



文献 CSTR:

32001.14.11-6035.noda.2022.0005.zh



文献 DOI:

10.11922/11-6035.noda.2022.0005.zh

数据 DOI:

10.57760/sciencedb.01767

文献分类: 地球科学

收稿日期: 2022-03-07

开放同评: 2022-03-08

录用日期: 2022-08-11

发表日期: 2023-06-30

2021 年武汉市主要城市公园数据集

廖一鸣^{1,2}, 王树竹¹, 常可¹, 秦畅¹, 邓濯缨¹, 吕争³, 周琪^{1*}

1. 中国地质大学(武汉)地理与信息工程学院, 武汉 430074

2. 浙江省城乡规划设计研究院, 杭州 310030

3. 中国资源卫星应用中心, 北京 100094

摘要: 城市公园数据是研究城市公园的空间分布、指导城市公园规划与管理的基础地理信息数据, 也是联合国可持续发展目标的评估指标之一。针对当前缺少开放的城市公园数据集, 以官方发布的《2021 年武汉市主要城市公园统计表》为参照, 基于 OpenStreetMap、POI 和谷歌影像等多源开放数据研制了 2021 年武汉市主要城市公园数据集。本数据集为 shp 格式, 包括城市公园的名称、面积、经纬度坐标和地址等属性信息; 数据现势性为 2021 年。采用相关性分析对比了基于本数据集与官方统计数据获取的城市公园面积, 相关系数高达 0.96, 表明本文所研制的数据集总体质量较好。基于多源开放数据获取城市公园, 不仅可以提高数据的可靠性, 还能减少费时、费力的人工采集工作, 对研制其他城市的公园数据集有借鉴意义。

关键词: 公园; 城市; 多源数据; OSM; POI

数据库(集)基本信息简介

数据库(集)名称	2021 年武汉市主要城市公园数据集
数据作者	廖一鸣、王树竹、常可、秦畅、邓濯缨、吕争、周琪
数据通信作者	周琪 (zhouqi@cug.edu.cn)
数据时间范围	2021 年
地理区域	湖北省武汉市
数据量	322 KB
数据格式	*.shp; *.xls;
数据服务系统网址	https://doi.org/10.57760/sciencedb.01767
基金项目	国家对地观测科学数据中心开放基金(NODAOP2021010); 国家自然科学基金(41771428)。
数据库(集)组成	本数据集包括 2021 年武汉市主要城市公园矢量数据和城市公园汇总表, 保存为一个压缩文件。

引言

城市公园是城市生态系统和城市景观的重要组成部分, 是满足城市居民的休闲需要, 提供休息、游览、锻炼、交流, 以及举办各种集体文化活动的场所。现有研究表明: 城市公园在保护生物多样性^[1]、降低城市噪声^[2-3]、减少城市热岛效

* 论文通信作者

周琪: zhouqi@cug.edu.cn

应^[4-5]和维持城市居民的身心健康^[6-7]等方面发挥着重要作用。2015 年,联合国可持续发展目标还提出:到 2030 年,全球要普遍建设包容、安全、有抵御灾害能力和可持续的城市和人类住区^[8],并把城市建成区中供所有人使用的开放公共空间(包括城市公园)作为评估指标之一。因此,城市公园数据不仅是指导城市规划与管理的基础地理信息数据,还是支撑联合国 2030 年可持续发展目标的重要数据来源。

