

# 不同样本方案下遗传元胞自动机的土地利用模拟及景观评价\*

冯永玖<sup>1\*\*</sup> 刘艳<sup>2</sup> 韩震<sup>1</sup>

(<sup>1</sup>上海海洋大学海洋科学学院, 上海 201306; <sup>2</sup>昆士兰大学地理、规划与环境管理学院, 澳大利亚布里斯班 4072)

**摘要** 利用元胞自动机 (cellular automata, CA) 模拟土地利用情景, 有助于理解其变化机理, 并为土地资源持续利用提供空间决策支持. 本文基于生物进化过程的遗传算法 (generic algorithm, GA) 将 CA 参数编码成为染色体, 在模拟结果与真实结果差异值的引导下, 通过选择、杂交和变异算子使最优的染色体得以遗传和保留, 从而建立智能优化的元胞自动机模型. 以浙江省嘉兴市 1992—2008 年土地利用变化为例, 分别利用 6% (66 个 · km<sup>-2</sup>) 和 3% (33 个 · km<sup>-2</sup>) 两种样本方案构建遗传 CA 模型进行土地利用变化模拟, 并通过混淆矩阵、Kappa 系数和景观指数对模拟结果进行评估. 结果表明: 遗传 CA 模拟结果能在数量、位置和景观格局上以超过 80% 的水平接近真实分类, 且大样本量构建的遗传 CA 的模拟精度更高; 2008 年的模拟精度和景观综合指数低于 2001 年, 表明遗传 CA 的模拟精度和景观综合指数随模拟时间而衰减.

**关键词** 土地利用模拟 元胞自动机 多重样本 遗传算法 景观评价

**文章编号** 1001-9332(2011)04-0957-07 **中图分类号** Q149; U945.23 **文献标识码** A

**Land use simulation and landscape assessment by using genetic algorithm based on cellular automata under different sampling schemes.** FENG Yong-jiu<sup>1</sup>, LIU Yan<sup>2</sup>, HAN Zhen<sup>1</sup> (<sup>1</sup>College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China; <sup>2</sup>School of Geography, Planning and Environmental Management, University of Queensland, Brisbane QLD 4072, Australia). -*Chin. J. Appl. Ecol.*, 2011, 22(4): 957-963.

**Abstract:** Simulating land use change scenarios with cellular automata (CA) can help to the policy makers in understanding the mechanisms of land change, and support the spatial decision-making for the sustainable use of land resources. Genetic algorithm (GA), an intelligent approach originally conceived from the biological process of evolution, has the capability of minimizing the difference between simulated and observed land use patterns with optimum chromosomes (*i. e.*, feasible CA parameters) obtained through a set of selection, crossover, and mutation operations. In this paper, GA-based CA model was developed, and applied to simulate the land use change in Jiaying City of Zhejiang Province in 1992-2008. This model was calibrated with 6% (66 samples · km<sup>-2</sup>) and 3% (33 samples · km<sup>-2</sup>) samplings, and the simulation results were evaluated based on confusion matrix, Kappa coefficient, and landscape metrics analysis. Over 80% of the land use features generated by the GA-based CA model matched the observed classification of land features geographically, and much higher simulation accuracy could be obtained with a larger sample. The simulation accuracy and the landscape metrics for 2001 were better than those for 2008, suggesting a tendency that the model's accuracy decreased over the simulating process.

**Key words:** land use simulation; cellular automata; multi-samples; genetic algorithm; landscape assessment.

\* 上海市教育委员会科研创新项目 (11YZ154)、上海高校选拔培养优秀青年教师科研专项基金项目 (SSC09018) 和上海海洋大学校博士启动基金暨环境工程重点学科基金项目 (A-2400-10-0134) 资助.

\*\* 通讯作者. E-mail: yjfeng@shou.edu.cn

2010-08-26 收稿, 2011-01-27 接受.

土地利用情景模拟能够为土地资源可持续利用提供技术手段,同时有助于提高区域空间、生态、环境和资源等的安全性<sup>[1-2]</sup>.地理元胞自动机(cellular automata, CA)因其具有“自下而上”(从底层交互而呈现全局结构)的能力,被广泛应用于土地利用变化的情景模拟<sup>[3-4]</sup>.研究认为,CA建模的核心问题是转换规则的确定.该规则由元胞邻域、空间变量、限制因素和随机变量等组成<sup>[5]</sup>.影响土地利用变化的空间变量通常存在严重的相关性,如何消除相关性并确定模型参数是CA转换规则的首要问题<sup>[6]</sup>.当变量相当复杂且存在多重相关性时,传统方法如多准则判别分析(MCE)<sup>[7]</sup>、Logistic回归<sup>[8-9]</sup>和偏最小二乘回归(PLS)<sup>[10]</sup>并不能完全消除其负面影响,通过这些方法获取的参数欠合理<sup>[11]</sup>.因此,近年来发展了多种智能方法,可自动获取空间直观模型的CA参数,如支持向量机<sup>[12]</sup>、神经网络<sup>[13]</sup>、蚁群智能算法<sup>[14]</sup>、粒子群智能算法<sup>[15]</sup>和普通遗传算法<sup>[16-18]</sup>等.

