

城市建成环境评估单元数据集（2022）研发

张淑杰, 李萌*, 淡铂洋, 韩靖北, 郝灵强

中国城市规划设计研究院, 北京 100044

摘要: 城市建成区、核心建成片区是进行城市建成环境评估的基本单元, 其大小直接影响建成环境指标的高低。本文利用多源大数据优势, 制定统一的划定方法和技术流程, 研发了北京、上海等 19 个城市的建成区和核心建成片区数据集 (2022)。对于建成区, 综合考虑不透水面占比、路网密度、POI 密度和人口密度, 并参考高分辨率遥感影像, 形成建成区的划定指标体系和方法, 在 500 m×500 m 网格尺度上科学快速的划定城市建成区范围。针对核心建成片区, 通过识别城市中心、较高密度的街镇、校核城市主要功能区和设施以及扣除大型非建设用地的开敞空间 4 个步骤, 划定重点城市的核心建成片区范围。通过统一数据来源和划定方法, 形成城市建成环境评估研究中具有横向可比较性的 2 个基本空间范围。数据集存储为.shp 格式, 由 16 个数据文件组成, 数据量为 2.23 MB (压缩为 1 个文件, 1.37 MB)。

关键词: 建成区; 核心建成片区; 大数据; 城市建成环境

DOI: <https://doi.org/10.3974/geodp.2025.04.05>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2025.04.05>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志 (中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2025.03.09.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2025.03.09.V1>.

1 前言

城市建成区作为集中承载城市基础公共服务设施的空间, 对其特征的刻画一直备受学者的关注。建成区和核心建成片区是进行城市建成环境评估的常用空间范围单元, 通过计算该范围内人口密度、建设强度、商业密度、交通可达性、土地利用混合度等指标, 评估城市建成环境在社会经济发展和空间形态方面的特征, 能够为城市体检、城市更新等工作提供客观准确的现状认知, 进而推动城市实现高质量发展目标。

城市建成区是集中进行城市开发建设, 落实城市主要功能布局的范围, 也是进行城市建设和城市管理的重点区域, 是对城市总体建成环境特征认知的宏观层次。建成区内部人口密度、公共设施服务水平明显高于周边农村地区, 自然环境和土地利用方式也与农村地

收稿日期: 2025-05-07; 修订日期: 2025-11-05; 出版日期: 2025-12-24

基金项目: 中华人民共和国科学技术部 (2022YFC3802802)

*通讯作者: 李萌, 中国城市规划设计研究院, limeng_go_for_it@163.com

数据引用方式: [1] 张淑杰, 李萌, 淡铂洋等. 城市建成环境评估单元数据集 (2022) 研发[J]. 全球变化数据学报, 2025, 9(4): 405-418. <https://doi.org/10.3974/geodp.2025.04.05>. <https://cstr.escience.org.cn/CSTR:20146.14.2025.04.05>.

[2] 张淑杰, 李萌, 淡铂洋等. 基于多源数据的城市建成环境评估单元数据集 (2022) [J/DB/OL]. 全球变化数据仓储电子杂志, 2025. <https://doi.org/10.3974/geodb.2025.03.09.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2025.03.09.V1>.

区具有很大的差异。作为中观分析单元,城市核心建成片区指位于城市中心区域、具有较高人口密度与建设强度、容纳核心城市功能与设施的建成区域,有助于深入剖析建成环境特征(如人口密度及建设强度等)。宏观尺度的建成区和中观尺度的核心建成片区划定范围的大小,直接影响建成环境指标的高低。本文基于多源大数据,利用统一的划定方法和划定流程,研发了我国19个城市的建成区边界和核心建成片区边界。

传统划定建成区的方法具有较强的主观性,城市规划者或城市建设管理者根据实际管理的需要,依据城市实际建设情况、公共设施和公用设施的服务范围,综合社会经济活动与城市中心联系的密切程度,以主观与客观相结合的方式确定城市建成区范围。人口密度作为反映经济活动强弱程度的重要指示特征,一度作为划分城镇的重要指标。然而,能够公开获取人口数据的统计单元一般比较大,而且人口统计周期长,难以及时客观地反应城镇范围的动态变化。

遥感数据及有关的大数据的出现,为研究建成区边界划定提供了新的数据支撑和技术思路。目前基于大数据提取建成区边界主要有单一要素法和多要素综合法两种。单一要素法,即通过对一种判定要素直接设定阈值或者构建组合指数后再设定阈值的方法提取建成区边界,常用的是夜间灯光遥感数据和兴趣点(Point of Interest, POI)数据。夜间灯光遥感数据能够直观综合地反映人类在地面上的社会经济活动,其幅面宽、时间分辨率高、免费获取的优势为城市建成区边界提取研究提供了很好的数据基础^[1,2]。POI数据能够在空间上反映城市结构和道路骨架的分布特征,能够反映城市集聚效应和规模效应。其密度的空间分布差异反映了不同的区域发展水平,与城市建成区分布高度相关,通过设置POI密度阈值可以提取城市建成区边界^[3]。为克服单一要素只能反映建成区某一方面特征而造成提取偏差的问题,有学者提出将夜间灯光数据与POI、路网等大数据相结合,以提高建成区提取精度^[4,5]。但由于夜间灯光数据不可避免的灯光“过饱和”和“外溢”现象,同时其空间分辨率较低,在建成区提取过程中往往会遗漏范围相对较小、与中心区有一定距离的开发区、矿区、科教园区、公园等区域,提取精度具有很大的局限性。本文在遵循建成区定义的基础上,综合统计数据、遥感解译数据、电子导航路网数据、POI兴趣点数据、百度LBS(Location Based Services)数据等多维度的判别因子,提出一种在网格尺度上科学快速提取建成区边界的方法。

