

中国长时间序列夜间灯光数据集的研发 (2000–2020)

仲晓雅^{1,2}, 闫庆武^{1,3*}, 李桂娥^{1,3}

1. 中国矿业大学中国资源型城市转型发展与乡村振兴研究中心, 徐州 221116;

2. 中国矿业大学环境与测绘学院, 徐州 221116; 3. 中国矿业大学公共管理学院, 徐州 221116

摘要: 夜间灯光影像能够反映地表夜间灯光亮度, 从而在一定程度上表征人类活动强度, 长时间序列的夜间灯光数据对于城市发展研究具有重要意义。基于 DMSP/OLS 第四版非辐射定标夜间年平均灯光强度数据和 NPP/VIIRS 月度数据, 经过对 DMSP/OLS 年度数据、NPP/VIIRS 月度数据进行数据预处理、校正和数据融合获得了中国 2000–2020 年的夜间灯光数据。该数据集主要包括基于 DMSP/OLS 年度数据和 EVI 数据修正获得的 2000–2013 年 EANTLI 年度数据集、2012 年 4 月至 2020 年 12 月 NPP/VIIRS 月度数据集、2012–2020 年 NPP/VIIRS 年度数据集和 2000–2020 年类 EANTLI 夜间灯光数据集, 其中 NPP/VIIRS 月度数据集和年度数据集的空间分辨率为 500 m, 其余数据空间分辨率为 1 km。数据集存储为.tif 格式, 由 750 个文件组成, 数据量为 2.21 GB (压缩为 6 个文件, 1.71 GB)。

关键词: DMSP/OLS; NPP/VIIRS; 夜间灯光遥感; 数据融合

DOI: <https://doi.org/10.3974/geodp.2022.03.12>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2022.03.12>

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.2022.06.01.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2022.06.01.V1>.

1 前言

夜间灯光数据具有在夜间探测地表发出的城市灯光、自然界火光等光源强度的能力, 从而体现城市夜间灯光的特征^[1]。夜间灯光数据不仅具有大尺度、多时相的特性, 还能够直观地表征人类活动强度, 全球卫星观测的夜间灯光数据已成为广泛使用的地理空间数据产品之一^[2]。其中 DMSP/OLS 和 NPP/VIIRS 数据作为最为常用的夜间灯光数据, 在城市扩张、空气污染和社会经济指标估算等方面都有广泛的应用^[3–5]。由于 DMSP/OLS 和 NPP/VIIRS 数据存在传感器不同、时空分辨率不同、数据处理方式不同、像元值含义不同

收稿日期: 2022-06-28; 修订日期: 2022-09-06; 出版日期: 2022-09-25

基金项目: 中央高校基本科研业务费专项资金资助 (2021ZDPY0205); 科技基础性工作专项项目 (2014FY110800); 国家自然科学基金项目 (42101459)

*通讯作者: 闫庆武, 中国矿业大学中国资源型城市转型发展与乡村振兴研究中心, yanqingwu@cumt.edu.cn

数据引用方式: [1] 仲晓雅, 闫庆武, 李桂娥. 中国长时间序列夜间灯光数据集的研发 (2000–2020) [J]. 全球变化数据学报, 2021, 6(3): 416–424. <https://doi.org/10.3974/geodp.2022.03.12>. <https://cstr.escience.org.cn/CSTR:20146.14.2022.03.12>.

[2] 仲晓雅, 闫庆武, 李桂娥. 中国长时间序列夜间灯光数据集 (2000–2020) [J/DB/OL]. 全球变化数据仓储电子杂志, 2022. <https://doi.org/10.3974/geodb.2022.06.01.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2022.06.01.V1>.