城市公园主要由城市中的公园绿地构成。遥感技术是获取城市绿地数据的主要方法。例如,刘文雅等基于高分二号卫星影像提取了北京市的城市绿地数据^[9];Rafiee 等则基于 Landsat TM 和 IRS LISS-III 影像对比了马沙德(伊朗)1987 和 2006 年的城市绿地变化^[10]。随着倾斜摄影和地面移动测量等新技术的发展,由百度和谷歌公司提供的街景地图也被用于城市绿地研究^[11]。Li 等基于谷歌街景地图评估了纽约曼哈顿的绿地覆盖率^[12]。然而,无论是基于遥感影像还是街景地图,均因缺少语义信息,通常只适用于获取城市绿地而非城市公园数据。众源地理数据 OpenStreetMap (OSM) 包含绿地、水体和运动设施等丰富的地理要素,被认为是获取城市公园的数据来源。Ludwig 和 Zipf 提出基于 OSM 数据的城市绿地制图方法^[13]。Liao 等对比分析了 6 种全球和区域性的开放地理信息数据(包括 OSM 数据),发现:基于 OSM 数据获取城市公园绿地的准确率、精确率、召回率和 F1 值等指标最高(80%以上)^[14]。然而,OSM 数据因由全球志愿者提供,不可避免地存在数据质量问题。如何基于 OSM 获取可靠的城市公园数据仍是亟待解决的难题。

为弥补上述不足,本研究以武汉市为例,以《2021 年武汉市主要城市公园统计表》^[15]发布的城市公园为参照,基于 OSM 获取城市公园矢量数据,并结合 POI 数据、谷歌(在线)影像、实地调查等为参考,补充与完善 OSM 数据的几何和属性信息,研制 2021 年武汉市主要城市公园数据集。本数据集可以应用于分析武汉市城市公园的空间分布、量化城市公园的相关指标,服务于城市规划、生态环境和生命健康等领域的科学研究。此外,本研究基于 OSM 等多源开放数据获取城市公园的思路,可以为获取其他城市的公园数据提供参考。

1 数据采集和处理方法

1.1 数据采集方法

本数据集使用的源数据有:

(1) 《2021 年武汉市主要城市公园统计表》

武汉市园林和林业局于 2022 年 3 月向公众发布了《2021 年武汉市绿化状况公报》,其中列出的《2021 年武汉市主要城市公园统计表》(本文简称“统计表”)(<https://flbook.com.cn/c/50a2pjxqCb#page/26>)提供了覆盖武汉市各辖区的 145 个城市公园的名称和面积信息。本研究参考该“统计表”研制相应的空间数据。

(2) OSM 数据

OSM 是最具代表性的众源地理信息项目之一,它由全球志愿者实时在线编辑与更新。OSM 数据为矢量格式(shp),包括道路、铁路、水系、建筑、土地覆盖和土地利用等丰富的地理要素。本研究从 Geofabrik 平台(<http://download.geofabrik.de/index.html>)下载了 2022 年 1 月的 OSM 数据,并获取与公园要素有关的矢量数据。

(3) POI 数据

POI 数据是指在地图上包含语义信息的坐标点,如超市、学校、医院和公园等。目前,多个电子地图平台为用户免费提供了获取 POI 数据的接口。本研究基于高德地图 Web 服务 API 爬取武汉市的“公园”POI 数据,得到保存为文本文件(csv)的属性信息(如“名称”、“地址”和“经纬度”等),再将经纬度信息空间化后得到 shp 格式的 POI 数据,即记录城市公园的位置信息。

(4) 谷歌影像数据

谷歌公司提供了反映真实地表特征的高分辨率卫星遥感影像,本文采用 2021 年谷歌(在线)影像作为目视解译城市公园的基础资料,辅助确定城市公园的边界或范围。

1.2 数据处理方法

图 1 为本文数据集的方法流程,具体如下:

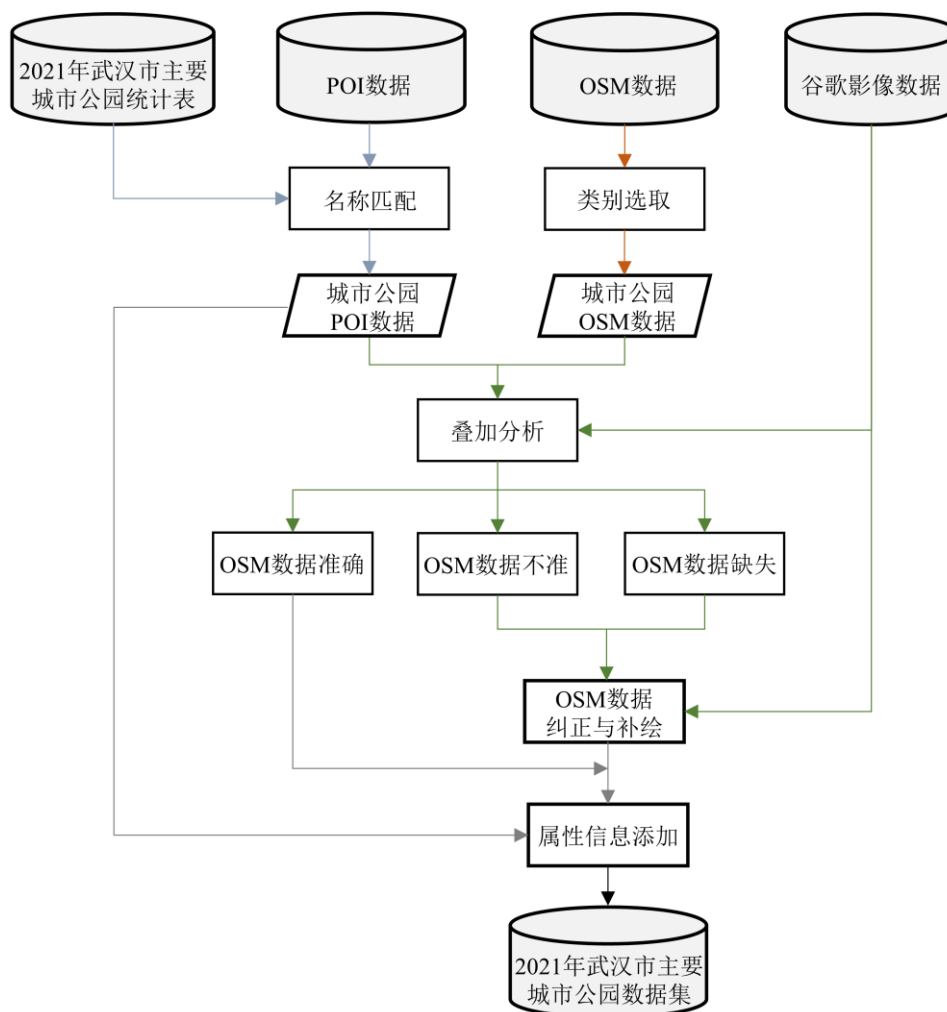


图 1 本文数据集的方法流程图

Figure 1 The workflow of our dataset

(1) 公园位置确定

由于高德平台的坐标系统为 GCJ-02 (火星坐标系),基于该平台采集的 POI 数据与公园的实际位置存在偏差。因此,首先将 POI 数据转换为 WGS84 坐标系。然后,基于 POI 数据的“名

称”和“统计表”的公园名称进行字段匹配，将匹配后的 POI 数据用于确定城市公园的位置。

(2) 公园几何数据获取

OSM 数据包括几何和属性信息。OSM 属性信息记录在“标签”中。选取“标签”为“landuse=park”和“poi=park”的 OSM 面状要素作为城市公园数据。

(3) 公园几何数据纠正与补绘

OSM 数据因由全球志愿者提供，不可避免地存在数据质量问题，因此需要进行质量检查。具体思路为：首先，叠加城市公园 POI 和 OSM 数据，筛选出与“统计表”对应的 OSM 城市公园数据；然后，叠加筛选后的 OSM 城市公园数据和谷歌影像，并检查 OSM 城市公园数据的准确性。可能有三种情况，即 OSM（几何）数据准确、不准或缺失：

1) OSM 数据准确：即 OSM 数据的几何边界或范围准确。判断依据为 OSM 数据包含植被、步道和水域等公园特有的要素，且 OSM 数据以外的区域主要是建筑和道路等要素。该类 OSM 数据无需纠正（图 2）。



