

文章编号: 1674-8085(2017)03-0070-06

基于地理加权回归的福建省公路通车里程数研究

林海燕, *李正红, 胡喜生, 郑小雪

(福建农林大学, 福建, 福州 350100)

摘 要: 以福建省 67 个市县为研究对象, 选取农民人均可支配收入、工业总产值、社会消费品零售总额为研究指标, 构建地理加权回归 (GWR) 模型, 并结合 ArcGIS10.0 和 GeoDA 等软件, 对福建省各市县的公路通车里程数的空间分布趋势、局部空间自相关和主要驱动因素进行了分析研究。结果表明, 农民人均可支配收入、工业总产值、社会消费品零售总额对公路通车里程数的影响程度在各市县地区存在差异, 呈现了空间非平稳性。各市县的公路通车里程数, 除了受地方政策等因素的影响外, 确实还与所处的地理环境有关, 总体上存在较为显著的空间相关性。

关键词: 地理加权回归; 公路通车里程数; 空间分布; 局部空间自相关; 空间相关性

中图分类号: U412.1

文献标识码: A

DOI:10.3969/j.issn.1674-8085.2017.03.012

RESEARCH ON HIGHWAY TRAFFIC MILEAGE OF FUJIAN PROVINCE BASED ON GEOGRAPHICAL WEIGHTED REGRESSION

LIN Hai-yan, * LI Zheng-hong, HU Xi-sheng, ZHENG Xiao-xue

(Fujian Agriculture and Forestry University, Fuzhou, Fujian 350108, China)

Abstract: Take 67 cities of Fujian Province as the research object, we select farmers per capita disposable income, gross value of industrial output, social total retail sales of consumer goods index for the study. Furthermore, we use geographically weighted regression (GWR) method based on the ArcGIS10.0 and GeoDA software. We also and introduce the geospatial elements, space highway mileage of all cities and counties in Fujian distribution trends. The local spatial autocorrelation and the main driving factors were analyzed. The results show that the difference of influence degree of per capita disposable income of farmers, the total industrial output value, the total retail sales of consumer goods in the city and county area and the spatial nonstationarity. The city and county highway traffic mileage are affected by the local policy and other factors, which are also related to the geographical environment, there is a significant spatial correlation on the whole.

Key words: geographically weighted regression; highway traffic mileage; spatial distribution; local spatial autocorrelation; spatial correlation

公路是交通运输系统的重要组成部分, 因其直达性和机动灵活等优点, 影响着社会生活的方方面面, 对经济和社会的发展起着重要的促进和保障作

用。上世纪 80 年代, 公路网规划工作逐渐开始被重视。在规划中, 一个重要指标是公路网总里程, 它是公路网等级结构优化的前提和基础。目前, 关

收稿日期: 2016-11-23; 修改日期: 2017-04-05

基金项目: 国家自然科学基金项目(41201100)

作者简介: 林海燕(1993-), 女, 福建莆田人, 硕士生, 主要从事交通运输规划与管理研究(E-mail:1760264747@qq.com);

*李正红(1964-), 女, 福建浦城人, 教授, 硕士生导师, 主要从事交通运输规划与管理等研究(E-mail:495273111@qq.com);

胡喜生(1979-), 男, 福建莆田人, 副教授, 硕士生导师, 主要从事 3S 技术及应用研究(E-mail:xshu@fafu.deu.cn);

郑小雪(1981-), 女, 福建福州人, 讲师, 博士, 主要从事管理决策, 物流工程技术方法等研究(E-mail:21079178@qq.com)。

于公路通车里程影响因素研究的相关文献主要运用普通线性回归方法。但线性回归模型假定了公路通车里程受各因素的影响强度与空间位置无关,忽略了空间误差自相关性,因此其研究结果只是“全局”的意义估算。

