

doi: 10.11830/ISSN.1000-5013.201704010



# 应用 ANN 与 CA 构建 GIS 模型的 城市非正规商业行为分析

吴保宽

(华侨大学 建筑学院, 福建 厦门 361021)

**摘要:** 基于神经网络算法(ANN)和元细胞自动机(CA),建构台湾逢甲夜市商圈的地理信息系统(GIS)模型,并对非正规摊贩商业行为进行数值模拟与实证分析.研究结果表明:非正规摊贩分布模拟结果为平均正确率 84.1%,而目标模式的平均正确率可达 94.2%;在环境因子影响性分析方面,街道位置类型是发生非正规摊贩最具影响性的环境因子,而背景建物宽度则是最不具影响性的环境因子;在环境因子相关性分析方面,街道位置类型因子是所有假设环境因子中最能吸引非正规摊贩聚集的正相关因子.

**关键词:** 城市非正规商业行为;神经网络算法;元细胞自动机;地理信息系统;环境因子

**中图分类号:** TU 17      **文献标志码:** A      **文章编号:** 1000-5013(2017)04-0497-06

## Analysis of Urban Informal Business Behavior Based on GIS Model Constructed by ANN and CA

WU Baokuan

(College of Architecture, Huaqiao University, Xiamen 361021, China)

**Abstract:** Based on the artificial neural network (ANN) algorithm and the cellular automation (CA), the geographic information system (GIS) model of the Fengjia night market in Taiwan was constructed, the commercial behavior of the informal street vendors was simulated. The results show that: the average simulation accuracy of the informal street vendor distribution is 84.1%, but the average simulation accuracy of target pattern is 94.2% high. In the environmental factor impact analysis, the street location type is the most influential environmental factor of the informal vendors. The background building width is the least influential environmental factor. In the environmental factor correlation analysis, the street position type factor is the positive correlation factor that is the most attractive environmental factor for attracting informal vendors.

**Keywords:** urban informal business behavior; artificial neural networks; cellular automation; geographic information system; environmental factors

数值模拟分析技术是城乡规划相关研究的重要发展领域.通过数值模型与研究对象的现实状态进行比对实验,能帮助研究者解析研究对象的深层知识<sup>[1-2]</sup>.近年来,由于地理信息系统(GIS)的发展,大幅提升数值模型技术的图像与数值信息的同步整合能力<sup>[3]</sup>.因此,各类数值模拟分析技术与 GIS 的合并应用成为近年城乡规划相关研究的重要趋势<sup>[4-5]</sup>,尤其是与各种人工智能算法结合的应用研究<sup>[6-8]</sup>.神经网络算法(ANN)<sup>[9-10]</sup>能依据研究者所观测环境因子与目标模式数据,通过训练程序建构研究对象的

**收稿日期:** 2016-10-17

**通信作者:** 吴保宽(1976-),男,讲师,博士,主要从事人工智能技术应用于城乡建筑设计与研究. E-mail: py1006@hqu.edu.cn.

**基金项目:** 国家自然科学基金资助项目(41401224);福建省自然科学基金资助项目(2016J01238)

专家知识数值模型<sup>[11]</sup>. 基于元细胞自动机(CA)原理建构的 GIS 模型,能以 CA 单元细胞格式与增生规则,体现城市非正规商业行为的碎形不规则发展特性,也能藉 GIS 技术记录各项环境观测数据,以图像同步展现各阶段研究成果<sup>[12]</sup>. 城市非正规商业行为是城市街道具有自发性的复杂行为,结合 ANN 与 CA 技术的 GIS 模型,不仅能模拟分析与描述城市非正规商业行为,也能预测其未来的发展趋势. 本文应用 ANN 和 CA 建构台湾逢甲夜市商圈的 GIS 模型,进行非正规商业行为的数值模拟,以及进行环境因子的分析实验,解析非正规摊贩在夜市街道上的分布状态,揭示各环境因子对此非正规商业行为的影响性与相关性.

