

新数据环境下的中国人居环境研究

龙瀛 郎崑

Human Settlement Study of China in the New Data Environment

LONG Ying¹, LANG Wei²

(1. School of Architecture, Tsinghua University, Beijing 100084, China; 2. Faculty of Construction and Environment, Hong Kong Polytechnic University, Hong Kong, China)

Abstract This paper is a systematic review about the new paradigm application in the urban system of China, which is presented by Long Ying, Wu Kang et al. (2014) as a “big model” for urban and regional research. These quantitative researches of urban systems of China from microscopic perspective include micro-level data construction, urban space development, urban spatial structure, ecological and environmental system analysis, and response from urban planning and design. This paper also made discussion and conclusion from research objects, methods, approaches, findings, and objectives, with an expectation of bringing new inspiration and reference for the future urban planning and development.

Keywords new data environment; big model; quantitative urban studies; new paradigm; Sciences of Human Settlements

摘要 随着由大数据和开放数据所构成的新数据环境的形成,以及日益成熟的计算能力与日臻完善的区域及城市分析和模拟方法,龙瀛、吴康等(Long, Wu et al., 2014)提出了“大模型”这种城市和区域研究的新范式,它是由大规模数据驱动的定量城市与区域研究工具,利用简单直接的建模方式,兼顾了大尺度和精细化模拟单元。本文是“大模型”研究范式在中国城市系统应用的系统综述,这些细粒度的中国人居环境研究囊括了微观层面基础数据构建、城市空间开发、城市空间结构、生态环境系统分析、城市规划及设计响应等方面,致力于对中国人居环境进行多维度解读。最后,就研究对象、研究方式和方法、研究发现以及研究目标进行了评述,以期对未来的城市规划与城市发展带来新的启示和参考。

关键词 新数据环境; 大模型; 定量城市研究; 新范式; 人居环境科学

1 引言

随着信息通信技术与物联网技术的发展,智能终端、射频识别(RFID)、无线传感器等装置产生的数据量与日俱增。同时,对互联网依赖性不断加强的城市社会经济活动,网络平台(主题网站、社交网站、搜索引擎等)也在产生着大量数据信息(龙瀛, 2014; Liu et al., 2015)。此外,各种政府和商业开放数据项目及志愿者地理信息项目(Volunteered Geographic Information, VGI)也在扩充着城市基础数据,这些数据共同形成了有别于传统抽样调研和静态统计数据的新数据环境(New Data Environment)。与传

作者简介

龙瀛, 清华大学建筑学院;

郎崑, 香港理工大学建筑与环境学院。

统数据相比,新数据环境主要呈现出精度高(以单个人或设施为基本单元)、覆盖广(不受行政区域限制)、更新快(每月、每日甚至每分钟更新)等特点。例如,传统数据多反映某一时刻或一段时间内城市所处的状态,只能覆盖有限的空间范围。而包括公交刷卡、出租车轨迹、信用卡交易记录、在线点评以及位置微博和照片等在内的新数据环境,则可以反映个人乃至整个城市短至每秒、长至多年的动态变化(Bagchi and White, 2004; 龙瀛等, 2012)。目前,在新数据环境下,我国开展的具有一定代表性的大范围细粒度的定量城市研究,详见龙瀛和刘伦的研究综述(龙瀛和刘伦, 2016)。

在国家出台的《国家新型城镇化规划(2014~2020年)》背景下,针对城市研究所面临的这些新的机遇和挑战,龙瀛、吴康等(2014)提出了“大模型”研究范式,试图在城市研究中兼顾覆盖大范围乃至全国范围的考量与精细化的数据处理分析。“大模型”是指由大规模数据驱动,兼顾大尺度和精细化计算单元的定量城市研究工具,代表了一种新的研究范式,其与传统的城市研究范式有较大不同(图1)^①。在传统数据环境下,受数据收集方法的限制,城市研究在研究覆盖范围和精细度上往往很难做到两者兼顾。“大模型”则兼顾了研究尺度和计算单元(大空间、细粒度)。而“大模型”的空间覆盖多为城市群或更大(全国)范围的大量城市,基本研究单元在物理空间维度多为地块、街区或街道等,在社会空间维度多为居民、家庭和企业个体等。“大模型”多采用传统计量分析思路,以更加简单直观的方式,通过覆盖所有的城市来缓解中小城市的技术和数据鸿沟,并致力于归纳城市系统的一般规律及地区差异,进而完善已有或提出全新的城市理论,最终实现支持规划设计和其他城市发展政策设计。

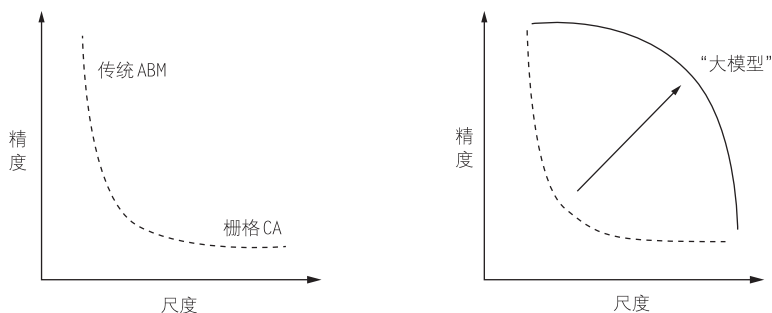


图1 传统城市区域研究模型与“大模型”对比

资料来源:龙瀛、吴康等(2014)。

新数据环境使“大模型”这一城市与区域研究的新范式的推广成为可能。自龙瀛及其合作者2014年提出“大模型”研究范式以来,大量的针对中国人居环境特别是城市系统的量化研究已经开展,涵盖微观层面基础数据构建、城市空间开发、城市空间结构、生态环境系统分析、城市规划及设计响应等多方面,总体框架如图2所示,六个方面的第一部分是细粒度基础数据重建,也是其他五个方面研究的基础,后五个方面分别对应中国城市发展的不同方面的重大问题。这些研究以人居环境质量为核心

心, 细粒度为研究中国人居环境提供了新视角, 以期对中国快速城镇化时期的人居环境质量进行全面的度量与监测, 为国家决策提供依据和保障。下文将简要介绍各个方面的典型案例, 并在最后从研究对象、研究方式和方法、研究发现以及研究目标等方面进行评述, 以期抛砖引玉, 给未来中国的城市研究、规划设计和政策制定带来新的启示与参考。

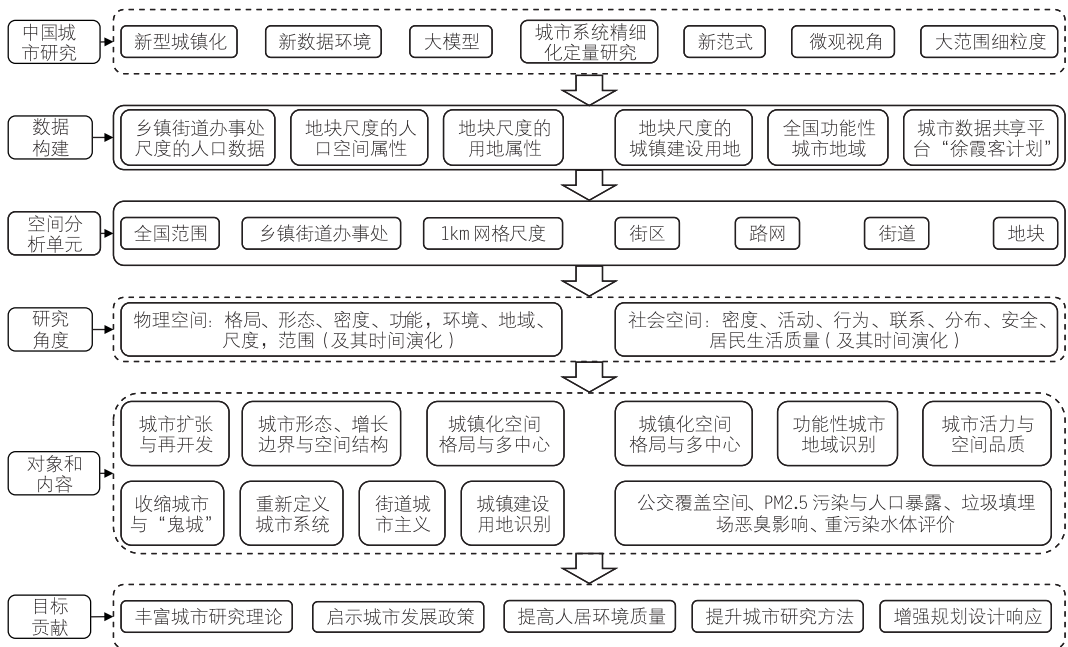


图2 “大模型”研究范式总体框架及其实证研究

2 微观层面基础数据构建

2.1 乡镇街道办事处尺度的人口密度

传统的人口空间格局研究多以省市或区县行政区作为分析单元, 空间分辨率较粗, 不易表达出精准的人口分布情况。毛其智等(2015)基于2000年第五次人口普查和2010年第六次人口普查的基础数据, 在国内较早建立了全国乡镇街道单元层级的人口数据库。首先利用Google API进行地址的空间化匹配, 得到2010年普查对应的43 536个乡镇街道单元点和2000年普查对应的50 518个乡镇街道单元点; 然后以2012年全国乡镇街道办事处级别的政区边界为基础, 对人口数据库的单元点进行匹配。结果显示, 2000年常住人口在全国39 007个乡镇街道的平均密度为873人/km², 到2010年则上升到977人/km²。在39 000余个乡镇街道单元中, 人口密度基本未变的仅为17 808个, 不到50%, 而大