英国地理学家 fotheringham 等提出了地理加权回归模型 (GWR), 该模型能够很好揭示空间异质性条件下的空间关系, 具有较为强大的空间数据局部分析能力, 目前已被广泛运用于经济学、地理学、生态学等众多领域的空间数据研究。但目前此方法在公路通车里程数的应用研究较少。鉴于其强大的空间分析能力, 本文通过借助 ArcGIS10.0 和 GeoDA 等软件, 构建模型对福建省各市县的公路通车里程数与当地的农民人均可支配收入、工业总产值、社会消费品零售总额之间的空间特征进行分析研究, 发现福建省公路通车里程数空间分布特点及规律, 这对发展公路建设及开展路网规划等工作具有一定的理论价值和实践指导价值^[1-9]。

1 数据的来源与研究方法

1.1 研究区域与数据来源

本文以福建省的 67 个市县区域、面积约 13 万 km^2 为研究区域。数据通过查询《福建省统计年鉴》, 收集 1999、2004、2009 和 2014 年等 4 个年份各市县的公路通车里程数、农民人均可支配收入、工业总产值、社会消费品零售总额数据集, 分析过程中对数据的处理主要选用 SAM3.0、GeoDa 以及 ArcGIS10.0 空间数据分析工具的相关模块。

1.2 研究方法——地理加权回归模型

1.2.1 基本原理

地理加权回归模型 (GWR) 是对普通线性回归 (OLS) 模型的扩展, 如下所示:

$$y_i = \beta + \sum_{k=1}^p \beta_k \chi_{ik} + \varepsilon_i \quad (1)$$

将地理位置引入回归参数中, 允许局部的参数估计, 扩展后的模型如下:

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) \chi_{ik} + \varepsilon_i \quad (2)$$

式中, y_i 和 $\chi_{i1}, \chi_{i2}, \dots, \chi_{ip}$ 为因变量和解释变

量 $\chi_1, \chi_2, \dots, \chi_p$ 在位置 (u_i, v_i) 处的观测值; 而 (u_i, v_i) 是第 i ($i=1, 2, \dots, n$) 个观测点的地理空间坐标, ε_i 是均值为 0, 方差为 σ^2 的误差项; $\beta_k(u_i, v_i)$ ($k=1, 2, \dots, p$) 是关于空间位置的 p 个未知函数, 与所在位置有关。地理加权回归模型对每一个观测点 (u_i, v_i) 的地理位置使用加权最小二乘法估算, 权重常选取距离衰减函数的某种形式。在模型中, 带宽的确定会直接影响到权重函数, 进而影响模型的精度, 所以带宽的选择也很重要。本文是利用 ArcGIS10.0 软件的地理加权建模模块, 实现数据与地图联接, 构建地理加权模型, 在 ArcGIS10.0 地理加权建模模块有多个权重函数和宽带选项, 本文选取高斯函数作为权重函数和采用 Akaike 信息量准则 (AIC) 来确定模型的带宽^[10-11]。

2 数据的空间自相关性检验与建模

空间自相关性检验可以检测各地区的公路通车里程数是否具有一定的空间关联性, 这将为引入地理加权回归模型 (GWR) 研究各解释变量 (农民人均可支配收入、工业总产值、社会消费品零售总额) 对因变量 (公路通车里程数) 的影响程度的可行性奠定理论基础^[12-14]。

2.1 趋势面分析

通过利用 ArcGIS10.0 的地统计分析模块 (geostatistical analyst) 的全局趋势面分析 (trend analysis) 功能, 对福建省各区域的公路通车路程数分布的趋势进行分析。结果 (见图 1) 表明, 1999 年, 坐标 x 方向上的投影, 是一条中间高两端数值低的开口朝下抛物线; y 方向, 是一条递增的曲线。这说明在东西方向上福建省公路通车里程数先增加后降低, 南北方向上由北向南逐渐增加, 即内陆地区的公路通车里程数大于东部沿海和西部地区, 南部公路通车里程数高于北部; 2004 年的趋势面, 横轴方向投影, 抛物线略微扁平, 说明内陆公路通车里程数稍微下降, 纵轴上, 曲线变得倾直, 表明由北向南增加的速度加快; 2009 年, 横坐标上趋势面的曲线发生明显的变化, 由开口向下抛物线变成至左向右递增的曲线, 纵坐标上趋势面的曲线增加

