

# 考虑空间自相关的贝叶斯事故预测模型

黄合来<sup>1</sup>, 邓雪<sup>2</sup>, 许鹏鹏<sup>1</sup>

(1. 中南大学 交通运输工程学院, 湖南 长沙 410075; 2. 西南交通大学 交通运输与物流学院, 四川 成都 610031)

**摘要:** 依托美国佛罗里达州 Hillsborough 县的数据建立区域安全预测模型. 将该县重新划分为 200, 500, 700 个交通安全分析小区; 提取小区层面路网特征数据、出行数据和其他影响因素, 考虑空间自相关特性, 建立贝叶斯空间模型; 分析小区层面影响因素的安全效应, 评价不同分区规模对安全因素效应的影响. 对比传统的泊松模型和泊松-对数正态模型, 贝叶斯空间模型具有更高的数据拟合度; 分区数目越多, 空间因素在随机因素中的比重越高; 同一种分区下, 路网特征变量的安全效应具有鲁棒性; 限速大于 56 km 的路段总长度是预测安全水平的主要指标.

**关键词:** 事故预测模型; 贝叶斯方法; 交通安全分析小区; 空间分析

**中图分类号:** U491.31

**文献标志码:** A

## Bayesian Crash Prediction Model Based on a Consideration of Spatial Autocorrelation

HUANG Helai<sup>1</sup>, DENG Xue<sup>2</sup>, XU Pengpeng<sup>1</sup>

(1. School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China; 2. School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China)

**Abstract:** A regional safety prediction model was proposed based on the data from Hillsborough County, Florida, USA. By regionalizing the county into 200, 500 and 700 traffic safety analysis zones, we developed a Bayesian spatial model with consideration of spatial autocorrelation to relate crash rate to zonal factors including road network, trip generation and so on. According to the model results, the relationships were investigated between traffic safety and zone-level factors, as well as the effects of varied zoning schemes on the estimation of factor effects. Results show that compared with the traditional Poisson model and Poisson-lognormal model, the Bayesian spatial model has a better model-fitting; the greater the total zone number is, the higher the spatial effects are; the factor estimates are robust given a specific zoning scheme; the most significant factor affecting zonal safety is

the total road length with speed limit over 56 km · h<sup>-1</sup>.

**Key words:** crash prediction model; Bayesian approach; traffic safety analysis zone; spatial analysis

传统交通安全实践局限于“打补丁式”的微观工程措施, 缺乏在交通系统规划过程中的前瞻性考虑. 通过在规划阶段对交通安全问题作前瞻性的分析来提高路网的安全水平非常重要. 当前, 新一代的交通规划方法, 即“交通安全规划”, 在发达国家尤其是美国成为一个重要研究热点<sup>[1]</sup>. 交通规划层面的影响因素与交通安全水平的关联方法, 即区域安全预测模型, 是“交通安全规划”理论的核心技术之一.

国外已经做了大量定量研究<sup>[2-6]</sup>来分析不同规划层面的影响因素与交通安全水平的关系, 交通分析小区 (traffic analysis zone, TAZ) 是主要的空间分析单元. 已有研究主要从区域人口经济特征<sup>[3]</sup>、土地利用特征<sup>[4]</sup>、区域交通特征<sup>[5]</sup>和区域路网特征<sup>[5-6]</sup>等方面进行分析. 就统计方法而言, 主要应用广义线性模型对 TAZ 事故频数与相关变量的关系进行建模<sup>[4]</sup>, 假设 TAZ 的事故分布在空间上不相关. 然而, 安全数据具有复杂的空间分布特征, 对空间特征的忽视将大大影响安全水平估计的准确性和鲁棒性. 为了克服广义线性模型的局限性, 空间统计方法<sup>[6]</sup>逐渐成为安全分析领域的一个前沿研究方向.

在国内, 文献<sup>[7]</sup>建议结合我国交通规划的工作流程考虑交通安全, 文献<sup>[8]</sup>探讨了城市总体规划和土地规划中影响因素的安全效应. 而从区域宏观层面反映路网特征、交通特征与交通安全之间关系的定量研究在我国几乎没有展开.

