



武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: 机器学习在城市空间演化模拟中的应用与新趋势
作者: 陈逸敏, 黎夏
DOI: 10.13203/j.whugis20200423
网络首发日期: 2020-08-31
引用格式: 陈逸敏, 黎夏. 机器学习在城市空间演化模拟中的应用与新趋势. 武汉大学学报(信息科学版). <https://doi.org/10.13203/j.whugis20200423>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI: 10.13203/j.whugis.20200423

机器学习在城市空间演化模拟中的应用与新趋势

陈逸敏^{1,2} 黎夏^{3,4}

1 中山大学地理科学与规划学院, 广东广州, 510275

2 广东省城市化与地理环境空间模拟重点实验室, 广东广州, 510275

3 华东师范大学地理科学学院, 上海, 200241

4 地理信息科学教育部重点实验室, 上海, 200241

摘要：城市模拟自 20 世纪 80—90 年代兴起后, 已成为城市研究的一种新范式, 体现了计算思维对于城市研究的深刻影响。城市模拟方法建立在元胞自动机 (cellular automata, CA) 和机器学习基础上, 形成了具有模拟城市复杂演化过程、实现多情景分析能力的城市 CA 模型。本文回顾了城市模拟的起源和发展, 在归纳城市 CA 一般结构的基础上, 讨论了机器学习方法在支持城市模拟方面的必要性和可行性, 并进一步综述了机器学习与 CA 在城市研究中的新趋势, 阐述了当前面临的主要挑战。

关键词：元胞自动机; 机器学习; 城市模拟

中图分类号：P208 **文献标志码：**A

全球已有超过 55% 的人口居住在城市, 而城市人口仍然在持续增长。城市集中了人类社会绝大多数的活动和财富, 城市的发展也对自然环境和生态系统形成了巨大影响。因此, 城市一直是地理学关注的重要研究对象。

城市研究经历了多种范式的转变。由于城市系统的复杂性, 城市系统的演化很难通过简单的数学方程式来表达。以元胞自动机 (cellular automata, CA) 为代表的模拟方法自 20 世纪 80-90 年代逐渐兴起, 被认为是研究复杂城市系统的有效手段。几乎是在同一时期, 以后向传播神经网络、决策树算法为代表的一系列机器学习方法被相继提出和完善, 机器学习开始全面兴起, 并迅速对地理信息科学产生影响。Stan Openshaw 的著作《Artificial intelligence in geography》^[1] 象征着机器学习在地理学领域所引发的重要变革。机器学习也很快被应用到 CA 建模与城市模拟, 极大促进了城市模拟研究的发展。城市模拟逐渐成为地理信息科学领域的一个重要研究方向。著名地理信息科学期刊

《International Journal of Geographical Information Science》于 2011 年发布的创刊 25 周年经典论文集中, 就有 4 篇论文是以城市模拟为主题^[2-5]。知名城市模型研究者 Keith Clarke 教授认为模拟方法提供了低成本、快速和安全的实验环境来研究真实城市中的种种复杂现象和问题, 同时也有助于发现新的结构、模式和过程^[6]。

不管是 CA 还是机器学习方法, 这两大城市模型支柱的建立均得益于计算机的发明和计算机科学的飞速发展, 同时它们也共同体现了“计算思维”这一别于过去的城市建模新理念。CA 的雏形最早可以追溯到 20 世纪 40—50 年代, 是由 Stanislaw Ulam 所建立的一个用于研究晶体的离散模型。John von Neumann 进一步提出了自动机的概念, 用于研究机器自我复制行为 (即用机器来制造机器)。随后 Stanislaw Ulam 和 John von Neumann 在研究流体运动中正式提出 CA 这一概念。50—60 年代是计算机基础理论发展的重要时期, 而 CA 也在这一时期逐渐被认为是一种通用计算模型^[7]。在 70 年代, John

项目资助：国家重点研发计划项目 (2019YFA0607201); 国家自然科学基金 (41871306)

