

基于空间自相关的东莞市主体功能区划分

林锦耀, 黎 夏

(中山大学地理科学与规划学院, 广东省城市化与地理环境空间模拟重点实验室, 广州 510275)

摘要: 中国“十一五”规划纲要明确提出, 根据资源环境承载能力、现有开发密度和发展潜力, 统筹考虑未来中国人口分布、经济布局、国土利用和城镇化格局, 将国土空间划分为优化开发、重点开发、限制开发和禁止开发四类主体功能区。至今已有不少学者开展主体功能区划分研究, 但所用方法工作量大, 或带有一定的主观性, 且未考虑同类功能区集聚或分散程度。因此提出利用遗传算法改进传统聚类方法, 自动划分主体功能区, 在划分过程中考虑区域的全局空间自相关特性, 使同类功能区在空间上呈集聚分布的格局。以近年来城市快速扩张的东莞市为例, 验证了此方法的可行性, 能简单有效地进行主体功能区划分。与常用的K-means聚类方法相比, 新方法划分结果更符合实际情况, 能进一步推广应用到其它地区主体功能区划分。

关键词: 主体功能区; 空间自相关; 遗传算法; 东莞

DOI: 10.11821/dlyj201402013

1 引言

城市化是全球普遍发生的重要社会经济现象^[1,2]。城市化推动了中国经济和社会的发展, 一定程度上改善了居民的生活条件。然而当前中国城市化渐渐脱离循序渐进的原则, 超出了正常的发展轨道^[3]。造成中国空间开发无序的重要原因是长期以来对空间布局方面的规划的忽视^[4]。中国的快速城市化给国土空间开发带来一系列的经济、人口、生态问题, 例如耕地面积大幅度减小^[5,6]、城乡和区域发展不协调^[7,8]、生态破坏严重等^[9]。因此国家在“十一五”规划纲要中提出将国土空间划分为四类主体功能区, 即优化开发区、重点开发区、限制开发区及禁止开发区, 以规范空间开发秩序, 形成合理的空间开发结构^[10]。

主体功能区的概念形成后, 国内学者相继对此开展研究。樊杰探讨了我国主体功能区划的科学基础^[4]; 朱传耿等利用空间分析及聚类方法划分江苏省新沂市主体功能区^[11]; 顾朝林等结合传统区划与多因子分析方法划分江苏省盐城市主体功能区^[12]; 丁于思等结合K-means聚类和层次聚类方法对湖南省进行主体功能区划分^[13]; Rao等利用神经网络进行广西都安县功能区划分^[14]等。然而, 《国务院关于印发全国主体功能区规划的通知》明确提出, 优化开发和重点开发区域应相对集中分布, 避免遍地开花。现有研究在区划过程中很少考虑同类功能区集聚与分散的程度。陆大道院士也曾指出, 在区域发展过程中, 大部分社会经济要素在“点”(各级居民点和中心城市)上集聚, 并由线状基础设施(交通干

收稿日期: 2013-07-15; 修订日期: 2013-11-23

基金项目: 国家自然科学基金项目(41371376)

作者简介: 林锦耀(1989-), 男, 广东广州人, 硕士, 从事地理模拟与优化等研究。E-mail: ljy2012@foxmail.com

通讯作者: 黎夏(1962-), 男, 教授, 博士生导师。E-mail: lixia@mail.sysu.edu.cn

线、水源通道等)联系在一起而形成“轴”。随着区域社会经济的进一步发展,“点—轴”必然发展到“点—轴—集聚区”。要使区域最佳发展,就要求以“点—轴”系统模式对社会经济客体进行组织^[15]。因此,很有必要在主体功能区划分过程中考虑集聚的因素。

空间自相关理论可以用来解决这一问题,它衡量同一个变量在不同空间位置上的相关性。空间对象具有自相关性,空间位置越靠近的现象就越相似。已有不少学者利用空间自相关原理分析各种地理现象^[16-18],其能很好地反映特定区域内地理要素的集聚与分散程度。根据资源环境承载能力、现有开发密度和发展潜力三项指标划分不同主体功能区的同时,也要满足区域的空间自相关程度较高。常用的区划方法难以达到要求,需要引进人工智能算法进行优化。遗传算法是基于达尔文“自然选择”及“适者生存”理论的全局性概率搜索算法,适用于解决各学科的复杂优化问题^[19,20]。遗传算法已被成功地运用在解决各种地理优化问题中^[21-24],因此,本文提出改进常用的K-means聚类方法,利用遗传算法解决主体功能区划分的优化问题。