等问题^[6]，长时间序列夜间灯光数据整合存在一定的难度。这些数据问题严重限制了夜间灯光数据的应用潜力，也会对基于夜间灯光数据研究的效果造成一定的影响。

本数据集收集了 2000–2013 年的 DMSP/OLS 数据和 2012 年 4 月至 2020 年 12 月的 NPP/VIIRS 数据，通过对 DMSP/OLS 年度数据处理及 NPP/VIIRS 月度数据进行年度数据合成，获得了 EANTLI 数据和 NPP/VIIRS 月度数据合成年度数据。通过融合 EANTLI 数据和 NPP/VIIRS 年度数据获得了长时间序列的类 EANTLI 数据。可以作为研究城市发展的有效工具，为城市规划、城市扩张、城市收缩和城市结构等领域提供基础数据支撑。

2 数据集元数据简介

《中国长时间序列夜间灯光数据集（2000–2020）》^[7]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。

表 1 《中国长时间序列夜间灯光数据集》元数据简表

条 目	描 述
数据集名称	中国长时间序列夜间灯光数据集（2000–2020）
数据集短名	NTLChina_2000-2020
作者信息	仲晓雅，中国矿业大学环境与测绘学院，851676389@qq.com 闫庆武，中国矿业大学中国资源型城市转型发展与乡村振兴研究中心， yanqingwu@cumt.edu.cn
地理区域	中国 数据年代 2000–2020 时间分辨率 月、年
数据格式	.tif 数据量 2.21 GB（压缩后为 1.71 GB） 空间分辨率 500 m、1,000 m
数据集组成	EANTLI 年度数据、NPP/VIIRS 月度数据、NPP/VIIRS 年度数据和类 EANTLI 年度数据
基金项目	教育部（2021ZDPY0205）；中华人民共和国科学技术部（2014FY110800）；国家自然科学基金（42101459）
数据计算环境	ArcGIS、Google Earth Engine、Origin
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 ^[8]
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS

3 数据研发方法

3.1 数据来源

本研究选用的 DMSP/OLS 数据来自 DMSP/OLS 第四版非辐射定标夜间年平均灯光强

度数据中 4 个不同传感器 (F14–F18) 的 2000–2013 年总计 22 期稳定夜间灯光数据进行实验, 下载自美国国家海洋与大气管理局 (NGDC) 下属的美国国家地球物理数据中心¹。本研究使用的 NPP/VIIRS 月度产品由科罗纳矿业学院的佩恩研究所的地球观测小组 (Earth Observation Group, EOG) 生产, 下载自 EOG 团队网站²。

本研究还使用了年均 EVI 数据 (增强植被指数), 用于对 DMSP/OLS 数据进行饱和校正, 减小 DMSP/OLS 数据的饱和效应。该数据在 MOD13A1 EVI 数据的基础上, 使用 Google Earth Engine 进行了年平均 EVI 数据合成。

3.2 算法原理

本研究主要基于 DMSP/OLS 和 NPP/VIIRS 夜间灯光数据, 通过分别对 DMSP/OLS 数据、NPP/VIIRS 月度数据进行数据预处理、校正和数据融合获得了覆盖 2000–2020 年的多种夜间灯光数据。

3.2.1 DMSP/OLS 数据处理

通过对 DMSP/OLS 数据进行数据预处理、相互校正与饱和校正获得了 2000–2013 年的 EANTLI 数据。其中, DMSP/OLS 数据的相互校正包括不同传感器之间的相互校正^[9]、多传感器年内融合和时间序列可比较校正^[10]。为缓解 DMSP/OLS 的过饱和问题, 采用 EVI 增强植被指数修正后的夜间灯光指数 (EANTLI)^[11]对相互校正后的 2000–2013 年的 DMSP/OLS 数据进行饱和校正。EANTLI 指数的公式见下式:

$$EANTLI = \frac{1 + (nNTL - EVI)}{1 - (nNTL - EVI)} \times NTL \quad (1)$$

式中, NTL 为 DMSP/OLS 像元值, $nNTL$ 为归一化的 NTL , EVI 为年平均 EVI 值。

3.2.2 NPP/VIIRS 数据处理

通过对 NPP/VIIRS 月度数据进行数据预处理、异常值处理和基于中位数的年度合成获得了 2012 年 4 月至 2020 年 12 月 NPP/VIIRS 月度数据和 2012–2020 年的 NPP/VIIRS 年度数据。其中数据的异常值处理包括 NPP/VIIRS 数据的噪声去除和极高值处理。