图 2 几何准确的 OSM 数据（韵湖公园）

Figure 2 The OSM data with precise geometric boundary (Yunhu Park)

2) OSM 数据不准：即 OSM 数据的几何边界或范围不准。该类 OSM 数据需要人工纠正。首先，基于谷歌影像判读 OSM 数据边界是否存在混分（如 OSM 数据边界内外均为植被，图 3a）；然后，通过实地调查确定公园的实际边界，并通过人工勾绘对 OSM 数据进行纠正（图 3b）。



图 3 人工纠正的前后对比结果

Figure 3 Comparison of OSM urban park data before and after manual correction

3) OSM 数据缺失：即城市公园对应的 OSM 数据缺失。该类 OSM 数据需要人工补绘。具体思路为：首先，基于 POI 数据确定城市公园的大致位置；然后，基于谷歌影像并结合实地调查确定城市公园的实际边界后补绘。

(4) 公园属性信息添加

为纠正和补绘后的 OSM 城市公园数据添加属性信息。首先，建立城市公园 POI 和 OSM 数据的空间对应关系；然后，将 POI 数据的“名称”和“地址”信息添加到 OSM 城市公园数据的属性表中；最后，计算每个城市公园的质心经纬度坐标，以及经过国家 2000 坐标系统投影后的面积，均添加到 OSM 城市公园数据的属性表中。

2 数据样本描述

2.1 数据集概述

本数据集为 2021 年武汉市主要城市公园数据集，数据存储在“2021 年武汉市主要城市公园数据集.zip”文件（大小为 322 KB）。该文件包括城市公园矢量数据和城市公园汇总表，分别保存为 shp 和 xls 格式。其中，城市公园矢量数据图层采用 WGS84 地理坐标系统，包含武汉市 145 个主要城市公园的地理位置、形状以及属性信息。图 4 为武汉市 145 个主要城市公园的空间分布图。

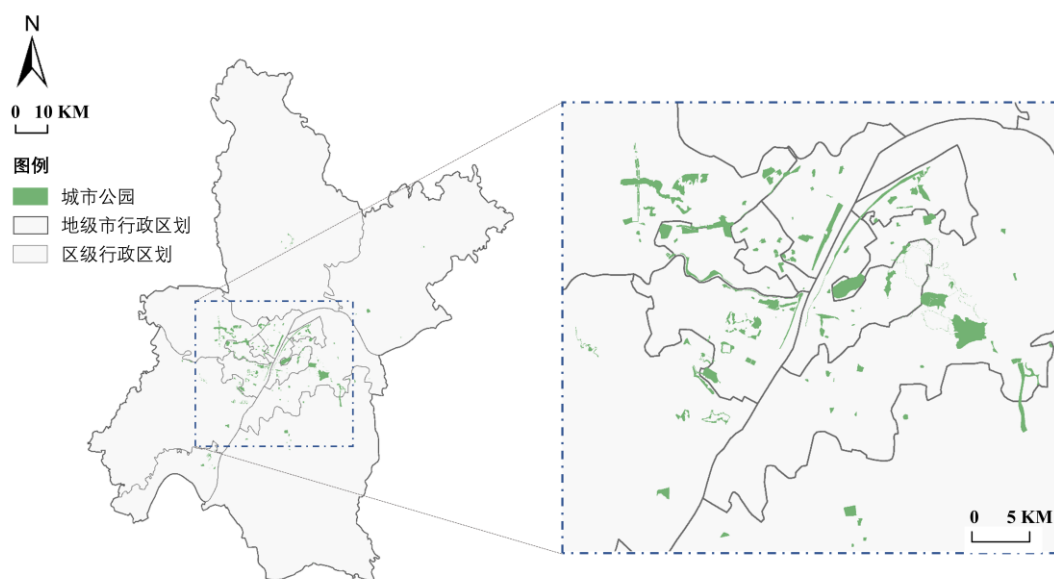


图 4 武汉市主要城市公园的空间分布图

Figure 4 The spatial distribution of major urban parks in Wuhan

2.2 数据属性表

城市公园矢量数据共包括 7 个属性字段，即公园编码、矢量类型、公园名称、公园质心的经度坐标、公园质心的纬度坐标、公园面积和公园地址（如表 1）。其中，公园编码（FID）是属性表中的唯一标识符。