将CA模拟结果与分类结果的累积差异值构建为目标函数,用以引导遗传算法(generic algorithm, GA)对CA参数的随机搜索,从而可构建基于遗传算法的CA模型.通过对土地利用的初始状态、末期状态和空间变量图层的采样可完成目标函数的构建.不同的采样方案将导致不同的目标函数,遗传搜索结束时的函数值也将不同,导致获取的CA参数不同<sup>[16,19]</sup>.在元胞自动机建模中,小样本量对计算资源的要求较低,却不能全面反映研究区域;大样本量能够较好地代表研究区域,但对计算资源的要求较高.基于遗传智能CA模型,以模拟结果与真实结果的平均差异值构建目标函数,可以统一目标函数值的量级,有利于不同采样方案的比较.本文应用基于遗传算法的地理元胞自动机模型,分别采用6%和3%两种样本方案,模拟1992—2008年浙江省嘉兴市土地利用变化过程,利用混淆矩阵、Kappa系数和景观指数测算模拟结果的精度,验证了遗传CA模型对土地利用分类的模拟效果,旨在推动模型建立者和决策者对土地利用变化动力学过程的理解,为土地资源可持续利用与空间决策支持提供科学依据和参考案例.

## 1 研究区域与研究方法

### 1.1 研究区概况

嘉兴市位于浙江省东北部、长江三角洲杭嘉湖平原腹地(30°21′—30°2′N, 120°18′—

121°16′E),是长江三角洲的重要城市之一.嘉兴市下辖南湖、秀洲两个行政区和平湖、海宁、桐乡3个市以及嘉善、海盐两个县,是上海通往杭州和湖州的交通要道,区位优势非常明显.本文以嘉兴市南湖区为中心,以向东12.5 km、向西10.3 km、向南9.7 km、向北9.1 km所覆盖的范围作为研究区域,是嘉兴市的核心区域,总面积429.55 km<sup>2</sup>.改革开放以来,嘉兴市经济快速发展,综合实力显著增强,并进入快速城市化时期,土地利用变化剧烈;尤其是2001—2008年间,嘉兴市城市建成区面积增加了1.5倍.利用智能化地理CA模型,能够准确揭示该市土地利用变化的规律,并模拟其未来情景.

利用1992、2001和2008年的TM/ETM<sup>+</sup>遥感影像,作为建模与检验模型正确性和有效性的原始数据,影像为760行、628列.以往的研究中,建模者为了兼顾运行效率,通常将空间分辨率为30 m的TM影像重采样为40 m或50 m,而本研究保留原始30 m空间分辨率.地理CA模型具有尺度(粒度)敏感性在对原始数据进行重采样时,不同的采样方法会导致不同的结果,甚至会破坏原始数据的结构<sup>[20]</sup>,并且较高分辨率的影像数据能够更好地呈现研究区域的真实状态,因此本文保留原始空间分辨率不变.

### 1.2 遗传元胞自动机模型

**1.2.1 土地利用变化的驱动因素** 地理CA模拟模型是在深入理解土地利用变化机理的基础之上构建的.一般情况下,土地利用演变的根源可概括为正距离、负距离、自然地理属性、社会经济环境和限制条件等100余种因素<sup>[7]</sup>.“正距离”因素对土地利用产生“推力”作用,如公路、铁路、居住区、工业区等;“负距离”对土地利用产生“阻力”作用,如耕地、水域、特殊用地等;自然属性指地形和地质条件,如高程、坡度和土质等;社会经济环境指对土地利用影响重大的经济、人口、政策等因素;限制因素指限制土地元胞发生状态改变的条件,如土地可开发总量和基本农田保护区等<sup>[8,21-22]</sup>.由于研究区域的具体情况不同,常以到市中心的距离( $d_{uc}$ )、到镇中心的距离( $d_{vc}$ )、到主要道路的距离( $d_{mr}$ )、到耕地的距离( $d_{pl}$ )和到水域的距离( $d_{wt}$ )作为建模需要考虑的主要因素<sup>[7,10]</sup>.本文在遗传CA模型中,基于正距离、负距离、自然地理属性、社会经济环境和限制条件这5个影响因素进行转换规则的参数挖掘和土地利用模拟.