1 研究应用技术

1.1 神经网络算法 ANN

ANN 算法具有许多特性,如目标优化能力、输入信息的误差容忍能力、因子分析能力、目标模式的预测能力等<sup>[10-11]</sup>. ANN 基于模式辨识与分类功能,在城乡规划相关研究领域有许多成果<sup>[6-8]</sup>. 多层 ANN 基于非线性技术难以解析目标模式数据与输入因子数据的关联性<sup>[13]</sup>,而单层 ANN 不仅能执行目标模式辨识与分类任务,也能通过 ANN 因子权重值分析与统计图表分析方法,理解各因子对目标模式的影响性及相关性. 文中采用的就是单层 ANN 技术,ANN 算法通过训练程序可降低输出数值与现实目标模式值的误差,以达成最佳目标模式分类结果. 完成训练程序的 ANN,最终模式输出值为

$$Y_j = \delta(\sum_{i=1}^n W_{i,j} + X_i + \beta). \tag{1}$$

式(1)中: $Y_j$  为 ANN 模式值输出结果; $X_i$  为第  $i$  项输入因子; $W_{i,j}$  为 ANN 因子权重值; $\delta$  为转换函数; $\beta$  为偏压值.

1.2 以 CA 原理建构的 GIS 模型

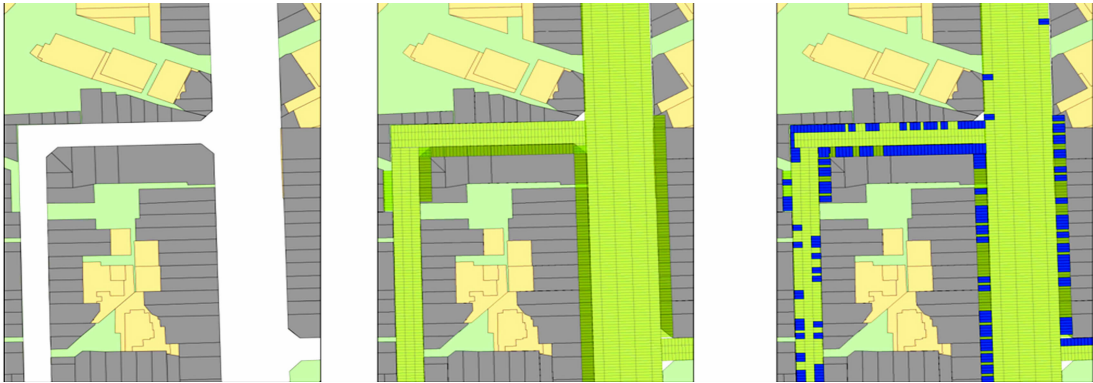
应用 GIS 进行城乡发展相关研究时,一般采用“宗地图”作为基础图形单元. 然而“宗地图”属于规范化基础土地轮廓,无法充分表达城乡发展过程的非正规行为模式. 因此,相关学者采用碎形图形(fractal graphics)作为应用 GIS 的基础图形单元,以表现城乡发展过程有如生命自组织的演化状态. 应用 CA 原理建构的 GIS 模型即属此类技术<sup>[9,12]</sup>. 以 CA 原理建构 GIS 模型具有有机性、相似性、自组织发展性、由下至上的发展逻辑等特点,很适合应用于研究城市非正规商业行为的数据记录、分析与成果展现等工作执行.

传统的 CA 程序有设定单元格式、设定单元模式状态、设定邻近关系、制定演化规则 4 项. 随著 CA 技术多元发展及与其他算法的结合应用,上述 CA 程序在操作上变得更加多元灵活<sup>[1-2,5]</sup>. 文中研究的 CA 应用操作主要在设定单元格式和设定单元模式状态两 CA 程序上,可藉 GIS 模型的建构过程进行设定;而设定邻近关系与制定演化规则两个 CA 程序,则能在 ANN 程序中以相关环境因子的方式进行设定,并由 ANN 最终判定各 CA 单元的模式状态.