部分乡镇街道单元的人口密度都有所降低或增加，其中约 22% 的乡镇街道人口密度表现为显著降低或增加（图 3）。密度降低的乡镇街道单元总计 12 840 个，总面积约 173.4 万 km^2 ，而密度增加的乡镇单元总计 8 359 个，总面积约 135.4 万 km^2 。这套微观数据是后续开展多项中国城乡研究工作的基础。

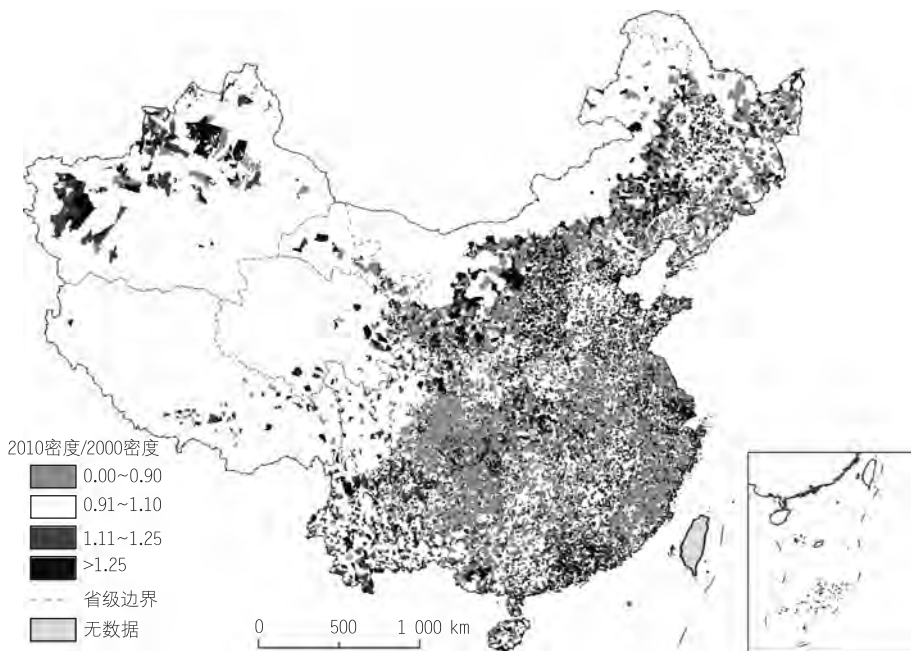


图 3 2000~2010 年中国乡镇街道办事处尺度的人口密度变化

资料来源：毛其智等（2015）。

2.2 地块尺度人口空间化与属性合成

由于长期以来我国人口方面的微观数据严重匮乏，因此人口微观样本的合成（population synthesis）是在我国进行精细化城市研究的重要工作环节（龙瀛、茅明睿等，2014）。例如，龙瀛等学者基于统计资料、小规模样本调查和常识性知识，合成了北京市全样本的居民个体数据，为开展精细化城市模拟提供了条件（龙瀛等，2011；Long and Shen, 2013）。同时，龙瀛和沈振江（Long and Shen, 2015）还利用覆盖全国的开放街道地图（OpenStreetMap, OSM）中的道路资料划分地块，并基于矢量元胞自动机（Cellular Automata, CA）模型，结合兴趣点（Points of Interest, POIs）数据推测地块是否为城市用地，并利用住房相关的网络签到数据来区别城市用地中的居住地块。最后，利用龙瀛和沈振江（Long and Shen, 2013）所开发的“Agenter”工具以及公开的区县尺度的人口分布数据，在地块尺度进行人口空间化，并对人口属性进行了合成（图 4）。

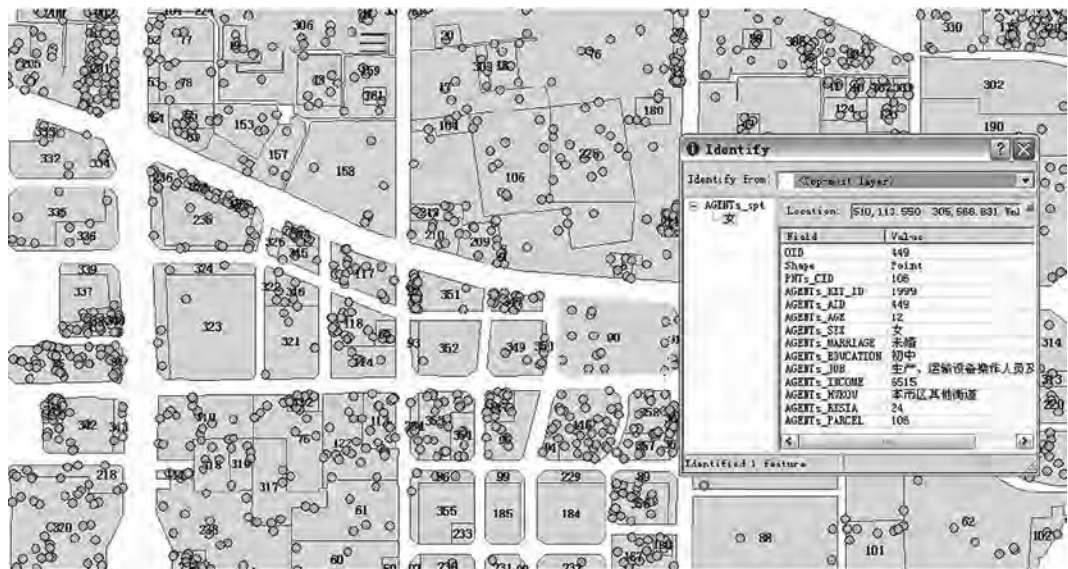


图4 北京市居民个体属性空间分布(部分)

资料来源:龙瀛和沈振江(Long and Shen, 2015)。

2.3 地块尺度的用地属性推导

由于传统收集和处理用地属性数据的方式耗时较长,而且用地现状数据在我国一般要求保密。针对国内城市用地现状图开放度不足的问题,刘行健和龙瀛(Liu and Long, 2015)提出了利用全国OSM路网和兴趣点数据自动化分地块并推导属性的方法,最终生成了全国297个城市的用地现状图,针对划分的各个地块,推导了主导用地功能、功能密度、功能混合度等指标(图5)。龙瀛和刘行健(Long and Liu, 2013)还对功能混合度进行了专题研究。

2.4 基于道路网络和兴趣点的城镇建设用地识别

城镇建设用地识别多基于遥感数据获得,需要大量的财力和人力,同时难以在短时间内针对较大区域识别精细化尺度(如地块、街区等)的城镇建设用地。因此,采用道路网和兴趣点识别城镇建设用地会是一种快速而有效的方法,即道路交叉口密度法。龙瀛、沈尧等(Long, Shen et al., 2016)首先利用路网数据来划定地块边界,依据国家道路等级宽度标准对不同等级道路生成2~30m(单边)不同宽度的缓冲区,进而切割出地块(包括城镇和非城镇地块);然后,基于矢量元胞自动机模型,根据地块的大小、POIs密度、区位特征等属性对初始地块进行筛选,识别出城镇地块;最后,借助

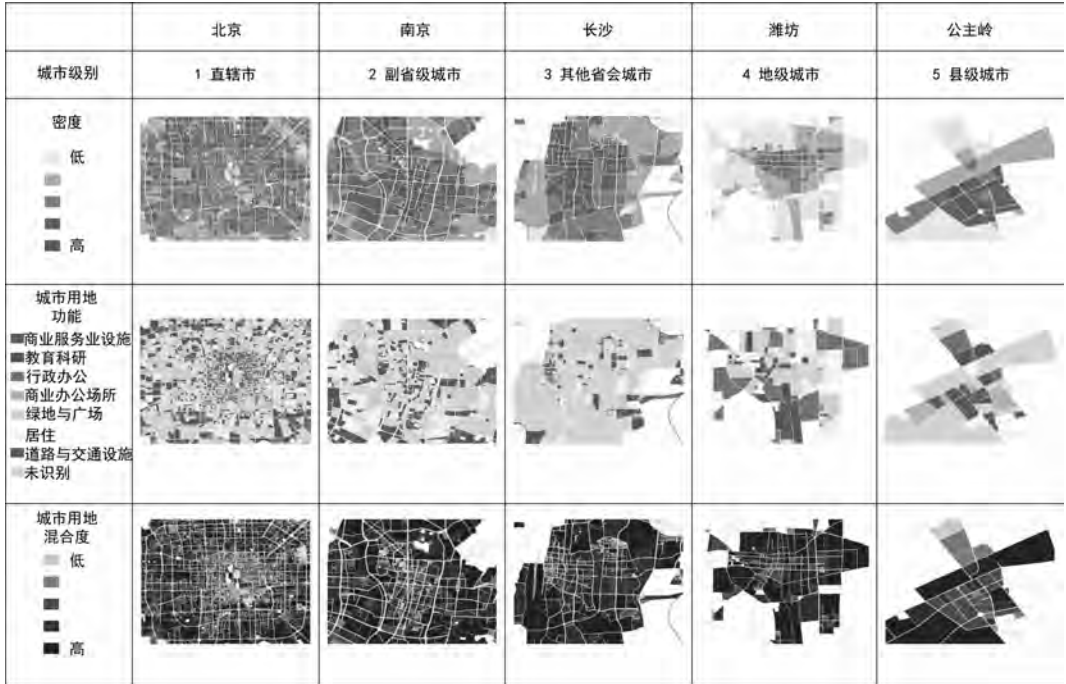


图5 利用 OSM 和 POIs 推导用地功能、密度和混合性

资料来源：刘行健和龙瀛 (Liu and Long, 2015)。

ArcGIS 中“aggregate polygon”功能对筛选出的城镇地块进行合并，从而提取出城镇建设用地范围。该模型在对城市地块赋值时同时考虑了地块的内生属性和邻近地块的属性，过程快捷、直接、准确，数据结果与传统方法相比具有高精度的特征。其与其他方法的对比详见图 6。

