选的贝叶斯时空模型探讨了城管事件在研究区域内相对风险的时空分布与变化趋势以及不同类别的POI对城管事件相对风险的影响。实验表明,贝叶斯时空模型的分析结果能够满足城管部门日常工作中关于城管事件智能分析与管理决策的需求。

但需要指出的是,贝叶斯时空模型在实际运用中也存在着一些局限。例如怎样选取最有效的超参数的先验分布需要进一步的探讨。作为协变量的有关经济、社会、环境等数据往往很难全面的获取。模型的空间尺度也可能存在一个可修改的区域单元问题(Modifiable Areal Unit Problem, MAUP)[29]。具体到本文中,模型的基本计算单元是分辨率为275 m×335 m的网格,城管事件数据的汇总统计,研究区域的空间自相关关系等重要的计算过程都依赖于这些基本计算单元。MAUP问题即为,若改变基本计算单元的形状或尺度,比如使用六边形网格或者将网格分辨率进行大尺度的变更[30],得到的结论可能会有差异。实际上这种尺度变换带来的尺度效应问题也是地理信息科学中的重要研究方向[31]。而考虑到本文的目的是探讨贝叶斯时空模型在城管事件数据分析上的适用性,因此有关以上这些方面的问题上将留待将来作进一步的深入研究。

参考文献(References):

[1] 王家耀.大数据时代的智慧城市[J].测绘科学,2014,39(5):3-7. [Wang J Y. Smart city in the era of big data[J]. Science of Surveying and Mapping, 2014,39(5):3-7.]

[2] 佘冰,朱欣焰,呙维,等.基于空间点模式分析的城市管理事件空间分布及演化:以武汉市江汉区为例[J].地理科学进展,2013,32(6):924-931. [She B, Zhu X Y, Guo W, et al. Spatial distribution and evolution of city management events based on the spatial point pattern analysis: A case study of Jianghan district, Wuhan city[J]. Progress in Geography, 2013,32(6):924-931.]

[3] 靖常峰,付艳丽,戴培培,等.近邻指数的城管案件全局分布模式分析[J].测绘科学,2017,42(6):76-81. [Jing C F, Fu Y L, Dai P P, et al. Spatial pattern analysis on global distribution of urban management case based on nearest index[J]. Science of Surveying and Mapping, 2017,42(6):76-81.]

[4] 伏家云,靖常峰,杜明义,等.参数优化DBSCAN算法的城管案件聚类分析[J].测绘科学,2018,43(8):135-140. [Fu J Y, Jing C F, Du M Y, et al. Urban management case spatial clustering analysis based on a parameter optimized DBSCAN algorithm[J]. Science of Surveying and Mapping, 2018,43(8):135-140.]

[5] Gelman A, Carlin J B, Stern H S, et al. Bayesian data analysis[M]. Boca Raton, FL: CRC Press, 2013.

[6] Ugarte M D, Adin A, Goicoa T, et al. On fitting spatio-temporal disease mapping models using approximate Bayesian inference[J]. Statistical Methods in Medical Research, 2014,23(6):507-530.

[7] Song C, Yang X, Shi X, et al. Estimating missing values in China's official socioeconomic statistics using progressive spatiotemporal Bayesian hierarchical modeling[J]. Scientific Reports, 2018,8(1):10055.

[8] Lawson A B. Bayesian disease mapping: Hierarchical modeling in spatial epidemiology[M]. Boca Raton, FL: CRC Press, 2013.

[9] Banerjee S, Carlin B P, Gelfand A E. Hierarchical modeling and analysis for spatial data[M]. Boca Raton, FL: CRC Press, 2014.

[10] 吴北平,杨典,王劲峰,等.利用贝叶斯时空模型分析山东省手足口病时空变化及影响因素[J].地球信息科学学报,2016,18(12):1645-1652. [Wu B P, Yang D, Wang J F, et al. Space-time variability and determinants of hand, foot and mouth in Shandong province: A Bayesian spatio-temporal modeling approach[J]. Journal of Geo-information Science, 2016,18(12):1645-1652.]

[11] Halonen J I, Blangiardo M, Toledano M B, et al. Is long-term exposure to traffic pollution associated with mortality? A small-area study in London[J]. Environmental Pollution, 2016,208:25-32.

[12] Osei F B, Stein A. Diarrhea morbidities in small areas: Accounting for non-stationarity in sociodemographic impacts using Bayesian spatially varying coefficient modelling[J]. Scientific Reports, 2017,7(1):9908.