的趋势呈现指数形式逐渐减缓,表明此期间东部沿海地区的公路建设发展极快,可能是福建省漳龙、三福、泉三等一些连接东部沿海和内陆地区的高速公路的建成通车,使得内陆和东部的公路通车里程数的差距逐渐减小;到了 2014 年,趋势面没有明显的转变趋于稳定,大致存在由西向东、由北向南逐渐递增的分布,这与现实福建省交通路网基本形成“两纵五横”主骨架是相符合的。总之,这四个时间点的趋势分布特征表明福建省的公路通车里程数存在某种空间聚集关系。

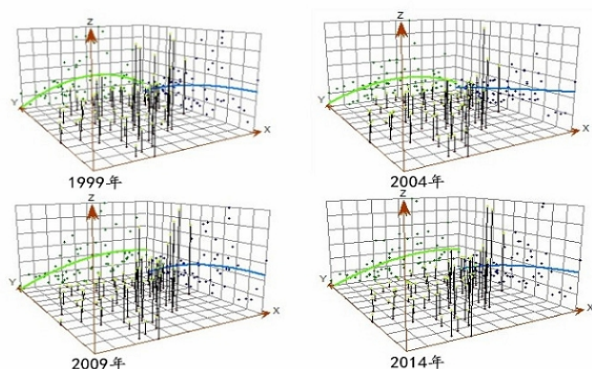


图 1 福建省公路通车里程数空间分布总体趋势

Fig. 1 Fujian highway mileage of the spatial distribution of the overall trend

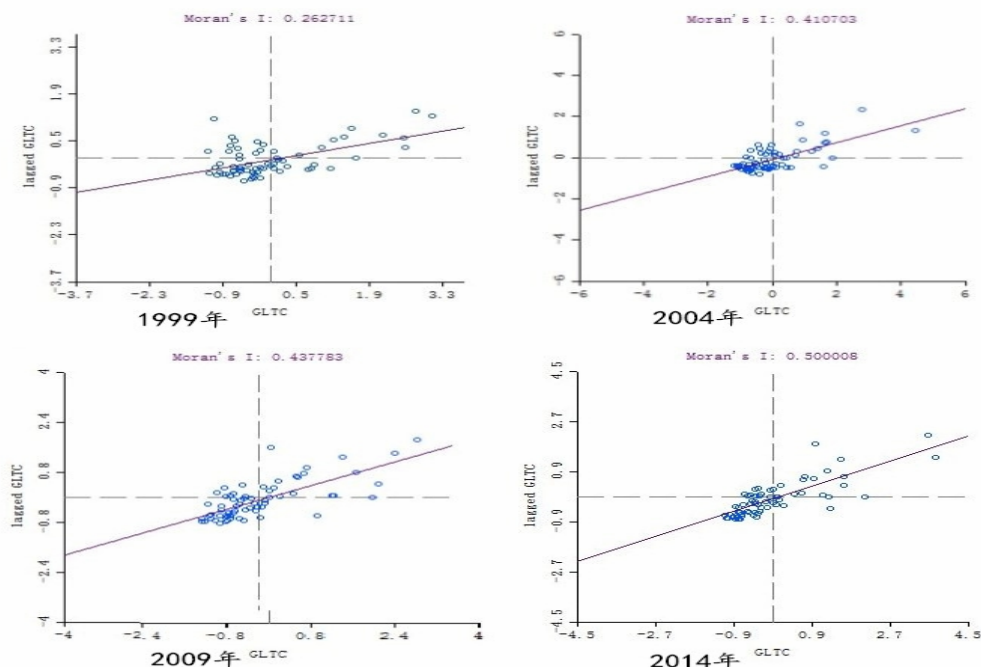


图 2 公路通车里程数 Moran 散点图

Fig.2 Highway mileage Moran scatterplot

2.2 局部空间自相关分析

全局的空间自相关分析可以实现从总体上观察各地区公路通车里程数的分布态势,而局部空间自相关分析能够进一步了解各地区通车里程数空间分布格局的演进过程。

2.2.1 Moran 散点图分析

Moran 散点图可以反映局部空间的异质性。散点图共有四个象限,如果观测值落于第一、第三象限,说明具有较强的空间正相关性;如果落于第二、第四象限,则存在较强的空间负相关性。本文运用空间数据分析 GeoDa 软件对数据进行处理,得到空间局部 Moran 散点图。

从 Moran 散点图(见图 2),可以看出各地区公路通车里程数不是均匀分布的,大多落于第一和第三象限,表明自 1999 年以来,公路通车里程数在局部空间存在较强的空间正相关。另外, Moran'I 指数均大于 0,四个时间点的 Moran'I 指数分别是 0.262711, 0.410703, 0.437783, 0.500008,莫兰数值呈逐年上升趋势,说明随着时间的递增,各地区的公路通车里程数存在的空间正相关性将越来越明显,区域分布不均匀。

2.3 模型的构建

根据以上空间自相关方法的分析,福建省公路

通车里程数的空间差异和变化是客观存在的。在此基础上,采用 SAM3.0 软件的 GWR 分析模块进行