2 研究对象及现场调研

文中研究实证材料为台湾逢甲夜市商圈的非正规摊贩群. 逢甲夜市是典型沿街发展夜市<sup>[14]</sup>,时常可见非正规摊贩占据街道两边的人行道及骑楼区域. 这些非正规摊贩群属不合法商业行为,然而其聚集所产生的热络氛围,是逢甲夜市成为知名夜市区的主要原因. 实验区域为逢甲夜市各街道区域. 为记录非正规摊贩的分布现况,将实验区域以 CA 网格单元形式建构 GIS 模型. CA 网格化过程,如图 1 所示. 图 1(c)中:深色标记表示被非正规摊贩所占据的 CA 单元.

实验区域共有 10 940 个 CA 单元. 经现场调研,其中被非正规摊贩占据 1 717 个 CA 单元(图 2)为文中研究目标模式(称 T1 模式并以数值 1 记录);另外的 9 223 个 CA 单元则是未被占据(称 T0 模式并以数值 0 记录). 由相关文献资料整理出 7 个环境因子,用于描述各 CA 单元所处位置与相关街道特征的属性与距离关系,如表 1 所示. 这些街道特征是有关学者归纳出具有吸引非正规摊贩聚集的潜在效应<sup>[14-15]</sup>,而所有环境因子数值都是经过归一化,设定在 0~1 之间. 有关算法的各程序都是以 MATLAB



(a) 原始街道状态 (b) CA 网格化街道区域 (c) 被非正规摊贩所占据的 CA 单元

图 1 以 CA 网格单元形式建构 GIS 模型的非正规摊贩分布

Fig. 1 Distribution of informal vendors by GIS model with CA grid formation

表 1 描述 CA 单元的环境因子

Tab. 1 Environmental factors to describe CA units

环境因子	因子说明
街道位置类型(SL)	CA 单元在街道的位置类型(骑楼、人行道、自行车道、慢车道、快车道等).
与十字路口距离(IntF)	CA 单元与最近十字路口的距离.
街道宽度(SW)	CA 单元所处街道的宽度.
背景建物宽度(BBW)	CA 单元背景建物的宽度.
与核心区 1 的距离(C1)	CA 单元与核心区 1 的距离;核心区 1 是逢甲大学的正门.
与核心区 2 的距离(C2)	CA 单元与核心区 2 的距离;核心区 2 是城市一级道路的交口.
与核心区 3 的距离(C3)	CA 单元与核心区 3 的距离;核心区 3 是城市一级与二级道路的交口.

语言所编写,GIS 模型则是以 ArcGIS 软件来操作执行.

3 数值模拟与现场调研比对

首先,进行 ANN 判定 CA 单元发生目标模式的 GIS 模拟结果. 有关训练 ANN 设定方面,是将所有 CA 单元分为 1/2 训练集和 1/2 测试集进行 ANN 训练程序. ANN 转换函数为 Logistic Sigmoid 函数,使 ANN 模式输出值限制在 0~1 之间. 完成 ANN 训练后,各 CA 单元可由式(1)产生最终 ANN 模式输出值,并依此来判定 CA 单元是否发生目标模式.

实验共执行 20 次 ANN 程序以测试 ANN 结果的稳定性,如表 2 所示. 由表 2 可知:实验模拟结果为平均正确率 84.1%,而 T1 目标模式的平均正确率可达 94.2%.