本文以近年来城市快速扩张的东莞市为例,根据2010年的统计资料,利用遗传算法自动划分各类主体功能区,在区划过程中考虑区域的全局空间自相关特性。研究结果表明,该方法能简单有效地进行功能区划分。与常用的K-means聚类方法相比,划分结果更符合区域的实际情况。

2 主体功能区划分方法

禁止开发区一般界限明确,不依赖于评价指标,不需要重新划分^[13]。因此,只需将区域划分为优化开发区、重点开发区以及限制开发区三类,再与禁止开发区图层叠加,就能得出主体功能区最终划分结果。

2.1 主体功能区划分指标体系

主体功能区是中国新近提出的概念,迄今为止还没有公认的主体功能区划分指标体系。而划分指标的选择关系到主体功能区划分的可操作性及科学性^[4,13]。本研究从东莞市的实际情况出发,遵循区划指标的科学性原则,以资源环境承载能力、现有开发密度和发展潜力三方面构建主体功能区划分指标体系(表1)。

2.2 基于聚类算法的划分方法

学者较多利用聚类算法划分主体功能区。聚类是指将对象的集合分组成为由类似的对象组成的多个类的过程。把 n 个对象 (x_1, x_2, \dots, x_n) 分成 c 类 (X_1, X_2, \dots, X_c) ,每类分别有 n_1, n_2, \dots, n_c 个对象,并求每类的聚类中心 C_i ,使得非相似性指标或距离指标的目标函数值达到最小^[25]。通常选欧氏距离作为非相似性指标,则目标函数可定义为:

$$G = \sum_{i=1}^c \sum_{j=1}^{n_i} \|x_j^i - C_i\|^2 \quad (1)$$

式中: x_j^i 表示第 i 类中的第 j 个对象。

利用聚类算法进行主体功能区划分,就是根据资源环境承载能力、现有开发密度和发展潜力三项指标,将指标值大小相近的区域划分为同一类功能区,以最大程度区分不同类型的功能区。但此方法只简单考虑各指标数值,存在一定的缺陷。

2.3 空间自相关

空间自相关理论可以用来解决功能区集聚与否的问题,分为全局空间自相关和局部空

表1 东莞市主体功能区划分指标体系

Tab. 1 Indicator system of Major Function Oriented Zoning (MFOZ) planning in Dongguan			
一级指标	二级指标	指标说明	原始数据来源
资源环境 承载能力	绿当量	指具有和森林基本相同的生态功能当量,可定义为: (区域森林面积+区域其它具有绿当量用地面积按“绿 量相当”原则折合成的森林面积)/区域土地总面积 ^[26]	东莞市城乡规划局
	环境容量	在人类生存和自然生态系统不致受害的前提下,某一 环境所能容纳的污染物的最大负荷量,包括大气环境 容量及水环境容量。分别采用A值法 ^[27] 和一维水质模 型 ^[28] 计算大气环境容量及水环境容量	东莞市环保局、水 务局
现有 开发密度	人均可利用土地资源	可利用土地的面积与总人口数的比值	东莞市城乡规划局
	人均可利用水资源	可利用水资源的体积与总人口数的比值	东莞市水务局
	地均GDP	单位面积土地创造的生产总值,衡量土地的使用效率	东莞市统计局
	人均GDP	在核算期内区域实现的生产总值与所属范围内的常 住人口的比值,衡量经济发展状况	东莞市统计局
	适宜建设用地利用率	区域现有建设用地面积占适宜建设用地面积的比例	东莞市城乡规划局
	城市化强度指数	衡量城市扩张强度,定义为:区域城市化过程中非城 市像元转化为城市像元的总数×100/(区域城市化前 的非城市像元总数×时间) ^[29]	东莞市城乡规划局
	人口密度	单位面积土地居住的人口数	东莞市统计局
	人口流动强度	暂住人口数占总人口数的比例,衡量区域吸纳流动人 口的规模	东莞市统计局
	公路网密度	现状交通干线长度与区域面积的比值,衡量区域道路 覆盖程度	东莞市交通运输局
	交通通达度	某个镇区到其它各个镇区所用时间的平均值	东莞市交通运输局
发展潜力	“十一五”期间GDP平 均增长率	GDP增长速度,衡量区域经济发展潜力	东莞市统计局