第一作者：陈逸敏, 博士, 副教授级, 主要从事城市计算与情景推演研究。chenym49@mail.sysu.edu.cn

通讯作者：黎夏, 博士, 教授, 长江学者。lixia@geo.ecnu.edu.cn

Conway 提出的“生命游戏”模型充分展示了 CA 如何通过极其简单的规则组合来“计算”出高度复杂的系统行为。Stephen Wolfram 在 80 年代深入研究了一维 CA 的行为并进行分类, 基于这些研究成果他撰写了《A New Kind of Science》一书。进入 80—90 年代, 随着地理信息系统 (GIS) 的快速发展和计算思维的引入, CA 开始被尝试用于城市研究^[8, 9]。CA 的离散特性和空间特性使其能够非常方便的与 GIS 相结合, 而 CA“以简单规则模拟复杂行为”的特点又进一步使其在研究复杂城市系统方面具有巨大优势。因此, 初等 CA 被逐步改造为城市 CA^[10, 11], 并成为地理计算和城市建模的标志性方法, 在城市增长模拟、城市发展情景分析、城市管控界线优化、城市影响评估等方面发挥重要作用^[12, 13]。

机器学习被认为起源于 20 世纪 60 年代, 其定义可概括为是一种计算机程序, 这种程序能让计算机从经验中提高解决某种任务的能力。从这个角度讲, 机器学习研究的初衷之一即是为了实现人工智能。机器学习方法的发展十分迅速, 衍生出了许多重要的分支, 例如神经网络所代表的“联结主义”(Connectionism)流派, 近年来由于深度学习的成功而盛行于许多研究领域。地理信息科学和城市研究对于机器学习方法的吸收和应用一直都十分积极。机器学习方法是解决空间选址、模型寻优、空间数据挖掘、模式识别、数据分类、特征筛选和提取等问题的有效方法。在城市模拟模型的发展历程中, 机器学习也起到了非常关键的作用。例如, 后向传播神经网络、决策树算法等重要算法刚刚提出不久就马上被结合到城市模拟中, 用于城市演化规则挖掘和多种城市规划策略的情景分析^[14-16]。机器学习后续的许多经典算法和重要成果也很快被用于城市模拟, 包括支持向量机^[17]、群集智能^[18]、集成学习^[19, 20]、迁移学习^[21]和深度学习^[22, 23]等。

在这一背景下, 本文首先归纳了城市 CA 的一般结构以说明机器学习方法与城市模拟结合的必要性和可行性。其次, 本文阐述了机器学习在解决城市演化规则这一城市模拟核心问题当中所起的作用及其原理。

最后, 本文综述了机器学习与城市 CA 研究的新趋势, 并讨论了目前所面临的主要挑战。

1 典型城市 CA 及其一般结构

大约自 20 世纪 90 年代中、后期开始, CA 被正式用于真实城市增长模拟与情景分析。随着研究的深入, 诞生了多个影响深远的城市 CA 模型, 它们分别是 SLEUTH^[2]、约束性 CA^[24]、MCE-CA^[25]、Logistic-CA^[4]、IF-THEN 规则 CA^[16]、ANN-CA^[26]以及 CBR-CA^[27]。这些模型体现了对城市过程建模的不同思路。其中, SLEUTH 模型将城市空间增长过程概括为四种类型 (自发增长、扩散增长、边缘增长和随道路蔓延), 可以为研究者提供更直观的城市演化规律信息。约束性 CA、MCE-CA 和 Logistic-CA 等模型, 则不约而同的将元胞由“非城市”状态转变为“城市”状态的概率表达为城市发展适宜性和邻域效应的函数, 这种形式也被许多研究所沿用。IF-THEN 规则 CA 和 ANN-CA 则利用了机器学习方法在处理复杂非线性关系方面的优势, 进一步提高了模型结果的精度和可靠性。CBR-CA 模型规避了城市系统时空异质性的规则化表达难题, 通过案例推理的方式, 直接利用“旧经验” (即样本数据组成的案例库) 来解决“新问题” (即预测某个位置是否由“非城市”状态转变为“城市状态”), 适用于空间范围较广、异质性较高的区域尺度城市增长模拟。

