

川滇黔接壤地区城镇化空间特征及驱动力分析

彭立^{1,2}, 刘邵权¹

(1. 中国科学院、水利部成都山地灾害与环境研究所, 成都 610041; 2. 中国科学院研究生院, 北京 100049)

摘要: 川滇黔接壤地区经济落后, 整体城镇化水平较低, 采用川滇黔接壤地区 2008 年各区县(市)的城镇化率作为测度指标, 运用探索性空间数据分析(ESDA)方法, 分析了本区域的城镇化水平的空间分布特征。通过对城镇化率的空间自相关系数 Moran's I 的计算, 发现各区县城镇化率的空间分布并非表现出完全随机性, 而是具有明显的空间自相关性, 具体表现为低值空间聚集。利用空间计量模型研究了城镇化水平的驱动机制。结果表明: 空间误差模型优于空间滞后模型和 OLS 模型; 相邻地区各因素的空间扩散效应对城镇化水平存在显著的影响, 其弹性系数为 0.416 1; 第三产业增加值是城镇化率的最主要驱动因子, 其弹性系数为 0.777 7。因此, 未来川滇黔接壤地区在推进城镇化时, 应考虑城镇化水平的空间相关性, 建立区域一体化开发政策, 并积极发展第三产业, 提高第三产业在国民经济中的比重。

关键词: 城镇化; 空间特征; 空间计量模型; 川滇黔接壤地区

中图分类号: F129.9

文献标志码: A

文章编号: 1003-2363(2012)04-0075-04

0 引言

城镇化是区域发展的重要推动力量, 尤其是对于西部地区来说, 普遍较低的城镇化水平反映了区域经济社会发展的落后性。城镇化过程实际上是一种自组织过程, 城镇化的演变、人口的城乡迁移都受某种自组织规律的支配, 在实践和空间上分别联系着自组织临界性和空间复杂性^[1]。基于地理学视角对城镇化空间格局的实证研究是建立在空间分析方法论基础上的^[2]。城镇化的表现也不单是简单的城乡人口结构的转化, 更重要的是其代表着产业结构及其空间分布结构的变化。作为四川、云南、贵州三省交界之地的川滇黔接壤地区有着独特的地理位置, 长久以来经济社会发展比较落后, 其城镇化的整体水平较低, 在空间分布上探索其城镇化的规律性和相关性有着重要的现实意义。本研究以川滇黔接壤地区作为实证研究对象, 在 ArcGIS 支持下, 运用空间数据探索分析方法(exploratory spatial data analysis, ESDA)和空间计量模型等方法, 揭示川滇黔接壤地区城镇化格局的空间规律和驱动因素, 可为跨省区域经济社会空间结构重构、协调区域城镇化发展提供决策依据。

1 研究方法

空间分析的独特贡献在于它借鉴相关社会科学的

方法和工具, 提供了准确认识、评价和综合理解空间位置和空间相互作用重要性的方法^[3]。本研究采用的是 ESDA, 其主要是通过对空间自相关的分析来揭示空间依赖性和异质性, 并基于空间自相关分析的结果进行空间计量建模。

1.1 空间数据探索分析

空间数据探索分析是一般数据探索分析的扩展, 具有一些针对空间数据特性的工具, 目的在于探测数据的空间属性, 并对下一步的数学建模具有重要价值。总体而言, 有两类空间数据探索分析方法, 一类为全局统计(global statistics), 主要探索某一属性在区域中的分布特性, 另一类为局域统计(local statistics), 通过对子区域中信息的分析, 探查区域信息变化是否平滑(均质)或存在突变(异质)。

1.1.1 全局空间自相关。全局空间相关测度可用来描述整个研究区域上所有空间对象之间的平均关联程度、空间分布模式及其显著性, 常用 Moran's I 和 Geary's C 进行检验^[4]。本研究采用的是 Moran's I 指数。Moran's I 的取值范围为 $[-1, 1]$, 大于零表明存在空间的正相关, 反之为负相关, 等于零则表明不存在空间相关性, 其计算公式如下:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{S^2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \quad (1)$$

