

中国城市 NDVI 十年度统计数据集 (250-m, 1990–2020) 的研发

刘海猛¹, 周天宇^{2*}, 勾 鹏²

1. 中国科学院地理科学与资源研究所, 北京 100101;
2. 南湖实验室大数据技术研究中心, 嘉兴 314002

摘 要: 植被覆盖是衡量生态环境质量和人居环境的重要指标, 然而目前还缺乏城市内部长时间序列的 NDVI 数据集。作者基于 Terre-MODIS NDVI、GIMMS NDVI 产品数据, 运用深度学习超分辨率的算法制作了 1990、2000、2010、2020 年中国 250-m 分辨率 NDVI 数据集。然后, 通过叠加不同时期城市的行政范围和实体地域范围, 提取不同城市范围内的 NDVI 平均值, 得到中国十年度及 361 个城市平均 NDVI 值统计数据集 (250-m, 1990–2020)。该数据集显示 1990–2020 年中国全境和城市范围的 NDVI 均呈先降低后升高的趋势, 但存在显著的空间异质性。该数据集可为城市生态环境治理、城市绿地规划建设、生态环境政策制定和政府绩效考核等提供基础数据支撑, 也可作为城市化与气候变化驱动下生态系统演化研究的基础数据。数据集存储为.tif 和.xlsx 格式, 空间分辨率为 250 m, 由 5 个文件组成, 数据量为 6.51 GB (压缩为 5 个文件, 1.83 GB)。

关键词: 植被覆盖; 城市化; NDVI; 城市实体地域; 深度学习; 中国

DOI: <https://doi.org/10.3974/geodp.2023.01.09>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2023.01.09>

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.2023.04.06.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2023.04.06.V1>。

1 前言

植被是生态系统的重要组成部分, 植被覆盖是衡量生态环境质量和人居环境的重要指标, 在全球环境变化研究中担任着“指示器”的角色^[1]。中国快速城市化进程极大地影响了植被覆盖^[2], 而植被覆盖降低反过来也作用于城市环境^[3], 加剧城市热岛效应^[4]、不利于空气污染物扩散^[5]、削弱区域碳汇能力^[6], 并对城市居民的健康福祉产生负面影响^[7]。让所

收稿日期: 2023-01-03; 修订日期: 2023-03-17; 出版日期: 2023-03-25

基金项目: 国家自然科学基金 (42171210); 中华人民共和国教育部 (22JJD790015)

*通讯作者: 周天宇, 南湖实验室大数据技术研究中心, zhoutianyu@nanhulab.ac.cn

数据引用方式: [1] 刘海猛, 周天宇, 勾鹏. 中国城市 NDVI 十年度统计数据集 (250-m, 1990–2020) 的研发[J]. 全球变化数据学报, 2023, 7(1): 65–74. <https://doi.org/10.3974/geodp.2023.01.09>.
<https://cstr.escience.org.cn/CSTR:20146.14.2023.01.09>.

[2] 刘海猛, 周天宇, 勾鹏. 中国十年度及 361 个城市平均 NDVI 值统计数据集 (250-m, 1990–2020) [J/DB/OL]. 全球变化数据仓储电子杂志, 2023. <https://doi.org/10.3974/geodb.2023.04.06.V1>.
<https://cstr.escience.org.cn/CSTR:20146.11.2023.04.06.V1>.

有市民享有足够的平等的城市绿地是联合国 2030 年可持续发展目标 11 的重要内容^[8]。

归一化植被指数 (Normalized differential vegetation index, NDVI) 是反映植被覆盖变化的主要指标之一。其中 GIMMS NDVI 和 MODIS NDVI 产品数据是最常用的 NDVI 数据集, 并已在区域和全球范围内广泛应用^[9, 10]。MODIS 数据被视为 GIMMS 数据的延续^[11], 但两者之间存在空间分辨率差距较大、时间覆盖不同等问题, 限制了对城市化过程中 NDVI 变化的正确评估和理解。已有的数据集多为栅格数据或城市行政范围内的数据, 且时间序列多为 2000 年之后^[12, 13], 缺乏以建设用地为主的城市实体地域范围内的长时间序列的高分辨率 NDVI 数据集。