间自相关。全局空间自相关描述某种现象的整体分布情况,判断此现象在区域内是否有集聚现象存在;局部空间自相关可以计算局部空间集聚性。根据“点—轴”理论可知,划分主体功能区只需要知道区域是否存在集聚现象及整体的集聚程度,而无需知道集聚发生的确切位置,因此只考虑区域的全局空间自相关特性。衡量全局空间自相关的指标有全局Moran's *I* 指数^[30]、全局Getis's *G* 指数^[31]以及全局Geary's *C* 指数^[32]等,其中最常用的是全局Moran's *I* 指数:

$$I=\frac{N\sum_i\sum_jw_{ij}(x_i-\bar{x})(x_j-\bar{x})}{(\sum_i\sum_jw_{ij})\sum_i(x_i-\bar{x})^2}$$

(2)

式中:*N*表示研究对象的数量;*w_{ij}*表示空间权重矩阵;*x_i*表示对象*i*的值; \bar{x} 表示所有对象的平均值。全局Moran's *I* 指数值在-1到1之间,大于0表示存在空间正相关,且越接近1集聚程度越高;等于0表示不存在空间自相关性,即随机分布。

2.4 基于遗传算法的划分方法

利用遗传算法改进传统聚类方法,划分主体功能区。算法随机生成初始解,然后不断迭代改进当前解,直到最后搜索到满意的解为止。在进化计算中,迭代计算过程采用了模拟生物体的进化机制,从一组解出发,采用类似于自然选择和有性繁殖的方式,在继承原

有优良基因的基础上，生成具有更好性能指标的下一代解的群体^[33-35]。首先需要将问题的解编码，并将染色体（ CM ）定义为待划分的三类主体功能区各自的聚类中心：

$$CM=[R_1,D_1,P_1,R_2,D_2,P_2,R_3,D_3,P_3]$$

(3)

式中： R 表示资源环境承载能力； D 表示现有开发密度； P 表示发展潜力。三种指标均为归一化后的无量纲量。

然后定义遗传算法的适应度函数 $f(x)$ ，用于评价每个染色体的适应度：

$$f(x)=w_G\times G+w_I\times(1-I)$$

(4)

式中： G 表示非相似性指标（式1）； I 表示三类功能区全局Moran's I 指数的平均值，下文将简要介绍其计算方法； G 、 I 均为归一化后的无量纲量； w_G 、 w_I 分别表示两者的权重，且 $w_G+w_I=1$ 。优化目标是求取适应度函数的最小值。

当遗传算法将区域划分为三类后，分别计算各类的全局Moran's I 指数。如计算优化开发区时，其他两类区域均看作非优化开发区。最后将三类的计算结果求平均值，并进行归一化处理，即为式4中的 I 。