**1.2.2 元胞自动机基础模型** 一般情况下,任意元胞 $k$ 在时刻 $t$ 从“未开发”状态转换为“已开发”状

态的全局概率可以表达为  $P'_k$ , 采用 Logistic 回归方程建立模型如下<sup>[7-9]</sup>:

$$P'_k = 1/[1 + \exp(-a_0 + a_1d_1 + \dots + a_md_m)] \times \frac{\sum_{n \times n} \text{cell}(s_k = \text{Dev})^t}{n \times n - 1} \times \text{con}(s_k^t = \text{Dev}) \times (1 + (-\ln\gamma))^\alpha \quad (1)$$

式中:  $P'_k$  为元胞  $k$  在时刻  $t$  的全局转换概率;  $1/[1 + \exp(-a_0 + a_1d_1 + \dots + a_md_m)]$  为受空间变量驱动的转换概率;  $d_1, d_2, \dots, d_m$  为空间变量;  $a_0$  为常数;  $a_1, \dots, a_m$  为各对应空间变量的权重;  $[\sum_{n \times n} \text{cell}(s_k = \text{Dev})^t]/(n \times n - 1)$  为元胞邻域的影响概率;  $\text{con}(s_k^t = \text{Dev})$  为元胞状态转化的限制条件;  $(1 + (-\ln\gamma))^\alpha$  为土地利用变化的随机影响因素;  $\gamma$  为  $0 \sim 1$  之间的随机实数;  $\alpha$  为随机因素大小的控制因子, 其值为  $[0, 10]$  之间的整数。

将元胞全局转换概率  $P'_k$  与设定的阈值  $P_{\text{thr}}$  ( $0 \sim 1$ ) 进行比较, 以确定下一时刻元胞  $k$  的状态是否发生转变。当土地元胞  $k$  的转换概率  $P'_k$  大于设定阈值  $P_{\text{thr}}$  时, 该元胞转化为已开发土地类型, 否则元胞保持未开发状态不变。

### 1.3 遗传智能 CA 挖掘模型

模拟实验表明, 基于 CA 的空间直观模拟结果与区域真实形态通常存在差异, 当两者无限接近时表明模型误差较小, 即计算所得的 CA 模型参数较好<sup>[18]</sup>。在 CA 全局转换概率中, 待定的常数和空间变量的权重  $a_0, a_1, \dots, a_m$  一般通过空间统计方法确定。但空间统计方法存在无法消除变量多重共线关系的不足, 故获取的模拟结果与真实形态不可避免地存在偏差。本文利用遗传进化计算方法, 自动挖掘和随机搜索获得模型误差(模拟值与真实值的差异)为最小值的 CA 参数组合, 从而达到 CA 优化建模的目的。

为了利用 GA 优化 CA 参数, 需要将 CA 建模问题映射到 GA 算法空间中, 即用数学模型将待解决的问题表达出来。根据模拟值与真实值的累积平均差异值, 构造 GA 待优化的目标函数  $F(a)$ :

$$F(a) = \frac{1}{n} \sum_{k=1}^n |P'_k - f_k| \quad (2)$$

式中:  $F(a)$  为模拟值与实际值之间的全局差异或误差,  $F(a)$  越小表明模型精度越高、获取的参数值越可靠;  $n$  为采集的样本点数量; 模拟值  $P'_k$  由式(1)计算获取;  $f_k$  为样本元胞  $k$  对应的真实值,  $f_k = 1$  表示土地已开发,  $f_k = 0$  表示土地未开发<sup>[18]</sup>。  $a = \{\alpha_0, \alpha_1,$

$\dots, \alpha_j, \dots, \alpha_m\}$ ,  $\alpha_j \in R$  为参数空间所有可行解(或状态)构成的集合, 目标函数  $F(a)$  取得最小值时对应的状态  $a^*$  即是 GA 挖掘所得的 CA 参数组合。通过遗传算法挖掘 CA 参数的模型由染色体编码与种群初始化、染色体评估函数以及选择、杂交和变异算子等几部分构成。

1) 染色体编码与种群初始化。在遗传 CA 模型中, 染色体即是问题的可行解, 其编码方式会对杂交和变异操作产生影响<sup>[19]</sup>。采用浮点数编码方法, 定义染色体的长度为有效解的变量长度, 染色体的每个个体相当于解的一维变量, 因此染色体编码可定义为:  $a = \{\alpha_0, \alpha_1, \dots, \alpha_j, \dots, \alpha_m\}$ ,  $\alpha_j \in R$ , 与 CA 参数的构成方式一致。在地理元胞自动机模拟中, 某特定空间变量的权重即是遗传算法(GA)的“基因”, 所有空间变量权重的组合就是一个可行的 CA 参数, 即 GA 的染色体“个体”; 一组可行的 CA 参数即是 GA 的“种群”。遗传算法在一个随机给定的初始种群中进行迭代搜索, 为了保证初始种群一定程度上的优良性, 可以采用在传统空间统计 CA 参数的附近产生初始种群。