本文分别收集了多个年份 Terre-MODIS NDVI 和 GIMMS NDVI 产品数据, 以 250-m 空间分辨率的 MODIS 数据为基准, 通过深度学习超分辨率的算法对 GIMMS 数据集进行空间降尺度, 以此延长高分辨率下 NDVI 的时空覆盖率, 得到 1990、2000、2010、2020 年中国 250-m 分辨率 NDVI 数据集。然后, 选择全球城市边界 (Global urban boundary, GUB) 数据集来厘定不同时期城市的实体地域范围, 并提取城市实体地域范围内的 NDVI 数值, 分别计算 1990、2000、2010 和 2020 年共 4 期中国 361 个城市行政范围和实体地域范围内 NDVI 平均值。该数据集可为城市生态环境治理、城市绿地规划建设、生态环境政策制定和政府绩效考核等提供基础数据支撑, 也可作为城市化与生态环境领域研究的基础数据。

2 数据集元数据简介

《中国十年度及 361 个城市平均 NDVI 值统计数据集 (250-m, 1990–2020)》^[14]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、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。

3 数据研发方法

3.1 数据来源

本研究选用的 GIMMS NDVI 数据集来自美国国家环境信息中心提供的第 3 版 NDVI 产品数据, 空间分辨率为 1/12 度 (约 8 km), 时间跨度为 1989–1991 年和 2000–2013 年。Terre-MODIS NDVI 产品数据来源于 NASA MODIS 陆地产品组根据统一算法开发的 MODIS 植被指数产品, 本研究选用的是 250-m 分辨率 16 天合成的 MOD13Q1 数据集, 以及 1-km 分辨率 16 天合成的 MOD13A2, 时间跨度均为 2000–2020 年。除此之外, 本文还选取了 Landsat 5 年际 NDVI 数据集、Aqua MODIS NDVI 数据集和 NOAA CDR AVHRR NDVI 数据集作为交叉验证数据。本研究中的城市行政区划数据为 2019 年, 来自国家基础地理信息中心, 涵盖了中国所有地级行政单元以及部分省份的直辖市。1990–2020 年城市实体地域范围来源于全球城市边界 (Global Urban Boundary, GUB) 数据集^[16]。该数据集结合了宏观尺度的核密度分析和微观尺度的邻域扩张算法, 对于城市外缘地区进行形态学处理, 采集了全球所有面积超过 1 km² 的城市及邻近居民点的边界。该数据集的城市轮廓中不仅包括不透水面, 也包括城市内部的绿地、水体等其他相关用地类型, 图 1 是截取的这个数据集中的北京市作为示例展示。

条 目	描 述
数据集名称	中国十年度及 361 个城市平均 NDVI 值统计数据集（250-m，1990–2020）
数据集短名	ChinaCitiesNDVI_1990_2020
作者信息	刘海猛 R-7364-2018, 中国科学院地理科学与资源研究所, liuhm@igsnr.ac.cn 周天宇, 南湖实验室大数据技术研究中心, zhoutianyu@nanhulab.ac.cn 勾鹏, 南湖实验室大数据技术研究中心, goupeng@nanhulab.ac.cn
地理区域	中国
数据年代	1990、2000、2010、2020 年
时间分辨率	年
空间分辨率	250 m
数据格式	.tif、.xlsx
数据量	6.51 GB
数据集组成	1990、2000、2010、2020 年中国 250-m 分辨率 NDVI 年平均数据; 1990、2000、2010、2020 年中国城市行政范围和实体地域范围 NDVI 年平均数据
基金项目	国家自然科学基金（42171210）；教育部人文社会科学重点研究基地重大项目（22JJD790015）
数据计算环境	Google Earth Engine; Python; 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	全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据集和通过《全球变化数据学报（中英文）》发表的数据论文。其共享政策如下：（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 ^[15]
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS

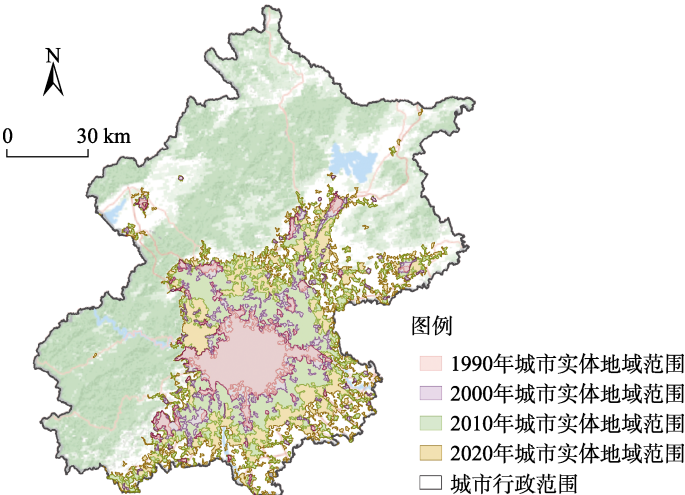


图 1 北京市城市行政范围和城市实体地域范围示意图

3.2 数据预处理

MODIS NDVI 数据被视为 GIMMS NDVI 数据的延续, 已有大量研究表明在一定的转换关系下, 两种数据结果具有极强的一致性。本文采用了 Fensholt 的换算公式(式 1)^[11], 将 GIMMS NDVI 数据值域转换至与 MODIS NDVI 一致。该公式基于 2000–2010 年期间, GIMMS NDVI 和 MODIS NDVI 每月观测值的线性回归趋势分析的回归系数所得, 并针对中国、加拿大、澳大利亚等国家开展了具有区域特点的模式拟合, 其长时间序列的结果具有一定的区域代表性^[11]。

GIMMS NDVI=1.39×MODIS NDVI-0.09

(1)

为进一步验证该换算公式在中国全域的适用性, 本文选取了位于不同区域的武汉、成都、昆明、杭州、沈阳 5 个城市, 计算 2010 年度 MODIS 数据在经过换算后与 GIMMS 数据的均方根误差 (RMSE), 以反映该公式在中国中部、西部、南部、东部和东北地区的适用性。结果如表 2 所示, MODIS NDVI 在经过式 1 换算后, 其与 GIMMS NDVI 的均方根误差平均在 0.024, RMSE 最高的城市武汉也仅为 0.034。因此, 本文认为该公式在中国范围内具有较好的适用性。

表 2 换算后 MODIS NDVI 和 GIMMS NDVI 在中国 5 个城市的均方根误差统计表

城市	武汉	成都	昆明	杭州	沈阳	平均值
RMSE	0.034	0.019	0.033	0.029	0.006	0.024

为尽可能弱化单一年份的植被指数可能受气象条件变化、极端气象灾害等因素的影响, 本数据集虽然重点选取了四个时间点 (1990、2000、2010、2020 年), 但每个时间点实际融合了 3 年的 NDVI 数据 (1989–1991 年、2000–2002 年、2009–2011 年和 2019–2021 年)。通过谷歌地球引擎平台 (Google Earth Engine, GEE) 对 GIMMS NDVI 数据和 MODIS NDVI 数据 (250-m) 进行子集提取、图像镶嵌、裁剪数据、像元融合、投影转换等预处理步骤, 并计算每个时间段内 NDVI 指数的平均值, 以代表相应时期植被生长的平均状态, 获得中国 GIMMS NDVI 数据集 (1990 年) 和 MODIS NDVI 数据集 (2000 年、2010 年、2020 年)。此外, 为给后续超分辨率模型提供训练样本, 本研究在 GEE 平台上采用同样的数据预处理步骤, 分别制作了 2000–2020 年期间 GIMMS NDVI、MODIS NDVI (1 km) 和 MODIS NDVI (250-m) 的年际尺度 NDVI 数据。

3.3 图像超分辨率深度学习算法

本研究基于 Lei 提出的一种局部-全局结合网络 (Local-Global Combined Networks, LGC-Net) 的图像超分辨率深度学习算法^[17], 分别使用 1-km 和 250-m 的 MODSI NDVI 数据集对 GIMMS NDVI 数据集进行二阶超分辨率图像重构, 实现 GIMMS NDVI 在不损失原有植被指数特征的情况下, 重采样至更高的空间分辨率。第一阶段将 2000–2010 年尺度的 GIMMS NDVI 以同年份 1-km MODIS NDVI 数据为训练对象, 实现 8-km 至 1-km 超分辨率模型训练, 并重构 1-km 分辨率下的 GIMMS NDVI 数据; 第二阶段将完成重构的 1-km GIMMS NDVI 图像以 250-m MODIS NDVI 数据为训练对象, 完成 1-km 至 250-m 超分辨率模型训练, 并将生成结果与原始 GIMMS NDVI 结果进行误差比对, 进一步优化超分辨率模