表 1 城市公园矢量数据的属性表说明

Table 1 The attribute table of urban park dataset

序号	字段名称	数据类型	字段描述
1	FID	string	公园编码
2	Shape	string	矢量类型
3	Name	string	公园名称
4	Lon	double	公园质心的经度坐标
5	Lat	double	公园质心的纬度坐标
6	Area	double	公园面积
7	Address	string	公园地址

3 数据质量控制与评估

以下将从完整性、位置精度和面积一致性 3 个方面对本文数据集的质量进行评估。

完整性：本研究基于 OSM 数据可以直接获取武汉市“公园统计表”中 106 个主要城市公园的数据，占总数（145 个）的 73%。其他的城市公园数据基于谷歌影像并结合实地调查补绘得到。因此，本文数据集包含“统计表”的所有城市公园，完整性达到 100%。

位置精度：本文数据集基于谷歌影像检查、纠正与补绘，其数据精度依赖谷歌影像的数据质量。现有研究表明：谷歌影像的平均位置精度小于 5 米^[16-17]。因此，本文数据集的位置精度相对较高。

面积一致性：将本文数据集的公园面积与“统计表”给出的统计面积进行相关性分析，得到相关系数为 0.96（图 5a），且该系数远高于基于 OSM 原始数据计算的城市公园面积与统计面积的相关系数（0.40，图 5b），表明本文数据集相较于 OSM 原始数据的质量有明显提升。

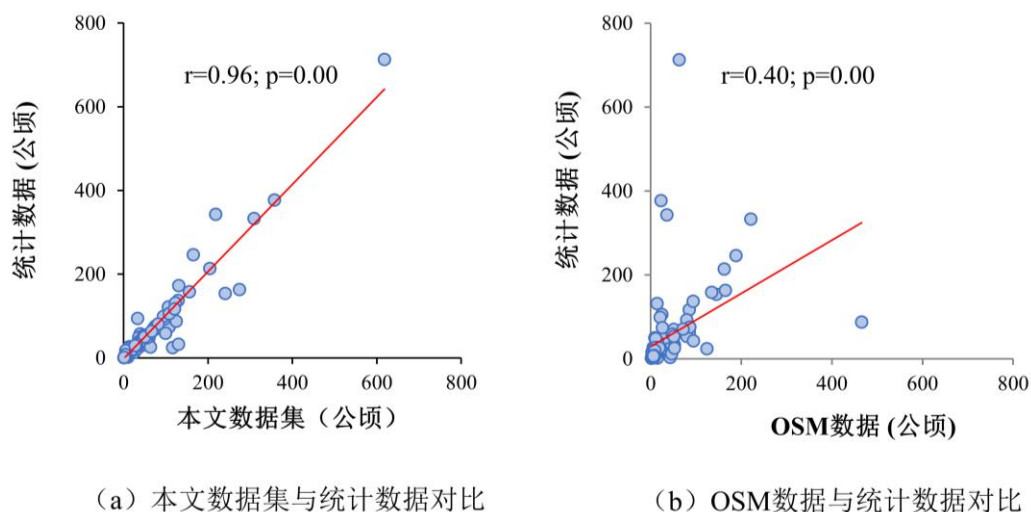


图 5 本文数据集 (a) 和 OSM 原始数据 (b) 与 (官方) 统计数据的城市公园面积比较

Figure 5 Comparison of urban park areas between (official) statistic results and our dataset (a) or original OSM dataset (b)