式中: $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$, 表示样本方差值; $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, 表示 X_i 的均值; n 表示研究地区的数目(本研究为 75 个); X_i 为观测值; w_{ij} 为研究对象 i, j 之间的空间邻接矩阵, 其目的是定义空间对象的相互邻接关系, 便

收稿日期: 2011-07-28; 修回日期: 2012-05-28

基金项目: 中国科学院知识创新工程重要方向项目(KZCX2-YW-333); 国家自然科学基金项目(41071350)

作者简介: 彭立(1983-), 男, 山东滕州人, 研究实习员, 博士研究生, 主要从事水土资源可持续利用研究, (E-mail) pengli@imde.ac.cn。

于把地理信息系统数据库中的有关属性放到所研究的地理空间上来对比。通常,空间邻接矩阵可以通过空间数据的拓扑属性如邻接性来构造,也可通过空间距离来构建。一般邻接矩阵 W_{ij} 为:

$$W_{ij} = \begin{cases} 1, & \text{当区域 } i \text{ 和 } j \text{ 区域相邻} \\ 0, & \text{当区域 } i \text{ 和区域 } j \text{ 不相邻} \end{cases} \quad (2)$$

计算出 Moran's I 之后,还需对其结果进行统计检验,一般采用 Z 检验:

$$Z(I) = \frac{I - E(I)}{\sqrt{\text{var}(I)}} \quad (3)$$

式中: $E(I)$ 表示数学期望; $\text{var}(I)$ 为总体方差。

1.1.2 局部空间自相关。空间联系局域指标^[5] (local indicators of spatial association, LISA) 是由全局空间 Moran's I 向局域或者单个空间研究对象的分解,可以表示某个位置上的观测值与周围区域单元观测值之间的联系,可以用于识别“热点区域”以及数据的异质检验^[6]。对于某个空间单元 i ,其计算公式为:

$$I_i = Z_i \sum W_{ij} Z_j \quad (4)$$

式中: I_i 为单元 i 的空间自相关指数; $Z_i = (X_i - \bar{X}) / \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$; $Z_j = (X_j - \bar{X}) / \sqrt{\frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2}$, Z_i, Z_j 都表示为观测值的标准差标准化形式; n 为研究单元个数(本研究为 75 个);其余字母含义同前文。

LISA 的 Z 检验为: $Z(I_i) = \frac{I_i - E(I_i)}{\sqrt{\text{var}(I_i)}}$, 式中: 字母含义同前文。

1.2 空间计量模型

空间相关性通常由两方面因素决定,一是地区样本观测值对相邻地区观测值存在的溢出效应,二是邻近地区关于因变量的误差冲击对本地区观测值产生的影响,因此,空间相关性主要体现在因变量的空间滞后项和空间误差项^[7]。本研究使用的空间计量模型主要是纳入了空间效应(空间相关和空间差异)的空间常数回归模型,包括空间滞后模型与空间误差模型 2 种。

1.2.1 空间滞后模型(spatial lag model, SLM)主要反映了样本观测值是如何通过空间机制作用于其他地区。其模型表达式为:

$$Y = \rho WY + X\beta + \epsilon \quad (5)$$

式中: Y 为因变量; X 为自变量; ρ 为空间滞后回归参数; WY 为空间滞后因变量; ϵ 为矢量形式的空间误差项。

1.2.2 空间误差模型(spatial error model, SEM)主要用来度量邻近地区因变量的误差冲击对本地区测察值的影响程度,SEM 模型强调空间扩散效应是外生冲击的结果^[8],因此,该模型中的空间相关作用存在于误差项当中。公式为:

$$Y = X\beta + \epsilon; \epsilon = \lambda W\epsilon + \mu \quad (6)$$

式中: ϵ 为矢量形式的空间误差项; λ 表示空间误差回归系数; W 为空间邻接矩阵; μ 为满足正态分布的随机误差向量。

因 SLM 模型和 SEM 模型反映的空间相关性是全局性的,应确保模型估计结果的无偏和有效性。采用普通最小二乘法(OLS) 估计时,则会造成估计结果的有偏或无效,因此,本研究主要选取极大似然法(log L) 进行估计^[8]。