在对 NPP/VIIRS 月度数据进行噪声处理时, 发现 NPP/VIIRS 月度数据暗像元 (亮度值低于 $1 \text{ nW} \cdot \text{cm}^{-2} \cdot \text{sr}^{-1}$ 的像元) 直方图中 95% 分割点处月度暗像元和年均暗像元数量的差值趋于一致。基于此, 提出了一种基于灵活阈值的方法进行噪声去除 (noise reduction based on flexible threshold, NRBFT), 即将影像暗像元直方图中的 95% 分割点作为噪声的最小阈值, 在此基础上去除暗像元中的噪声 (见式 2)。

$$DN_{(n,i)} = \begin{cases} 0 & DN_{(n,i)} < TH \\ DN_{(n,i)} & \text{其他} \end{cases} \quad (2)$$

式中, $DN_{(n,i)}$ 为 NPP/VIIRS 数据的像元辐亮度, TH 为噪声的最小阈值 (影像暗像元直方图中的 95% 分割点)。

以 2020 年 12 月的 NPP/VIIRS 影像为例 (见图 1), 当阈值选取 95% 分割点时, 2020 年 12 月暗像元数量与 2020 年平均数量的差值趋于一致, 能够剔除绝大部分的噪声; 图 1

¹ 美国国家地球物理数据中心网站. <https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>.

² EOG 团队网站. <https://eogdata.mines.edu/products/vnl/>.

（b）显示了 2020 年 12 月 NPP/VIIRS 暗像元的空间分布，直方图 95%分割点以内的暗像元包含了大部分的噪声，而其余暗像元大多为正常暗像元分布在亮像元周边。

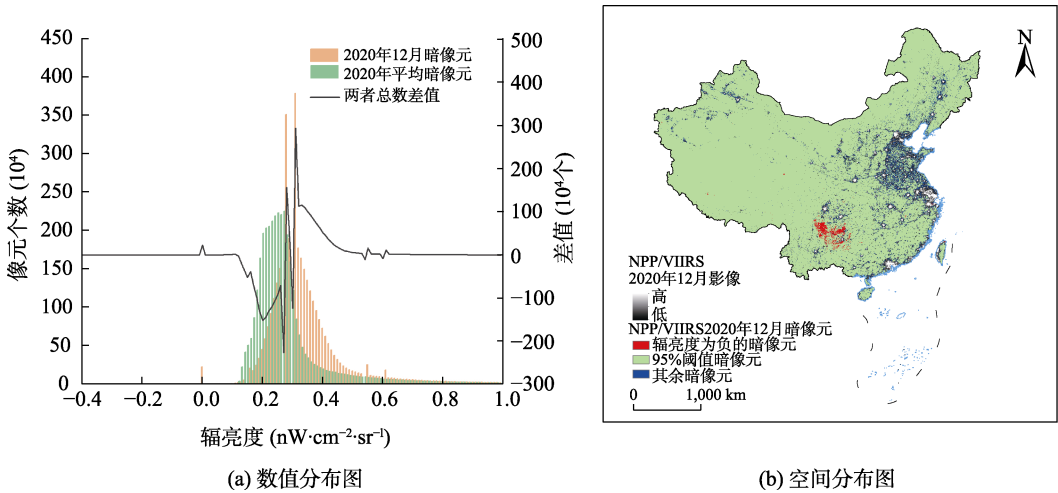


图 1 中国 2020 年 12 月 NPP/VIIRS 影像暗像元分布图
（注：图 1b 参照审图号 GS2019（1822）的标准地图制作）

由于高反射率表面影响使得 NPP/VIIRS 月度数据中部分像元亮度值极高，不符合实际情况，因此需要对影像中的极高值进行消除。本研究中，假设其他地区的夜间灯光亮度不超过大城市中心（上海市、北京市和广州市）的最大夜间灯光亮度^[12]，使用 5×5 的均值滤波对亮度值超过极值的异常像元进行动态剔除。图 2 显示了上海市、北京市和广州市的逐月最高值以及极高值处理时采用的极值参考值。