表 2 ANN 目标模式模拟结果

Tab. 2 ANN simulation results of target pattern

序号	T1 群 误判数量	总误判 数量	T1 群 正确率	总正 正确率	序号	T1 群 误判数量	总误判 数量	T1 群 正确率	总正 正确率
1	99	1 743	0.942	0.841	11	101	1 743	0.941	0.841
2	101	1 740	0.941	0.841	12	99	1 732	0.942	0.842
3	99	1 736	0.942	0.841	13	98	1 744	0.943	0.841
4	100	1 744	0.942	0.841	14	99	1 735	0.942	0.841
5	99	1 742	0.942	0.841	15	97	1 735	0.944	0.841

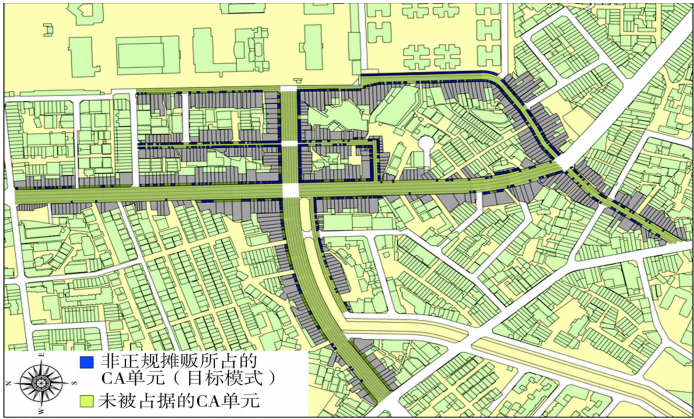


图 2 目标模式分布的现况调查结果

Fig. 2 In-situ investigation of distribution of target pattern

续表 2

Continue table

序号	T1 群 误判数量	总误判 数量	T1 群 正确率	总正 确率	序号	T1 群 误判数量	总误判 数量	T1 群 正确率	总正 确率
6	98	1 741	0.943	0.841	16	101	1 744	0.941	0.841
7	99	1 746	0.942	0.840	17	99	1 742	0.942	0.841
8	98	1 743	0.943	0.841	18	98	1 742	0.943	0.841
9	101	1 747	0.941	0.840	19	98	1 744	0.943	0.841
10	98	1 737	0.943	0.841	20	99	1 747	0.942	0.840
					平均	99.050	1 741.350	0.942	0.841

ANN 判定 CA 单元发生目标模式的 GIS 模拟结果,如图 3 所示.图 3 中:深色 CA 单元式是 ANN 所判定的目标模式.将图 3 与图 2 进行比对,观察 ANN 模拟结果与调研现况的相似性.平均模拟正确率代表环境因子设定(表 1)对目标模式的掌握与描述能力,下一实验阶段将以此模拟表现为基础,进行环境因子的分析工作.

4 环境因子分析结果

ANN 模式分类逻辑源于数学回归技术,单层 ANN 的训练过程是一种线性函数的最优求解过程.完成训练的 ANN 函式其因子权重值(式(1)的  $W_{i,j}$ ),表示 ANN 在多维空间进行模式分类时,对各输入因子的依赖程度<sup>[11,16]</sup>.因此,对单层 ANN 因子权重值进行分析,能观察出哪个环境因子对目标模式具有较高影响性.表 3 为已训练 ANN 的因子权重值,其数值绝对值越高者,代表其所对应的环境因子对目标模式的影响性越高.各环境因子其 ANN 因子权重绝对值高低排列如下:SL,C2,C3,C1,SW,IntF,BBW.由此可知:SL 是所有假设环境因子中,对非正规摊贩的发生最具影响性;而 BBW 则是最不具影响性的环境因子.