2.5 功能性城市地域识别

城市的行政边界往往不能真正地反映出城市的实际大小和劳动就业及经济活动的实际情况。然而，长久以来中国尚没有公认的功能性城市地域 (Functional Urban Area, FUA) 的划定标准，此外，以往研究多以单一区域或城市作为研究对象，并没有针对中国整个城市系统的探讨。为此，笔者及其合作者将 2010 年中国乡镇街道办事处尺度的人口密度数据 (详见第 2.1 节) 用于识别城市的内核，结合 2014 年全国的公交线路和站点资料来反映研究单元之间的联系，最终得到全国地级及以上城市的功能性城市地域，其中又细分为城市边缘区和城市中心区 (图 7)。该研究是首次尝试对全国范围内的城市功能地域予以识别和分析。

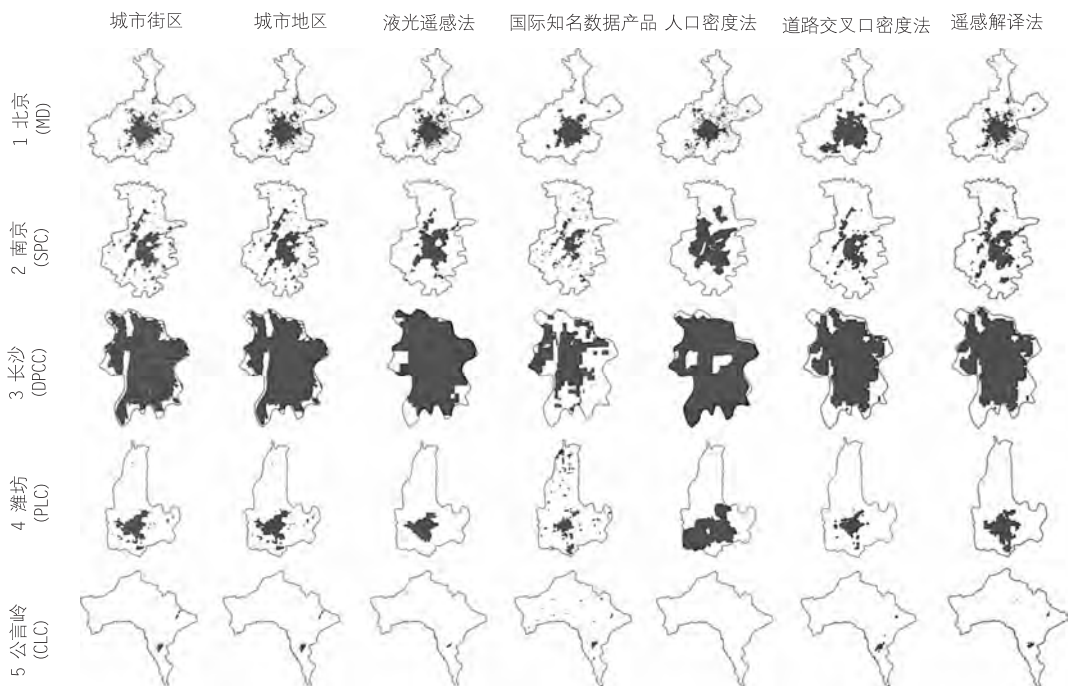


图6 不同方法对中国五座城市城镇建设用地的界定结果
资料来源：龙瀛、沈尧等 (Long, Shen et al., 2016)。

2.6 “徐霞客计划”

和城市有关的数据正日益增多，且涉及人们生活的各个方面。城市研究者们正在尝试从非正统城市数据中加深对我们城市的理解。但由于很多数据过于精确，往往涉及隐私和数据持有者的核心利益，为此，龙瀛等发起了一种城市数据共享平台“徐霞客计划” (SinoGrids, <http://www.beijingcitylab.com/projects-1/14-sinogrids/>) (Zhou and Long, 2016)。该计划相比中国已有的共享数据计划，更倾向于反映社会空间的数据，致力于将散布在各个研究者手中的全国微观数据汇总到 1km 网格尺度。尺度上来说，1km 是既能够开展城市间区域分析也能够进行城市内部研究的尺度。该计划为潜在的数据贡献者提供了指南和工具，协助其将大规模的微观数据汇总到 SinoGrids 平台，进而形成一个“众包”的中国大数据和开放数据平台。目前，“徐霞客计划”已经征集到多个图层，涵盖人口密度、道路交叉口数量、公共设施数、Flickr 照片数、微博签到数等。

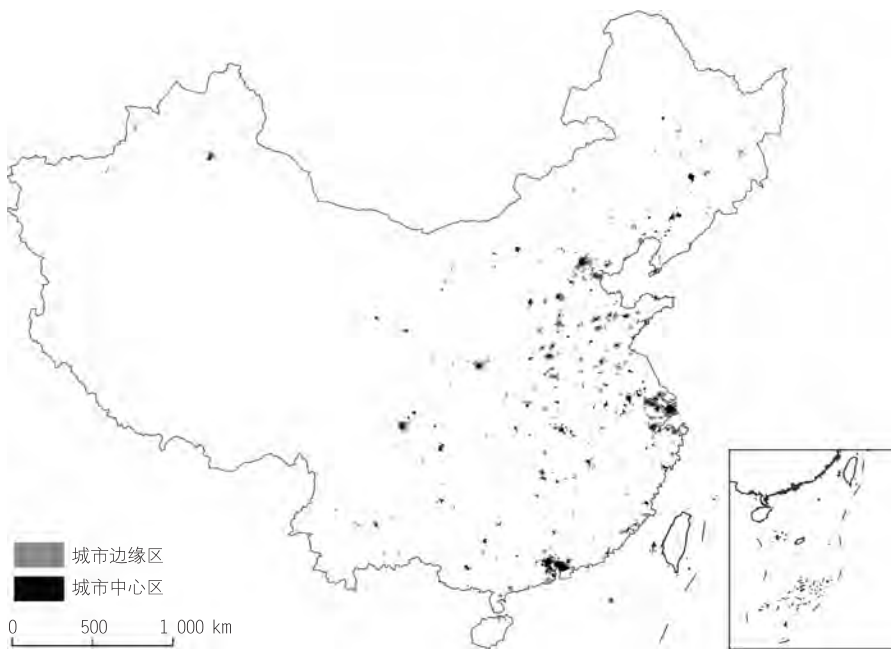


图7 全国地级及以上城市的功能性城市地域识别结果

3 城市空间开发

3.1 城镇化空间格局

针对我国城镇化空间格局的界定问题，周一星和史育龙（1995）率先提出了建立中国城市实体地域的概念，包括城市统计区、城镇统计区和城镇型居民区。为了进一步确定我国的城镇化空间格局，毛其智等（2015）根据前文提到的数据和方法，首先将城镇化地区的门槛密度定为 $1\ 000\ \text{人}/\text{km}^2$ （2010年全国人口密度为 $977\ \text{人}/\text{km}^2$ ）；其次，采用周一星和史育龙（1995）提出的 $2\ 000\ \text{人}/\text{km}^2$ 的平均密度标准划定城市统计区；最后，参考日本的人口集中地区（Densely Inhabited District, DID）概念，作为我国高密度城镇化地区（ $4\ 000\ \text{人}/\text{km}^2$ ）的识别依据。通过如上的人口密度界定，识别了三种不同密度下的中国微观尺度上的城镇化空间格局。

3.2 城市扩张与再开发

在中国过去 30 余年的快速城镇化过程中，空间扩张是城市开发的主要表现形式。龙瀛（2015）根据 1980 年代末和 2010 年中国的城镇建设用地遥感影像解译数据，分析出城镇建设用地扩张的结果，

即 1980~2010 年中国 280 余个地级以上城市的城镇建设用地面积发生了扩张。此外,通过中国 2000~2013 年的土地出让数据(共 34 169 km²)发现,在过去十余年间,存量开发在我国不同规模的城市开发比例中仅占 18%~35%(总计平均为 24%),而且多分布于主要城市的中心城区。长时间(2000~2013 年)和大规模(全国)的历史数据表明,存量开发短时间内难以成为中国城市开发的主体形式。

由于中国不同城市所处的发展阶段差异较大,中西部大量城市和东部欠发达的中小城市依然处于工业化与城市化快速扩张的中期阶段,存量和增量开发成本差距很大。因此,存量开发不应当“一刀切”地成为所有城市开发的主体形式。当前,存量规划、用地零增长似乎更适用于超大型城市,对于广大仍处于快速城市化阶段的中小城市,存量和增量开发需要根据具体城市的社会经济水平采用均衡发展策略。