核心建成片区的概念接近于内城、城市中心区等概念,对其划定方法的研究更是多源。从划定的最小单元来看有街区、道路和行政边界,按照街区划分主要应用于建筑形体的规划管控^[6,7],道路划分主要用于对中心区域的划分^[8],行政边界由于具有边界稳定、客观性较强的优势,同时常作为人口、经济等要素的统计和管理单元,应用更为广泛。常青等以行政边界作为圈层依据,对北京和其他国际城市在各个圈层的尺度和密度进行对比研究^[9];王浩锋等以街镇行政单元边界为依据划定密度分布的基本单元^[10]。在划定因子方面,史北祥对各种界定标准进行了分析,包括空间肌理、路网密度、人口密度等,并提出“公共服务设施指数法”等具体测算方法^[11,12]。本文基于街道乡镇边界,综合利用在地活动人口密度、建设强度、商业中心与办公中心等主要功能区,以及城市级大型公共设施的空间分布等判定因子,划定19个城市的核心建成片区边界。

2 数据集元数据简介

《基于多源数据的城市建成环境评估单元数据集（2022）》^[13]的名称、作者、地理区域、数据年代、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表1。

表1 《基于多源数据的城市建成环境评估单元数据集（2022）》元数据简表

条 目	描 述
数据集名称	基于多源数据的城市建成环境评估单元数据集（2022）
数据集短名	BuiltUpUnits
作者信息	张淑杰，中国城市规划设计研究院，115417592@qq.com 李萌，中国城市规划设计研究院，limeng_go_for_it@163.com 淡铂洋，中国城市规划设计研究院，browndby@163.com 韩靖北，中国城市规划设计研究院，hanjingbei@163.com 郝灵强，中国城市规划设计研究院，1175513356@qq.com
地理区域	中国19个城市：北京、上海、广州、深圳、天津、重庆、济南、青岛、南京、杭州、郑州、武汉、长沙、成都、昆明、西安、哈尔滨、沈阳、大连
数据年代	2022年
数据格式	.shp
数据量	2.23 MB
数据集组成	19个城市建成区边界和核心建成片区边界数据
数据计算环境	ArcGIS
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲11号100101，中国科学院地理科学与资源研究所
数据共享政策	（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循10%引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的10%，同时需要对摘取的数据记录标注数据来源 ^[14] 。
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS, GEOSS, PubScholar, CKRSC

3 数据研发方法

本研究以建成区与核心建成片区的定义为基础，充分利用多源大数据优势，综合考虑城市人口、建筑、路网、POI、用地等要素，制定了统一的划定方法和技术流程，并生产了北京、上海等19个城市的建成区和核心建成片区数据集。具体数据来源见表2。

3.1 算法原理

3.1.1 建成区划定原理

根据《城市规划基本术语标准》^[17]的规定，城市建成区是指城市行政区内实际已成片开发建设、市政公用设施和公共设施基本具备的地区。针对定义中的“实际成片开发建设”

表2 数据来源说明

数据名称	数据类型	数据来源	年份
不透水面 ^[15]	栅格	参照 2022 年卫星影像数据 (获取自 BIGEMAP), 对清华大学地球系统科学系宫鹏教授团队生产的 2017 年 10 m 分辨率全球土地覆盖产品进行校正, 得到目标城市 2022 年的不透水面数据 (FROM-GLC10, 2017)	2022
路网 ¹	矢量 (线)	四维图新	2022
POI (Point of Interest) ¹	矢量 (点)	四维图新	2022
AOI (Area of Interest) ¹	矢量 (面)	四维图新	2022
人口数量 ²	矢量 (点)	LBS 人口数据	2022
建筑面 ^[16]	矢量 (面)	广东省城市化与地理环境空间模拟重点实验室刘小平教授团队生产的三维建筑数据集 (Building height of Asia in 3D-GloBFP, 2020)	2020
高分辨率遥感影像 ³	栅格	BIGEMAP	2022

“市政公用设施和公共设施基本具备”两个重要特征, 充分利用多源大数据进行表达。其中 (1) 对“已成片开发建设”区域的识别: 利用 10-m 分辨率遥感影像解译获得的不透水面数据, 提取人工建造形成的硬质地表; (2) 对“市政公用设施基本具备”的判定: 选择导航电子地图的道路网数据 (含乡镇村等级以上道路), 通过道路网密度因子反映部分市政公用设施是否具备。同时, 通过道路的空间分布也可以判断图斑与中心城区是否连接成片; (3) “公共设施基本具备”判定: 选择导航电子地图的 POI 数据, 通过对教育、医疗、文化、体育等主要公共设施空间分布的分析, 评估不同区域公共服务设施建设水平。最后, 将百度 LBS 在地活动人口密度和高分辨率遥感影像作为校核因子, 对初步划定的准建城区进行修正。

3.1.2 核心建成片区划定原理

核心建成片区是指城市内部承担核心职能、呈现高密度开发态势的中心性建成区域。其内涵主要体现为以下 3 点: 在区位上, 处于地理中心或被普遍认可的城市核心区域; 在形态上, 表现出较高的建设强度和人口密度; 在功能上, 汇聚了商业、金融、行政及历史文化等核心功能, 是高等级公共服务设施的密集区。在划定核心建成片区时, 充分考虑其定义中各项要点的表达。首先识别建成区范围内的地理中心和认知中心, 形成中心区域; 同时测算各街道、乡镇单元的密度情况, 识别较高密度范围; 之后验证核心功能区和主要设施是否已被高密度范围涵盖; 最后识别大型水域、郊野公园等大型生态空间, 不纳入计算范围 (图 1)。

3.2 技术路线

3.2.1 建成区划定技术路线

建成区划定的技术路线可概括为: 首先提取地表覆盖类型为不透水面的区域, 并将其

¹ 四维图新. <https://www.seewayai.com/>.

² 百度慧眼. <https://huiyan.baidu.com/>.

³ BIGEMAP. <http://www.bigemap.com/>.