尽管这些模型各有特点, 但其一般结构却是一致的 (图 1)。模型的输入数据主要包括城市土地利用观测数据和一系列影响城市增长的驱动力因素, 利用这些数据来获取城市演化规则。对应于不同的规则挖掘方法 (如逻辑回归方法、机器学习方法等), 城市演化规则的表达大致具有三种形式: (1) 从城市发展适宜性换算为发展概率, (2) 以一套“IF-THEN”规则来表达, (3) 除上述两种形式以外的其他表达方式 (如神经网络结构, 或者案例推理形式)。不管采用哪种形式, 这些规则均作用于全局, 且一般是确定性的。与之相对的, 是作用于局部, 且动态变化的邻域效应影响, 它所反映的是城市空

间演化过程中存在的路径依赖效应和正反馈作用^[5]。全局性规则和局部的邻域效应共同决定了元胞是否在下一个时刻发生(土地利用)状态的变化。

模型以迭代的方式模拟城市空间演化的过程,模型的误差则通过比较模拟结果与观测数据的一致性程度来进行评估^[28]。其中,通过全局性规则获得的发展概率,可以用接收者操作特征方法(Receiver Operating Characteristic;简称 ROC)或其改进版本总体操作特征方法(Total Operating Characteristic;简称 TOC)^[29]来检验;模型输出的空间模拟结果则可以利用 Kappa 指数^[30]、Figure-of-merit 指数^[31]和景观指数组合^[32]等进行检验。

在上述结构中,机器学习方法的作用在于从城市土地利用观测数据和驱动力因素中获取全局性的规则。许多经验研究均认为在获取规则方面,机器学习方法相对于常规的统计方法具有更高的准确度。

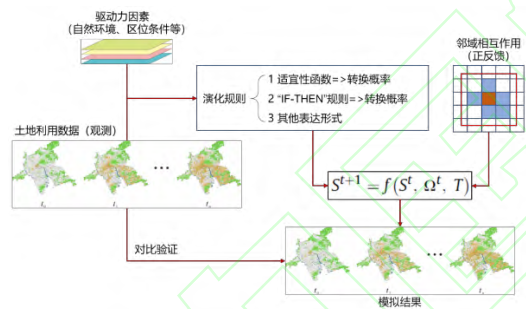


图1 城市CA模型的一般结构

Fig.1 General structure of Urban CA Model

2 利用机器学习方法挖掘城市演化规则

相比常规的统计方法,机器学习方法通常不会假定数据满足某种分布,因此具有更高的灵活性。机器学习方法也被认为更擅长处理数据中的非线性关系^[26]。这些特性使得机器学习方法被许多研究所青睐。具体到城市增长模拟研究,机器学习方法的作用在于为城市CA模型提供演化规则。尽管不同的机器学习方法在理念、结构和技术实现上各有特点,但在获取城市演化规则方面具有高

度相似性:将演化规则挖掘还原为分类问题。

分类问题一般描述为:给定数据集 $X = \{x_1, x_2, \dots, x_n\}$, 每个数据点 x_n 均拥有 m 个特征 f_m , 且 x_n 的类别(也称作标签)是类别集 $C = \{c_1, c_2, \dots, c_k\}$ 中的一个,则分类就是建立一个映射 F 使得 $F(x_n) \Rightarrow c_k$, 即利用映射 F 实现数据 x_n 的分类。这一逻辑也适用于城市演化规则的获取。其中,类别集一般由土地利用类别或状态构成(例如:{城市, 非城市}, 或者{转变为城市状态, 保持非城市状态}), 特征 f_m 则用影响城市增长的驱动力因素来表示,并通过随机采样的方式建立数据集。最后,训练某种机器学习方法来建立映射 F 。但与常规机器学习分类稍有不同的是,在获得映射 F 之后,通常是选择输出数据 x_n 属于类别 c_k 的概率 p_k (例如,元胞 x_n 属于“转变为城市”的概率)而非直接输出概率最大的类别。在城市CA中,一般由概率 p_k 和邻域效应共同影响元胞是否发生状态变化。