2 研究区概况和数据获取

川滇黔接壤地区包括四川省西南部、云南省东北部和贵州省西北部 14 个地、市、州的 75 个县(区、市),总面积 19 万多 km^2 , 分别占西南 3 个省的 16.9%。该区是我国西南的一大老少边穷地区,它处于西南的几何中心,依托重庆、成都、昆明和贵阳四大城市,以及成昆、贵昆、内昆及川黔铁路。该区曾是国家“三线”建设中生产重点布局地区,也是全国国土开发规划中 17 个重点开发区之一(即攀西—六盘水资源综合开发区),长江沿岸产业带开发最上游地区。川滇黔接壤地区城镇化水平整体较低,除个别区县(攀枝花市的东区、西区,六盘水市的钟山区)城镇化率达到 50% 以上外,其余 72 个区县都在 50% 以下,其中 49 个区县的城镇化率在 20% 以下(图 1)。空间分析的范围是川滇黔接壤地区,包括 75 个区县,以区县为基本分析单位。空间数据从 1 : 400 万的国家基础地理信息数据中提取出以县域为基本尺度的川滇黔接壤地区行政边界,城镇化数据均以 2008 年为时间节点,来自相关省市统计年鉴和统计公报,并先在 ArcGIS 软件中进行数据匹配。

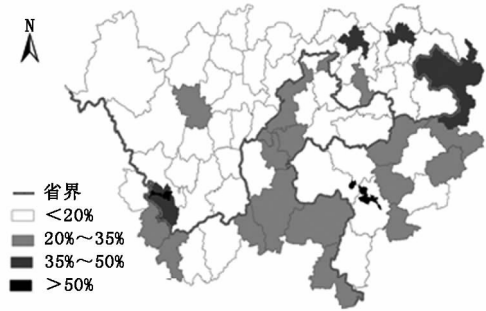


图 1 川滇黔接壤地区城镇化水平分级格局

Fig. 1 The spatial pattern of urbanization level in Chuan-Dian-Qian boundary area

3 空间分析结果

经计算,全局空间自相关指数 Moran's $I = 0.2538$, 检验值 $Z = 3.42$, 通过 $p < 0.01$ 水平下的显著性检验。城镇化率呈现正的空间自相关特性,即从川滇黔接壤地区整体上看,城镇化水平相似的县域(高—高或低—低)在空间上呈集聚趋势。利用局部 Moran 指数公式,计算各区县城镇化率的局部 Moran's I 值,得到各区县城镇化率空间分异状态的 Moran 散点图(图 2)。

Moran 散点图以笛卡尔直角坐标体系为表现形式,分析全省耕地资源的空间相关性,横坐标为城镇化率,纵坐标为空间权重矩阵加权后的城镇化率,直观反映出研究变量与空间滞后的关系,即检测局部空间的异质性^[9]。第一至四象限点分别表示某空间单元与相邻单元的城镇化率呈现高值与高值集聚、高值与低值集聚、低值与高值集聚、低值与低值集聚的空间关系。从图中可以看出各区县单元分布在第三象限的最多,再次说明了城镇化率更多呈现正相关,具体更多地表现在低值聚集的状态,城镇化率较低的地区相对集中分布,形成连片的低值区。根据 Moran 散点图中各象限的区县不同类型在 ArcGIS 中做出其分布图,并将计算出的各区县显著性 LISA 值 ($p \leq 0.05$) 也标示在图上,得到图 3。