定义为“建成地区”；在建成地区内设置一定的阈值，将道路网密度和 POI 密度高于阈值的区域判定为“准建成区”；基于百度网格人口密度和高分辨率遥感影像对准建成区进行校核修正，获得城市“建成区”边界（图 2）。

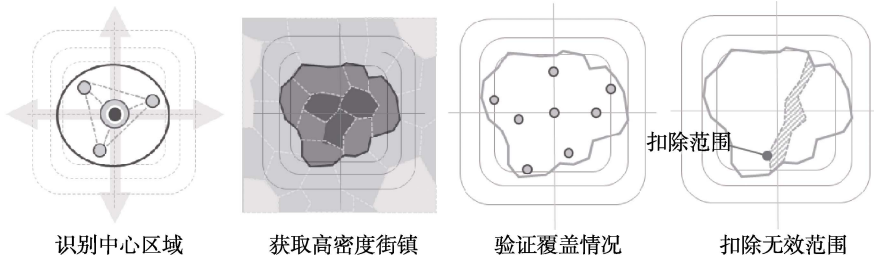


图 1 核心建成片区划定原理示意图

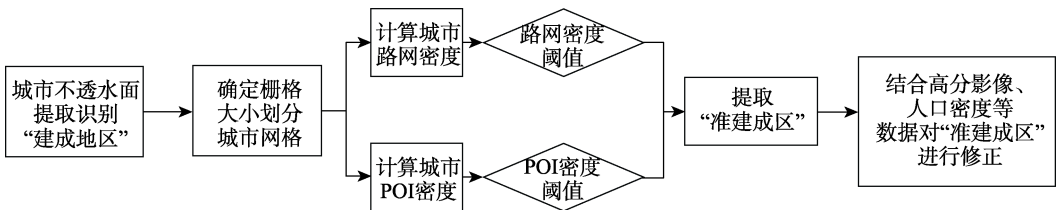


图 2 建成区划定技术路线图

以北京市为例，建成区提取过程如下：

第一步：识别“建成地区”。

在城市市辖区范围内，以高分遥感影像解译的土地覆盖数据为基础，将类型为不透水面的区域作为“建成地区”。依据清华大学地球系统科学系利用 AI 技术生产的全球土地覆盖产品（FROM-GLC10），北京市共有建成地区 2,560 km²（图 3）。

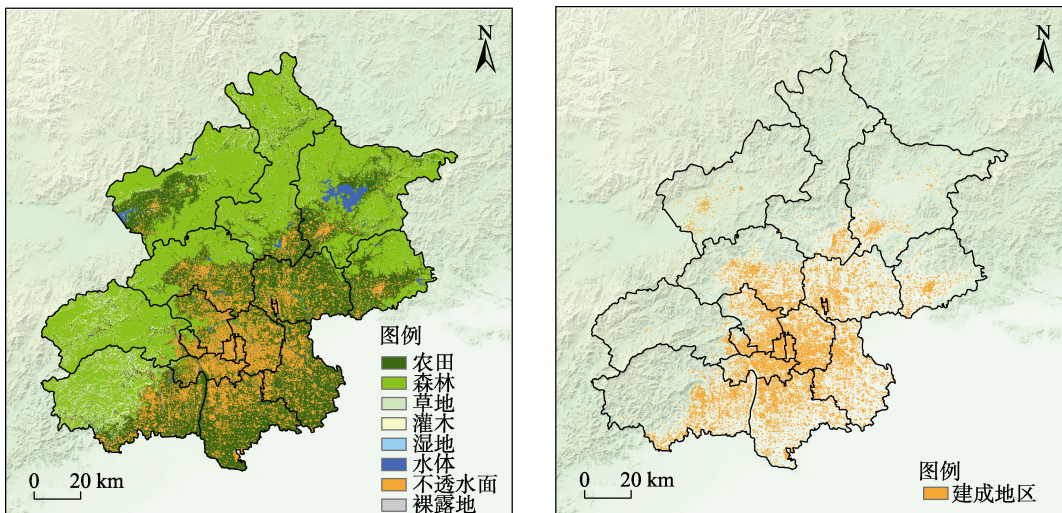


图 3 北京市土地覆盖类型（左）和“建成地区”（右）分布图

第二步：识别“准建成区”。

(1) 将市辖区范围划分为均匀的 $500\text{ m} \times 500\text{ m}$ 网格。网格的大小对建成区面积和边界形状有一定影响，网格过大容易导致带状、指状形态的建成区，以及中小规模的城市建成区面积误差过大，或者遗漏与城市中心组团不相连的开发区、矿区等城市用地斑块；网格过小，虽然能够提高精细程度，但斑块过于破碎会给后续处理和分析工作带来不必要的麻烦，并且在太小的斑块内难以对基础设施和公共服务设施的建设情况进行评价。本文提取建成区的最小单元设置为 500 m 栅格。

(2) 确定城市网格尺度道路网密度和 POI 密度阈值，阈值的大小直接决定建成区边界的空间位置和面积大小。考虑到 19 个城市的地理区位、形态特征、城镇化水平、经济社会发展水平等特征均具有较大的差异，各城市应当设置不同的阈值。本文借鉴夜间灯光遥感数据提取建成区最广泛使用的突变检测法确定道路网密度和 POI 密度的阈值。该方法基于形态学特征认为城市区域会保持整体几何形状的完整性，在阈值逐渐升高的过程中，城市建成区边界会沿着边缘缩小，但是当分割阈值达到某一个点时，建成区边界不再沿着边缘缩小，而是从内部破碎，产生新的较小的多边形，进而导致建成区边界周长突然增加，并将这个点定义为阈值点^[1]。本文认为路网密度与 POI 密度在由中心城区向外围郊区过渡的过程中同样存在一个突变点（即城市建成区边界），这个突变点即为该因子的阈值。经计算，北京市划定建成区方法中路网密度阈值为 4.1 km/km^2 ，POI 密度阈值为 45 个/km^2 （图 4）。