型的精度。最后，将已完成数据预处理的 1990 年 GIMMS NDVI 数据输入这个二阶超分辨率模型，重构 250-m 分辨率的 1990 年 GIMMS NDVI 数据。LGC-Net 的算法原理如图 2 所示：（1）利用多层卷积，获取不同层的主要特征表示；（2）在局部-全局信息结合部分，通过将不同层的结果级联起来，实现一种多叉结构；（3）最后在重建部分对图像进行更高分辨率的重构。

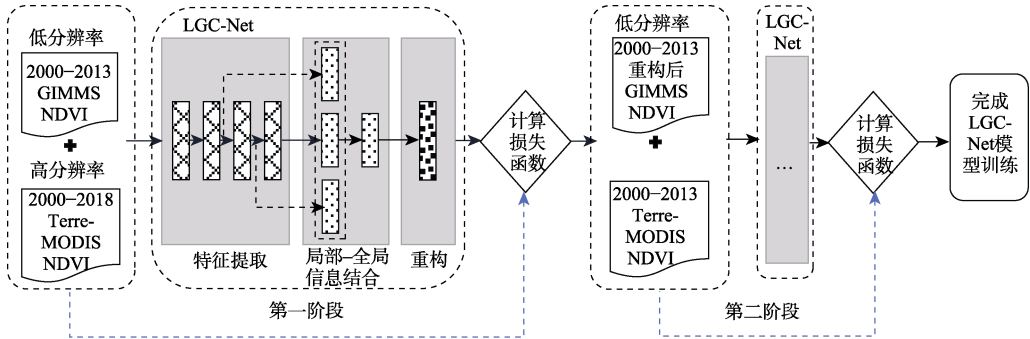


图 2 LGC-Net 图像超分辨率深度学习算法流程图

3.4 技术路线

本数据集的主要研发过程如图 3 所示。首先，按照前文数据预处理的步骤，分别对 2000–2013 年度 GIMMS NDVI 数据集、2000–2013 年度 Terre-MODIS 数据集（1-km 和 250-m）进行数据预处理，并将这些数据作为训练集输入 LGC-Net 深度学习模型进行训练，获得图像超分辨率模型。然后，将预处理后的低分辨率 1990 年 GIMMS NDVI 数据集输入至该模型，进行图像第二阶段重构，获得重构后的 250-m 分辨率 GIMMS NDVI 数据集。对重构后的 GIMMS 和 MODIS NDVI 数据集进行整理、配准和统一坐标投影，并进行数据精度验证，得到中国全域 1990、2000、2010 和 2020 年 250-m NDVI 数据集。最后，将全国地级行政区矢量数据、全国 GUB 矢量数据与 NDVI 数据集进行空间叠加，裁剪并提取城市行政边界和实体边界内所有的像素点并计算他们的 NDVI 均值，得到中国城市 250-m 分辨率植被覆盖数据集（1990–2020 年）。

4 数据结果与验证

4.1 数据集组成

《中国十年度及 361 个城市平均 NDVI 值统计数据集（250-m，1990–2020）》的组成内容、数据大小、数据类型和命名方式如表 3 所示，主要包括 4 个栅格数据和 2 个统计表格数据。

4.2 数据结果

4.2.1 1990–2020 年中国 NDVI 的时空变化

图 4 展示了 1990–2020 年 250-m 分辨率下中国全域 NDVI 分布。去除水体影响部分，1990、2000、2010 和 2020 年中国范围 NDVI 平均值分别为 0.316、0.299、0.287、0.297，呈现一个先下降后上升的变化趋势，即中国总体的国土绿度在 1990–2010 年呈现下降趋势，