了模拟计算,得到模型相关拟合参数。

表 1 GWR 与 OLS 模型拟合参数比较
Table 1 Result comparisons between GWR (Local model)
and OLS (Global model)

Coefficients	1999		2004		2009		2014	
	GWR	OLS	GWR	OLS	GWR	OLS	GWR	OLS
Sigma	0.018	0.14	0.027	0.18	0.061	0.263	0.067	0.279
AIC	-69.873	-66.389	-41.421	-33.114	11.653	18.181	18.473	26.085
R ²	0.466	0.392	0.654	0.58	0.632	0.563	0.745	0.691
Adjusted R ²	0.432	0.373	0.632	0.566	0.609	0.549	0.728	0.681
P-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0	<0.001

从表 1 可以看出,首先,GWR 模型的 AIC 值小于 OLS 模型,根据 Fotheringham 等人的评价标准,如果 GWR 模型的 AIC 值比 OLS 模型数值相差大于 3,即使考虑 GWR 模型的各种复杂性,GWR 模型也比 OLS 模型的应用性更佳;其次,决定系数 R² 四年的数值,GWR 模型都超过 OLS 模型,表明了 GWR 模型的解释性能更好。通过两模型相关拟合参数之间的对比,结果表明,GWR 模型的模拟结果更为准确,通过了相应的检验。

3 公路通车里程数的驱动因素结果分析

地理加权回归模型可以模拟空间位置变化的条件下因变量(公路通车里程数)和解释变量(农民人均可支配收入、工业总产值、社会消费品零售总额)的相互作用,并根据各解释变量的回归系数,分析各个解释变量对公路通车里程数的影响程度。本文运用 ArcGIS10.0 软件的空间分析工具——地理加权模型,进行建模,建模后将各解释变量的回归系数通过选取自然间断点分级法将各年份回归系数分为五类(红、蓝、绿、黄、灰,影响程度红色区域最大,灰色最小)显示在地图上,从而根据地图进行分析各解释变量对因变量(公路通车里程数)的空间影响。