本文基于美国佛罗里达州 Hillsborough 县的数据进行区域安全建模研究. 基于安全同质性原则, 将该县重新划分为不同数量规模的交通安全分析小区

(traffic safety analysis zone, TSAZ),提取小区层面数据,引入空间条件自回归先验信息建立贝叶斯空间模型,分析区域安全影响因素;评价不同分区规模对安全因素效应的影响,为交通安全规划提供技术支撑和理论依据。

## 1 数据准备

### 1.1 交通安全分析小区

区划问题是交通规划中一个最为基础的问题,

传统交通分析小区的划分鲜有考虑安全相关因素,而现存大部分区域安全研究均以交通分析小区为基本单元。TSAZ 的划分在兼顾交通规划层面因素的前提下,考虑事故特征的区域同质性。本文基于 Full-order-CLK 算法<sup>[9]</sup>,按照严重事故、道路长度和总出行生成的聚集程度(权重分别为 0.50,0.25,0.25)获得 Hillsborough 县的多种分区方案,并基于 200,500 和 700 个 TSAZ 分区方案建立事故预测模型,分区方案如图 1 所示。

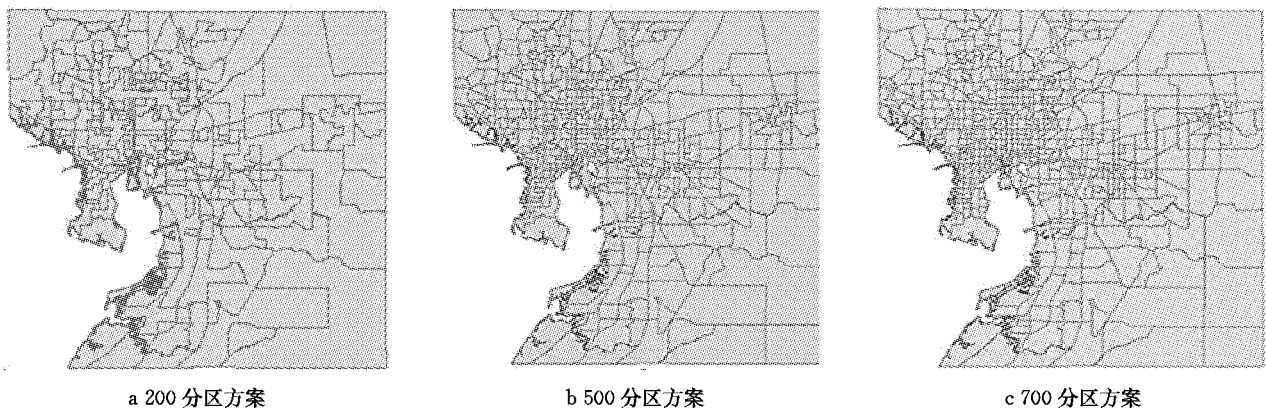


图1 佛罗里达州 Hillsborough 县重新分区图

Fig.1 The new zoning map of Hillsborough County, Florida, USA

### 1.2 数据收集

采用佛罗里达州 Hillsborough 县所有州立道路 3 年的数据(2005—2007 年),包括事故数据、道路特征数据、出行数据以及空间分布特征数据。

事故数据从佛罗里达州运输部事故分析报告系统中获得,其提供的最重要的信息是事故发生的小区归属以及伤害严重性等级。对于事故伤害的严重程度,分为 5 个等级:1 为没有受伤,2 为轻伤,3 为非致残伤害,4 为致残伤害,5 为交通死亡。