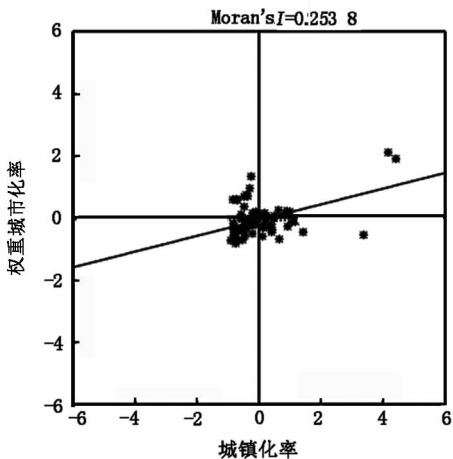


图 2 川滇黔接壤地区城镇化率 Moran 散点图

Fig. 2 Moran scatter plots of urbanization rate in Chuan-Dian-Qian boundary area

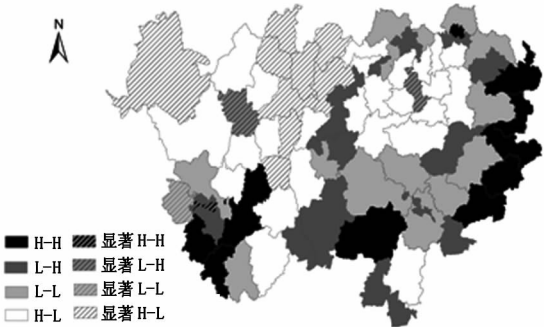


图 3 川滇黔接壤地区城镇化率 LISA 集聚图

Fig. 3 LISA cluster map of urbanization rate in Chuan-Dian-Qian boundary area

分布在第一象限(高-高)的有 12 个区县,分布在第三象限的有 34 个区县,两者共占川滇黔接壤地区县域总数的 61.33%,城镇化水平呈现出一定程度的两极分化。相对于 H-H 区, L-L 区范围明显更广,主要分布于川滇黔接壤地区西北部和中北部,并且连片出现,这些区域成为城镇化率低值聚集区不足为怪,主要源于该区域恶劣的自然地理条件及长期落后的发展基础。城镇

化的高值区域分布比较离散,空间集聚性不强,还没有形成较大范围连片集中分布的城镇化高值区,唯一显著的 ($p \leq 0.05$) 是攀枝花市的东区和西区,它们是整个川滇黔接壤地区经济社会发展水平最高的区域,也是城镇化水平最高的区域。H-L 区分布特征最为分散,呈镶嵌状分布,这也体现了即使在城镇化水平较低的地区,一定的行政区域内(比如一个地级市)也存在城镇化水平的差异,即经济、资源和人口的相对富集造成的局部极化差异。

4 空间计量模型估计结果及分析

由于川滇黔接壤地区城镇化率整体上存在一定的空间正相关性,因此,地区之间在空间上相互独立的假定不能成立,从而如果采用普通最小二乘法来估计分析城镇化率的驱动作用^[10],其结果是有偏的,在处理这些带有空间特性的数据时必须考虑空间自相关所带来的影响。采用空间滞后模型和空间误差模型进行空间计量分析可以较好地反映空间效应(表 1)。当然,为了比较,同时进行了 OLS 估计。

区域经济结构差异直接导致了区域经济发展水平不同,进一步影响了区域城镇化水平。为揭示城镇化率差异的驱动力的深层机理,根据数据可得性、相关性、整体性、代表性等原则选取 4 个主要影响因子,包括人均 GDP、人口密度、第二产业增加值(亿元)和第三产业增加值(亿元)。

从拟合优度 R^2 和极大似然估计量来看,空间误差模型和空间滞后模型明显优于最小二乘法模型,再次证明了考虑空间效应的优越性和必要性,避免了模型估计偏差。对于空间误差模型和空间滞后模型之间的优劣比较来说,空间误差模型的拟合优度 R^2 和极大似然估计值大于空间滞后模型的,且从赤池信息量和施瓦茨信息量可以看出,空间误差模型均小于空间滞后模型。综上所述,空间误差模型为最优。