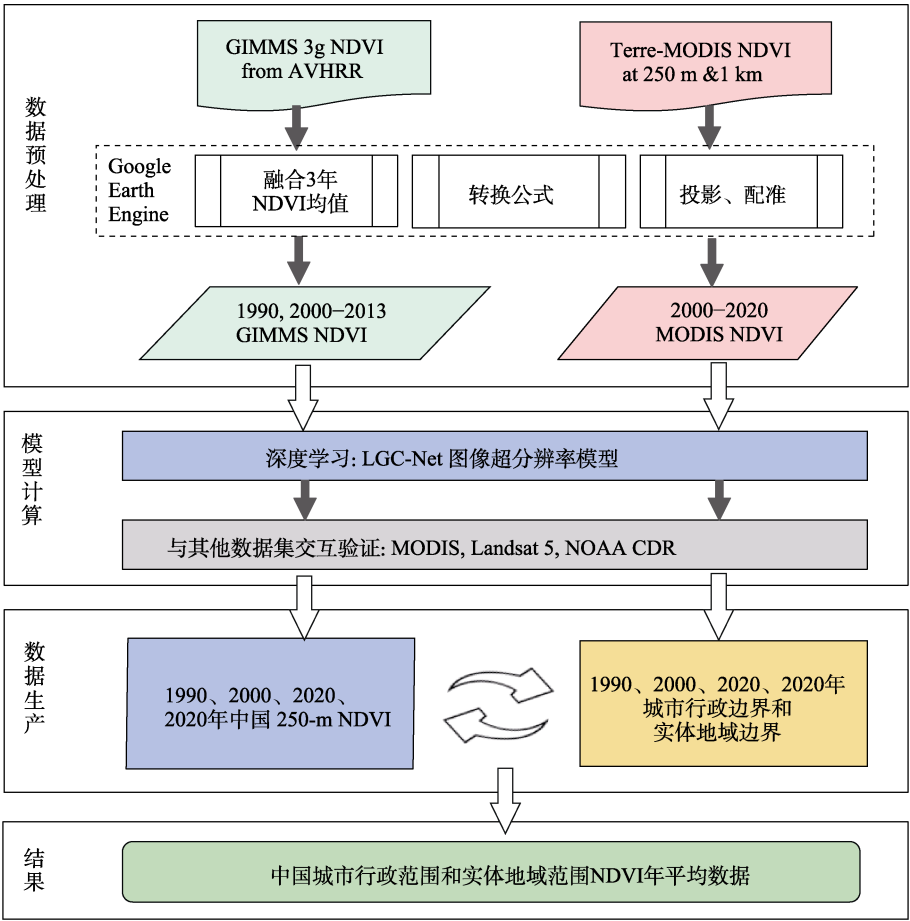


图 3 技术路线流程图

表 3 中国十年度及 361 个城市平均 NDVI 值统计数据集（250-m，1990–2020）包含数据列表

数据名称	对应文件	数据大小	数据格式
1990、2000、2010、2020 年中国 250-m 分辨率 NDVI 年平均数据	China_1990_2020_NDVI	6.51 GB	.tif
1990、2000、2010、2020 年中国城市行政范 围和实体地域范围 NDVI 年平均数据	ChinaCities_1990_2020_NDVI	81 KB	.xlsx

但近 10 年改善明显。从空间格局来看，在 1990 至 2020 年间，中国植被指数空间分布格局相对稳定。受自然地理条件和气候因素影响，NDVI 值总体呈现从东南沿海向西北内陆地区逐级递减。高值（ $\text{NDVI}>0.6$ ）主要分布在中国东南丘陵、沿海地区和南部的云贵高原地区；中值（ $0.6>\text{NDVI}>0.3$ ）主要分布在中国东部的华北平原、西部的四川盆地和东北部分地区；低值（ $\text{NDVI}<0.3$ ）主要分布在西部的青藏高原、塔里木盆地，以及北部的内蒙古高原。对比图 4a 和图 4d 可以发现，从 1990–2020 年，植被绿度在南方大部分地区和黄土高原地区普遍增强，长三角城市群、京津冀城市群、珠三角城市群、山东半岛城市群的城镇密集区植被绿度减少明显，其他植被绿度减少明显的区域主要分布的西部地区。