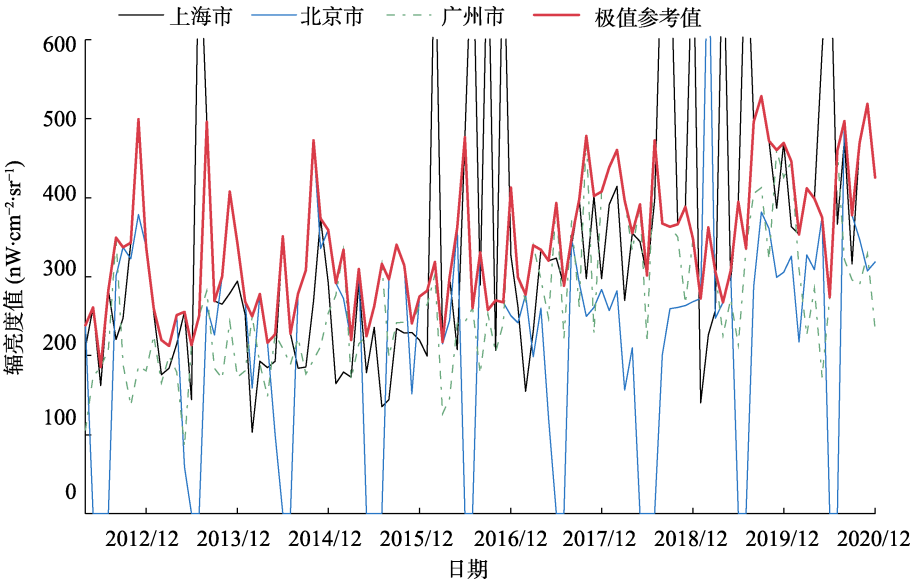


图 2 NPP/VIIRS 月度影像逐月最高值

3.2.3 长时间序列夜间灯光数据融合

传统研究在对不同传感器夜间灯光数据进行融合时，难以兼顾两种数据的准确性以及像元值的动态范围。本研究对 EANTLI 数据和 NPP/VIIRS 数据，进行了长时间序列夜间灯光数据融合：首先，基于核密度对 NPP/VIIRS 年度合成数据进行空间重采样^[13]，统一不同夜间灯光数据的空间分辨率；其次通过对数变换、拟合模型构建和指数变换，统一 EANTLI 数据和 NPP/VIIRS 数据的像元亮度值；最后，对 2000–2020 年的夜间灯光数据进行时间序列连续校正^[10]，获得长时间序列的类 EANTLI 数据。

通过对数变换，EANTLI 和 NPP/VIIRS 数据具有类似“S”形曲线的对应关系（见图 3），因此使用具有“S”型曲线特征的 logistic 模型进行拟合，其中 Boltzmann 函数（见式 3）拟合效果最佳。由于 2012 年的 NPP/VIIRS 月度数据存在 1–3 月的数据缺失，为保证拟合结果的有效性，使用 2013 年对数变换后的 EANTLI 和 NPP/VIIRS 数据进行非线性拟合，获得拟合参数。

$$\log EANTLI_i = A + \frac{A - B}{1 + e^{(\log NPP/VIIRS_i - C)/D}} \tag{3}$$

式中， $\log NPP/VIIRS$ 为对数变换后 NPP/VIIRS 像元值， $\log EANTLI_i$ 为使用对数变换后 NPP/VIIRS 像元值通过 Boltzmann 函数拟合获得的类 EANTLI 数据的像元对数值。A、B、C、D 为 Boltzmann 函数中的参数，拟合后分别为-1.395,1、6.966,3、1.125,7 和 1.304,4。