2) 染色体评估函数。利用评估函数  $Eval(A)$  评价染色体的优劣, 规定  $Eval(A)$  值越小, 则对应的染色体越优<sup>[19]</sup>。

$$Eval(A) = F(a) \quad (3)$$

式中:  $Eval$  为评价函数;  $A$  为染色体;  $F$  为差异函数;  $a$  为可行的 CA 参数。

3) 种群的选择、杂交和变异。选择算子是根据染色体评估值, 对种群中的个体进行优胜劣汰操作, 在 CA 参数挖掘中保留父代群体中评估值低的个体, 遗传到下一代群体。杂交算子指两个相互配对的染色体按某种方式相互交换部分个体, 形成两个新的个体, 从而产生进化程度更高的种群。变异算子是随机地对基因进行强制变化, 改变搜索方向、扩大搜索空间, 可提高遗传算法的局部搜索能力。基因交叉的百分数由基因交叉概率控制, 交叉概率越高, 收敛于最优解区域的速度越快。

## 2 CA 参数的挖掘

通过 ENVI 软件对获取的 3 组遥感影像进行预处理, 主要包括图像增强、几何校正和图像分类, 利用 ArcInfo 软件的空间分析功能获取空间距离图层。在 VB.NET 编程环境下, 结合 GIS 嵌入式组件 ArcGIS Engine 开发了模拟软件包来执行样本选取、地理 CA 模拟和结果评估等功能。元胞自动机参数的

遗传算法获取功能则利用 Matlab 开发完成,并集成在前述 CA 模拟软件包中。

为了挖掘 CA 参数,需要从影响因素图层、初始状态图层和末期状态图层中分别获取一定比例的样本量,有研究认为 20% 比较适合<sup>[18]</sup>。实际模拟中,各案例采用的空间尺度不尽相同,因此以采样密度(即单位面积采集的样本数量)作为标准更客观。从目标函数的构造可知,由于不同样本量产生的累积误差值存在很大差异,造成挖掘所得的 CA 参数值存在一定差异。因此,本文利用两种采样方案作为遗传 CA 参数挖掘和模型建立的基础,采样量分别为 6% 和 3%,对应的采样密度分别是 66 和 33 个·km<sup>-2</sup>。

在遗传 CA 模型中,遗传算法的参数设置如下:1)选择算子:采用锦标赛选择法(tournament selection)从种群中随机挑选一定数目的个体,将最好的个体选作父代染色体,这个过程重复  $n$  次完成个体的选择,选择参数为竞赛规则,其值在  $[2, n]$ ;2)杂交算子:采用单点算子,在个体编码串中随机设置一个杂交点,然后在该点相互交换两个配对个体的部分基因;3)变异算子:采用换位变异算子,即在染色体上随机产生 2 个变异点,交换两点上的基因;4)初始种群 50,迭代次数 500,CA 参数的上下界为  $[-1.6, 1.6]$ ,初始染色体为  $[0.01, 0.01, 0.01, 0.01, 0.01, 0.01]$ 。利用上述优化策略,对嘉兴市 6% 和 3% 样本量构成的染色体评估函数进行优化,以挖掘 CA 参数(图 1)。

从评估值的改善过程来看,GA 收敛速度较快,3% 样本在 315 次迭代时已获取到最佳评估值(0.21),而 6% 样本在 350 次迭代时才获得最佳评

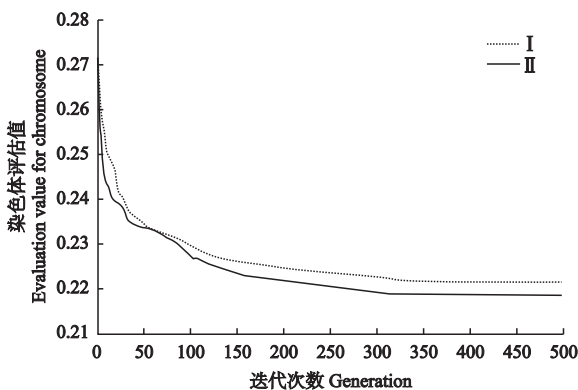


图 1 利用遗传算法挖掘 CA 参数的过程

Fig. 1 Process of mining CA parameters with genetic algorithm.