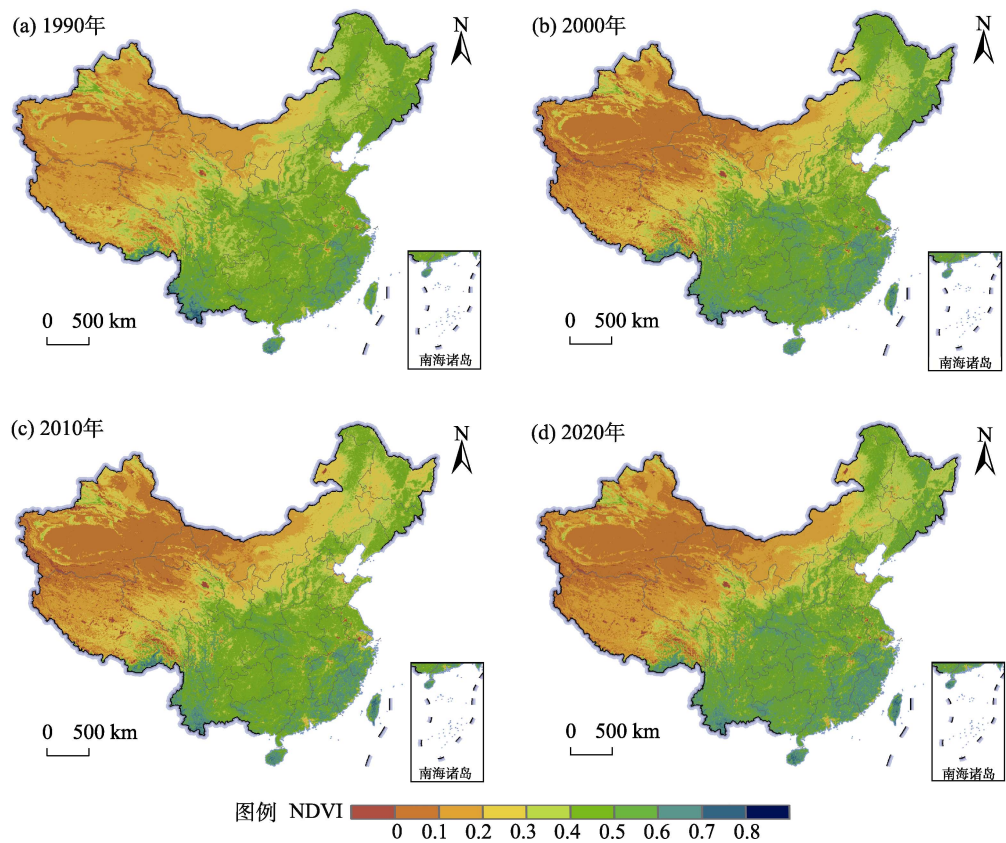


图 4 1990–2020 年中国 NDVI 时空格局图（依据审图号 GS(2019)1823 号标准地图制作）

4.2.2 1990–2020 年中国城市行政范围和实体地域范围内 NDVI 的时空变化

图 5 展示了 1990–2020 年中国 361 个城市行政范围和实体地域范围 NDVI 的时空变化。城市行政范围内 NDVI 的时空分布及变化趋势与上文栅格尺度的特征基本一致，但城市实体地域范围的 NDVI 并不完全符合栅格尺度的规律，例如在长三角、珠三角、长江中游城市群的部分城市尽管行政区域内 NDVI 平均值比较高，但城市实体地域范围 NDVI 值比较低；新疆的和田市、昆玉市在行政区域内 NDVI 值比较低，但城市实体地域范围 NDVI 值却比较高。除了西北部分绿洲型城市之外，城市行政范围的 NDVI 值均大于城市实体地域范围的值。1990、2000、2010 和 2020 年全国城市实体地域范围 NDVI 平均值分别为 0.321、0.278、0.264、0.290，可见近十年全国的城市绿度有显著提升。

4.3 数据结果验证

由于 NDVI 数值缺乏地面站点监测直接证据，我们将不同 NDVI 产品数据统一到相同的空间分辨率和投影坐标系中，分析通过深度学习模型生成的本数据集与其他数据产品之间的相关性来验证本数据集的时空可靠性。本文以北京市为验证区域，大小约为 16,392.99 km²，对应 343,967 个分辨率为 250 m 的像元。将本研究生成的 2010 年 250-m NDVI 数据与 Aqua-MODIS NDVI、Landsat 5 NDVI、NOAA NDVI 三个数据产品进行逐像素的对比分析，分别计算相关系数、 R^2 、均方根误差（RMSE）等。从图 6 结果来看，

250-m NDVI数据结果与其他NDVI产品数据均有较强的相关性,相关系数分别为0.991,4、0.842,5、0.790,3。总体来看,本文生产的 250-m NDVI 数据集整体精度较高,数据结果具有一定的可靠性。