相关路网特征数据和出行数据主要来源于佛罗里达州运输部。基于 GIS 地图小区划分边界对数据进行集计处理,获得小区层面的变量。

(1)路段总长度:每个 TSAZ 内道路中心线总里程数。

(2)分类的路段长度:每个 TSAZ 内路段中心线公里数按速度限制分类,① 限速 24 km(15 英里)的路段长度;② 限速 40 km(25 英里)的路段长度;③ 限速 56 km(35 英里)的路段长度;④ 限速 105 km(65 英里)的路段长度。其中,限速 40 km 和 56 km 的路段是主要类型,分别占路段总长度的 76.86%和 13.01%。为避免变量冗余问题,将所有分类归为限速低于 56 km 的路段长度和限速大于或等于 56 km

的路段长度 2 类。

(3)交叉口数量:每个 TSAZ 的交叉口总数。

(4)总出行发生量:每个 TSAZ 总的出行发生量。

(5)总出行吸引量:每个 TSAZ 总的出行吸引量。

(6)基于 TSAZ 的出行发生量和吸引量,包括从家去工作,从家去上学,从家去购物,从家去社交活动,从家去其他,非从家去工作,非从家去其他的出行发生量和吸引量;轻型/重型卡车,出租车,飞机的出行发生量和吸引量;其他 TSAZ 到目标 TSAZ 的出行发生量和吸引量。

使用 ArcGIS 软件自动生成相邻小区的空间邻接矩阵。

### 1.3 数据前处理和变量选择

定义严重性等级为 4(致残伤害)和 5(交通死亡)的事故为严重事故,其余为非严重事故。分别调查与各 TSAZ 有关的总事故和严重事故,作为预测模型的响应变量。

在因素较多且因素间存在广泛多重共线性的情况下,模型解释变量的筛选具有一定的挑战性。模型的数据拟合度和简约程度是两个重要的变量选择标

准.通过数据探索性分析和对变量的逐一筛选,选定3类重要的安全影响因素,包括路段长度、交叉口数量和出行量.路段长度是反映交通基础设施规模的重要指标之一;交叉口是交通事故多发点,在控制了路段长度的前提下,采用交叉口密度(小区内交叉口总数/小区道路总长度)反映交通状况的复杂程度.此外,出行量是交通事故发生的直接机会变量,也是传统交通规划的重要数据来源.本研究选用总出行发生量和吸引量作为重要解释变量,同时,由于家庭和工作出行数据是一个重要的出行类型,也在模型中重点考虑.

为了避免总出行量和工作出行量的强相关问题,将其分别加入模型,即同一种分区、同一种响应变量下,建立两个事故预测模型,每个模型包含5个解释变量.具体模型变量如表1所示.

表1 模型变量表

Tab.1 Variables used in the models

总事故模型一 严重事故模型三	总事故模型二 严重事故模型四
交叉口密度	交叉口密度
限速<56 km路段长度	限速<56 km路段长度
限速≥56 km路段长度	限速≥56 km路段长度
总出行发生量	家庭-工作出行发生量
总出行吸引量	非家庭-工作出行发生量

## 2 模型建立

### 2.1 模型结构

#### 2.1.1 泊松模型

由于泊松模型能够契合道路交通事故的零散性、随机性以及事故频次为非负整数的特点,被广泛应用于事故预测中.假定在2005—2007年3年中各分析小区内事故频次 $Y_i$ 服从泊松分布,则第 $i$ 个小区发生 $k$ 次事故的概率为

$$P(Y_i = k) = \frac{\lambda_i^k}{k!} \exp(-\lambda_i) \quad (1)$$

式中: $\lambda$ 为泊松分布的强度参数,反映该小区内某时段平均发生事故频次.

将 $\lambda$ 进行对数处理,建立与解释变量的线性关系,作为模型的连接函数,即

$$\ln(\lambda_i) = \beta_0 + \beta X_i \quad (2)$$

式中: $\beta_0$ 和 $\beta$ 为待估系数; $X_i$ 为解释变量的向量.

#### 2.1.2 泊松-对数正态模型

泊松模型假定事故均值与方差相等.然而,事故数据往往具有过度离散性,其方差显著大于均值<sup>[3]</sup>.