3.3 自然城市视角下的城市扩张

已有的城市扩张研究的主要数据源是统计局数据与遥感影像解译数据,如前所述,这些数据不适用于大地域范围的精细化尺度的频繁检测。龙瀛等借鉴瑞典地理学家江斌所提出的自然城市(Natural Cities)的概念,对侧重物理空间开发的城市扩张概念进行了扩展,利用 2009 年和 2014 年覆盖全国的夜光遥感、道路交叉口、兴趣点和位置微博数据,分别从物理、形态、功能和社会四个维度分析了中国城市扩张。结果显示,在全国范围内,从自然城市面积和发育成熟度两个衡量标准来看,四个维度的扩张是不均衡的,程度依次递减,即 2014 年城市扩张区域相比 2009 年城市地域,对应着较大地块的物理空间开发、偏低的城市功能发展与极低的人类活动强度增加,且这一现象在全国所有城市群内均存在。片面地追求城市发展规模而忽视当前城市客观发展规律,已经造成大量城市功能和活动严重落后于城市物理空间的开发。

3.4 地块尺度的城市扩张模拟

囿于数据和技术,传统的大尺度城市增长模拟多对应公里尺度的网格单元,大尺度精细化的城市增长模拟还较为少见。为此,龙瀛、吴康等(Long, Wu et al., 2014)建立了针对覆盖全国所有城市的地块尺度的城市增长模型(Mega-Vector-Parcels Cellular Automata Model, MVP-CA),对全国 654 个城市 2012~2017 年的城市空间增长过程进行地块尺度的模拟。该模型包括三个宏观模块、地块生成模块和矢量 CA 模块。其中宏观模块基于各个城市的历史阶段城市增长信息以及国家空间发展战略,对未来五年内城市增长的速度进行情景分析(基准发展情景的模拟结果见图 8);地块生成模块则直接利用全国的真实路网进行划定,后续具有拓展地块细分(parcel/block subdivision)的功能;矢量 CA 模块是 MVP-CA 模型的核心模块,它在上两个模块的基础上,针对每个划定的地块单元,结合全国兴趣点数据,考虑每个地块的大小、紧凑度、区位特征、功能密度等属性,使用约束性矢量 CA 方法对未来的地块尺度的城镇开发进行模拟。

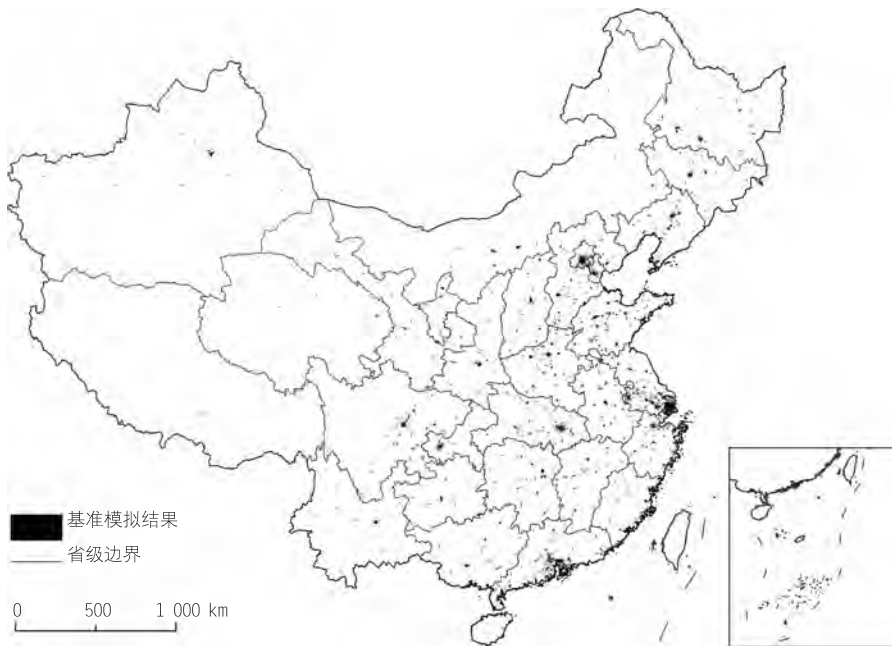


图8 全国所有城市基准发展情景下的城市增长模拟结果

资料来源：龙瀛、吴康等（Long, Wu et al., 2014）。

3.5 “鬼城”识别

中国“鬼城”不乏媒体报道，但系统梳理中国“鬼城”情况的研究仍然比较有限。龙瀛和吴康（2016）对此进行了初步探讨。该研究利用某大型互联网公司匿名的反映用户活动的大数据，对每个城市2000年以前和以后的城镇建设用地分别进行人类活动强度评价。他们认为，当城市新开发地区的人类活动强度偏低且与老城区差异显著的时候，即可以视之为“鬼城”。研究发现，在485个数据较为全面的城市中，有389个城市新区的人类活动强度显著低于老区。在地级及以上城市中，中国排名前20位的“鬼城”主要分布在东北和山东（图9）。而新区人类活动强度高于老区的96个城市，多为县级市、较小规模或新设立的地级市。这些城市的新区相比老区具有较高的人类活动强度，一方面源于新区完善的开发和功能完善；另一方面也源于老区人类活动强度不高。总体来看，城市的行政级别越高，新、旧区的人类活动差别越大；越是低等级的中小城市（尤其是年轻的小城市），不存在明显的城市集聚中心，新、旧中心人类活动强度差异越小。

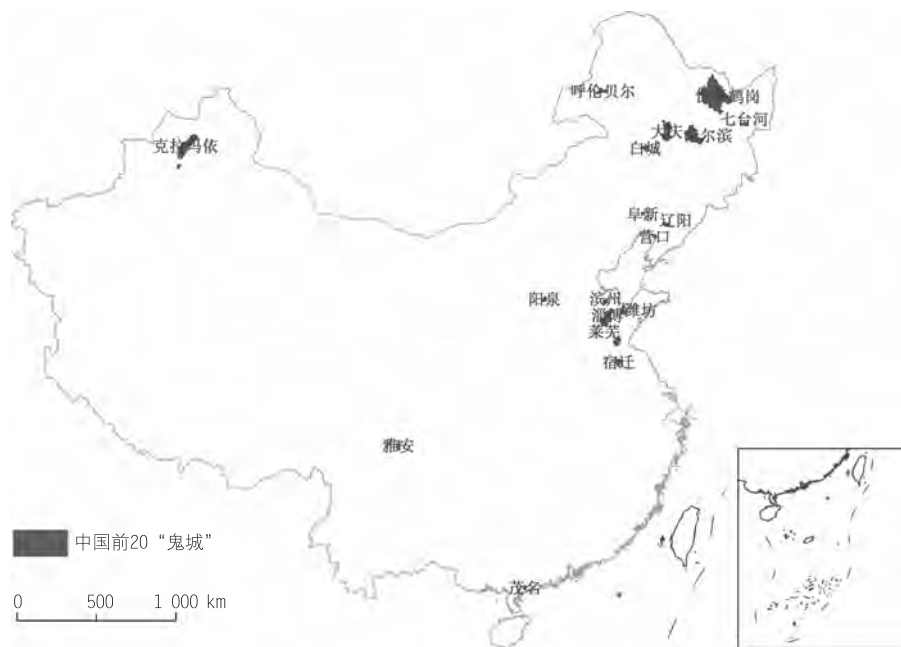


图9 中国排名前20位的鬼城分布

资料来源：龙瀛和吴康（2016）。

4 城市空间结构

4.1 多中心城市空间结构

城市空间结构从根本上决定了一个城市的特性，中国大量的城市在城市规划中都明确指出要打造多中心的城市空间结构。在快速城市化的背景下，以往关于城市空间结构的研究都是基于传统静态数据，例如人口普查、家庭出行调查等，都是5~10年才进行一次更新的数据，落后于中国的快速城市化过程。此外，已有研究也多针对一个城市开展，还没有针对所有城市进行全面分析的研究。为此，笔者及其合作者利用传统数据和新数据，在乡镇街道办事处尺度从多中心角度系统考察了中国所有城市的城市空间结构。针对多中心城市，识别出城市副中心的具体位置以及副中心之间和主中心与次中心之间的联系。此研究主要是通过人类活动密度来确定城市的多中心分布，并评价现实多中心与规划多中心的偏差。研究结果显示，在选取的284个城市中，70个城市具有明显的副中心，其中一半城市仅具有一个副中心（图10）。