I: 3% 样本量 3% samples; II: 6% 样本量 6% samples.

估值(0.22),可见大样本量产生目标函数的收敛速度稍慢,且获取的评估值更大。根据染色体评估函数的定义,最佳评估值即是传统方法计算所得的元胞转换概率与实际转换概率之间的平均差值,可见评估值越小则误差越小。但影响评估值的因素较多,包括目标函数和遗传优化策略等。由于本文采取了相同的优化策略,因此影响评估值的主要因素为目标函数,而目标函数主要由样本量决定,所以 6% 样本量获取的评估值大于 3% 样本量。由表 1 可以看出,3% 和 6% 样本量获取的 CA 参数在量级上非常接近,在“正距离”和“负距离”的判断上均比较准确,但在具体数值上还存在一定差异。差异较明显的参数为到镇中心距离( $a_2$ )、到主要道路距离( $a_3$ )和到耕地距离( $a_4$ ),3% 样本量认为, $a_2$  和  $a_4$  对嘉兴市土地利用变化的作用较大,而 6% 样本量认为  $a_3$  的作用较大。根据嘉兴市 1992 年以来的土地利用变化情况,该区域城市建成区和农村居民点的增长主要发生在主要道路的两侧和交通便捷之地,但并不否认镇中心作为次要增长点的有效作用。这与 6% 样本量计算的结果较符合。

### 3 结果与分析

在地理 CA 模拟程序包中,利用循环测试方法设置阈值( $P_{thr}$ )和模拟次数(Iter)的组合。首先,遗传 CA 模型中阈值选择范围设置为  $[0.42, 0.92]$ ,以 0.42 为初始值,经过一定模拟次数产生 2001 和 2008 年模拟结果,并与分类结果进行比较,通过总体精度和 Kappa 系数评价模拟结果的优劣;然后,以 0.04 的步长增加阈值,通过一定模拟次数得到模拟结果。该系列模拟结果中,一定存在  $P_{thr}$  和 Iter 的最佳组合,可得到最佳模拟结果。通过测试发现, $P_{thr} = 0.58$  与 Iter 为 9 和 16 分别可得到 2001 和 2008 年的最佳模拟结果(图 2)。

本研究重点在于检测和模拟土地利用从非城市

表 1 基于不同样本量获取的遗传 CA 参数

Table 1 CA parameters achieved by GA with different samples

样本方案 Sampling schemes	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
3% 样本 3% samples	0.61	0.81	0.57	0.20	-0.45	-0.79
6% 样本 6% samples	0.55	0.64	0.38	0.81	-0.28	-0.80

$a_0$ : 常数 Constant;  $a_1$ : 到市中心的距离 Distance to city center;  $a_2$ : 到镇中心的距离 Distance to town center;  $a_3$ : 到主要道路的距离 Distance to main road;  $a_4$ : 到耕地的距离 Distance to cropland;  $a_5$ : 到水体的距离 Distance to water.

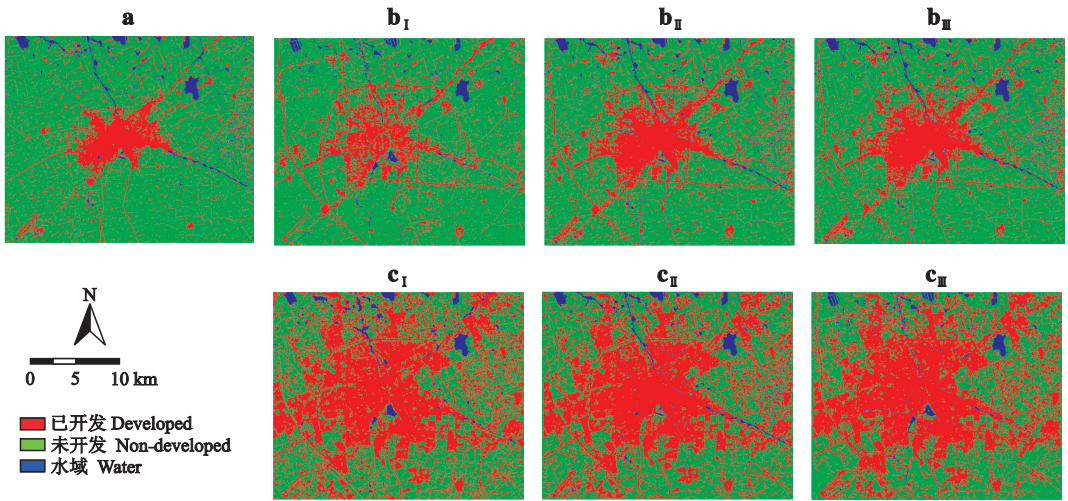


图2 利用遗传 CA 模型模拟的嘉兴市 2001 和 2008 年土地利用变化

Fig. 2 Simulation of urban land use changes of Jiaxing City in 2001 and 2008 using the GA based on CA model.

a) 1992; b) 2001; c) 2008. I: SVM 分类 SVM classifier; II: 3% 样本量模拟 Simulated result with 3% samples; III: 6% 样本量模拟 Simulated result with 6% samples.