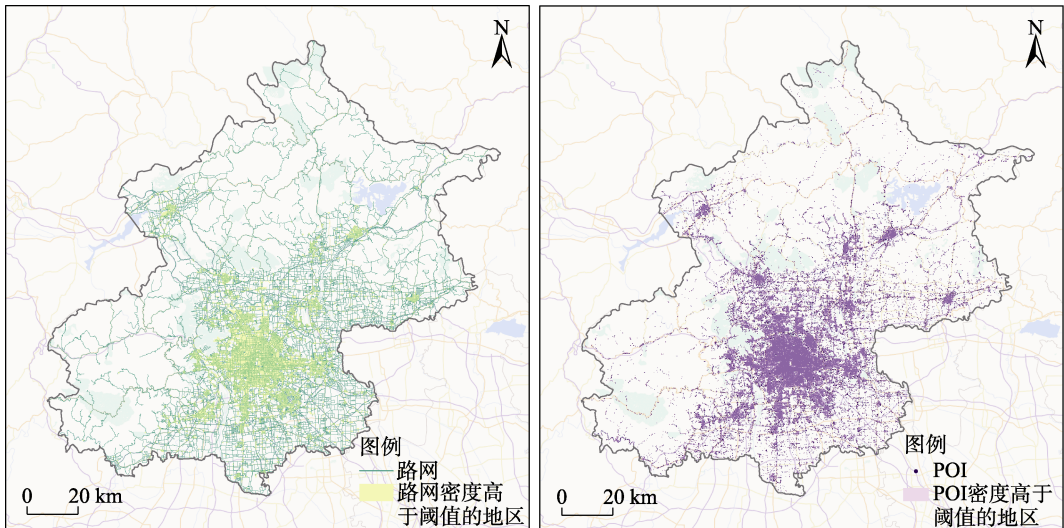


图 4 北京市路网密度高于阈值的地区（左）和 POI 密度高于阈值的地区（右）分布图

(3) 将市辖区“建成地区”内道路网密度或 POI 密度高于其对应阈值的网格判定为“准建成区”，并对“准建成区”进行多边形合并、去噪或填充处理。首先提取道路网密度、POI 密度中任一指标高于其对应阈值的栅格，并转换成矢量面；接着对矢量多边形做合并处理，形成“准建成区”斑块；之后删除“准建成区”斑块中面积低于 2 km^2 （即小于 8 个格网组成）的细碎多边形，消除面积低于 10 km^2 （即小于 40 个格网组成）的空洞，得到较为规范的“准建成区”（图 5）。

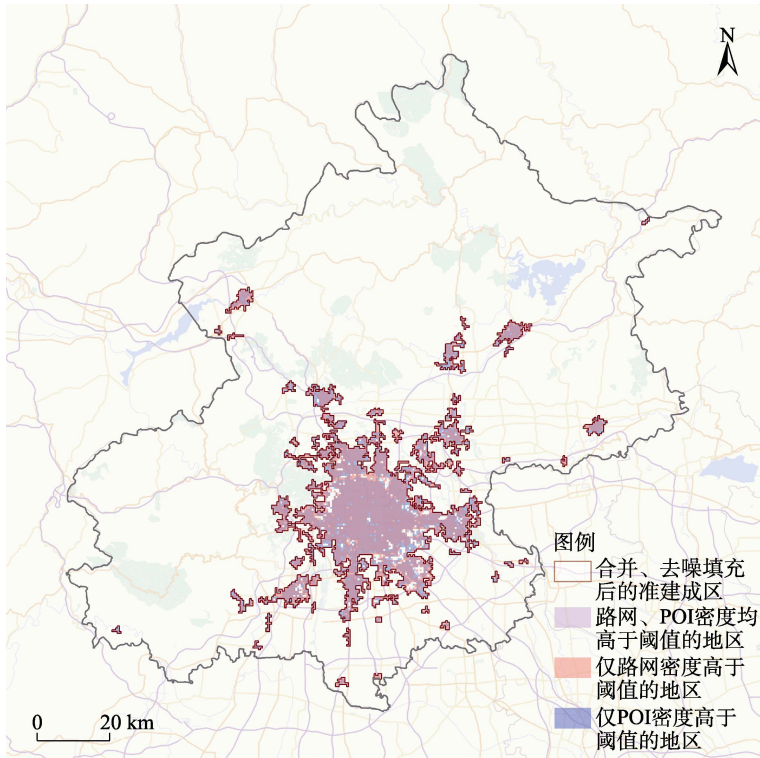


图5 合并、去噪、填充后的准建成区分布图

第三步：对“准建成区”进行修正，形成城市建成区边界（图6）。

（1）计算城市网格尺度人口密度。虽然《城市规划基本术语标准》中城市建成区的定义并不包含对城市人口数量的描述，但无论是成片开发建设的土地，还是完备的市政公用设施，都是服务于人的。人口数量是反映城市建成状况的重要指标之一。本文基于百度LBS数据计算城市 $500\text{ m} \times 500\text{ m}$ 网格尺度人口密度指标，作为校核因子对城市建成区边界进行修正。

（2）通过整合高分辨率遥感影像、百度网格人口密度、城市路网和生态廊道等数据修正准建成区边界，同时人工去除大面积水域与位于城市边缘的山体等非建设用地，以获取更准确的建成区范围。

3.2.2 核心建成片区划定技术路线

核心建成片区划定的最小单元是街道、乡镇级别行政区划边界，为保证不同城市间建成环境指标的可比性，并借鉴国内外类似概念（如北京核心区 92.5 km^2 ^[18]、巴黎都市圈核心区 105 km^2 ^[19]），将面积控制在 $90\text{--}160\text{ km}^2$ 左右。具体的划定流程包括以下4个步骤（图7）：

第一步：识别城市中心区域。城市中心包括两种类型，一种是客观上的几何中心，即地理中心；一种是主观上大众对城市的认知中心。对于地理中心，可通过城市建成区范围的几何中心（通常为环形道路的中心、主要干道的交汇路口）进行识别。认知中心即识别能够代表城市中心形象的代表性公共设施和公共区域，通常为商业区、文化设施、公园广