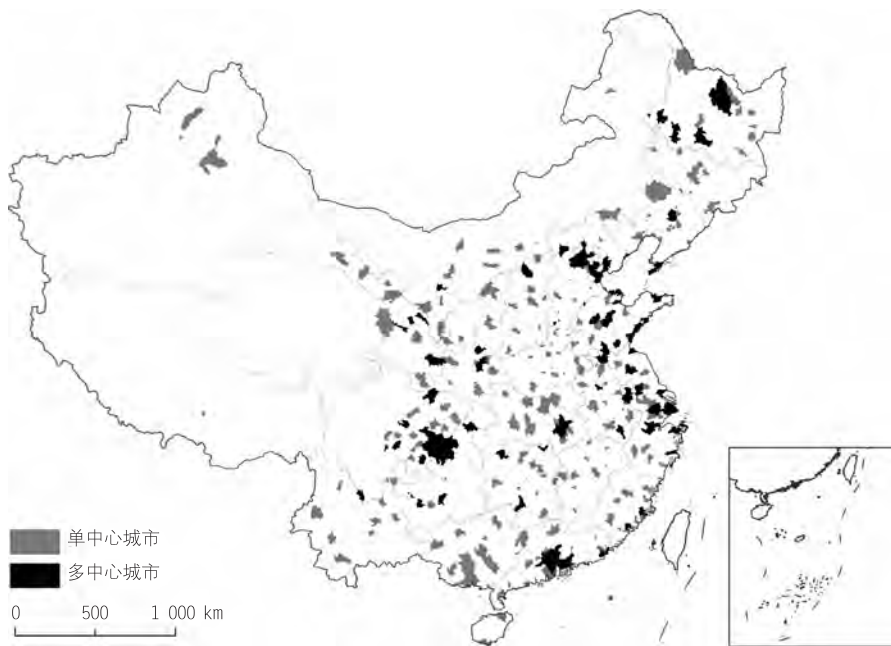


图 10 全国城市的多中心城市空间结构评价

4.2 城市形态评价

中国的快速城镇化伴随着显著的土地利用变化，在其过程中驱动和影响城市空间分散与集聚的因素，在最近几年得到了广泛的关注。由于其自上而下和自下而上的双重性质，中国城市呈现出更加复杂和多样的城市土地利用格局。笔者及其合作者利用 2011 年全国道路网和兴趣点数据，基于刘行健和龙瀛 (Liu and Long, 2015) 共享的城市地块及其城镇土地使用性质数据，结合 ArcGIS、Python、MatLab 和 SPSS 等分析工具，对中国 60 个具有代表性的内地城市及香港进行了城市用地结构的评价。此研究提供了一种有效的新型综合城市空间分析和评价方法，即通过采用多种测度以城市用地结构 (Urban Land Use Pattern) 为系统分析的核心考察对象，更好地揭示了中国在城镇化进程中复杂的城市空间特性，提高了我们对城市用地结构的理解，并且有助于城乡规划与治理的变革。研究者在空间熵和相异性两个基本指标基础之上，将城镇土地使用模式分成三综合类 (住宅、商业和公共)，结合元胞自动机 (CA) 建模，系统地评估了样本城市的城市扩张情况和用地混合程度 (图 11)。结果表明，城市不仅呈现出独特的空间碎片差异，并且在迅速扩张的同时，变得较为分散。大中小城市各自的扩张动力来源于不同的主导土地开发类别和地方优先政策，并且城市形态的形成机制跟城市的人口规模和城市面积大小有密切关系，而其中政府仍然对城市形态的形成有着重要的影响。

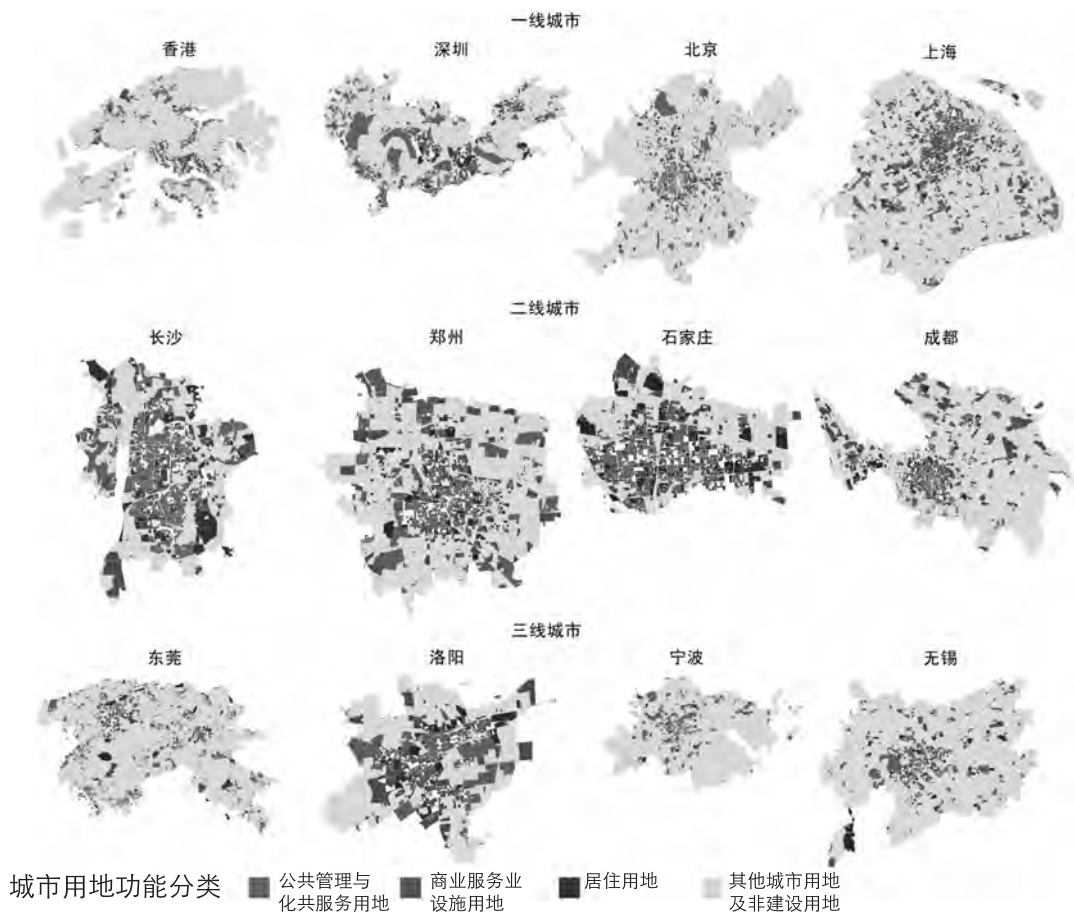


图 11 中国主要城市用地结构

4.3 公交站点服务范围及空间特征评价

为了找到中国城市公交服务的一般模式和规律，最终揭示中国城市系统的空间发展活动规律，李苗裔和龙瀛（2015）基于全国 313 个主要城市的 867 263 个公交站点数据，以 300m、500m 和 800m 服务半径计算出了每个城市城镇建设用地区域内公交站点服务的覆盖率（城市的公交站点覆盖范围和城镇建设用地面积之比），探讨了其空间特征，并进行了城市间的横向比较。其中，全国 281 个地级及以上城市的 500m 范围公交站点覆盖率的平均值为 64.4%。进一步，基于公交站点覆盖的空间特征，该研究将 313 个城市聚合为五类，同时利用 Flickr 照片、位置微博和兴趣点数据，对公交站点 500m 服务范围内的人的活动及设施情况进行分析。结果显示，尽管仅有 75.6% 的城镇建设用地范围有公交

站点服务覆盖,但该范围内包括了94.4%的设施和超过92%的人类活动,部分没有公交站点服务的地区(24.4%),其设施配套和人类活动水平都较低(10%以下)。可以看出,我国城市公交站点布局,基本满足了大多数人的活动需要和设施需求(图12)。

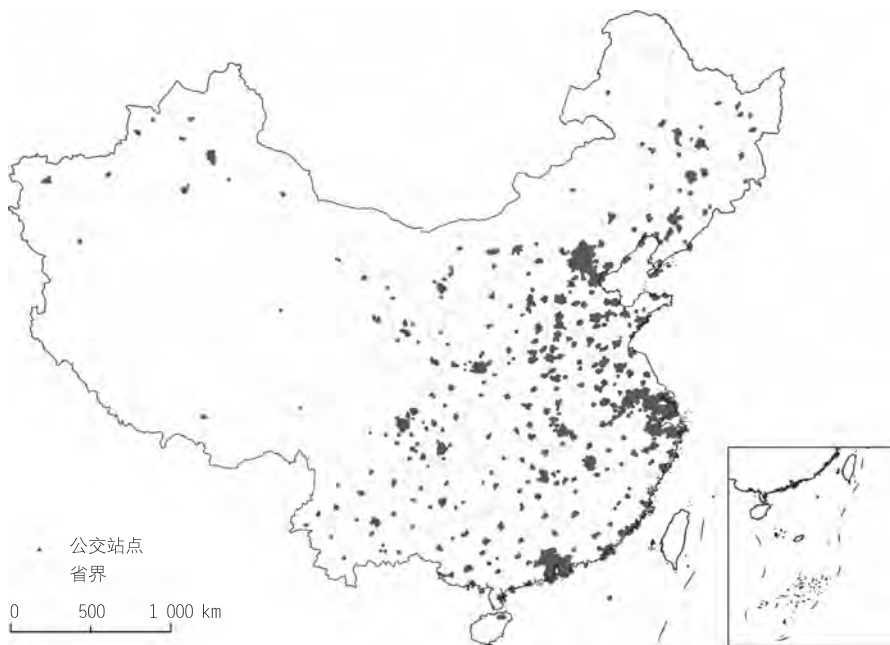


图12 中国城市公交站点分布
资料来源:李苗裔和龙瀛(2015)。

5 生态环境系统分析

5.1 PM_{2.5} 污染与人口暴露评价

PM_{2.5} 污染已成为我国亟须解决的任务,对其进行系统的分析是解决问题的关键一步。为此,龙瀛、王江浩等(Long, Wang et al., 2014)收集了2013年4月8日至2014年4月7日每日的PM_{2.5} 浓度值,采集范围覆盖了全国190个城市共计945个空气质量监测点。除了地面监测站数据外,还利用中分辨率成像光谱仪大气气溶胶厚度(MODIS AOD)数据对PM_{2.5} 进行插值补充,以弥补部分地区监测站稀疏的问题。另外,结合2010年乡镇街道办事处尺度的人口数据(详见第2.1节),评价了PM_{2.5} 污染的人口暴露风险。总的来看,人口密度越大,全年暴露天数越多,暴露强度越大。研究还发现,654个城市中,25个城市空气质量达标,仅占3.8%。654个城市的平均达标天数比例为