泊松-伽马模型与泊松-对数正态模型通过引入模型残差随机项,容许事故过度离散特征的存在,成为两个主流的事故预测模型的基本结构,本文采用泊松-对数正态模型进行研究.在泊松-对数正态模型中,假定第 $i$ 个小区的事故频次 $Y_i$ 服从泊松分布

$$Y_i \sim P(\lambda_i) \quad (3)$$

泊松-对数正态回归模型在基本泊松模型的连接函数中加入一个反映数据过度离散特征的残差项 $\theta_i$ ,并假定其服从均值为零的正态分布,即

$$\ln(\lambda_i) = \beta_0 + X_i\beta + \theta_i \quad (4)$$

$$\theta_i \sim N(0, 1/\tau_\theta) \quad (5)$$

式中: $\tau_\theta$ 为精度参数,等于方差的倒数.

#### 2.1.3 贝叶斯空间模型

空间相关广泛地存在于相邻的空间小区中,对影响事故频次因素的估计效果的精确度有显著影响.贝叶斯空间模型能有效地拟合空间相关特征<sup>[3]</sup>.本研究在泊松-对数正态模型的基础上,引入条件自回归模型<sup>[10]</sup>反映相邻小区间可能的空间关联.假设 $Y_i$ 是第 $i$ 个 TSAZ 的3年平均事故率,模型如下:

$$Y_i \sim P(\lambda_i)$$

$$\ln(\lambda_i) = \beta_0 + X_i\beta + \theta_i + \phi_i \quad (6)$$

在贝叶斯模型估计中,为 $\beta$ 和 $\theta_i$ 设定无信息先验分布

$$\beta \sim N(0, 1/000) \quad (7)$$

$$\theta_i \sim N(0, 1/\tau_\theta) \quad (8)$$

$\phi_i$ 满足条件自回归模型的先验分布

$$\phi_i \sim N(\bar{\phi}_i, 1/\tau_i) \quad (9)$$

式中: $\bar{\phi}_i = \frac{\sum_{i \neq j} \phi_j \omega_{ij}}{\sum_{i \neq j} \omega_{ij}}$ ,  $\omega_{ij}$ 为邻接矩阵 $\omega$ 的元素,  $\omega_{ij} =$

$$\begin{cases} 1, & \text{TSAZ}_i \text{ 和 } \text{TSAZ}_j \text{ 相邻} \\ 0, & \text{TSAZ}_i \text{ 和 } \text{TSAZ}_j \text{ 不相邻} \end{cases}; \tau_i = \frac{\tau_c}{\sum_{i \neq j} \omega_{ij}}, \tau_c \text{ 为}$$

先验条件自回归模型的精度参数.

本研究采用贝斯特等人建立的先验分布形式以保证模型收敛<sup>[3]</sup>.

$$\tau_\theta \sim \gamma(0.001, 0.001) \quad \tau_c \sim \gamma(0.1, 0.1) \quad (10)$$

式中: $\tau_\theta$ 和 $\tau_c$ 分别表示由空间效应和随机效应导致的过度离散特征,其比值 $\alpha$ 能反映数据过度离散特征的结构,可由下式估计:

$$\alpha = \frac{s(\phi)}{s(\theta) + s(\phi)} \quad (11)$$

式中: $s$ 为标准差.

### 2.2 模型仿真结果评价标准

采用贝叶斯偏差信息准则(deviance

information criterion, DIC)对待选模型进行科学评价. DIC 准则综合考虑模型拟合优度和模型复杂性,其计算方法如下<sup>[11]</sup>:

$$\text{DIC} = \bar{D} + p_D$$

(12)

式中: $\bar{D}$ 用来衡量模型拟合优度; $p_D$  为有效参数,用来衡量模型复杂性.

一般地, DIC 越小,模型越优. DIC 差值大于 5 则表明模型之间具有显著的差异.

2.3 模型参数估计方法

采用 WinBUGS 软件<sup>[12]</sup>对模型进行参数估计. 使用 3 条链进行迭代,通过查看历史记录检查其收敛情况. 确保所有参数收敛以后,舍去前 1 000 次迭代结果. 为减少自相关性,在剩下的样本中每 10 个中抽取 1 个,从 1 001 次开始进行 10 000 次迭代.