4 数据价值

本研究研制了武汉市主要城市公园的数据集，数据采集与整理、质量控制过程严格，可为准确评估城市公园的规模、揭示城市公园的分布特征提供数据支撑，为规划与设计新的城市公园和改善城市公园的服务水平提供决策支持，还可以与人口、经济、环境和健康等其他数据相结合，研究城市公园与社会经济、生态环境和生命健康等因素的关系。此外，该数据集基于 OSM 等多源开放数据获取城市公园的方法能有效提高数据质量和采集效率，对研制其他城市的公园数据集有借鉴意义。因此，本数据集可能被城市规划、地理、生态和医学等领域的专家学者所关注。

5 数据使用方法和建议

以分析武汉市城市公园的使用率和可达性^[18-19]为例，本文数据集的可能使用方法如下：

（1）城市公园使用率分析

城市公园使用率是衡量居民访问城市公园频次的指标，用于评价城市公园的使用情况^[18]。例如，通过叠加城市公园数据与新浪微博签到数据，可以计算城市公园的新浪微博签到次数，再通过计算签到次数与公园面积的比值即得到城市公园使用率。分析该指标可以为改造城市公园、增强城市公园的活力和吸引力提供依据。

（2）城市公园可达性分析

城市公园可达性是衡量居民从居住地到达城市公园难易程度的指标，用于评价城市公园的布局是否合理^[20]。例如，基于城市公园数据、人口数据和道路网数据计算城市居民到达城市公园的最近距离，再通过计算一定距离内城市公园覆盖的人口数量，分析不同城市公园的服务水平，为新的城市公园选址提供理论依据。