图1为该方法的总体流程图。

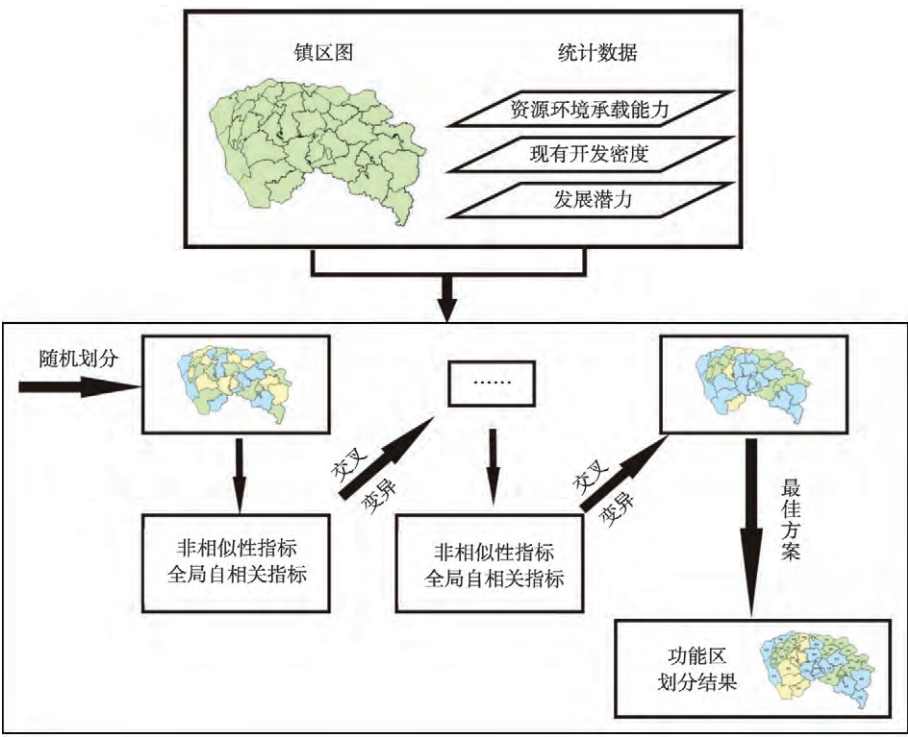


图1 利用遗传算法划分主体功能区流程图

Fig. 1 Flowchart of MFOZ planning based on genetic algorithms

3 东莞市主体功能区划分

3.1 区域概况

东莞市位于广东省中南部、珠三角走廊的中心区域，北邻广州，南接深圳，且邻近港

澳。东莞市行政区总面积为2465 km²，辖4个区、28个镇。改革开放以来，东莞市的区位优势及良好的投资环境使其成为珠三角地区外商投资的重心，经济快速增长，城市迅猛扩张，但同时也带来了一系列的城市问题。为了引导城市用地与空间资源的合理配置，东莞市政府划定了生态控制线，对保障城市基本生态安全有重要意义。结合东莞市实际情况，本研究将生态控制线内的区域以及虎门近代史迹与历史人文景观区划分为禁止开发区。

3.2 遗传算法参数设置

利用遗传算法解决优化问题还需要设置遗传参数。本研究将个体数设置为50，交叉概率为0.90，变异概率为0.01。同时采用精英选择策略，一直保留最好的解，确保进化过程中不会被淘汰掉。

为更客观地确定适应度函数（式4）的权重，本文还进行了模型敏感性分析，检验 w_G 、 w_I 分别取不同值时该方法的效果（表2）。

当 w_G 取1， w_I 取0时，与聚类方法相同，虽然非相似性指标值较小，但划分结果近似随机分布；当 w_G 取0.75， w_I 取0.25时，划分结果有一定的集聚程度，但仍有改善空间；当 w_G 取0.25， w_I 取0.75时，划分结果集聚程度较高，但部分非相似性指标值差异较大的邻接区域被划分为同一类功能区；当 w_G 取0， w_I 取1时，划分结果没有实际意义。综上，当 w_G 、 w_I 均取0.5时，能较好地同时满足两个条件。

从遗传算法最佳适应度值变化曲线（图2）可以看出算法效率高，收敛速度较快。

3.3 实验结果

优化开发区其区域竞争力较强，现有开发密度已经很高，但是资源环境承载力开始减弱，人口聚集度较高，可利用土地资源处于紧缺状态，未来发展潜力逐渐减弱；对于重点开发区的判断标准为现有开发密度较低，但未来发展潜力巨大，具备较好的交通条件和经济基础；而限制开发区为竞争力一般，而且人口聚集程度一般，未来发展潜力较低的区域，该区域内生态保护区较多或者资源较贫乏。

根据以上判断标准，分别利用K-means聚类方法及遗传算法将东莞市各镇区划分为优化开发区、重点开发区、限制开发区三类（图3b、图3a）。通过两图对比可看出，遗传算法划分结果的功能区集聚程度比聚类方法高，且更为合理。如聚类方法将东城、厚街、虎门划分为重点开发区，而遗传算法将三者划分为优化开发区，更符合东莞市的实际情况。从图3a可看出，各类主体功能区集聚程度较高，优化开发区集中在主城区及南北连线，符合“点—轴—集聚区”模式。另有三大重点开发区及两大限制开发区。

东莞市人民政府于2011年发布的《东莞市国民经济和社会发展第十二个五年规划纲

表2 模型敏感性分析
Tab. 2 Sensitivity analysis by using different combinations of w_G and w_I

方案	$w_G=1$,	$w_G=0.75$,	$w_G=0.5$,	$w_G=0.25$,
指标	$w_I=0$	$w_I=0.25$	$w_I=0.5$	$w_I=0.75$
G	0.3116	0.3391	0.3565	0.3752
I	0.2204	0.2838	0.3202	0.3633