从空间误差模型估计结果看,空间误差弹性系数 λ 为 0.416 1,通过 1% 显著性检验,说明相邻地区城镇化率的空间效应影响存在于误差项中,不仅包含地区间城镇化率的相互作用,还存在于各种复杂的空间因素中,因此,各区县城镇化水平的落后不仅决定于其本身,也受周围区县的影响。从自身驱动因子来看,人均 GDP、人口密度、第二产业增加值、第三产业增加值的弹性系数均为正,即对城镇化水平有正效应。第二产业增加值和第三产业增加值的弹性系数分别是 0.417 3 和 0.777 7,其中,第三产业增加值这一因子的弹性系数最大,说明第三产业增加值是城镇化水平最重要的驱动因子,这也符合了国内外城镇化的相关研究结论。发展经济学指出,随着资本密集化程度的提高和科学技术的进步,现代工业部门创造的就业机会已越来越少,大量的农村劳动力将转移到城镇商业、服务业等第三产业领

域。美国、日本等国经济发展实践也得出这样的结论，

表 1 模型估计值及检验结果

Tab. 1 Estimation and test results, assessed with the 3 models

变量	最小二乘法		空间滞后模型		空间误差模型	
	系数	T 检验	系数	Z 检验	系数	Z 检验
常数项	-0.346 5***	-3.603 8	-0.377 1***	-4.029 8	-0.324 1***	-3.770 5
人口密度	0.000 2***	4.767 1	0.000 2***	4.035 9	0.000 3***	5.631 8
人均 GDP	0.049 4**	2.608 8	0.049 5***	2.735 2	0.029 4	1.485 3
第二产业增加值	0.396 7***	3.312 5	0.393 9***	3.457 1	0.417 3***	3.829 4
第三产业增加值	0.878 1***	4.087 0	0.906 3***	4.403 0	0.777 7***	4.095 0
空间弹性系数	—	—	0.156 9	1.533 3	0.416 1***	3.501 2
拟合优度	0.724 0		0.733 2		0.764 7	
极大似然估计量	74.396 0		5.403 9		78.359 7	
赤迟信息量	-138.792 0		-138.808 0		-146.719 0	
施瓦茨信息量	-127.205 0		-124.903 0		-135.131 9	

说明：极大似然估计量、赤迟信息量和施瓦茨信息量 3 个指标是对多个模型拟合效果进行检验的指标；***，** 分别表示在 1%，5%水平下显著。

即城镇化水平与第三产业发展的相关性高于与第二产业发展的相关性，第三产业是城镇化的最大推动力。因此，本地区落后的城镇化水平和本地产业结构的不合理、第三产业不发达密切相关。

5 结语

从空间分析的结果来看，城镇化率的低水平正相关，可能是区县间的相互联系和相互作用的结果，也可能解释为区域本底条件和历史发展基础的相似性。这也说明作为一个研究单元，川滇黔接壤地区虽然跨四川、云南、贵州 3 个省，但从区位、资源、环境、发展历史等综合因素来讲，其区域的整体性较好。采用统计学方法研究某一区域发展问题时，需要考虑到区域作为研究样本事实上的“不独立”特性，考虑所处的区域大背景。

采用空间计量模型较好地揭示了各驱动因子对城镇化率的影响机制，体现了城镇化率分布的空间相关性，整个估计结果更为可靠。除了本身的各驱动因子的作用外，相邻地区各因素的空间误差效应也对城镇化率存在显著的影响，其弹性系数为 0.416 1，因此，各区县城镇化水平的落后不仅决定于其本身，也受周围区县的影响。在城镇化水平自身驱动因子中，第三产业增加值的弹性系数最大，说明第三产业是城镇化水平最重要的驱动因子，提高城镇化水平最重要的方向是产业结构的升级。国内外研究表明，产业结构遵循着从第一产业向第二产业再逐步向第三产业的方向递进，即产业结构高级化趋势。从各国城镇化进程发展结果来看，发展第三产业是城镇化的必由之路。