因此,城市演化规则的获取即转化为映射 F 及概率 p_k 的获取。类似的,映射 F 也具有上文提到的三种表达形式:城市发展适宜性函数、“IF-THEN”规则和其他形式。若采用城市发展适宜性函数来换算城市发展概率,则机器学习方法的任务是确定适宜性函数的最优系数。现有的研究大多采用 Wu^[4] 的 Logistic-CA 所采用的适宜性函数形式,并通过遗传算法^[33]、支持向量机^[17]、粒子群算法^[18]等来优化适宜性函数的系数。因此,这些研究所建立的模型可以认为是 Logistic-CA 的变种。“IF-THEN”规则是经典的决策树算法所采用的映射表达形式。由决策树算法生成的“IF-THEN”规则集合可以直接控制城市CA模型^[16]。与此类似,蚁群算法^[34]等群集智能算法也可以用于生成“IF-THEN”规则集合。除了这两种形式以外,其他机器学习算法如神经网络、随机森林或其他集成分类器,以及近年来兴起的深度学习等方法,具有更为复杂的结构,“黑箱”的特点也更为明显。

3 机器学习与城市模拟的新趋势

相比过去许多研究聚焦于对机器学习

方法应用于城市研究的可行性探索,近十年来,一个新的趋势是尝试将机器学习方法和城市 CA 相结合来解决具体的城市问题。其中,机器学习方法和城市 CA 模型大多作为一个更大的分析框架中的一部分,用以生成不同条件假设下的城市空间格局。典型的研究问题包括:城市管控界线设计、城市地块更新,以及未来不同城市演化路径的生态与气候影响。

城市生态控制线、基本农田保护区、城市增长边界等管控界线等管控界线设立,从空间建模的角度来看其实质是空间优化问题,即在满足一定的社会、经济约束下实现某种规划目标的最大化/最优化。例如,城市生态控制线的规划通常需要考虑最大化的覆盖生态环境质量最好或者生态脆弱地区,同时满足连通性、紧凑性等形态约束。因此,以往的研究采用智能体建模或者群集智能算法来解决这类空间优化问题^[35, 36]。此外,这些空间优化模型还可以与城市 CA 结合形成耦合模型,来探索城市动态演化下的生态控制线优化方案^[37]。类似的,城市增长边界规划也可以基于城市 CA 情景模拟来实现^[38],其核心思想是利用城市 CA 来获取城市演化的历史趋势和规律,并用以预测未来的城市布局,在此基础上利用 GIS 基本的几何分析功能完成增长边界的划定。近年来也有研究者尝试将生态控制线与城市增长边界同时纳入情景模拟之中^[39]。

另外,我国许多城市也逐渐改变了以扩张为主的增长方式,转为开展城市内部更新。为了适应这一趋势,研究者建立了面向地块对象的城市 CA 模型^[20, 23, 40],并通过集成学习方法、随机森林、深度学习等方法来获取不同城市土地利用类型之间的复杂相互作用关系,实现地块尺度的城市土地利用变化模拟,有助于为城市内部更新提供决策支持。

在更大的时空尺度上,城市对于生态环境和气候变化的影响越来越受到研究者的重视。因此,基于机器学习的城市 CA 也逐渐成为大尺度城市演化影响研究的重要模型之一。例如,最近的一项研究基于神经网络和城市 CA 建立了对接 IPCC 共享社会经济路径 (SSP) 的未来全球城市演化情景,