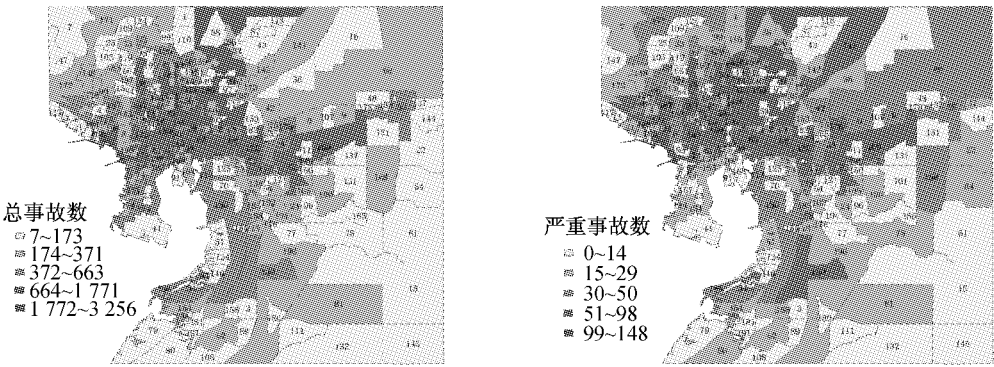


图 2 佛罗里达州 Hillsborough 县 2005—2007 年事故地理分布图(以 200 个 TSAZ 为例)

Fig.2 Crash distribution maps of Hillsborough County, Florida from 2005 to 2007(200 TSAZ as an example)

口经济特征<sup>[3]</sup>、土地使用功能<sup>[4]</sup>等因素密切相关. 西北部海岸地区经济发展水平高,人口相对密集,事故频次较高;而东南部地区的事事故频次较低.

3.2 结果分析

3.2.1 DIC 评价

如表 2 所示,泊松-对数正态模型和贝叶斯空间

3 模型结果

3.1 事故说明

如图 2 所示,以 200 个 TSAZ 为例,说明佛罗里达州 Hillsborough 县 2005—2007 年事故频次地理分布特点. 该县共发生 57 694 起事故,最低为东北部第 112 个 TSAZ 的 7 起,最高为东部第 161 个 TSAZ 的 3 256 起,事故频率标准差为 369. 47,变化范围较大;共发生 4 854 起严重事故,在 TSAZ 内变化范围从 0~148 起,标准差为 23. 14,其中第 190 个 TSAZ 严重事故频次最大,第 15 个 TSAZ 内的严重事故数为 0. 西北部地区事故相对集中,而东南部地区事故相对较少. 以往研究表明,事故的空间分布与区域人

模型的  $p_D$  值大于泊松模型而 DIC 值远小于泊松模型,说明前两者模型结构较为复杂,且拟合效果大大优于泊松模型;泊松-对数正态模型与贝叶斯空间模型相比,两者的 DIC 值较为接近,仅从 DIC 值无法得出最优模型. 从表 3 和表 4 可知,随着分区数目增多,空间因素比例  $\alpha$  逐渐增大且统计显著,说明贝叶

表 2 基于 DIC 值的模型评价

Tab.2 Model evaluation on the basis of DIC

模型	200 分区		500 分区		700 分区	
	$p_D$	DIC	$p_D$	DIC	$p_D$	DIC
泊松模型一	6	149 827	6	20 556	6	23 748
泊松模型二	6	14 7321	6	20 839	6	23 947
泊松-对数正态模型一	192	1 630	488	3 983	673	5 327
泊松-对数正态模型二	199	1 754	487	3 979	668	5 321
贝叶斯空间模型一	197	1 467	485	3 978	669	5 321
贝叶斯空间模型二	197	1 466	484	3 976	692	5 340
泊松模型三	6	3 148	6	4710	6	5 698
泊松模型四	6	3 156	6	4 731	6	5 731
泊松-对数正态模型三	213	1 344	450	2 834	544	3 575
泊松-对数正态模型四	209	1 350	443	2 829	539	3 570
贝叶斯空间模型三	203	1 336	480	2 855	556	3 577
贝叶斯空间模型四	207	1 337	403	2 788	539	3 559