在今后的研究中，我们将扩充本文数据集，如提供不同年份的武汉市主要城市公园数据集；并尝试基于本文方法研制其他城市的公园数据集。

致谢

感谢武汉市园林和林业局、武汉市自然资源和规划局，以及高德和谷歌公司提供数据支持。

数据作者分工职责

廖一鸣（1996—），女，广西壮族自治区柳州市人，硕士研究生，研究方向为空间数据处理与分析。主要承担工作：论文撰写与制图。

王树竹（1998—），女，宁夏回族自治区吴忠市人，硕士研究生，研究方向为空间数据处理与分析。主要承担工作：数据处理及质量控制。

常可（1997—），女，山西省临汾市人，硕士研究生，研究方向为空间数据处理与分析。主要承担工作：数据处理及质量控制。

秦畅（2000—），男，河南省开封市人，本科生。主要承担工作：数据处理及质量控制。

邓濯缨（2002—），男，湖北省襄阳市人，本科生。主要承担工作：数据处理及质量控制。

吕争(1984—), 男, 湖北省武汉市人, 高级工程师, 研究方向为遥感数据处理。主要承担工作: 数据质量控制。

周琪(1984—), 男, 湖北省武汉市人, 博士, 副教授, 研究方向为空间数据处理与分析。主要承担工作: 研究方案设计与论文修改。

参考文献

- [1] CORNELIS J, HERMY M. Biodiversity relationships in urban and suburban parks in Flanders[J]. *Landscape and Urban Planning*, 2004, 69(4): 385–401. DOI: 10.1016/j.landurbplan.2003.10.038.
- [2] COHEN P, POTCHTER O, SCHNELL I. The impact of an urban park on air pollution and noise levels in the Mediterranean city of Tel-Aviv, Israel[J]. *Environmental Pollution*, 2014, 195: 73–83. DOI: 10.1016/j.envpol.2014.08.015.
- [3] XING Y, BRIMBLECOMBE P. Traffic-derived noise, air pollution and urban park design[J]. *Journal of Urban Design*, 2020, 25(5): 590–606. DOI: 10.1080/13574809.2020.1720503.
- [4] CAO X, ONISHI A, CHEN J, et al. Quantifying the cool island intensity of urban parks using *ASTER* and *IKONOS* data[J]. *Landscape and Urban Planning*, 2010, 96(4): 224–231. DOI: 10.1016/j.landurbplan.2010.03.008.
- [5] YAN H, WU F, DONG L. Influence of a large urban park on the local urban thermal environment[J]. *Science of the Total Environment*, 2018, 622/623: 882–891. DOI: 10.1016/j.scitotenv.2017.11.327.
- [6] BROWN G, SCHEBELLA M F, WEBER D. Using participatory GIS to measure physical activity and urban park benefits[J]. *Landscape and Urban Planning*, 2014, 121: 34–44. DOI: 10.1016/j.landurbplan.2013.09.006.
- [7] MULLENBACH L E, MOWEN A J, BAKER B L. Assessing the relationship between a composite score of urban park quality and health[J]. *Preventing Chronic Disease*, 2018, 15: E136. DOI: 10.5888/pcd15.180033.
- [8] United Nations General Assembly. Transforming our world: The 2030 agenda for sustainable development [C]. New York: United Nations, 2015.
- [9] 刘文雅, 岳安志, 季珏, 等. 基于 DeepLabv3+语义分割模型的 GF-2 影像城市绿地提取[J]. *国土资源遥感*, 2020, 32(2): 120–129. DOI: 10.6046/gtzyyg.2020.02.16. [LIU W Y, YUE A Z, JI J, et al. Urban green space extraction from GF-2 remote sensing image based on DeepLabv3+semantic segmentation model[J]. *Remote Sensing for Land & Resources*, 2020, 32(2): 120–129. DOI: 10.6046/gtzyyg.2020.02.16.]
- [10] RAFIEE R, SALMAN MAHINY A, KHORASANI N. Assessment of changes in urban green spaces of Mashad city using satellite data[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2009, 11(6): 431–438. DOI: 10.1016/j.jag.2009.08.005.
- [11] Wang, R., Helbich, M., Yao, Y., Zhang, J., Liu, P., Yuan, Y., & Liu, Y. (2019). Urban greenery and mental wellbeing in adults: Cross-sectional mediation analyses on multiple pathways across different greenery measures. *Environmental research*, 176, 108535. <https://doi.org/10.1016/j.envres.2019.108535>
- [12] LI X J, ZHANG C R, LI W D, et al. Assessing street-level urban greenery using Google Street View

and a modified green view index[J]. *Urban Forestry & Urban Greening*, 2015, 14(3): 675–685. DOI: 10.1016/j.ufug.2015.06.006.

[13] LUDWIG C, ZIPF A. Exploring regional differences in the representation of urban green spaces in OpenStreetMap[J]. *Geographical and Cultural Aspects of Geo-Information: Issues and Solutions*, 2019: 10.

[14] LIAO Y M, ZHOU Q, JING X Q. A comparison of global and regional open datasets for urban greenspace mapping[J]. *Urban Forestry & Urban Greening*, 2021, 62: 127132. DOI: 10.1016/j.ufug.2021.127132.

[15] 武汉市园林和林业局, 2021 年武汉市主要城市公园 (绿化广场) 统计表[R], 《2021 年武汉市绿化状况公报》, 2021:44. [Wuhan Landscape and Forestry Bureau. Statistical Table of Main Urban Parks (Green Square) in Wuhan City in 2021[R]. 2021 Wuhan Greening Status Bulletin, 2021:44.]

[16] MOHAMMED N Z, GHAZI A, MUSTAFA H E. Positional accuracy testing of Google Earth[J]. *International Journal of Multidisciplinary Sciences and Engineering*, 2013, 4(6): 6-9.

[17] WANG X W, WANG F. The precision of google earth map analysis with the coordinates of igs stations[J]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2020, XLII-3/W10: 1053–1056. DOI: 10.5194/isprs-archives-xlii-3-w10-1053-2020.

[18] 李方正, 董莎莎, 李雄, 等. 北京市中心城绿地使用空间分布研究: 基于大数据的实证分析[J]. *中国园林*, 2016, 32(9): 122–128. [LI F Z, DONG S S, LI X, et al. Spatial distribution research on the use of green space in the central city of beijing—based on the empirical analysis of big data[J]. *Chinese Landscape Architecture*, 2016, 32(9): 122–128.]