注：表中 G 为归一化后的值， I 为三类功能区原始值的平均值。

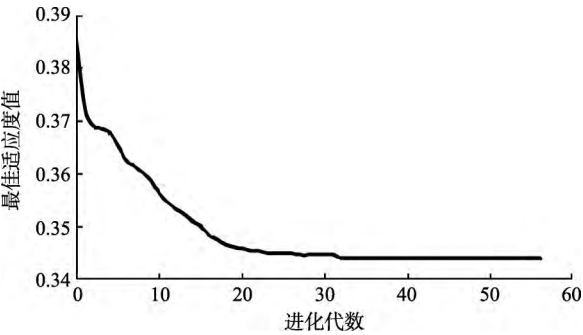


图2 遗传算法最佳适应度值变化曲线
Fig. 2 The relationship between the best fitness value and generation

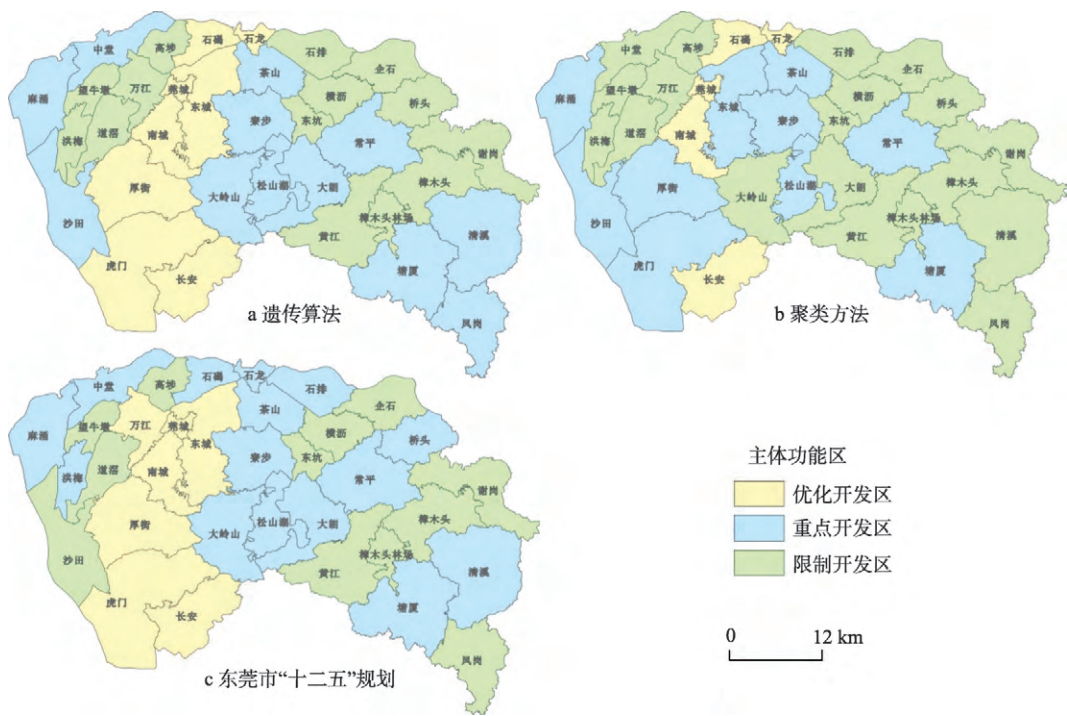


图3 东莞市主体功能区划分结果对比

Fig. 3 Comparison of MFOZ planning results in Dongguan

要》中提到：主城区（莞城、南城、东城、万江）以及厚街—虎门—长安为优化提升区；松山湖—生态园—长安新区及大岭山—大朗组团、麻涌—洪梅—中堂组团、石龙—石碣—石排组团、常平—桥头—樟木头组团、塘厦—清溪组团为重点拓展区^[36]。对比本研究划分结果与规划纲要结果图（图3c），发现两者吻合程度较高，符合产业集聚及区域协调发展的大方向。进一步计算规划纲要结果的非相似性指标及全局 Moran's *I* 指数，分别为 0.3432、0.3046，与本研究结果（0.3565、0.3202）较为接近。上述表明，本文提出的方法切实可行，可进一步推广应用到其它地区的主体功能区划分。最后，再将禁止开发区图层与上一步遗传算法的划分结果叠加，得出东莞市四类主体功能区最终划分结果（图4）。