向城市的转变,因此图 2 将土地类型分为已开发、未开发和水域 3 类。利用遥感软件 ENVI 的 SVM 分类器对 1992 年遥感影像进行分类的结果,作为模拟的初始状态(图 2a);利用遥感软件 ENVI 的 SVM 分类器对 2001 年和 2008 年遥感影像分类的结果,分别作为 2001 年(图 2b<sub>I</sub>)和 2008 年(图 2c<sub>I</sub>)的基准图;利用遗传元胞自动机得到 3% 和 6% 样本量下 2001 年(图 2b<sub>II</sub>和图 2b<sub>III</sub>)和 2008 年(图 2c<sub>II</sub>和图 2c<sub>III</sub>)的模拟结果。

研究认为,地理 CA 的模拟结果检验通常有逐点对比和整体对比两种方法。前一种方法是将模拟结果和实际分类进行叠加,通过逐点对比计算其精

度,即计算混淆矩阵<sup>[23]</sup>;后一种方法是检验模拟结果的空间格局与实际空间格局的符合程度<sup>[18]</sup>。

本文通过模拟图层的叠加对比,评价了 3% 和 6% 两种样本量在 2001 和 2008 年模拟结果的精度(表 2),结果显示,对于 3% 样本量构造的遗传 CA 模型,2001 年总体精度(80.1%)高于 2008 年(78.9%),而 2001 年表征一致性的 Kappa 系数(0.51)小于 2008 年(0.60);对于 6% 样本量构造的遗传 CA 模型,2001 年总体精度(83.9%)大于 2008 年(82.5%),而 2001 年的 Kappa 系数(0.61)小于 2008 年(0.66)。表明遗传 CA 模拟的总体精度和 Kappa 系数随模拟时间而衰减。6% 样本量优化的遗

表 2 遗传 CA 空间直观模拟结果的混淆矩阵

Table 2 Confusion matrices between classifications from remote sensing and simulation outcomes generated by the GA based on CA model

年份 Year	样本量 Samples	遥感分类结果(像元) Remote sensing based classification (pixels)			生产者精度 Producer's accuracy (%)	使用者精度 User's accuracy (%)	总体精度 Overall accuracy (%)	Kappa 系数 Kappa coefficient	
		土地类型 Land type	A	B					C
2001	3%	A	73696	51607	1833	71.2	58.0	80.1	0.51
		B	27208	302863	3774	83.5	90.7		
		C	2587	8142	5570	49.8	34.2		
	6%	A	85481	44245	1895	82.6	64.9	83.9	0.61
		B	15151	309668	3939	85.4	94.2		
		C	2859	8699	5343	47.8	31.6		
2008	3%	A	204077	37639	1964	78.8	83.8	78.9	0.60
		B	48610	166063	3847	80.6	76.0		
		C	6197	2400	6483	52.7	43.0		
	6%	A	220797	36968	2392	85.3	84.9	82.5	0.66
		B	34413	167196	4288	81.1	81.2		
		C	3674	1938	5614	45.7	50.0		

A: 已开发 Developed; B: 未开发 Non-developed; C: 水域 Water.

表 3 遗传 CA 模拟结果与分类结果的景观格局比较

Table 3 Comparison of landscape pattern between simulation outcomes derived by the GA based on CA model and classified results

景观指数 Landscape index	2001			2008		
	分类结果 Classified result	3% 样本 3% samples	6% 样本 6% samples	分类结果 Classified result	3% 样本 3% samples	6% 样本 6% samples
NP	4413	4804	4813	4801	5013	4917
PD	14.79	16.10	16.13	16.54	17.27	16.94
LSI	33.02	39.44	37.32	47.69	43.00	45.11
FRAC	1.40	1.44	1.43	1.41	1.46	1.43
CONTAG	41.45	33.34	34.98	21.70	23.45	22.57
AI (%)	90.8	88.9	89.5	86.5	87.9	86.8
$E_{land}$ (%)	-	12.5	9.9	-	5.9	3.1

NP: 斑块数量 Patch number; PD: 斑块密度 Patch density; LSI: 景观形状指数 Landscape shape index; FRAC: 斑块分维数 Patch fractal dimension; CONTAG: 景观蔓延度指数 Landscape spread index; AI: 聚合性指数 Cohesiveness index;  $E_{land}$ : 景观综合评价指数 Landscape comprehensive evaluation index.