[19] 马琳, 陆玉麒. 基于路网结构的城市绿地景观可达性研究: 以南京市主城区公园绿地为例[J]. *中国园林*, 2011, 27(7): 92–96. DOI: 10.3969/j.issn.1000-6664.2011.07.022. [MA L, LU Y Q. Research on landscape accessibility of city green space based on traffic network—case study of the garden green space of the city proper of Nanjing[J]. *Chinese Landscape Architecture*, 2011, 27(7): 92–96. DOI: 10.3969/j.issn.1000-6664.2011.07.022.]

[20] LONG X J, CHEN Y J, ZHANG Y H, et al. Visualizing green space accessibility for more than 4, 000 cities across the globe[J]. *Environment and Planning B: Urban Analytics and City Science*, 2022, 49(5): 1578–1581. DOI: 10.1177/23998083221097110.

论文引用格式

廖一鸣, 王树竹, 常可, 等. 2021 年武汉市主要城市公园数据集[J/OL]. *中国科学数据*, 2023, 8(2). (2023-06-30). DOI: 10.11922/11-6035.noda.2022.0005.zh.

数据引用格式

廖一鸣, 王树竹, 常可, 等. 2021 年武汉市主要城市公园数据集[DS/OL]. *Science Data Bank*, 2022. [2022-05-18]. DOI:10.57760/sciencedb.01767.

A dataset of major urban park of Wuhan in 2021

LIAO Yiming^{1,2}, WANG Shuzhu¹, CHANG Ke¹, QIN Chang¹, DENG Zhuoying¹, LV Zheng³, ZHOU Qi^{1*}

1. School of Geography and Information Engineering, Wuhan 430074, P.R. China

2. Zhejiang Urban and Rural Planning Design Institute, Hangzhou 310030, P.R. China

3. China Centre for Resources Satellite Data and Application, Beijing 100094, P.R. China

*Email: zhouqi@cug.edu.cn

Abstract: Urban park data have been widely applied to urban planning and management. The availability of urban park has also been viewed as one of the evaluation indicators of the UN's sustainable development goals. However, currently there is still a lack of urban park datasets that are open to the public. To fill this gap, this study aims to produce a dataset of major urban parks of Wuhan in 2021. This dataset was produced based on multi-source data, including OpenStreetMap, POI and Google Earth image, with the official *Statistical Table of Major Urban Parks of Wuhan in 2021* as a reference. This dataset is in the format of ESRI shapefile, covering the name, area, latitude and longitude coordinates and address of the city parks in the year of 2021. We found that the correlation coefficient between the areas of urban parks for our dataset and the official statistic results is up to 0.96, which confirms the reliability and accuracy of our dataset. The approach of using multi-source data for acquiring urban park data boasts the advantage in reducing time-consuming and labor-intensive manual work; more importantly, it may also be used as a reference in acquiring urban park data of other cities.

Keywords: park; urban; multi-source data; OSM; POI

Dataset Profile

Title	A dataset of major urban park of Wuhan in 2021
Data corresponding author	ZHOU Qi (zhouqi@cug.edu.cn)
Data authors	LIAO Yiming, WANG Shuzhu, CHANG Ke, Qin Chang, DENG Zhuoying, LV Zheng, ZHOU Qi
Time range	2021
Geographical scope	Wuhan City, Hubei Province
Data volume	322 KB
Data format	*.xls; *.shp
Data service system	< https://doi.org/10.57760/sciencedb.01767 >
Sources of funding	National Earth Observation Data Center (NODAOP2021010); National Natural Science Foundation of China (41771428).
Dataset composition	This dataset includes vector data and summary tables of major urban parks in Wuhan (2021) saved as a compressed file (A dataset of major urban park of Wuhan in 2021.zip).