4 结论与讨论

至今主体功能区划分仍未有一套公认的方法。现有方法工作量大，或带有一定的主观性，且未考虑同类功能区集聚与分散的程度，划分结果有一定的缺陷。本文基于“点—轴”理论，提出利用遗传算法改进传统聚类方法，自动划分主体功能区，并在划分过程中考虑区域的全局空间自相关特性，使同类功能区在空间上呈集聚分布的格局。将该方法应用到近年来城市快速扩张的东莞市，与常用的 K-means 聚类方法相比，新方法能快速、有效地划分主体功能区，划分结果更符合区域的实际情况，为其他地区的主体功能区划分提供了一种新的思路。

根据东莞市主体功能区最终划分结果，结合多中心网络发展策略，集聚发展的核心区域—主城区以优化开发为主体功能，成为区域性服务和消费中心；整合虎门及长安空间资

- [8] Sicular T, Ximing Y, Gustafsson B, et al. The urban-rural income gap and inequality in China. *Review of Income and Wealth*, 2007, 53(1): 93-126.
- [9] Hubacek K, Guan D, Barrett J, et al. Environmental implications of urbanization and lifestyle change in China: Ecological and water footprints. *Journal of Cleaner Production*, 2009, 17(14): 1241-1248.
- [10] 中华人民共和国中央人民政府. 中华人民共和国国民经济和社会发展第十一个五年规划纲要. http://www.gov.cn/ztl/2006-03/16/content_228841.htm, 2006-03-16. [The Central People's Government of the People's Republic of China. The Eleventh Five-Year Plan for National Economic and Social Development of the People's Republic China. http://www.gov.cn/ztl/2006-03/16/content_228841.htm, 2006-03-16.]
- [11] 朱传耿, 仇方道, 马晓冬, 等. 地域主体功能区划理论与方法的初步研究. *地理科学*, 2007, 27(2): 136-141. [Zhu Chuangeng, Qiu Fangdao, Ma Xiaodong, et al. The theories and methods of Major Function Regionalization. *Scientia Geographica Sinica*, 2007, 27(2): 136-141.]
- [12] 顾朝林, 张晓明, 刘晋媛, 等. 盐城开发空间区划及其思考. *地理学报*, 2007, 62(8): 787-798. [Gu Chaolin, Zhang Xiaoming, Liu Jinyuan, et al. The development spatial regionalization of Yancheng City and relevant consideration. *Acta Geographica Sinica*, 2007, 62(8): 787-798.]
- [13] 丁于思, 高阳, 周震虹. 基于混合聚类的湖南主体功能区划分研究. *经济地理*, 2010, 30(3): 393-396. [Ding Yusi, Gao Yang, Zhou Zhenhong. Combinational clustering based principal function regionalization of Hunan Province. *Economic Geography*, 2010, 30(3): 393-396.]
- [14] Rao Y, Hu B. Functional Zoning of land use in Karst region based on neural network. *Geomatics for Integrated Water Resources Management (GIWRM)*, 2012 International Symposium on. 2012: IEEE.
- [15] 陆大道. 关于点-轴空间结构系统的形成机理分析. *地理科学*, 2002, 22(1): 1-6. [Lu Dadao. Formation and dynamics of the "Pole-Axis" spatial system. *Scientia Geographica Sinica*, 2002, 22(1): 1-6.]
- [16] Conway D, Li C Q, Wolch J, et al. A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values. *The Journal of Real Estate Finance and Economics*, 2010, 41(2): 150-169.
- [17] Lian J, Gong H, Li X, et al. The analysis of economic spatial characteristics of Beijing-Tianjin-Hebei metropolitan region based on GIS. 17th International Conference on Geoinformatics. 2009. IEEE.
- [18] Overmars K, De Koning G, Veldkamp A. Spatial autocorrelation in multi-scale land use models. *Ecological Modelling*, 2003, 164(2): 257-270.
- [19] Golberg D E. *Genetic algorithms in search, optimization, and machine learning*, MA: Addison-Westey, Readomg, Mass, 1989.
- [20] Holland J H. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Ann Arbor: MIT press, 1992.
- [21] Li X, Yang Q, Liu X. Discovering and evaluating urban signatures for simulating compact development using cellular automata. *Landscape and Urban Planning*, 2008, 86(2): 177-186.
- [22] Li X, Yeh A G O. Integration of genetic algorithms and GIS for optimal location search. *International Journal of Geographical Information Science*, 2005, 19(5): 581-601.
- [23] Stewart T J, Janssen R, van Herwijnen M. A genetic algorithm approach to multiobjective land use planning. *Computers & Operations Research*, 2004, 31(14): 2293-2313.
- [24] Tseng M H, Chen S J, Hwang G H, et al. A genetic algorithm rule-based approach for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2008, 63(2): 202-212.
- [25] Eisen M B, Spellman P T, Brown P O, et al. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*, 1998, 95(25): 14863-14868.
- [26] 刘艳芳, 明冬萍, 杨建宇. 基于生态绿当量的土地利用结构优化. *武汉大学学报: 信息科学版*, 2002, 27(5): 493-498. [Liu Yanfang, Ming Dongping, Yang Jianyu. Optimization of land use structure based on ecological green equivalent. *Geomatics and Information Science of Wuhan University*, 2002, 27(5): 493-498.]
- [27] 徐芙蓉, 施介宽. A值法研究大气总量控制的环境质量达标保证率. *四川环境*, 2003, 22(2): 70-73. [Xu Furong, Shi Jiekuan. The environmental quality standard guarantee ratio for A-value method in the research of total air pollutant emission control. *Sichuan Environment*, 2003, 22(2): 70-73.]
- [28] 韩进能. 河流一维水质模型在水环境容量计算方面的应用. *环境科学与技术*, 1995, 18(4): 43-45. [Han Jinneng. The