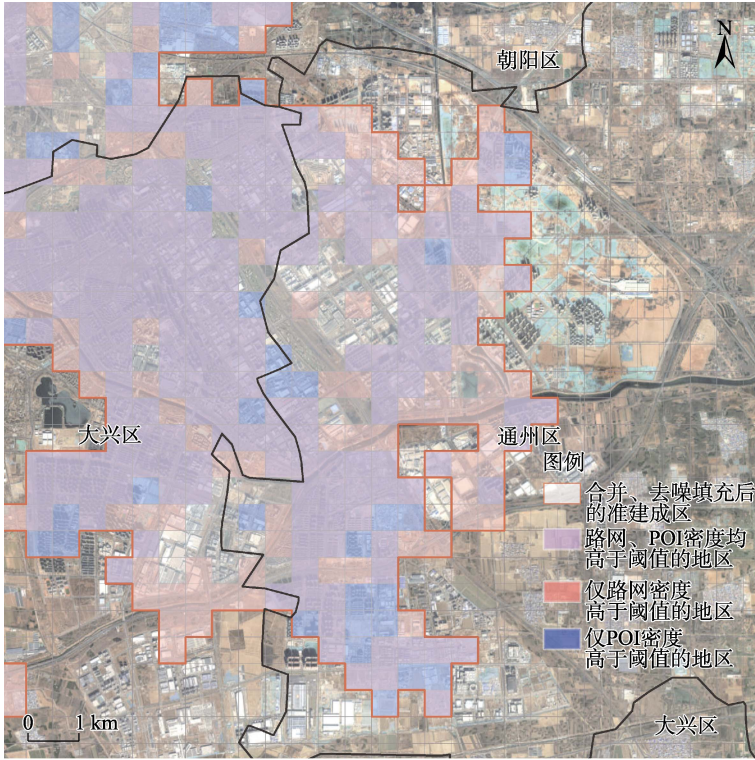


图6 基于校核因子的准建成区修正样例图

场等作为城市中心。

第二步：识别密度较高的街镇。基于百度网格人口密度数据、建筑基底数据，计算城市建成区范围内各街道或乡镇单元的建设强度和人口密度指标。通过指标空间分布情况，识别较高密度范围，作为核心建成片区的空间范围基础。

第三步：校核城市主要功能区和设施。验证城市的主要功能区和主要城市级设施是否已被第二步识别的高密度范围囊括。一般考察城市的商业中心、商务中心、行政中心、历史街区和大型公共设施的布局情况，但是与城市中心区域距离较远的城市副中心（如北京通州副中心、天津滨海新区等）在本文的核心建成片区划定中暂不纳入。

第四步：扣除大型开敞空间。识别大型水域、郊野公园等建设强度低、且人们较少到达和使用的生态空间，从核心建成片区范围中扣除。其中，城市公园是城市居民日常使用的空间，仍纳入到城市核心建成片区的计算范围内。

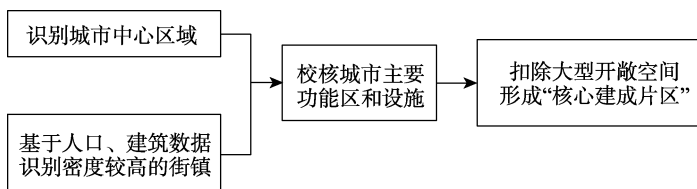


图7 核心建成片区划定技术路线

4 数据结果与验证

4.1 数据集组成

本文基于多源大数据，利用统一的划定方法和划定流程，生产了我国 19 个城市的建成区边界和核心建成片区边界。数据集存储为.shp 格式，在属性表中列出了建成区、核心建成片区所属的城市名称、省名称与区划代码信息（表 3）。数据集由 2 个数据文件组成，数据量共为 2.23 MB（压缩为 1 个文件，1.37 MB）。

表 3 数据集属性表信息

字段名称	字段说明
FID	序号
Shape	“面”要素类型
CityName_C	城市中文名称
CityName_E	城市英文名称
CityCode	城市行政区划代码
ProName_C	省份中文名称
ProName_E	省份英文名称

4.2 数据结果

本文划定的 19 个城市建成区和核心建成片区结果如图 8 所示。

本文分析了 19 个城市建成区及核心建成片区在各区县单元中的分布情况（图 9）。分析结果显示，不同城市内部区县单元的城市化水平存在显著差异，主要体现在建成区覆盖程度（建成区面积占比）和城市核心功能集聚程度（核心建成片区面积占比）两个维度上。

（1）高度城市化的核心区域

部分区县单元在建成形态和核心功能布局上已高度成熟，其建成区与核心建成片区均占据绝对主导地位。典型代表包括：北京东城区、西城区（建成区占比 100%，核心建成片区占比近 100%），天津和平区（100%，100%），西安新城区（100%，99.6%）、碑林区（100%，100%），上海虹口区（97.7%，97.3%），青岛市南区（96.4%，95.9%）、市北区（97.9%，97.3%）。这些区县单元的建成区与核心建成片区占比普遍超过 95%，是所在城市承载核心功能的关键区域。

（2）建成区成熟但核心功能集聚不足的区域

部分区县显示出建成区覆盖率高但核心建成功能相对薄弱的特征。成都市武侯区的建成区占比为全市最高（96.7%），但核心建成片区占比（56.9%）明显偏低。这一现象表明，这些区域在空间建成形态上已高度城市化，但在承载和集聚城市核心功能方面仍有发展空间。

（3）建成区面积占比较低的区域

部分城市的区县单元建成区占比较低（低于 50%），核心建成片区的占比则更低，反映了城市建成区、核心建成片区分布受自然地理条件、城市发展规划等因素影响的特征。以哈尔滨为例，南岗区的建成区和核心建成片区占比均为本城市最高，但也仅有 42.2%