70.96%。全国 8.27 亿人口所生活的地区，一年内 PM_{2.5} 超过国家标准 ($75\mu\text{g}/\text{m}^3$) 的时间为 3 个月，其中 2.23 亿人口所居住区域 PM_{2.5} 超标半年，对应的国土面积为 34.8 万 km^2 。

5.2 城市形态对 PM_{2.5} 的影响识别

继 PM_{2.5} 污染与人口暴露评价，笔者及其合作者从人口、经济、用地、交通、气候、其他污染物等方面，分别针对全部城市选择了 11 个变量，针对地级市选择了 18 个变量，分析和揭示了城市形态对 PM_{2.5} 的影响。该研究利用第 5.1 节获得的 PM_{2.5} 数据以及《中国城市统计年鉴 2013》，采用分层线性回归模型 (Hierarchical Linear Models, HLM)，分两个层次逐步讨论了城市形态对中国所有城市 PM_{2.5} 的影响。结果显示，大城市和特大城市的 PM_{2.5} 年均值更高，全局上 PM_{2.5} 集聚分布趋势不显著，而局部聚集现象显著。城市自身因素可能发挥着更重要的影响，然而绿地比例高的城市不一定就意味着较高的空气质量，公共交通服务好的城市空气质量较好。但与西方研究结论不同，建成区人口密度对 PM₁₀ 和 PM_{2.5} 却有着显著的正向影响，即人口密度高可能会导致空气质量的恶化。因此，对特大城市、大城市来说，应该适当疏解城市功能，避免城市密度过高。

5.3 垃圾填埋场的恶臭影响评价

在中国，垃圾填埋场恶臭影响是形成邻避效应 (Not in My Back Yard, NIMBY) 的重要原因之一。为了揭示垃圾填埋场的邻避效应并提出相应的规划对策，蔡博峰等 (Cai et al., 2015) 基于全国 1 955 个垃圾填埋场 (不包括台湾、香港和澳门)，利用 FOD 模型计算每个垃圾填埋场的恶臭气体排放量，之后利用点源连续高斯模型作为恶臭气体扩散模型，针对每个垃圾填埋场逐一计算其恶臭排放和扩散范围 (图 13)。然后根据垃圾填埋场的影响范围，利用高空间分辨率人口密度、兴趣点和位置微博等数据，评估垃圾填埋场恶臭影响的人口、敏感单位和人群活动。研究发现，全国垃圾填埋场恶臭影响的人口达到 1 228 万人，其中受影响的敏感人群 (儿童+老人) 达到 264 万人，受影响的敏感单位 (学校和医院) 达到 7 818 个，受影响的人群活动占全国总人群活动的 1.82%。

5.4 重污染水体识别

在我国的实际工作中，由于对重污染水体和黑臭水体的判别缺乏明确标准与识别手段，相关规定难以操作和有效落实，这些都给计划实施和监督考核造成了严重障碍。为此，石峰和龙瀛 (2015) 以问题为导向，首次尝试采用互联网开放信息大数据，即对互联网媒体曝光最多、群众投诉议论最多的污染水体进行数据搜索和统计分析，找出最受关注的污染水体，得到全国重污染水体和黑臭水体的总体分布情况，直接反映民意诉求。首先，研究选取了全国 1 461 条河流进行自动检索，在统计与分析后，得出河流污染与黑臭问题的民众和媒体关注度数据；然后，利用全国三、四级河流分布数据和电子地图信息，选取河流名称、河流位置及“污染、水污染、重污染、黑臭、水质恶化”等关键字段；

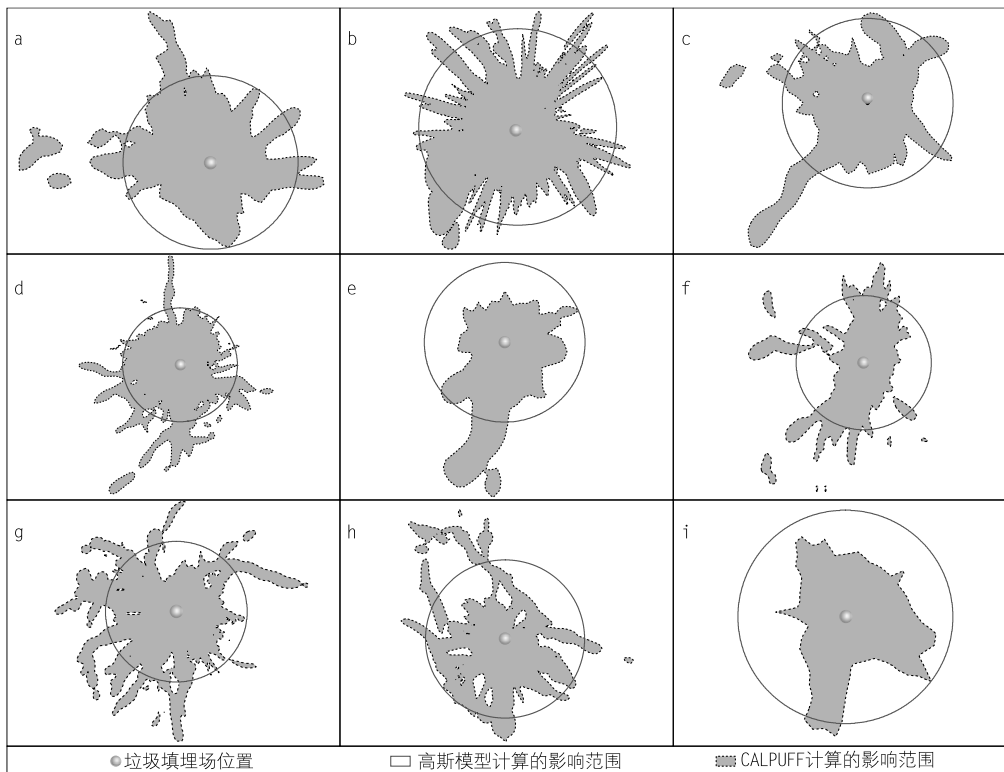


图 13 CALPUFF 模型和地面点源连续高斯模型影响范围比较

注：a-北京六里屯垃圾填埋场；b-上海老港生活垃圾处置场；c-南昌市麦园垃圾处理场；d-广州兴丰生活垃圾卫生填埋场；e-深圳市下坪固体废弃物填埋场；f-重庆长生桥垃圾卫生填埋场；g-成都市固体废弃物卫生处置场；h-西安江村沟垃圾填埋场；i-西宁沈家沟垃圾填埋场。

最后，通过基于百度搜索引擎的大数据分析，锁定如河北漕河等具有一定代表性的重污染河流，并对全国 1 400 余条河流按照网上受关注程度进行了分级。此方法可以不受监测条件、布点方案的限制，作为现行监测系统的有益补充。

6 城市规划与设计响应

6.1 城市增长边界评价

城市增长边界 (Urban Growth Boundaries, UGBs) 是我国城市规划编制的核心成果，也是规划实施评价关注的重点内容。我国城市规划中的规划城镇建设用地区域，通常被认为是中国的城市增长边界 (龙瀛等, 2009; Long, Gu et al., 2012; Long, Han et al., 2013)。我国已有的城市增长边界

实施效果评价工作，多从评价物质空间开发与城市增长边界的一致性 (conformity) 入手，且多针对一个城市，少有研究对中国各个城市的城市增长边界实施成效进行系统评价。为此，龙瀛及其合作者搜集了全国超过 200 个城市的正在实施的城市总体规划图，从中提取了城镇建设用地范围即规划 UGBs，并将其数字化为 GIS 图层。之后，将各个城市的 UGBs 与利用遥感观测到的城市扩张进行叠加分析，最后计算各个城市开发的合法率，这一正在开展中的研究为城市间的横向对比提供了可能。除了物质空间角度的评价外，还可以从社会空间维度进行评价。例如，龙瀛、韩昊英等 (Long, Han et al., 2015) 利用包括公交地铁刷卡数据、出租车轨迹、位置微博和照片在内的多源新数据，详细评价了北京 UGBs 的实施效果，结果显示，虽然有大量的非正式开发分布在 UGBs 之外，但 UGBs 所包含的区域内则容纳了 95% 以上的城市活动和移动轨迹。