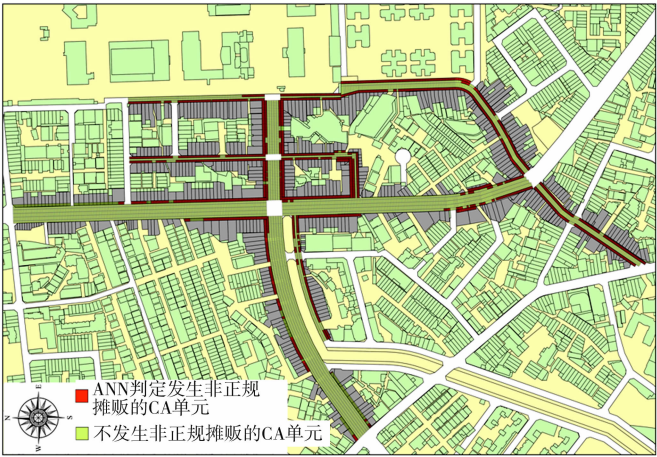


图 3 以 GIS 展现 ANN 模拟结果

Fig. 3 ANN simulation result demonstrated by GIS

表 3 以 ANN 因子权重值分析环境因子的影响性

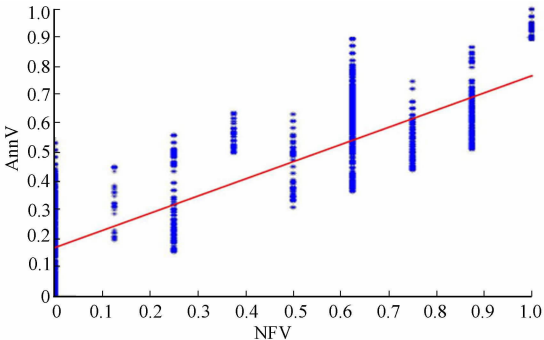
Tab. 3 Impact analysis of environmental factors by ANN weight values

环境因子	SL	IntF	SW	BBW	C1	C2	C3
原始 ANN 因子权重值	0.503 7	0.112 2	-0.127 0	-0.026 1	-0.173 2	0.197 4	-0.183 6
绝对值后的 ANN 因子权重值	0.503 7	0.112 2	0.127 0	0.026 1	0.173 2	0.197 4	0.183 6

环境因子的相关性分析工作,则须搭配统计图表分析方法的散点图和皮尔森相关系数来进行.所有 CA 单元的 ANN 模式输出值与各环境因子变数值的散点图分析结果,如图 4 所示.图 4 中:AnnV 为 ANN 模式输出值;NFV 为经过归一化的环境因子变数值.各环境因子变数值与 ANN 模式输出值的皮尔森相关系数分析结果,如表 4 所示.

由表 4 及图 4 可得,正相关环境因子为:SL,IntF, BBW 和 C24 个环境因子,其余则为负相关因子.由以上 ANN 因子权重值分析与散点图、皮尔森相关系数分析后可得:街道位置类型因子是所有假设环境因子当中最能吸引非正规摊贩聚集的正相关因子.

比较表 3,4 可知:SW(街道宽度)与 BBW(背景建物宽度)在表 3 属对目标模式影响性偏低群,但在表 4 属相关系数偏高群.进一步由图 4 发现,这是因为此两数据在散点图上的分布极不均匀造成的.因此,高相关