发现 2050 年代将成为全球城市演化的重要时间拐点,许多国家将在 2050 年代之后面临城市收缩的压力^[41]。另外一项研究模拟了 2015 至 2050 年全球主要城市区域的城市用地增长^[42],发现城市用地的增长将引起夏季日间和夜间气温升高 0.5℃–0.7℃,是全球温室气体排放引起的气温升高幅度的 0.5 至 2 倍。国内学者也针对我国国内区域间社会经济结构差异巨大的特点,利用随机森林算法、城市 CA 和多区域投入产出分析方法建立了中国 SSP 城市遥关联(Tele-connection)情景模拟模型,分析了区域间产业结构紧密联系背景下本地城市增长对其他地区的生态环境和资源消耗影响^[43]。

4 讨论与总结

城市模拟是地理信息科学和城市研究的重要方法。借助机器学习方法挖掘城市演化规则,驱动城市 CA 等模拟模型,可以为城市空间规划和未来城市发展路径探索提供决策支持。尽管机器学习方法已在城市模拟领域获得了巨大成功,但仍然面临一些挑战。

1)用于训练机器学习模型的样本和知识是否具有时空可迁移性。例如,国内不同地区的城市,其产业结构和发展阶段各不相同。从一个区域中训练的模型、挖掘的知识是否能够或多大程度上适用于另外一个区域的城市建模。对于这一问题的解答不仅有助于改进建模技术,也能够推进对于城市系统演化的理解。目前相关的研究较为缺乏,尽管已有学者尝试应用迁移学习方法来开展实验^[21],但其时空跨度较小,仍不足以提供较为全面的理解。

2)利用机器学习方法获得的知识存在解释性难题。机器学习方法非常擅长解决分类、目标识别、参数寻优等问题,但其规则的解释性较差,很难被解读和“转译”为更直观的信息,直接影响了研究者对于城市系统内在机理的认知。如何利用机器获取更具人文特征的知识,而不仅仅是“机器的知识”,仍然是一个难以解决的问题。

3)尽管机器学习与城市 CA 已经在空间

规划、全球变化影响等重要应用上取得进展,城市模拟模型仍然有待进一步发展和改良。目前,对于影响城市系统演化的关键要素如人口、活动、交通和土地利用等的过程模拟依然无法实现完整的耦合,缺少通用、统一的城市系统模型。由于不同城市过程的主体、机制和时空粒度等各不相同,实现这些过程的耦合建模涉及到一个根本问题,即如何建立城市系统的基本单元,以整合不同时空粒度的要素并进一步表达这些基本单元的空间相互作用。未来对于这一关键问题的探索,可能需要借鉴已经被用于模拟全球变化的“地球系统耦合模式”,尝试建立面向城市的“通用城市系统耦合模式”。

参考文献

- [1] Openshaw S, Openshaw C. Artificial Intelligence in Geography [M]. 1997: John Wiley & Sons, Inc.
- [2] Clarke K, Gaydos L. Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore [J]. *International Journal of Geographical Information Science*, 1998, 12(7): 699-714
- [3] Li X, Yeh A G O. Modelling sustainable urban development by the integration of constrained cellular automata and GIS [J]. *International Journal of Geographical Information Science*, 2000, 14(2): 131-152
- [4] Wu F. Calibration of stochastic cellular automata: the application to rural-urban land conversions [J]. *International Journal of Geographical Information Science*, 2002, 16(8): 795-818
- [5] Brown D G, et al. Path dependence and the validation of agent-based spatial models of land use [J]. *International Journal of Geographical Information Science*, 2005, 19(2): 153-174
- [6] Clarke KC. Why simulate cities? [J]. *GeoJournal*, 2014, 79(2): 129-136
- [7] Smith III A R. Cellular automata complexity trade-offs [J]. *Information and Control*, 1971, 18(5): 466-482
- [8] White R, Engelen G. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns [J]. *Environment and Planning A*, 1993, 25: 1175-1175
- [9] Batty M. Cellular automata and urban form: a primer [J]. *Journal of the American Planning Association*, 1997, 63(2): 266-274
- [10] Batty M. Urban evolution on the desktop: simulation with the use of extended cellular automata [J]. *Environment and Planning A*, 1998, 30: 1943-196
- [11] Torrens P, O'Sullivan D. Cellular automata and urban simulation: where do we go from here? [J]. *Environment and Planning B: Planning and Design*, 2001, 28(2): 163-168
- [12] Santé I, 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
- [13] Li X, et al. Experiences and issues of using cellular automata for assisting urban and regional planning in China [J]. *International Journal of Geographical Information Science*, 2017, 31(8): 1606-1629
- [14] Li X, Yeh A G O. Calibration of cellular automata by using neural networks for the simulation of complex urban systems [J]. *Environment and Planning A*, 2001, 33(8): 1445-1462
- [15] Li X, Yeh A G O. Neural-network-based cellular Automata for Realistic and Idealized Urban Simulation [J]. *Journal of Geographical Sciences*, 2002, 57(2): 159-166 (黎夏, 叶嘉安. 基于神经网络的单元自动机 CA 及真实和优化的城市模拟 [J]. *地理学报*, 2002, 57(2): 159-166)
- [16] Li X, Yeh A G O. Data mining of cellular automata's transition rules [J]. *International*