3.1 农民人均可支配收入对公路通车里程数影响空间变异特征

从图 3 来看,1999 年农村居民人均可支配收入水平对公路通车里程数总体影响呈正向作用(系数 β_1 全为正数);2004 年回归系数表现有正有负绝

大部分为负效应,可见随着经济的快速增长,各地收入差距逐渐缩小,农村居民人均可支配收入对公路通车里程数的影响逐渐减小。但 2009 年总体转为正效应;2014 年又呈现有正有负。从空间分布上来看,1999 年回归系数呈现从南到北的递增;2004 年表现从西到东的递增,其中福州、泉州、厦门(红色区域)周围地区较高,向外回归系数逐渐减小,说明了该年份农村居民人均可支配收入水平在福州、泉州、厦门地区对公路通车里程数影响高于其它地区;2009 年和 2014 年的回归系数都高值点北上西移,集中在南平各地区,农村居民人均可支配收入水平回归系数呈现由西北向东南方向逐渐递减趋势,反映了西北地区农村居民人均可支配收入水平对公路通车里程数的贡献高于其它地区。总而言之,4 个年份回归系数大体呈波动变化,对公路通车里程数影响波幅比较宽,可见福建大部分地区的公路通车里程数是受农村居民人均可支配收入水平正负双重作用的影响。

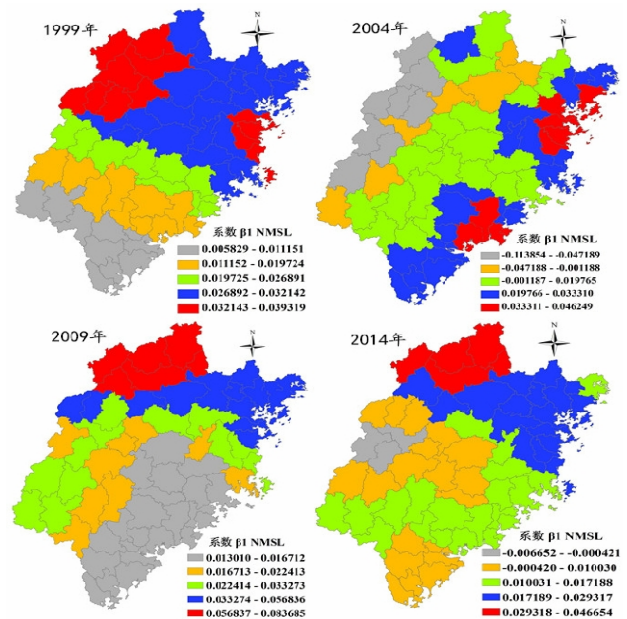


图 3 农民人均可支配收入系数估计的空间分布
Fig.3 Farmers per capita disposable income of the coefficient estimates of the spatial distribution

3.2 工业总产值对公路通车里程数影响空间变异特征

一个地区的工业发展水平是影响其公路建设的重要因素,图 4 展示了工业总产值对公路通车里程数影响的空间分布特征。从图中可以看出,1999 年工业总产值回归系数全部为正数,高值区(红色区域)主要分布在福建西北的南平和三明的若干个

地区, 低值区(灰色区域)聚集在福建东南的漳州和泉州一带, 系数空间分布呈现由西北向东南递减的趋势; 2004 年系数表现为有正有负, 其分布特征呈现高值区出现在南平、三明的部分地区, 然后系数由高值区向四周变小。值得注意的是泉州的部分地区转变成高值区, 这可能是实行改革开放和台湾局势的缓和, 经济建设中心逐渐由内陆向沿海地区转移, 沿海泉州的晋江, 石狮等的工业发展较为活跃, 从而对公路通车里程数的贡献较大; 2009 年回归系数的空间分布格局发生了较大的变化, 表现为福建西部向东部逐渐变小, 另外系数全部为正数, 数值与 1999 年、2004 年相比有较大幅度的增长, 说明该年份工业总产值对公路通车里程数的影响作用增强, 这可能是 2004 年, 海峡西岸经济区战略的提出, 海峡西岸经济区的主体福建省开始加大投资力度, 力促经济发展, 为工业发展注入生机和活力, 改善基础设施。到 2014 年, 系数出现较大幅度的下降, 表现为有正有负但大体上为正效应, 反映了随着工业水平的不断发展, 各地区工业水平差距逐渐缩小, 对公路通车里程数影响减小了。纵向比较这四个时间点, 总体上系数对公路通车里程数的作用为正效应, 即工业总产值越高, 公路通车里程数越大, 极少数地区工业发展步伐快慢不一以至于对公路通车里程数的影响程度产生差异。