表 3 总事故模型参数估计  
Tab.3 Estimation for total crash models

变量	200 分区			500 分区			700 分区		
	P	P-LN	BS	P	P-LN	BS	P	P-LN	BS
模型一									
交叉口密度	<b>0.057</b>	<b>0.235</b>	-0.101	0.116	<b>0.162</b>	<b>0.156</b>	<b>0.088</b>	<b>0.123</b>	<b>0.206</b>
限速<56 km 路段长度	<b>-1.795</b>	-0.153	<b>-0.918</b>	<b>-0.131</b>	<b>-0.160</b>	0.048	<b>-0.163</b>	<b>-0.160</b>	0.028
限速≥56 km 路段长度	<b>1.421</b>	<b>0.211</b>	<b>0.911</b>	<b>0.188</b>	<b>0.273</b>	<b>0.296</b>	<b>0.218</b>	<b>0.279</b>	<b>0.335</b>
总出行发生量	<b>1.480</b>	0.145	-1.367	<b>0.251</b>	<b>0.354</b>	<b>0.198</b>	<b>0.251</b>	<b>0.321</b>	<b>0.178</b>
总出行吸引量	<b>-1.352</b>	-0.075	1.019	<b>0.053</b>	0.051	0.044	<b>0.016</b>	0.029	0.010
$\alpha$			<b>0.317</b>			<b>0.569</b>			<b>0.597</b>
模型二									
交叉口密度	<b>0.054</b>	<b>0.284</b>	-0.141	<b>0.118</b>	<b>0.168</b>	0.155	<b>0.090</b>	<b>0.125</b>	<b>0.205</b>
限速<56 km 路段长度	<b>-1.786</b>	-0.145	<b>-0.952</b>	<b>-0.108</b>	<b>-0.135</b>	0.048	<b>-0.139</b>	<b>-0.137</b>	0.044
限速≥56 km 路段长度	<b>1.438</b>	<b>0.150</b>	<b>1.017</b>	<b>0.189</b>	<b>0.274</b>	<b>0.301</b>	<b>0.219</b>	<b>0.284</b>	<b>0.337</b>
家庭-工作出行发生量	<b>0.957</b>	0.057	0.443	<b>0.193</b>	<b>0.275</b>	<b>0.167</b>	<b>0.206</b>	<b>0.265</b>	<b>0.140</b>
非家庭-工作出行发生量	<b>-0.924</b>	<b>0.231</b>	-0.260	<b>0.164</b>	<b>0.201</b>	<b>0.127</b>	<b>0.133</b>	<b>0.169</b>	<b>0.104</b>
$\alpha$			<b>0.266</b>			<b>0.577</b>			<b>0.608</b>