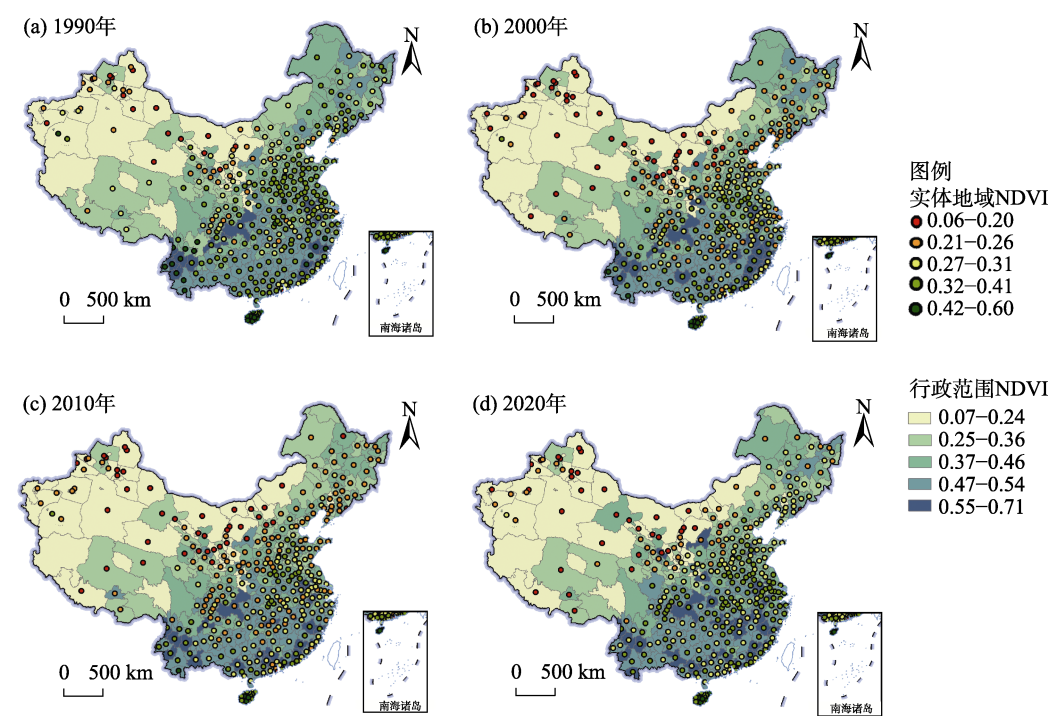


图 5 1990–2020 年中国城市行政范围和实体地域范围 NDVI 时空变化
(依据审图号 GS(2019)1823 号标准地图制作)

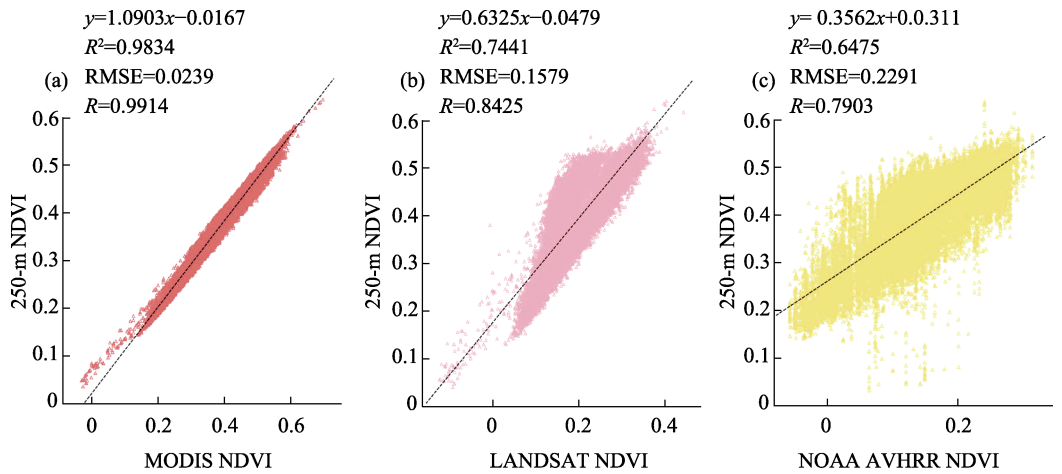


图 6 模型拟合 NDVI 与其他 NDVI 产品数据交叉验证结果图 (a、b、c 的交叉验证对象分别为:
Aqua-MODIS NDVI、Landsat 5 NDVI 和 NOAA NDVI)

5 讨论和总结

本文基于30年的GIMMS NDVI数据和中高时空分辨率的Terre/MODIS NDVI数据，结合GUB城市边界数据集，运用Google Earth Engine平台和Python，采用图像超分辨率算法和交叉验证的方法，制作了1990–2020年中国城市250-m分辨率植被覆盖数据集。与已有的NDVI产品数据相比，本数据集将原先中高分辨率NDVI的观测时间从20年延长至30年，并通过叠加城市实体地域边界数据，计算了城市建成实体地域范围的NDVI。之所以本数据集中的NDVI选择了年度平均值，而没有采用NDVI年最大值，是考虑到年度平均值更能反映植被对人类福祉的全年贡献，更适应于在人与自然耦合研究、城市生态环境整体评价、城市化环境效应、城市生态治理等方向的应用。