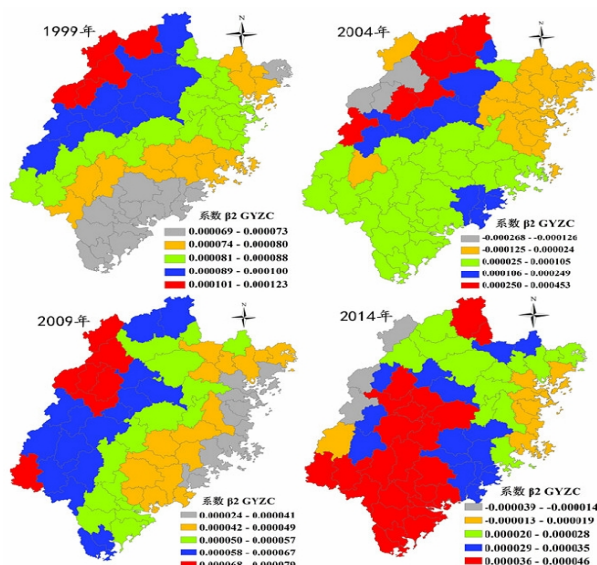


图 4 工业总产值系数估计的空间分布

Fig.4 Industrial output value of the coefficient estimates of the spatial distribution

3.3 社会消费品零售总额对公路通车里程数影响空间变异特征

图5所示, 绝大部分地区的社会消费品零售总额回归系数为负值, 这表明社会消费品零售总额越大, 公路通车里程数越小。1999—2014的4个年份的回归系数的高值区(红色区域)和低值区(灰色区域)空间表现出相对明显的空间分布特征。1999 年所有地区回归系数为负, 高值区主要分布在福建南部的漳州、龙岩一带, 低值区集中于南平等地区; 2004 年绝大多数地区回归系数为负, 少数地区上升为正, 高值区向西北方向转移, 低值区分散于南平、三明和沿海地区的厦门和泉州的几个点上。2009 年回归系数继续偏移, 高值区出现在福建东西两侧若干个地区, 而龙岩和南平的部分地区变成了最小值区域; 到2014 年回归系数出现增长呈现为有正有负但总体为负, 分布格局有了较大的变化, 高值区只剩福建西部一带若干个地方, 低值区则聚集在南部漳州地区。从这四个年份的GWR模型回归系数变化可以看出, 回归系数虽表现为小幅度的上升, 但对公路通车里程数影响仍为负面作用, 且低值区的数值的绝对值有减小的趋势, 说明抑制作用将逐渐减落。

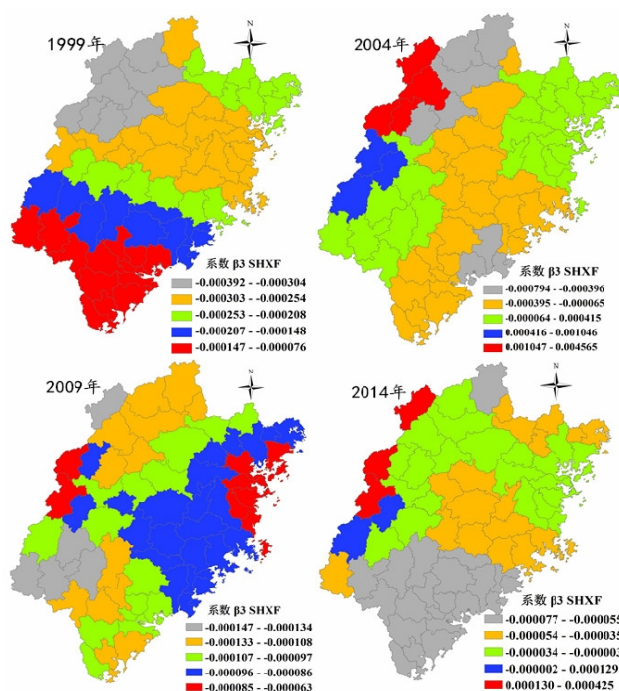


图 5 社会消费品零售总额系数估计的空间分布

Fig.5 Total retail sales of social consumer goods of the coefficient estimates of the spatial distribution