注:P 为泊松模型;P-LN 为泊松-对数正态模型;BS 为贝叶斯空间模型;粗体数值为 95%统计显著。

表 4 严重事故模型参数估计  
Tab.4 Estimation for severe crash models

变量	200 分区			500 分区			700 分区		
	P	P-LN	BS	P	P-LN	BS	P	P-LN	BS
模型三									
交叉口密度	<b>0.032</b>	0.077	0.113	-0.024	-0.014	<b>0.108</b>	-0.033	-0.039	<b>0.185</b>
限速<56 km 路段长度	<b>-0.332</b>	<b>-0.537</b>	-0.312	<b>-0.172</b>	<b>-0.178</b>	-0.076	<b>-0.152</b>	<b>-0.136</b>	-0.033
限速≥56 km 路段长度	<b>0.405</b>	<b>0.695</b>	<b>0.631</b>	<b>0.284</b>	<b>0.374</b>	<b>0.355</b>	<b>0.312</b>	<b>0.406</b>	<b>0.377</b>
总出行发生量	<b>0.166</b>	<b>0.417</b>	0.135	<b>0.123</b>	<b>0.149</b>	0.092	<b>0.196</b>	<b>0.212</b>	<b>0.103</b>
总出行吸引量	<b>0.040</b>	-0.155	-0.06	<b>0.051</b>	0.034	0.050	0.023	0.019	0.036
$\alpha$			<b>0.542</b>			<b>0.563</b>			0.727
模型四									
交叉口密度	<b>0.035</b>	0.079	0.116	-0.018	-0.009	<b>0.109</b>	-0.029	-0.036	<b>0.181</b>
限速<56 km 路段长度	<b>-0.300</b>	<b>-0.487</b>	-0.270	<b>-0.162</b>	<b>-0.168</b>	-0.081	<b>-0.135</b>	<b>-0.118</b>	-0.044
限速≥56 km 路段长度	<b>0.395</b>	<b>0.700</b>	<b>0.634</b>	<b>0.286</b>	<b>0.377</b>	<b>0.364</b>	<b>0.314</b>	<b>0.410</b>	<b>0.389</b>
家庭-工作出行发生量	<b>0.106</b>	0.224	0.056	<b>0.102</b>	<b>0.122</b>	0.096	<b>0.168</b>	<b>0.173</b>	<b>0.106</b>
非家庭-工作出行发生量	<b>0.101</b>	0.011	-0.017	<b>0.087</b>	<b>0.087</b>	0.066	<b>0.100</b>	<b>0.110</b>	<b>0.071</b>
$\alpha$			<b>0.547</b>			<b>0.556</b>			<b>0.704</b>

注:P 为泊松模型;P-LN 为泊松-对数正态模型;BS 为贝叶斯空间模型;粗体数值为 95%统计显著。

斯空间模型能够更好地拟合安全水平空间相关性特征。

从表 2 还可以看出,同一种模型,随着分区数目增多,DIC 值显著增大,模型拟合度降低,说明本文所采用的宏观区域交通指标变量在区域面积较大时对事故率的预测更为准确。

### 3.2.2 变量分析

对于总事故,无论分区数目多少,泊松模型中的解释变量全部显著。而贝叶斯空间模型的显著性解释变量数目较少,当 TSAZ 的数目为 200 时,显著性解释变量只有 2 个(限速低于 56 km 的路段和限速大于等于 56 km 的路段长度);当 TSAZ 的数目分别为 500 和 700 时,贝叶斯空间模型的显著性解释变量增多且趋于稳定,除限速低于 56 km 的路段长度和总的出行吸引量外,其余变量全都显著。

对于严重事故,泊松模型的显著性解释变量数目最多。在贝叶斯空间模型中,当 TSAZ 数目为 200

时,显著性解释变量仅有 1 个(限速大于等于 56 km 的路段长度);当 TSAZ 数目为 500 时,显著性解释变量有 2 个(限速大于等于 56 km 的路段长度和交叉口密度);当 TSAZ 数目为 700 时,显著性解释变量增多且趋于稳定,除限速低于 56 km 的路段长度和总的出行吸引量外,其余变量全都显著。

交叉口:在大多数模型中,交叉口密度都是显著正相关,尤其是当分区数目足够多时。模型稳定后,给定路段长度、出行数据,交叉口数目越多,事故发生的频次越高。

路段长度:限速低于 56 km 的路段长度在贝叶斯空间模型中几乎都不显著,而其在泊松模型和泊松-对数正态模型中几乎都是显著负相关;限速大于等于 56 km 的路段长度在所有模型中都是显著正相关,且在同一模型中其系数最大,说明高速路段长度是预测道路安全水平的主要指标。

出行数据:总出行吸引量在泊松-对数正态模型

和贝叶斯空间模型中都不显著;同一模型中,总出行发生量的系数总是大于总出行吸引量的系数,说明出行发生量对事故发生的预测能力更强;同一模型中,从家里去工作的出行发生量的系数总是大于非从家里去工作的出行发生量的系数,说明以家为基准的工作出行量有更好的事故预测能力;大多数出行数据的系数都为正,说明出行量与事故发生成正比例关系。