传 CA 模型模拟结果的总体精度高于 3% 样本量所构建模型,且前者 Kappa 系数表征的一致性程度较高。在地理 CA 建模过程中,建议选取尽可能大的样本量,以更好地代表整个研究区域,同时又要兼顾模型运行效率及对计算资源的要求,在两者之间取得平衡。

为了多角度评价模拟结果,进一步从景观格局视角分析土地利用空间结构的差异。通过 6 个景观指数[斑块数量(NP)、斑块密度(PD)、景观形状指数(LSI)、斑块分维数(FRAC)、景观蔓延度指数(CONTAG)和聚合性指数(AI)<sup>[24]</sup>]多角度评价模拟结果的景观特征,并与实际分类结果的景观指数进行对比。同时,为了评价模拟结果,本文提出了综合评价指标:

$$E_{land} = \sqrt{\frac{1}{n} \sum_{i=1}^n [(S_i - C_i)/C_i]^2} \quad (4)$$

式中: $E_{land}$ 为模拟结果与实际分类结果的综合景观差异指数; $n$ 为景观指数的数量; $S_i$ 为模拟结果的第*i*种景观指数; $C_i$ 为第*i*种景观指数的实际值。

2001—2008 年,嘉兴市土地利用总斑块数量(NP)和斑块密度(PD)均呈增加态势,土地利用结构趋于复杂化;表征斑块形状复杂程度的景观形状指数(LSI)增大,表明斑块破碎化程度增加;表征斑块自相似和复杂程度的斑块分维数(FRAC)呈增加趋势,显示景观形态趋于复杂;景观蔓延度指数(CONTAG)呈减小趋势,说明景观中不同斑块类型的团聚程度越来越高;聚合性指数(AI)呈减少趋势,表明相同斑块类型的聚集程度越来越低。2001 年 3% 和 6% 样本量的  $E_{land}$  值分别为 12.5% 和 9.9%,2008 年 3% 和 6% 样本量的  $E_{land}$  值分别为 5.9% 和 3.1%,表明 6% 样本量产生的模拟结果与

分类结果差异较小,即 6% 样本量构造的遗传 CA 模型较好。因此,在构建遗传 CA 模型时,不仅要考虑遗传智能的优化策略,更要考虑被优化的目标函数,该函数通过样本数据产生,以致样本数据对智能优化和参数挖掘均有较大影响,需要多少样本量应该根据研究区域而定。

#### 4 结 语

地理元胞自动机模型在土地利用、城市生长和生态格局等具有空间特征的复杂系统中广泛应用,不仅能追溯其历史格局,也能根据不同限制条件预测其未来情景。然而,在构建 CA 模型过程中,涉及的空间变量复杂且具有多重共线性关系,普通方法获取的模型所产生的模拟结果与真实分类结果之间存在差异。这种差异大则表明模型性能较差,而差异小则表明模型性能较优。根据 CA 模型的构建原理,可以将累积差异值通过离散的数学函数表达出来。利用生物信息学中的遗传智能算法,以上述函数为优化目标,利用其函数值评价染色体的优劣,并逐步寻找到最小差异值,经过遗传算法自动搜索,能够随机优化并挖掘地理 CA 模拟模型。本文以嘉兴市为例,随机获取 6% 和 3% 的样本点构建目标函数并挖掘 CA 参数,模拟了该区域 1992—2008 年的土地利用变化过程,并利用混淆矩阵、Kappa 系数和景观指数对模拟结果进行评价,结果表明,在既定遗传优化策略的情况下,与小样本量获取的模拟结果相比,大样本量的遗传 CA 模型的精度较高,景观格局更接近遥感分类结果。