- applications of one-dimensional water quality model in calculation of water environment capacity. *Environmental Science & Technology*, 1995, 18(4): 43-45.]
- [29] 何春阳, 史培军, 陈晋, 等. 北京地区城市化过程与机制研究. *地理学报*, 2002, 57(3): 363-371. [He Chunyang, Shi Peijun, Chen Jin, et al. Process and mechanism of urbanization in Beijing area. *Acta Geographica Sinica*, 2002, 57(3): 363-371.]
- [30] Moran P A P. Notes on continuous stochastic phenomena. *Biometrika*, 1950. 17-23.
- [31] Getis A, Ord J. The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 1992, 24(3): 189-206.
- [32] Geary R C. The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 1954, 5(3): 115-146.
- [33] Booker L B, Goldberg D E, Holland J H. Classifier systems and genetic algorithms. *Artificial Intelligence*, 1989, 40 (1): 235-282.
- [34] Myung H, Kim J H. Hybrid evolutionary programming for heavily constrained problems. *BioSystems*, 1996, 38(1): 29-43.
- [35] 李敏强. 遗传算法的基本理论与应用. 北京: 科学出版社, 2002. [Li Minqiang. *The Basic Theory and Applications of Genetic Algorithms*. Beijing: Science Press, 2002.]
- [36] 东莞市人民政府. 东莞市国民经济和社会发展第十二个五年规划纲要. <http://www.dg.gov.cn/business/htmlfiles/cndg/s1271/201105/338947.htm>, 2011-04-25. [The People's Government of Dongguan. *The Twelfth Five-Year Plan for National Economic and Social Development of Dongguan*. <http://www.dg.gov.cn/business/htmlfiles/cndg/s1271/201105/338947.htm>, 2011-04-25.]

MFOZ planning of Dongguan based on spatial autocorrelation by using genetic algorithms

LIN Jinyao, LI Xia

(School of Geography and Planning, and Guangdong Key Laboratory for Urbanization and Geo-simulation,
Sun Yat-sen University, Guangzhou 510275, China)

Abstract: According to the 11th Five-Year Plan (2006-2010), land should be divided into four types of Major Function Oriented Zones (MFOZ), namely optimal development zone, key development zone, restricted development zone and prohibited development zone, based on bearing capacity of resources and environment, existing development intensity and potential of development. So far, numerous researchers have already conducted some related studies. Still, these somewhat subjective and labor-consuming methods seldom concern the cluster or dispersion degree of the same certain zone. This research tried to improve traditional clustering-based method by using genetic algorithms. This automatic method which is then applied to MFOZ planning of Dongguan concerns global spatial autocorrelation of the region. The comparison demonstrates that this simple and effective method has better performance than commonly-used clustering-based methods. The proposed method can be further applied to MFOZ planning of other regions.

Key words: MFOZ; spatial autocorrelation; genetic algorithms; Dongguan