6.2 收缩城市及其规划设计对策

伴随着快速城镇化，我国局部城市出现了人口收缩现象，阻碍了城镇化的健康发展和资源的有效利用。以此为出发点，龙瀛和吴康 (Long and Wu, 2016) 利用五普 (2000 年) 和六普 (2010 年) 中的乡镇街道办事处尺度人口数据，对全国 654 个城市的人口变化进行了分析，其中地级及以上城市采用 2012 年市辖区范围，县级市由于数据可获得性等原因，采用 2012 年市域范围。分析结果显示，中

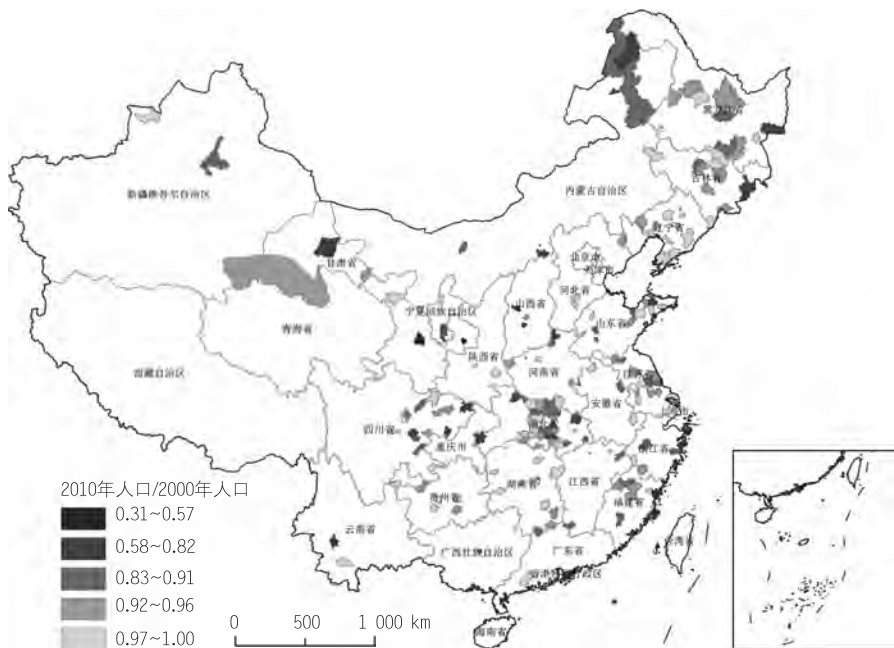


图 14 中国收缩城市分布

资料来源：龙瀛、吴康等 (2015)。

国 180 个城市发生了人口总量/密度的下降,即存在着 180 个收缩城市(图 14)。龙瀛、吴康等(2015)还进行了收缩城市的分类和影响因素分析,并针对中国收缩城市自身的特点,提出初步的对策建议和规划设计应对,以及中国收缩城市的深度探测、典型收缩城市研究、规划应对手段以及研究网络的发展四个方面构成的中国收缩城市研究的研究框架。

6.3 数据增强设计

龙瀛和沈尧(2015)提出的数据增强设计(Data Augmented Design, DAD)是以定量城市分析为驱动,通过数据分析、建模、预测等手段,为规划设计的全过程提供调研、分析、方案设计、评价、追踪等支持工具,以数据实证提高设计的科学性,并激发规划设计人员的创造力。DAD 利用简单直接的方法,充分整合新旧数据源,强化规划设计中方案生成或评估的某个环节,易于推广到大量场地,同时兼顾场地的独特性。DAD 属于继计算机辅助设计(Computer Aided Design, CAD)、地理信息系统(Geographical Information System, GIS)和规划支持系统(Planning Support System, PSS)之后的一种新的规划设计支持形式。DAD 实际增强的是对城市实体的精确理解,对实体组织和其效应间复杂关系的准确把握以及对空间创造积极影响的切实落实。

6.4 街道城市主义

龙瀛所倡导的街道城市主义(Street Urbanism)是在认识论层面上认识城市的一种方式,在方法论上是建立以街道为个体的城市空间分析、统计和模拟的框架体系。在现有的数据增强设计的框架下,街道城市主义将吸收已有设计师、评论家和学者对街道的思考与认识,并结合成熟理论将其成果用于设计实践。在过去的若干年里受到数据和城市发展阶段的限制,地块多是城市研究的核心,日益成为城市研究的基本单元和日常规划的管理对象,而作为骨骼起到支撑作用的街道遭受了大量的忽视,相对而言更多的关注是来自于设计师和社会观察家的关注(偏向于定性的认知)。在新数据环境下,新型城镇化提出以人为本的城镇化,城市管理和规划走向精细化、智慧化,中国的部分城市还发生城市收缩现象,以及城市生活空间从地块转向城市的种种现状,引导着城市研究者开始关注街道视角的城市研究,这就是提出街道城市主义的初衷。同时,街道城市主义并不否定地块主义的作用,而是希望街道能真正起到骨骼的支撑作用,连接作为肌肉的地块与城市,使城市迸发出真正的活力。

7 新数据环境下的中国城市研究述评

7.1 研究对象应回归客观的城市系统

长期以来,中国对于“城市”的界定一直存在着行政地域(城市管辖权对应的空间范围)和实体

地域（城市建成区范围）的“二元性”割裂，造成多数城市研究的对象不是城市，而是区域；部分针对中国城市系统的研究，针对的对象不够全面，例如中国目前除了官方认可的 653 个不同等级的城市外（2014 年口径），大量的县城和发育完善的镇如按照国际惯例亦属于城市。为此，龙瀛利用道路交叉口数据重新定义了中国的城市系统。研究显示，如果以 100 个交叉口作为最小的城市门槛，则中国有 4 629 个城市，其中 3 340 个位于现有城市的市区边界之外（图 15）。这些被忽略城市的快速扩张、人口收缩与空置现象，因为游离于决策者、学者和统计资料的视野之外，更加值得关注。通过重新定义城市和构建中国城市系统，有望更加客观地认识中国的城市系统，引导城市研究工作关注更为接近实际的城市系统。

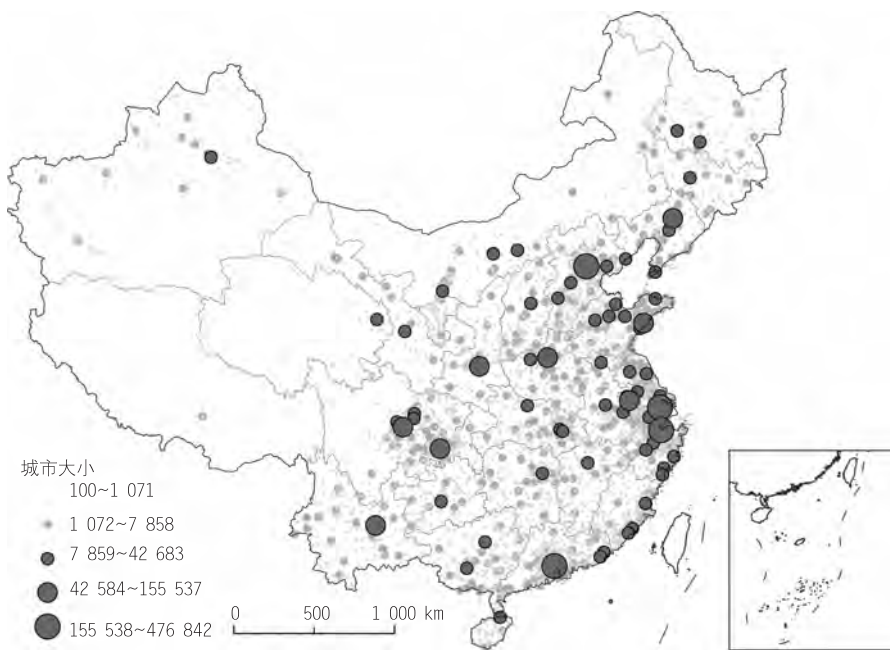


图 15 重新定义的中国城市系统

注：城市大小为道路交叉口数量。

资料来源：龙瀛和吴康（2016）。

7.2 研究方式和方法正经历四方面变革

新数据环境为城市研究带来了新的机遇，除了在数量上的更好保障研究外，在本质上也在促进城市研究发生变革。龙瀛和刘伦（2016）、龙瀛和刘伦（Long and Liu, 2015）基于龙瀛研究团队所开展的大量定量城市研究工作，首先探讨了大数据和开放数据形成的新数据环境及国内外定量城市研究概况，然后围绕典型案例重点对当前定量城市研究的四项变革及相关实践展开讨论，最后提出相关思

考。该研究认为，新数据环境推动了定量城市研究的四大变革：①空间尺度上由小范围高精度、大范围低精度到大范围高精度的变革（即本文所主要探讨的“大模型”）；②时间尺度上由静态截面到动态连续的变革；③研究粒度上由“以地为本”到“以人为本”的变革；④研究方法上由单一团队到开源众包的变革。在变革的同时，当前定量城市研究也面临着数据有偏、多现状研究少远景判断、多客观认识少规划启示以及规划理论和学科发展等相关问题。