川滇黔接壤地区虽然分属云贵川三省，但山水相连，自然地理环境相近，区位条件一致，社会经济发展水平总体相近，资源丰富且相互配套。按照市场经济和资源优化配置原则，接壤地区必须打破行政区间封闭的状态，整合优化区域开发，才能提高城镇化水平，适应区域发展和西部大开发的战略需要。因此，将川滇黔接壤地

区作为一个区域整体来开发^[11]，有其特定的内涵和战略思考。加快本区的综合开发，促进地区社会经济的持续发展，对促进长江沿岸产业带建设，以及东、中、西部协调发展具有重大意义。

参考文献：

[1] 陈彦光. 城镇化:相变与自组织临界性[J]. 地理研究, 2004,23(3):301—311.

[2] 马晓冬,马荣华,徐建刚. 基于 ESDA-GIS 的城镇群体空间结构[J]. 地理学报,2004,59(6):1048—1057.

[3] Anselin L. The Future of Spatial Analysis in the Social Sciences[J]. Geographic Information Sciences, 1999,5(2):67—76.

[4] 孟斌,王劲峰,张文忠,等. 基于空间分析方法的中国区域差异研究[J]. 地理科学,2005,25(4):393—400.

[5] Anselin L. Local Indicators of Spatial Association-LISA [J]. Geographical Analysis,1995,27(2):93—115.

[6] Anselin L, Syabri I,Kho Y. GeoDa: An Introduction to Spatial Data Analysis[J]. Geographical Analysis, 2006, 38(1):5—22.

[7] 文继群,濮励杰,张润森. 耕地资源变化的空间计量及其驱动力分析——以江苏省为例[J]. 长江流域资源与环境,2011,20(5):628—634.

[8] 周国富,连飞. 中国地区 GDP 数据质量评估——基于空间面板数据模型的经验分析[J]. 山西财经大学学报,2010,32(8):17—24.

[9] 彭颖,陆玉麒. 成渝经济区经济发展差异的时空演变分析[J]. 经济地理,2010,30(6):912—917.

[10] 龙爱华,徐中民,王新华,等. 人口、富裕及技术对 2000 年中国水足迹的影响[J]. 生态学报,2006,26(10):3358—3365.

[11] 陈治谏. 川滇黔接壤地区总体发展战略研究[J]. 长江流域资源与环境,1994,3(3):193—199.

Spatial Characteristic of the Urbanization Level and Its Driving Forces in Chuan-Dian-Qian Boundary Area

Peng Li^{1,2}, Liu Shaoquan¹

(1. *Chengdu Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu 610041, China*; 2. *Graduate School of Chinese Academy of Sciences, Beijing 100049, China*)

Abstract: The whole urbanization level of Chuan-Dian-Qian boundary area is still low. Using principal component analysis and clustering ESDA method, the paper reveals spatial characteristic of the urbanization rate of 75 counties in Chuan-Dian-Qian boundary area. The spatial linkage of counties is described with a spatial weights matrix. There are significant positive spatial autocorrelation (Moran's I) of the urbanization rate in Chuan-Dian-Qian boundary area. That is, the relatively low level county tends to be located nearby other low level county more often than expected due to random chance. The driving mechanism of urbanization rate is analyzed using spatial econometric model. Some results are drawn as follows: (1) For this case, spatial error model is better than spatial lag model and OLS model. (2) There was an obvious significant spatial diffusion effect in the urbanization level in various factors in adjacent areas, and the elastic coefficient was 0.416 1. (3) The added value of tertiary industry is the main driving factor and its elastic coefficient was 0.777 7. Therefore, we should take the spatial correlation into consideration in the future to improve urbanization level rapidly, and establish an integrative regional development policy, and develop the tertiary industry to raise the proportion of the tertiary industry in the national economy.

Key words: urbanization; spatial characteristic; spatial econometric model; Chuan-Dian-Qian boundary area