(a) SL

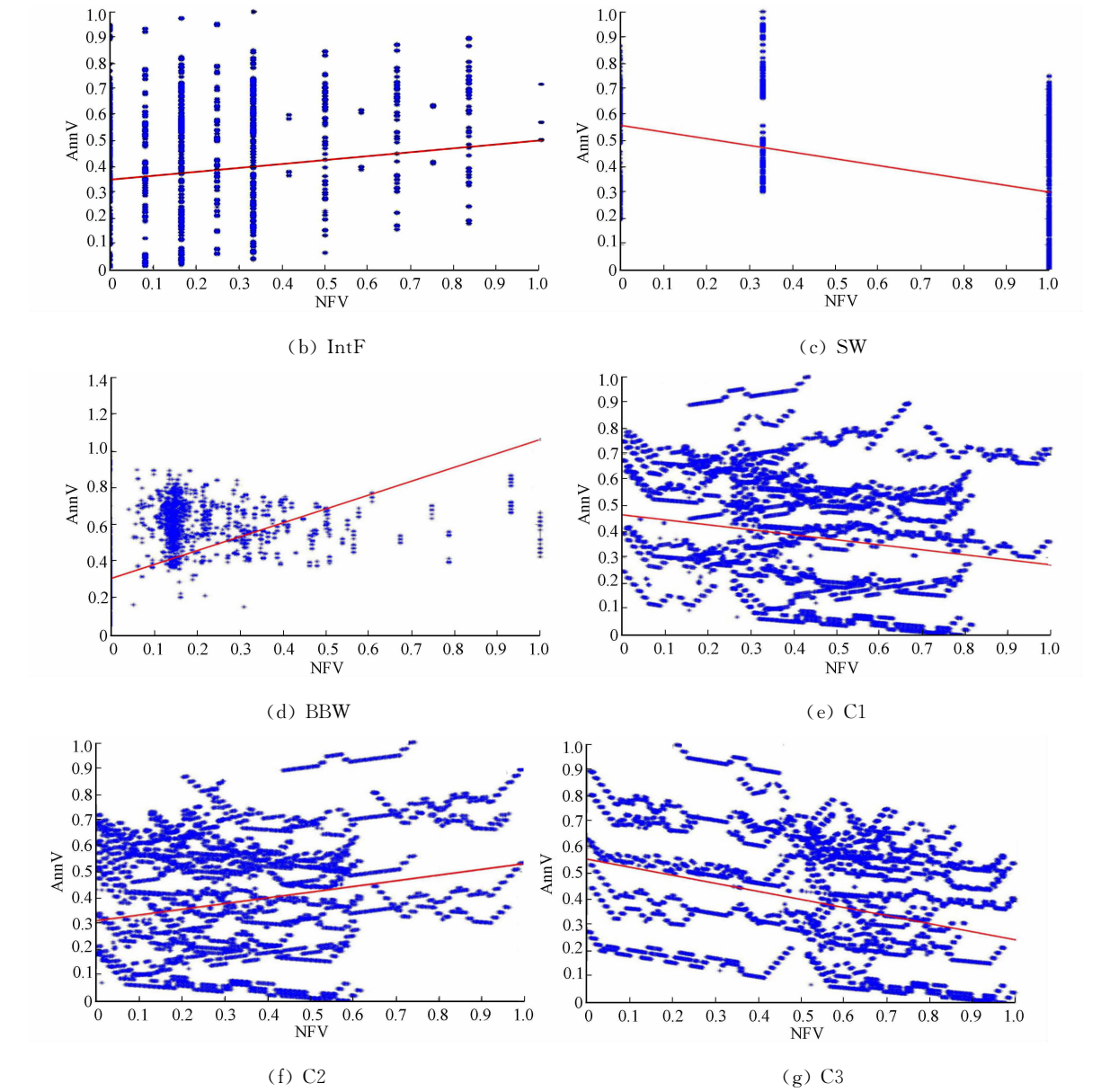


图 4 应用散点图来进行各环境因子的相关性分析

Fig. 4 Correlation analysis of each environmental factor by scatter diagram

表 4 各环境因子变数值与 ANN 模式输出值的皮尔森相关系数分析

Tab. 4 Correlation analysis of Pearson correlation coefficient between different kinds of environmental factor values and ANN outputs of target pattern

环境因子	SL	IntF	SW	BBW	C1	C2	C3
与 ANN 模式输出值的相关系数	0.856 7	0.142 1	−0.454 0	0.537 6	−0.178 6	0.209 5	−0.329 7

系数并不一定表示该环境因子对目标模式具有较高影响性,具有数学回归理论支持的 ANN 因子权重值分析,才能客观判断环境因子对目标模式的影响性. 然而,上述环境因子分析方法的共同使用与对比分析,才能全面挖掘环境因子信息.

另外,由 ANN 因子权重值分析辅以散点图、皮尔森相关系数的分析结果,可作为下一步修正环境因子实验假设的依据. 如 BBW 是对目标模式最不具效益的环境因子,因此,在未来的环境因子实验假设可考虑以其他因子替换.