4 结论与展望

通过综合运用空间方法分析并构建模型,对福建省的公路通车里程数进行研究,结果显示公路通车里程数的驱动因素确实存在较为明显的空间相关性,且各驱动因素对各个地区的公路通车里程数的作用也不一致。

根据研究结果表明,在开展公路网规划工作时也应考虑不同地区的工业化发展状况、社会消费水平、城镇化建设进度等因素的影响,实现公路交通建设与社会经济等发展相协调,发挥最大的调控作用。据了解,福建省正开展海峡西岸经济区高速公路网规划的工作,建议在规划中应重视经济扩散效应区带动作用,加强地区联系和协作,在已建成的“两纵五横”的交通网络的基础上,早日实现“三纵、八横、三环、三十三联”的交通辐射格局,推进“大港口,大通道,大物流”建设为各地区往来提供便利的交通条件。

参考文献:

- [1] 张生瑞,周伟. 公路网规划理论与方法[M]. 北京:中国铁道出版社,2009:3-6.
- [2] 张先叶. 指数平滑法在辽宁省公路运输里程预测中的应用[J]. 物流技术,2015,34(8):159-161.
- [3] 张萌,张生瑞,王超深. 基于改进的 GM(1,1)模型的公路里程预测[J]. 山东交通学院学报,2009,17(3):31-34.
- [4] 杨玉建. 黄河三角洲粮食产量和典型影响因素的地理加权回归模型研究[J]. 中国农业工程学会 2011 年学术年会论文集,2011:3.
- [5] 赵山春,牛惠民. 基于熵的干线公路通车里程组合预测[J]. 交通科技与经济,2010,12(1):25-27.
- [6] 张琰,梅长林. 基于地理加权回归的我国中东部城市商品房价格的空间特征分析[J]. 数理统计与管理,2012,31(5):898-905.
- [7] 胡明星,吴江,朱选. 基于地理加权回归的南京市居住用地分布规律研究[J]. 东南大学学报:英文版,2015,31(1):149-156.
- [8] 佚名. 福建举行高速公路通车里程突破 3000km 暨宁武高速公路政和至武夷山段通车典礼[J]. 城市道桥与防洪,2012(12):119-199.
- [9] 江瑞明. 浅谈福建省高速公路交通量发展趋势[J]. 海峡科学,2010(11):99-101.
- [10] Brunsdon C, Fotheringham A S, Charlton M E. Geographically weighted regression: a method for exploring spatial nonstationarity[J]. Geographical analysis, 1996, 28(4): 281-298.
- [11] Brunsdon C, Fotheringham A S, Charlton M. Some notes on parametric significance tests for geographically weighted regression[J]. Journal of Regional Science, 1999, 39(3): 497-524.
- [12] 何清,陈楠,张开洲. 基于GWR模型的福建县域人口老龄化影响因素分析[J]. 贵州大学学报:自然科学版,2014,31(5):129-135.
- [13] 艾福利,庞西磊,汤庆园,等. 基于地理加权回归模型的浙江省台风经济损失影响因子分布规律研究[J]. 北京师范大学学报:自然科学版,2013(1):61-67.
- [14] 高新才,高新雨,丁绪辉. 基于 ESDA-GIS 的六盘山片区县域经济空间格局演变研究[J]. 西北师范大学学报:自然科学版,2014,50(4):111-116.