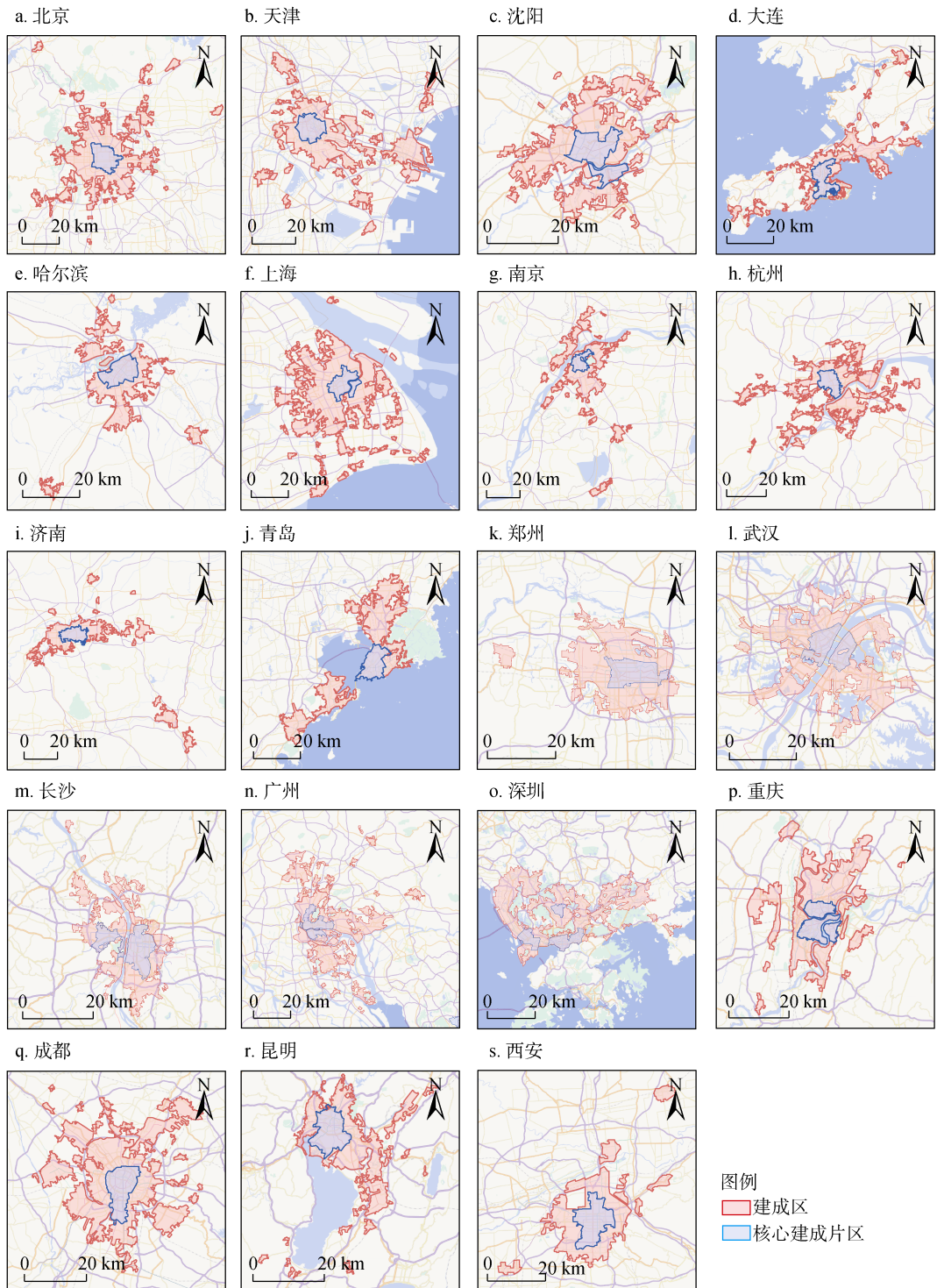


图8 北京等19个城市建成区、核心建成片区空间范围图(2022)