7.3 研究发现体现为四方面中国城市化悖论

在经历了 30 多年的经济高速增长和土地快速开发后，中国经济步入了“新常态”，并确立了“新型城镇化”的推进战略。然而通过分析覆盖整个中国的既涵盖物理空间也囊括社会活动的精细化数据，龙瀛和吴康（2016）发现中国的城市化在快速发展过程中表现出不容忽视的四个方面的悖论，即快速空间扩张与过度呼吁存量规划之间的悖论、局部人口收缩与规划人口膨胀之间的悖论、中高强度建筑开发与低密度人类活动之间的悖论、行政地域的城市和作为实体地域的城市之间的悖论。针对中国的城市化，客观地认知这些悖论，并对面向存量的城乡规划法和规划编制办法进行深入研究，对中国的城市建设与发展至关重要。在下一个阶段的城市化过程中，如果不注意这些存在的问题，中国的诸多宏伟的城市群战略、城市规划策略（如存量规划）以及重大基础设施投资，可能会面临不能实现或浪费的局面。

7.4 研究目标应回归提高居民生活质量

李克强总理在总结关于国务院 2015 年《中华人民共和国国民经济和社会发展第十三个五年规划纲要》工作报告中强调，城市规划采用大数据的根本目的是提高城市居民生活质量。在过去，多数社会科学研究依赖于问卷调查，而如今利用新数据来体现“人文关怀”是我们同样重点关注的方向。例如，面向决策者，使其能够制定合理的决策；面向开发商，使其建设好的城市开发项目；面向规划师、设计师，使其能制订好的可以实施的规划设计方案；个别研究则是直接针对居民，围绕间接或是直接的以提高居民生活质量而展开的。例如，龙瀛、孙立君等（2015）介绍了其利用公共交通刷卡数据所开展的大量研究，如通过识别居住地特征和就业行为特征，进而优化线路、定制公交选线，提高公交通勤舒适性等，以及通过关注个体出行行为，以此来提高公共交通系统的运行效率和人性化服务水平，提高特定人群的生活质量，如学生、城市贫困者、极端出行乘客等。在 PM_{2.5} 研究中，通过对污染的时间及空间分布特征分析和相应人口暴露的评估，为公众和政府决策提供了充分的依据。

8 结语

在国家不断着眼提升新型城镇化、提高基本公共服务均等化水平、促进社会公平的同时，国务院

于2015年11月针对《中华人民共和国国民经济和社会发展第十三个五年规划纲要》再次强调要开门编规划,利用好互联网、大数据等技术平台,广泛听取各方意见,真正做到聚众智编制规划,最终达到提高人们生活质量的目標。伴随着新数据环境的逐渐发展成型,定量城市研究的新范式使得城市研究与规划实践的结合由宏观总体规划层面扩展至中、微观的详细规划和城市设计层面。在传统的城市研究中,城市体系中的城市大多抽象为“点”,侧重考查城市间的相互作用和联系。而大范围精细尺度上的大样本量城市分析,不再是孤立的研究各个城市,除了考虑城市内部的发展动态,还关注城市间的联系“网络”。

本文所探讨的“大模型”范式指导下的中国城市系统化研究的诸多案例中,有老问题新探索,也有新问题新探索。新数据环境的形成,使得这些覆盖全国范围细粒度的人居环境多维度探索在如今成为可能。国际上也正涌现出越来越多的覆盖多个城市的精细化定量城市研究工作,考虑到中国未来信息通信技术的大力发展和对城市开发建设品质追求的日益提升,“大模型”研究范式将在中国城市系统化研究中起到更大的作用,它将对中国城乡规划科学化起到积极提升的作用,也有望推动我国人居环境科学的大力发展。

致谢

本文所介绍的案例来自多位合作者,在此表示感谢。

注释

① 关于城市模型的讨论,详见刘伦等(2014)。

参考文献

- [1] Bagchi, M., White, P. 2004. What Role for Smart-card Data from Bus Systems? *Municipal Engineer*, Vol. 157, No. 1.
- [2] Cai, B., Wang, J., Long, Y. et al. 2015. Evaluating the Impact of Odors from the 1955 Landfills in China Using a Bottom-up Approach. *Journal of Environmental Management*, Vol. 164, No. 12.
- [3] Liu, X., Long, Y. 2015. Automated Identification and Characterization of Parcels (AICP) with OpenStreetMap and Points of Interest. *Environment and Planning B: Planning & Design*, In press.
- [4] Liu, X., Song, Y., Wu, K. et al. 2015. Understanding Urban China with Open Data. *Cities*, Vol. 47, No. 9.
- [5] Long, Y., Gu, Y., Han, H. 2012. Spatiotemporal Heterogeneity of Urban Planning Implementation Effectiveness: Evidence from Five Master Plans of Beijing. *Landscape and Urban Planning*, Vol. 108, No. 2.
- [6] Long, Y., Han, H., Lai, S. et al. 2013. Urban Growth Boundaries of the Beijing Metropolitan Area: Comparison of Simulation and Artwork. *Cities*, Vol. 31, No. 4.
- [7] Long, Y., Han, H., Tu, Y. et al. 2015. Evaluating the Effectiveness of Urban Growth Boundaries Using Human Mobility and Activity Records. *Cities*, Vol. 46, No. 8.
- [8] Long, Y., Liu, X. 2013. How Mixed is Beijing, China? A Visual Exploration of Mixed Land Use. *Environment & Plan-*

ning A, Vol. 45, No. 12.

- [9] Long, Y., Liu, L. 2015. Big/Open Data in Chinese Urban Studies and Planning: A Review. *ISOCARP Review*, Vol. 11.
- [10] Long, Y., Shen, Z. 2013. Disaggregating Heterogeneous Agent Attributes and Location from Aggregated Data, Small-scale Surveys and Empirical Researches. *Computers, Environment and Urban Systems*, Vol. 42, No. 11.
- [11] Long, Y., Shen, Z. 2015. Population Spatialization and Synthesis with Open Data. *Geospatial Analysis to Support Urban Planning in Beijing*. Springer International Publishing.
- [12] Long, Y., Shen, Y., Jin, X. 2016. Mapping Block-level Urban Areas for a Large Geographical Area. *Annals of the Association of American Geographers*, Vol. 106, No. 1.
- [13] Long, Y., Wang, J., Wu, K. et al. 2014. Population Exposure to Ambient PM_{2.5} at the Subdistrict Level in China. SSRN 2486602.
- [14] Long, Y., Wu, K. 2016. Shrinking Cities in a Rapidly Urbanizing China. *Environment and Planning A*, Vol. 48, No. 2.
- [15] Long, Y., Wu, K., Mao, Q. 2014. Simulating Urban Expansion in the Parcel Level for All Chinese Cities. *ArXiv Preprint*, 1402.3718.
- [16] Zhou, Y., Long, Y. 2016. SinoGrids: A Practice for Open Urban Data in China. *Cartography and Geographic Information Science*, in press.
- [17] 李苗裔、龙瀛：“中国主要城市公交站点服务范围及其空间特征评价”，《城市规划学刊》，2015年第6期。
- [18] 刘伦、龙瀛、麦克·巴蒂：“城市模型的回顾与展望——访谈麦克·巴蒂之后的新思考”，《城市规划》，2014年第8期。
- [19] 龙瀛：“城市大数据与定量城市研究”，《上海城市规划》，2014年第5期。
- [20] 龙瀛：“高度重视人口收缩对城市规划的挑战”，《探索与争鸣》，2015年第6期。
- [21] 龙瀛：“街道城市主义——新数据环境下城市研究与规划设计的新思路”，《时代建筑》，2016年第2期。
- [22] 龙瀛、韩昊英、毛其智：“利用约束性CA制定城市增长边界”，《地理学报》，2009年第8期。
- [23] 龙瀛、刘伦：“新数据环境下定量城市研究的四个变革”，《国际城市规划》，2016年（已接受，待刊登）。
- [24] 龙瀛、茅明睿、毛其智等：“大数据时代的精细化城市模拟：方法、数据和案例”，《人文地理》，2014年第3期。
- [25] 龙瀛、沈尧：“数据增强设计——新数据环境下的规划设计回应与改变”，《上海城市规划》，2015年第2期。
- [26] 龙瀛、沈振江、毛其智：“城市系统微观模拟中的个体数据获取新方法”，《地理学报》，2011年第3期。
- [27] 龙瀛、孙立君、陶遂：“基于公共交通智能卡数据的城市研究综述”，《城市规划学刊》，2015年第3期。
- [28] 龙瀛、吴康：“中国城市化的几个现实问题：空间扩张、人口收缩、低密度人类活动与城市范围界定”，《城市规划学刊》，2016年第2期。
- [29] 龙瀛、吴康、王江浩：“中国收缩城市及其研究框架”，《现代城市研究》，2015年第9期。
- [30] 龙瀛、吴康、王江浩等：“大模型：城市和区域研究的新范式”，《城市规划学刊》，2014年第6期。
- [31] 龙瀛、张宇、崔承印：“利用公交刷卡数据分析北京职住关系和通勤出行”，《地理学报》，2012年第10期。
- [32] 毛其智、龙瀛、吴康：“中国人口密度时空演变与城镇化空间格局初探——从2000年到2010年”，《城市规

划》，2015年第2期。

[33] 石峰、龙瀛：“基于互联网大数据技术的重污染水体识别研究”，http://mp.weixin.qq.com/s?__biz=MjM5ODI3ODQ3Ng==&mid=400514011&idx=1&sn=50eebd544de038c802a320de1eccc44&scene=1&srcid=0130iyAWkD95qS6fcxr6od0Z#wechat_redirect，2015年。

[34] 周一星、史育龙：“建立中国城市的实体地域概念”，《地理学报》，1995年第4期。