[13] 关雪峰,曾宇媚.时空大数据背景下并行数据处理分析挖掘的进展及趋势[J].地理科学进展,2018,37(10):1314-1327. [Guang X F, Zeng Y M. Research progress and trends of parallel processing, analysis, and mining of big spatiotemporal data[J]. Progress in Geography, 2018,37(10):1314-1327.]

[14] 刘玲,李钢,杨兰,等.深圳市快递自提点的空间分布特征与影响因素[J].地球信息科学,2019,21(8):1240-1253. [Liu L, Li G, Yang L, et al. Spatial distribution characteristics and influencing factors of the delivery sites in Shenzhen[J]. Journal of Geo-information Science, 2019,21(8):1240-1253.]

[15] Vatcheva K P, Lee M J, McCormick J B, et al. Multicollinearity in regression analyses conducted in epidemiologic studies[J]. Epidemiology (Sunnyvale, Calif), 2016,6(2):227.

[16] Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics[J]. Annals of the Institute of Statistical Mathematics, 1991,43(1):1-20.

[17] Blangiardo M, Cameletti M. Spatial and spatio-temporal Bayesian models with R-INLA[M]. London: John Wiley and Sons, 2015.

[18] Besag J. Spatial interaction and the statistical analysis of lattice systems[J]. Journal of the Royal Statistical Society: Series B(Methodological), 1974,36(2):192-225.

[19] Lichstein J W, Simons T R, Shriner S A, et al. Spatial autocorrelation and autoregressive models in ecology[J]. Ecological Monographs, 2002,72(3):445-463.

[20] Song C, He Y, Bo Y, et al. Risk assessment and mapping of hand, foot, and mouth disease at the county level in mainland China using spatiotemporal zero-inflated Bayesian hierarchical models[J]. International Journal of Environmental Research and Public Health, 2018,15(7):1476.

[21] Wang Y, Lai Y, Du Z, et al. Spatiotemporal distribution of hand, foot, and mouth disease in Guangdong province, China and potential predictors, 2009-2012[J]. International Journal of Environmental Research and Public Health, 2019,16(7):1191.

[22] 王劲峰,廖一兰,刘鑫.空间数据分析教程(第二版)[M].北京:科学出版社,2019. [Wang J F, Liao Y L, Liu X. 2019. Spatial data analysis tutorial(second edition)[M]. Beijing: Science Press, 2019.]

[23] Schrodle B, Held L. Spatio-temporal disease mapping using INLA[J]. Environmetrics, 2011,22(6):725-734.

[24] Spiegelhalter D J, Best N G, Carlin B P, et al. Bayesian measures of model complexity and fit[J]. Journal of the Royal Statistical Society:Series B(Statistical Methodology), 2002,64(4):583-639.

[25] Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations[J]. Journal of the Royal Statistical Society: Series B(Statistical Methodology), 2009,71(2):319-392.

[26] Richardson S, Thomson A, Best N, et al. Interpreting posterior relative risk estimates in disease-mapping studies[J]. Environmental Health Perspectives, 2004,112(9):1016-1025.

[27] 林文棋,陈会宴,谢盼,等.基于多源数据的北京市朝阳区人口时空格局评估与预测[J].地球信息科学学报,2018,20(10):1467-1477. [Lin W Q, Chen H Y, Xie P, et al. Spatial-temporal variation evaluation and prediction of population in Chaoyang district of Beijing based on multisource data[J]. Journal of Geo-information Science, 2018,20(10):1467-1477.]

[28] KnorrHL. Bayesian modelling of inseparable space-time variation in disease risk[J]. Statistics in Medicine, 2000,19(17-18):2555-2567.

[29] Openshaw S. The modifiable areal unit problem//Concepts and Techniques in Modern Geography[M]. Norwich: Geo Books, 1984.

[30] 彭建,徐飞雄.不同格网尺度下的黄山市生境质量差异分析[J].地球信息科学学报,2019,21(6):887-897. [Peng J, Xu F X. Effect of grid size on habitat quality assessment: A case study of Huangshan city[J]. Journal of Geo-information Science, 2019,21(6):887-897.]

[31] 李志林,王继成,谭诗腾,等.地理信息科学中尺度问题的30年研究现状[J].武汉大学学报·信息科学版,2018,43(12):2233-2242. [Li Z L, Wang J C, Tan S T, et al. Scale in Geo-information Science: An overview of thirty-year development[J]. Geomatics and Information Science of Wuhan University, 2018,43(12):2233-2242.]