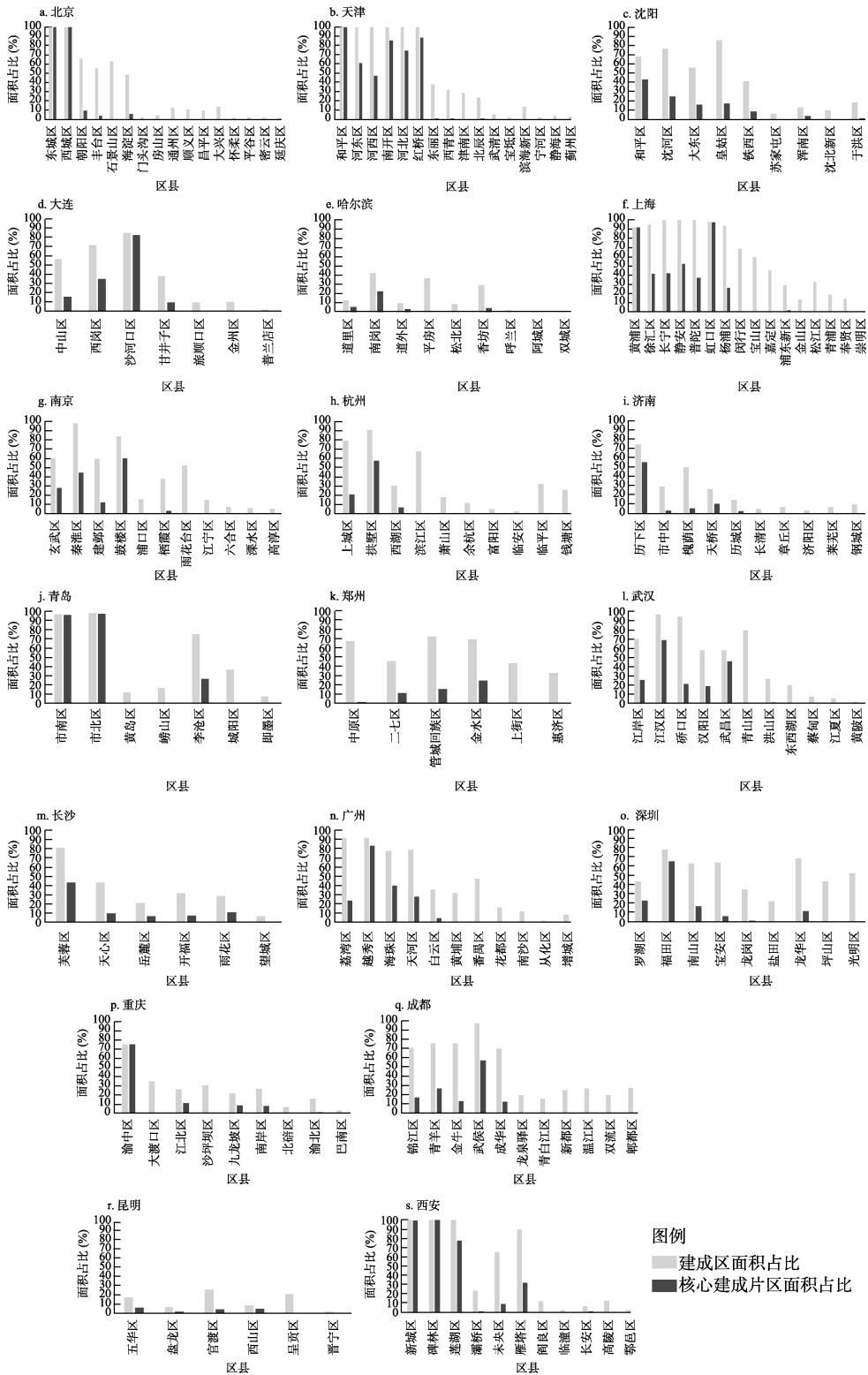


图 9 北京等 19 个城市区县单元建成区、核心建成片区面积占比统计图

和 22.2%（区内含大面积农田）；昆明建成区占比最高的官渡区仅 25.5%（区内多为山地和水域）。

分析表明，不同区县单元在承载城市实体空间（建成区）和核心功能空间（核心建成片区）方面存在显著层级差异。这种差异深刻反映了城市内部发展不均衡、功能分布梯度及自然环境约束效应，为理解城市空间结构与功能布局提供了重要依据。

4.3 数据结果验证

为验证建成区划定结果的准确性，本研究将 19 个城市的建成区与住建部 2022 年第三方城市体检所使用的建成区数据进行详细对比（图 10）。结果表明，二者在建成区的空间中心位置和整体形态上表现出高度一致性；在面积大小上亦基本吻合，但不同城市呈现出一定差异。具体而言，天津、南京、长沙、昆明、哈尔滨、沈阳、济南和杭州的建成区面积与参考数据的差值小于 1%，展现出最佳的吻合度；而武汉、青岛、重庆、郑州的面积差值则超过 10%，差异相对显著；其余城市的面积差值在 1%–10% 之间，差异水平适中。面积差异的主要原因是两种算法在城市内部及边缘面积较大的绿地、水体的判定上存在差异。

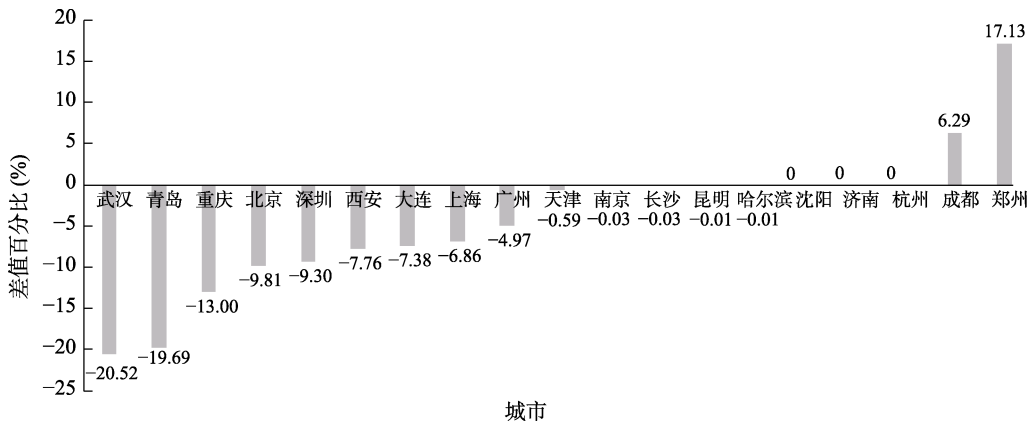


图 10 建成区结果验证图

由于资源环境、历史文化和发展水平等因素存在差异，各城市的建成区面积差异悬殊，例如上海、北京的建成区面积超过 1,500 km²，而昆明、哈尔滨、长沙等则小于 500 km²。这种显著的尺度差异导致直接比较各建成区间的设施覆盖、人口密度等城市发展评价指标不尽合理。然而城市的核心区域（尤其是具有较高密度并承载核心功能与设施的区域）面积差异通常较小（图 11）。因此，在该尺度上比较城市发展水平，能有效弥补基于建成区整体比较的局限性。由于缺乏官方公布的核心建成区标准数据集，为评估划定结果的准确性，本文采用标准差作为辅助判据。划定结果显示：19 个案例城市核心建成片区面积标准差为 22.06，远低于建成区面积标准差（394.63），表明划定结果满足期望的空间特征（即核心区面积差异小），验证了方法的有效性。