-
- Journal of Geographical Information Science*, 2004, 18(8): 723-744
- [17] Yang Q S, Li X, Shi X. Cellular automata for simulating land use changes based on support vector machines [J]. *Computers & Geosciences*, 2008, 34(6): 592-602
- [18] Feng Y J, et al. Modeling dynamic urban growth using cellular automata and particle swarm optimization rules [J]. *Landscape and Urban Planning*, 2011, 102(3): 188-196
- [19] Li X, Liu X P, Gong P. Integrating ensemble-urban cellular automata model with an uncertainty map to improve the performance of a single model [J]. *International Journal of Geographical Information Science*, 2015, 29(5): 762-785
- [20] Chen Y M, Liu X P, Li X. Calibrating a Land Parcel Cellular Automaton (LP-CA) for urban growth simulation based on ensemble learning [J]. *International Journal of Geographical Information Science*, 2017, 31(12): 2480-2504
- [21] Li X, et al. Knowledge transfer and adaptation for land-use simulation with a logistic cellular automaton [J]. *International Journal of Geographical Information Science*, 2013, 27(10): 1829-1848
- [22] He J, et al. Mining transition rules of cellular automata for simulating urban expansion by using the deep learning techniques [J]. *International Journal of Geographical Information Science*, 2018, 32(10): 2076-2097
- [23] Zhai Y, et al. Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata [J]. *International Journal of Geographical Information Science*, 2020, 34(7): 1475-1499
- [24] White R, Engelen G, Uljee I. The use of constrained cellular automata for high-resolution modelling of urban land-use dynamics [J]. *Environment and planning B*, 1997, 24: 323-344
- [25] Wu F, Webster C J. Simulation of land development through the integration of cellular automata and multicriteria evaluation [J]. *Environment and planning B*, 1998, 25: 103-126
- [26] Li X, Yeh A G O. Neural-network-based cellular automata for simulating multiple land use changes using GIS [J]. *International Journal of Geographical Information Science*, 2002, 16(4): 323-343
- [27] Li X, Liu X P. An extended cellular automation using case-based reasoning for simulating urban development in a large complex region [J]. *International Journal of Geographical Information Science*, 2006, 20(10): 1109-1136
- [28] Tong X, Feng Y J. A review of assessment methods for cellular automata models of land-use change and urban growth [J]. *International Journal of Geographical Information Science* 2020, 34(5): 866-898
- [29] Pontius Jr R G, Si K. The total operating characteristic to measure diagnostic ability for multiple thresholds [J]. *International Journal of Geographical Information Science*, 2014, 28(3): 570-583
- [30] van Vliet J, Bregt A K, Hagen-Zanker A. Revisiting Kappa to account for change in the accuracy assessment of land-use change models [J]. *Ecological modelling*, 2011, 222(8): 1367-1375
- [31] Pontius R, et al. Comparing the input, output, and validation maps for several models of land change [J]. *The Annals of Regional Science*, 2008, 42(1): 11-37
- [32] Chen Y M, et al. Modeling urban land-use dynamics in a fast developing city using the modified logistic cellular automaton with a patch-based simulation strategy [J]. *International Journal of Geographical Information Science*, 2014, 28(2): 234-255
- [33] Li X, Yang Q S, Liu X P. Discovering and