本数据结果表明，1990–2020年中国全境和城市范围的NDVI均呈现先降低后升高的趋势，但区域间存在明显的异质性。虽然城市化进程使得城市周边大量植被覆盖区被转变为不透水地表，但城市建成范围内的NDVI在近十年呈现明显提升。这些结果表明城市绿地建设、退耕还林、植树造林、自然保护区建设等生态环境改善措施和修复工程已取得明显的成效，城市化正在给植被带来更多积极的影响。本数据集可以为后续进一步揭示更长时间序列下，中国植被覆盖对气候变化、人类活动和其他驱动因子的响应机理提供基础数据，也可以为指导政府部门评估区域生态环境质量、制定生态保护措施和区域可持续发展提供数据支撑。

作者分工：刘海猛对数据集的开发做了总体设计，撰写数据论文和数据可视化；周天宇采集和处理了数据，设计了模型和算法，撰写数据论文；勾鹏做了数据验证。

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] Duveiller, G., Hooker, J., Cescatti, A. The mark of vegetation change on Earth's surface energy balance [J]. *Nature Communications*, 2018, 9(1): 679.
- [2] Zhang, L., Yang, L., Zohner, C. M., et al. Direct and indirect impacts of urbanization on vegetation growth across the world's cities [J]. *Science Advances*, 2022, 8(27): eabo0095.
- [3] 王静, 周伟奇, 许开鹏等. 京津冀地区城市化对植被覆盖度及景观格局的影响 [J]. *生态学报*, 2017, 37(21): 7019–7029.
- [4] Zhou, D., Zhao, S., Zhang, L., et al. The footprint of urban heat island effect in China [J]. *Scientific Reports*, 2015, 5(1): 11160.
- [5] Liu, H., Cui, W., Zhang, M. Exploring the causal relationship between urbanization and air pollution: evidence from China [J]. *Sustainable Cities and Society*, 2022, 80: 103783.
- [6] Liu, X., Wang, S., Wu, P., et al. Impacts of urban expansion on terrestrial carbon storage in China [J]. *Environmental Science & Technology*, 2019, 53(12): 6834–6844.
- [7] Cunsolo, W. A., Harper, S. L., Ford, J. D., et al. Climate change and mental health: an exploratory case study from Rigolet, Nunatsiavut, Canada [J]. *Climatic Change*, 2013, 121: 255–270.

-
- [8] Cheng, Y., Liu, H., Wang, S., *et al.* Global action on SDGs: policy review and outlook in a post-pandemic era [J]. *Sustainability*, 2021, 13(11): 6461.
- [9] Zhou, D., Zhao, S., Zhang, L., *et al.* Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities [J]. *Remote Sensing of Environment*, 2016, 176: 272–281.
- [10] Zhou, T., Liu, H., Gou, P., *et al.* Conflict or Coordination? Measuring the relationships between urbanization and vegetation cover in China [J]. *Ecological Indicators*, 2023, 147: 109993.
- [11] Fensholt, R., Proud, S. R. Evaluation of earth observation based global long term vegetation trends—comparing GIMMS and MODIS global NDVI time series [J]. *Remote sensing of Environment*, 2012, 119: 131–147.
- [12] 徐洋, 杨雅萍. 1982–2020 年中国 5-km 分辨率逐月 NDVI 数据集[DS]. 2022.
- [13] 董金玮, 周岩, 尤南山. 2000–2020 年中国 30 米年最大 NDVI 数据集[DS]. 2021.
- [14] 刘海猛, 周天宇, 勾鹏. 中国十年度及 361 个城市平均 NDVI 值统计数据集 (250-m, 1990–2020) [J/DB/OL]. 全球变化数据仓储电子杂志, 2023. <https://doi.org/10.3974/geodb.2023.04.06.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2023.04.06.V1>.
- [15] 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).
- [16] Li, X., Gong, P., Zhou, Y., *et al.* Mapping global urban boundaries from the global artificial impervious area (GAIA) data [J]. *Environmental Research Letters*, 2020, 15(9): 094044.
- [17] Lei, S., Shi, Z., Zou, Z. Super-Resolution for remote sensing images via local–global combined network [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(8): 1243–1247.