## 5 主要结论

应用 ANN 与 CA 技术的 GIS 模型进行以下两阶段实验操作:1) 以 ANN 模式辨识与 CA 细胞增

生形态来模拟逢甲夜市非正规摊贩的分布;2) 通过 ANN 因子权重值分析与散点图、皮尔森相关系数等分析方法,解析各环境因子对于发生非正规摊贩的影响性与相关性. 通过实验,非正规摊贩分布模拟结果为平均正确率 84.1%,而目标模式的平均正确率可达 94.2%.

在环境因子影响性分析方面,SL 是发生非正规摊贩最具影响性的环境因子,BBW 是最不具影响性的环境因子. 在环境因子相关性分析方面,与发生非正规摊贩现象呈正相关因子为:SL,IntF,BBW 与 C2,其余为负相关因子. 因此,街道位置类型因子是所有假设环境因子中最能吸引非正规摊贩聚集的正相关因子. 上述环境因子的分析结果,可进一步作为未来修改环境因子实验假设的依据.

## 参考文献:

- [1] THAPA R B,MURAYAMA Y. Urban growth modeling of Kathmandu metropolitan region, Nepal[J]. Computers Environment and Urban Systems,2011,35(1):25-34.
- [2] SANT I,GARC A A M,MIRANDA D,*et al.* Cellular automata models for the simulation of real-world urban processes: A review and analysis[J]. Landscape and Urban Planning,2010,96(2):108-122.
- [3] BOLSTAD P. GIS Fundamentals: A first text on geographic information systems[M]. 3rd ed. White Bear Lake, Minn: Eider Press,2008:15-36.
- [4] 胡永进. 地理信息系统在城市规划中的应用[J]. 江苏城市规划,2010(5):32-37.
- [5] BATTU M. Modelling urban development with geographical information systems and cellular automata[M]. Boca Raton: CRC Press,2009:11-27.
- [6] PIJANOWSKI B C,TAYYEBI A,DOUCETTE J,*et al.* A big data urban growth simulation at a national scale: Configuring the GIS and neural network based land transformation model to run in a high performance computing (HPC) environment[J]. Environmental Modelling and Software,2014,51:250-268.
- [7] GREKOUSIS G,MANETOS P,PHOTIS Y N. Modeling urban evolution using neural networks, fuzzy logic and GIS: The case of the Athens metropolitan area[J]. Cities,2013,30(1):193-203.
- [8] BASSE R M,OMRANI H,CHARIF O,*et al.* Land use changes modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale[J]. Applied Geography,2014,53:160-171.
- [9] HAUPT S E,LAKSHMANAN V, MARZBAN C, *et al.* Environmental science models and artificial intelligence [C]// Artificial Intelligence Methods in the Environmental Sciences. Netherlands: Springer,2009:3-13.
- [10] DUDA R O,HART P E,STORK D G. Pattern classification[M]. 2nd ed. New York: Wiley,2001:25-61.
- [11] BISHOP C M. Neural networks for pattern recognition[M]. New York: Oxford University,1995:76-93.
- [12] YASSEMI S,DRAGICEVICĀ,SCHMIDT M. Design and implementation of an integrated GIS-based cellular automata model to characterize forest fire behaviour[J]. Ecological Modelling,2008,210(1/2):71-84.
- [13] ZHANG G P. An investigation of neural networks for linear time-series forecasting[J]. Computers and Operations Research,2001,28(12):1183-1202.
- [14] HSIEH A T,CHANG J. Shopping and tourist night markets in Taiwan[J]. Tourism Management,2006,27(1): 138-145.
- [15] YATMO Y A. Perception of street vendors as out of place urban elements at day time and night time[J]. Journal of Environmental Psychology,2009,29(4):467-476.
- [16] MARZBAN C. Basic statistics and basic AI: Neural networks[C]// Artificial Intelligence Methods in the Environmental Sciences. Netherlands: Springer,2009:15-47.

(责任编辑: 陈志贤      英文审校: 方德平)