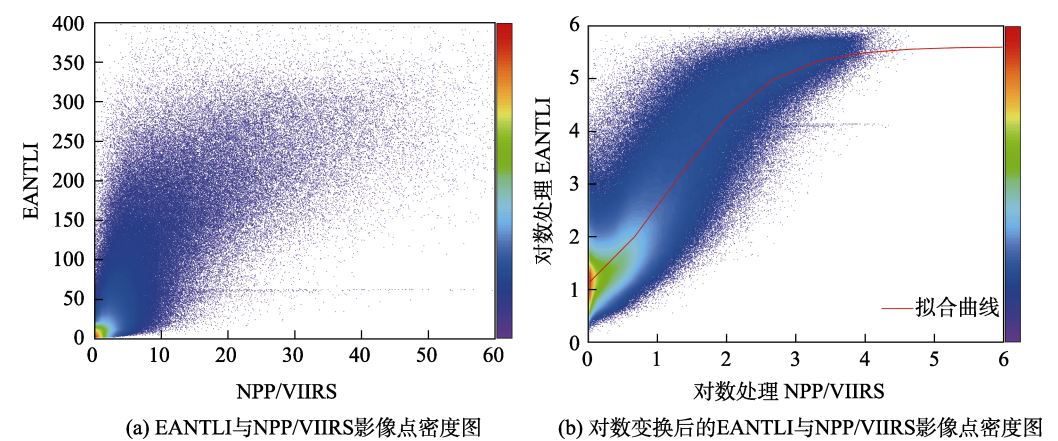


图 3 中国 2013 年夜间灯光影像点密度图

3.3 技术路线

如图 4 所示，数据集的主要研发过程为：首先，对 2000–2013 年的 DMSP/OLS 影像以及 2012–2020 年的 NPP/VIIRS 月度影像进行数据处理，通过对 DMSP/OLS 影像进行数据预处理、相互校正与饱和校正获得年度 DMSP/OLS 数据（年度 EANTLI 数据）；同时，对 NPP/VIIRS 影像进行数据预处理、异常值处理以及基于中位数的年度合成获得年度 NPP/VIIRS 数据；最后，对年度 EANTLI 数据和 NPP/VIIRS 数据进行数据融合，通过对二者进行空间分辨率的整合、拟合函数构建以及连续校正，获得长时间序列夜间灯光数据（类 EANTLI 数据）。

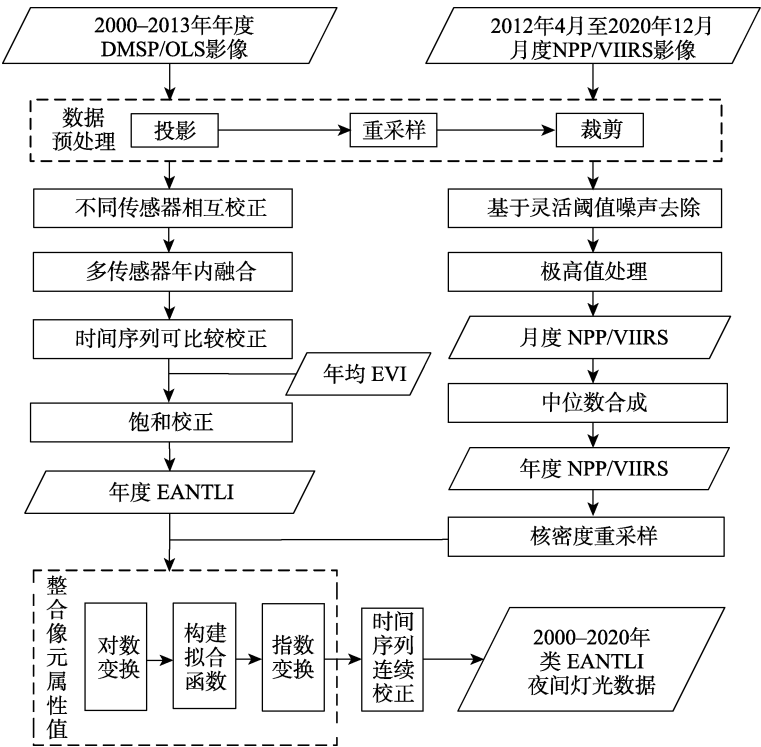


图 4 数据研发的技术路线图

4 数据结果与验证

4.1 数据集组成

中国长时间序列夜间灯光数据集主要包括 2000–2013 年 EANTLI 年度数据集、2012 年 4 月至 2020 年 12 月 NPP/VIIRS 月度数据集、2012–2020 年 NPP/VIIRS 年度数据集和 2000–2020 年类 EANTLI 夜间灯光数据集（表 2）。