- evaluating urban signatures for simulating compact development using cellular automata [J]. *Landscape and Urban Planning*, 2008, 86(2): 177-186
- [34] Liu X P, et al. A bottom-up approach to discover transition rules of cellular automata using ant intelligence [J]. *International Journal of Geographical Information Science*, 2008, 22(11-12): 1247-1269
- [35] Chen Y M, et al. An agent-based model for optimal land allocation (AgentLA) with a contiguity constraint [J]. *International Journal of Geographical Information Science*, 2010, 24(8): 1269-1288
- [36] Li X, et al. Coupling urban cellular automata with ant colony optimization for zoning protected natural areas under a changing landscape [J]. *International Journal of Geographical Information Science*, 2011, 25(4): 575-593
- [37] Li X, et al. Concepts, methodologies, and tools of an integrated geographical simulation and optimization system [J]. *International Journal of Geographical Information Science*, 2011, 25(4): 633-655
- [38] Liang X, et al. Delineating multi-scenario urban growth boundaries with a CA-based FLUS model and morphological method [J]. *Landscape and Urban Planning*, 2018, 177: 47-63
- [39] Chen Y M, et al. Simulating urban growth boundaries using a patch-based cellular automaton with economic and ecological constraints [J]. *International Journal of Geographical Information Science*, 2019, 33(1): 55-80
- [40] Yao Y, et al. Simulating urban land-use changes at a large scale by integrating dynamic land parcel subdivision and vector-based cellular automata [J]. *International Journal of Geographical Information Science*, 2017, 31(12): 2452-2479
- [41] Chen G Z, et al. Global projections of future urban land expansion under shared socioeconomic pathways [J]. *Nature Communications*, 2020, 11(1): 1-12
- [42] Huang K N, et al. Projecting global urban land expansion and heat island intensification through 2050 [J]. *Environmental Research Letters*, 2019, 14(11): 114037
- [43] Chen Y M, et al. Tele-connecting China's future urban growth to impacts on ecosystem services under the shared socioeconomic pathways [J]. *Science of The Total Environment*, 2019, 652: 765-779

Applications and New Trends of Machine Learning in Urban

Simulation Research

CHEN Yimin^{1,2} LI Xia^{3,4}

¹ School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China

² Guangdong Key Laboratory for Urbanization and Geo-simulation, Sun Yat-sen University, Guangzhou 510275, China

³ School of Geographic Sciences, East China Normal University, Shanghai 200241, China

⁴ Key Lab of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China

Abstract : Urban simulation research originated between the 1980s and 1990s. Today urban simulation has become a new paradigm of urban research, which is an

important outcome of computational thinking in urban research. Urban simulation methods are usually based on cellular automata (CA) and machine learning. A series of urban CA models have been developed to simulate complex urban evolution processes and associated multi-scenario analysis. This article reviews the origin and progress of urban simulation research. With the discussion of urban CA's general structure, this article explains the necessity and feasibility of machine learning methods to support urban simulation. Furthermore, this article reviews the integration of machine learning and CA in urban research, and also discusses its new trends and emerging challenges.

Keywords: cellular automata (CA); machine learning; urban simulation

First author: CHEN Yimin, PhD, associate professor, specializes in urban computation and scenario simulation.
Email: chenym49@mail.sysu.edu.cn

Corresponding author: LI Xia, PhD, professor, specializes in land use change modeling.
Email: lixia@geo.ecnu.edu.cn

Foundation support: The National Key Research and Development Program of China (2019YFA0607201), and National Natural Science Foundation of China (41871306).