最后,从空间因素比例 $\alpha$ 的估计参数来看,随着分区数目增加,空间因素比例 $\alpha$ 增加。当分区数目足够多时,有必要考虑相邻区域的空间相关性,而文中提出的贝叶斯空间模型能较好拟合空间相关性;同时,在贝叶斯空间模型中,随着分区数目增多,显著变量数目也增多。然而,如前所述,在使用宏观区域交通指标变量预测安全水平时,在区域较大的情况下数据拟合程度降低。因此,实际中分区规模要适中,以平衡拟合度和变量显著性,既达到统计上的要求又保证工程应用的可行性。

## 4 结论

(1)对比传统的泊松模型和泊松-对数正态模型,贝叶斯空间模型具有更高的数据拟合度。

(2)分区数目越多,空间因素在随机因素中的比例越高。

(3)同一分区下,路网特征变量(交叉口密度、限速低于56 km的路段长度和限速大于等于56 km的路段长度)的安全效应具有鲁棒性。

(4)限速大于等于56 km的路段长度是预测安全水平的主要指标。

本文就交通安全规划的核心技术问题——区域安全建模进行了研究,基于美国佛罗里达州 Hillsborough 县的数据,在 TSAZ 层面针对事故均值提出了能够有效拟合事故空间相对特征的区域安全建模研究提供技术基础。

## 参考文献:

- [1] Washington S P, Schalkwyk V I, Mitra S, et al. Incorporating safety into long-range transportation planning [R]. Washington D C: Transportation Research Board, 2006.
- [2] Huang H, Abdel-Aty M. Multilevel data and Bayesian analysis in traffic safety [J]. Accident Analysis and Prevention, 2010, 42(6): 1556.
- [3] Huang H, Abdel-Aty M, Darwiche A L. County-level crash risk analysis in Florida: Bayesian spatial modeling [J]. Transportation Research Record, 2010, 2148, 27.
- [4] Guevara F L D, Washington S P, Oh J. Forecasting crashes at the planning level: simultaneous negative binomial crash model applied in Tucson, Arizona [J]. Transportation Research Record, 2004, 1897: 191.
- [5] Haynes R, Jones A, Kennedy V, et al. District variations in road curvature in England and Wales and their association with road-traffic crash [J]. Environment and Planning: A, 2007, 39(5): 1222.
- [6] Hadayeghi A, Shalaby A, Persaud B. Development of planning-level transportation safety models using full Bayesian semi parametric additive technique [J]. Journal of Transportation Safety and Security, 2010, 2(1): 45.
- [7] 缪玉玲, 陈小桐, 李峰, 等. 引入交通安全对交通规划的新思考[C]//第二届全国公路科技创新高层论坛论文集. 北京: 当代中国出版社, 2004: 58-63.  
MIAO Yuling, CHEN Xiaotong, LI Feng, et al. The new thinking of introducing traffic safety on transportation planning [C]//Papers of the Second National Highway Science and Technology Innovation Forum. Beijing: Contemporary China Publishing House, 2004: 58-63.
- [8] 唐珏琳. 城市规划中的交通安全因素分析[J]. 公路与汽运, 2007(4): 47.  
TANG Juelin. Factor analysis of traffic safety in urban planning [J]. Highways and Automotive Applications, 2007(4): 47.
- [9] Guo D. Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP) [J]. International Journal of Geographical Information Science, 2008, 22(7): 801.
- [10] Besag J. Spatial interaction and the statistical analysis of lattice systems [J]. Journal of Royal Statistical Society Series B, 1974, 36(2): 192.
- [11] Spiegelhalter D, Best N, Carlin B P, et al. Bayesian measures of model complexity and fit [J]. Journal of Royal Statistical Society B, 2002, 64(4): 583.
- [12] Spiegelhalter D J, Thomas A, Best N G, et al. WinBUGS version 1.4.1 user manual [M]. Cambridge: MRC Biostatistics Unit, 2003.

[1] Washington S P, Schalkwyk V I, Mitra S, et al. Incorporating