根据既定区域计算 CA 模型参数受到遗传优化策略和样本量的影响,并不能确定是否适合于其他区域。但通过遗传智能优化地理 CA 模型的方法和

框架,对各区域的土地利用变化和城市生长均具有普适性。

#### 参考文献

- [1] Liu X-P (刘小平), Li X (黎夏), Peng X-J (彭晓鹏). Niche-based cellular automata for sustainable land use planning. *Acta Ecologica Sinica* (生态学报), 2007, **27**(6): 2391-2403 (in Chinese)
- [2] He CY, Okada N, Zhang QF, et al. Modeling urban expansion scenarios by coupling cellular automata model and system dynamic model in Beijing, China. *Applied Geography*, 2006, **26**: 323-345
- [3] Seppelt R, Voinov A. Optimization methodology for land use patterns using spatially explicit landscape models. *Ecological Modelling*, 2002, **151**: 125-142
- [4] Liu Y-L (刘耀林), Liu Y-F (刘艳芳), Ming D-P (明冬萍). Urban expansion model based on cellular automata of decision-making with grey situation. *Geomatics and Information Science of Wuhan University* (武汉大学学报·信息科学版), 2004, **2**(1): 7-13 (in Chinese)
- [5] Zhou C-H (周成虎), Ou Y (欧阳), Ma T (马廷), et al. Theoretical perspectives of CA-based geographical system modeling. *Progress in Geography* (地理科学进展), 2009, **28**(6): 833-838 (in Chinese)
- [6] Kathleen PB, Elena GI. Spatially explicit micro-level modelling of land use change at the rural-urban interface. *Agricultural Economics*, 2002, **27**: 217-232
- [7] Wu F, Webster CJ. Simulation of land development through the integration of cellular automata and multi-criteria evaluation. *Environment and Planning B: Planning and Design*, 1998, **25**: 103-126
- [8] Wu F. Calibration of stochastic cellular automata: The application to rural-urban land conversions. *International Journal of Geographical Information Science*, 2002, **16**: 795-818
- [9] Wu G-P (吴桂平), Zeng Y-N (曾永年), Zou B (邹滨), et al. Simulation on spatial land use patterns using autoLogistic method: A case study of Yongding County, Zhangjiajie. *Acta Geographica Sinica* (地理学报), 2008, **63**(2): 156-164 (in Chinese)
- [10] Feng Y-J (冯永玖), Tong X-H (童小华), Liu M-L (刘妙龙). Modelling urban growth with geographical cellular automata and partial least squares regression. *Journal of Tongji University* (Natural Sciences) (同济大学学报·自然科学版), 2010, **38**(4): 608-612 (in Chinese)
- [11] Santé I, García AM, Miranda D, et al. Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning*, 2010, **96**: 108-122
- [12] Yang Q-S (杨青生), Li X (黎夏). Cellular automata for simulating land use changes based on support vector machine. *Journal of Remote Sensing* (遥感学报), 2006, **10**(6): 836-846 (in Chinese)
- [13] Li X, Yeh AGO. Neural-network based cellular automata for simulating multiple land use changes using GIS. *International Journal of Geographical Information Science*, 2002, **16**: 323-343
- [14] Liu X, Li X, Liu L, et al. A bottom-up approach to discover transition rules of cellular automata using ant intelligence. *International Journal of Geographical Information Science*, 2008, **22**: 1247-1269
- [15] Feng Y-J (冯永玖), Tong X-H (童小华), Liu Y (刘艳), et al. A particle swarm intelligence based cellular model for urban morphology evolution modelling: A case study in Jiading District of Shanghai. *Journal of Geo-Information Science* (地球信息科学学报), 2010, **12**(1): 17-25 (in Chinese)
- [16] Karafyllidis I. Acceleration of cellular automata algorithms using genetic algorithms. *Advances in Engineering Software*, 1999, **30**: 419-437
- [17] Feng C (冯春), Ma J-W (马建文), Dai Q (戴芹), et al. A cellular automata model based on genetic algorithm simultaneously retrieving factor weights. *Geo-Information Science* (地球信息科学), 2005, **7**(1): 71-75 (in Chinese)
- [18] Li X, Yang QS, Liu XP. Genetic algorithms for determining the parameters of cellular automata in urban simulation. *Science in China Series D: Earth Sciences*, 2007, **50**: 1857-1866
- [19] Sivanandam SN, Deepa SN. Introduction to Genetic Algorithms. Berlin: Springer, 2009
- [20] Kocabas V, Dragicevic S. Assessing cellular automata model behaviour using a sensitivity analysis approach. *Computers, Environment and Urban System*, 2006, **30**: 921-953
- [21] Liu Y. Modelling Urban Development with Geographical Information Systems and Cellular Automata. Boca Raton: CRC Press, 2008
- [22] Han J, Hayashi Y, Cao X, et al. Application of an integrated system dynamics and cellular automata model for urban growth assessment: A case study of Shanghai, China. *Landscape and Urban Planning*, 2009, **91**: 133-141
- [23] Pontius RG, Huffaker D, Denman K. Useful techniques of validation for spatially-explicit land-change models. *Ecological Modelling*, 2004, **179**: 445-461
- [24] Turner MG, Gardner RH. Quantitative Methods in Landscape Ecology. New York: Springer-Verlag, 1991

作者简介 冯永玖,男,1981年生,博士,讲师.主要从事遥感与GIS应用、生态安全评估与空间直观模拟研究. E-mail: yjfeng@shou.edu.cn

责任编辑 杨弘