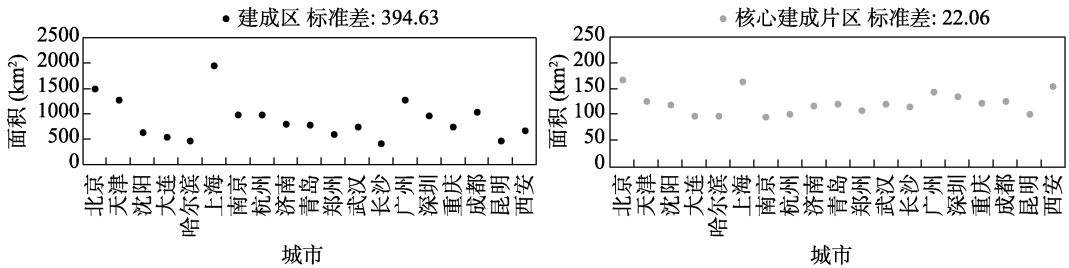


图 11 核心建成片区结果验证图

5 讨论和总结

城市建成区和核心建成片区是评估城市建成环境指标的常用空间边界，也是进行城市管理、规划建设、城市更新和城市体检等多项工作的基础空间边界。传统的划定方法以人工为主，主观性太强。随着新型大数据的出现和空间数据挖掘技术的进步，建成区划定方法逐步得到改进和完善，提取的准确度和精确度也在不断提高。本文提出的多源数据融合方法，综合利用高分辨率遥感影像、道路网、兴趣点、百度 LBS 人口密度等数据，在格网尺度上高效识别城市建成区及核心建成片区边界。该方法借助大数据优势，为建成环境评估研究提供了精准的空间基础。基于多源数据的城市建成区和核心片区划定方法具有同样的数据来源、相同的划定思路和方法、统一的技术流程，能够保证提取的多个城市边界范围横向可对比、纵向可比较、年度可更新。

然而，本研究在建成区边界的划定方法上仍存在一定的局限性。核心问题在于未能完全响应建成区定义中“市政公用设施基本具备”这一关键维度。尽管本文充分考虑了道路网这一重要的市政基础设施，但对于同等重要的市政管线（如给排水、燃气、电力、通信管网等）却未能有效纳入考量。主要受限于两方面因素：一是不同城市的市政管线数据电子化、标准化程度差异显著；二是此类数据属于敏感数据，尚无公开的数据集。因此当前划定结果在严格意义上对建成区定义的诠释尚有不足。此外，本研究依据的清华大学地球系统科学系利用 AI 技术生产的全球土地覆盖产品（FROM-GLC10）未进行真实性检验，其中存在的问题尚需进一步研究。

未来，将进一步拓宽数据源，同时积极引入人工智能技术，以提升数据处理、特征提取及边界判定的自动化水平，构建更为精准、高效且标准化的城市建成环境评估数据支撑体系。

作者分工：张淑杰对数据集的开发做了总体设计，并撰写了论文的主体部分。李萌对城市建成区边界进行了提取，同时参与论文的写作。淡铂洋对文中的基础数据进行了处理，同时参与论文的写作。韩靖北对核心建成片区的边界进行了划定。郝灵强对数据集和论文进行了订正和修改。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

- [1] 舒松, 余柏菡, 吴健平等. 基于夜间灯光数据的城市建成区提取方法评价与应用[J]. 遥感技术与应用, 2011, 26(2): 169–176.
- [2] 杨艺, 孙文彬, 韩亚辉. 基于 DMSP/OLS 灯光亮度组合值的城市建成区提取方法[J]. 国土资源遥感, 2020, 23(3): 39–48.
- [3] 许泽宁, 高晓璐. 基于电子地图兴趣点的城市建成区边界识别方法[J]. 地理学报, 2016, 71(6): 928–939.
- [4] 王晓艳, 李成名, 陈宪冬等. 夜间灯光融合多源数据的建成区提取方法[J]. 遥感信息, 2021, 36(5): 114–123.
- [5] 郑洪哈, 桂志鹏, 栗法等. 夜间灯光数据和兴趣点数据结合的建成区提取方法[J]. 地理与地理信息科学, 2019, 35(2): 25–32.
- [6] 黄宁, 徐志红, 徐莎莎. 武汉市城市建设用地强度管控实证研究与动态优化[J]. 城市规划学刊, 2012(3): 96–101.
- [7] 金探花, 杨俊宴, 王德. 从城市密度分区到空间形态分区: 演进与实证[J]. 城市规划学刊, 2018(4): 34–40.
- [8] 唐婧娴, 龙瀛. 特大城市中心区街道空间品质的测度——以北京二三环和上海内环为例[J]. 规划师, 2017, 33(2): 68–73.
- [9] 常青, 徐勤政, 杨春等. 北京新总规建设用地减量调控的思考与探索[J]. 城市规划, 2017, 41(11): 33–40.
- [10] 王浩锋, 施苏, 饶小军. 城市密度的空间分布逻辑——以深圳市为例[J]. 城市问题, 2015(8): 22–32.
- [11] 史北祥. 亚洲城市中心区极核结构现象的定量研究[D]. 南京: 东南大学, 2015.
- [12] Imhoff, M. L., Lawrence, W. T., Stutzer, D. C., *et al.* A technique for using composite DMSP/OLS “City Lights” satellite data to accurately map urban areas [J]. *Remote Sensing of Environment*, 1997, 61(3): 361–370.
- [13] 张淑杰, 李萌, 淡铂洋等. 基于多源数据的城市建成环境评估单元数据集 (2022) [J/DB/OL]. 全球变化数据仓储电子杂志, 2025. <https://doi.org/10.3974/geodb.2025.03.09.V1>.
- [14] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策 [OL]. <https://doi.org/10.3974/dp.policy.2014.05> (2017年更新).
- [15] Gong, P., Liu, H., Zhang, M., *et al.* Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017 [J]. *Science Bulletin*, 2019, 64: 370–373.
- [16] Che, Y., Li, X., Liu, X., *et al.* 3D-GloBFP: the first global three-dimensional building footprint dataset [J]. *Earth System Science Data Discussions*, 2024, 16(11): 18.
- [17] 中华人民共和国建设部. 城市规划基本术语标准(GB/T 50280—98)[S]. 北京: 中国建筑工业出版社, 1998.
- [18] 北京市规划和自然资源委员会. 北京城市总体规划(2016年—2035年)[Z]. 北京: 北京市人民政府, 2017.
- [19] 倪金城. 大巴黎都市圈城区与郊区融合的纽带: 巴黎市域(郊)铁路[J]. 城市轨道交通, 2021(2): 26–31.