表 2 数据集文件表

文件夹名称	命名方式	数据说明	数据格式	数据数量	数据量
EANTLI	EANTLI_yyyy	yyyy 年的 EANTLI 数据	.tif	14 条	122 MB
NPP_VIIRS 月 度数据	NPP_VIIRS_yyyymm	yyyy 年 mm 月的 NPP/VIIRS 数据	.tif	105 条	1.26 GB
NPP_VIIRS 年 度数据	NPP_VIIRS_yyyy	yyyy 年的 NPP/VIIRS 数据	.tif	9 条	121 MB
类 EANTLI	EANTLI_Like_yyyy	yyyy 年的类 EANTLI 数据	.tif	21 条	219 MB

注：每条数据包含 5 个文件。其中.tif 为影像文件，.tfw 为影像坐标文件，tif.ovr 为金字塔文件，.tif、.aux、.xml 为影像辅助文件，.tif、.xml 为影像文本信息。

4.2 数据结果

为体现夜间灯光变化趋势，对中国六大区域的夜间灯光亮度总和和亮像元总数进行分区统计。由图 5(a)可知，除了东北地区和华北地区，其余四个区域的夜间灯光亮度总和都

呈现出持续上升的趋势，其中经济水平较为发达的华东地区夜间灯光亮度总和上升最为显著。东北地区和华北地区的夜间灯光总和都在 2011–2014 年期间出现了一定程度的下降，这可能是因为在在此期间北部地区部分城市出现了人口流失的现象，城市夜间活力下降。由图 5(b)可知，除了东北地区，其他五个区域的亮像元总数都呈现出增长的趋势，其中西北地区 and 西南地区在 2012–2013 年期间亮像元数显著增加，这可能是西部开发所兴起的基础设施建设所致，另外，由于 NPP/VIIRS 原始数据的分辨率高于 DMSP/OLS 数据，因此可能保留了更多西部地区夜间灯光面积较小的像元亮度。东北地区的亮像元总数出现多次较大幅度的下降，这表明该地区不仅存在人口流失，还存在城市收缩的现象，部分区域的夜间灯光变暗甚至消失。

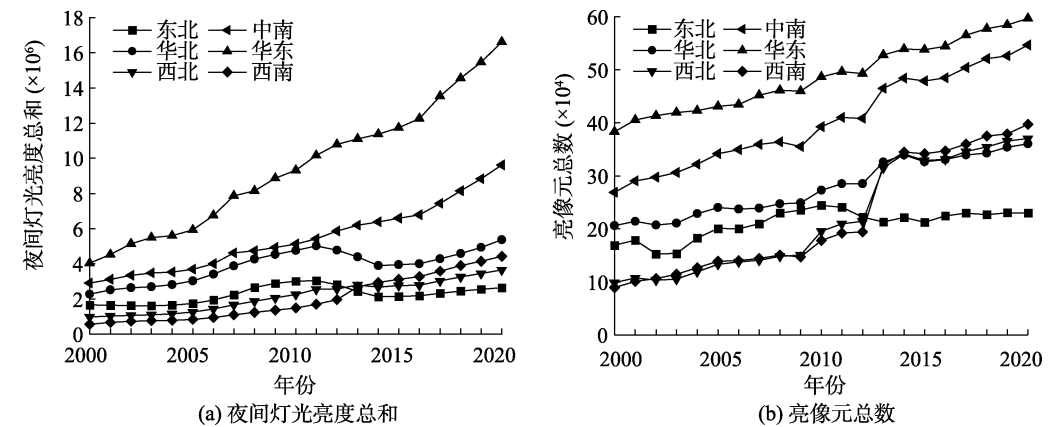


图 5 中国六大区域夜间灯光年际变化

4.3 数据结果验证

将本研究获得的长时间序列夜间灯光数据（类 EANTLI 数据）与 Chen^[14]等工作进行了对比。由于类 EANTLI 数据和类 NPP/VIIRS 数据的空间分辨率不同，因此在对二者进行对比评价时，应更加关注数值的变化而不是数值大小。由图 6 可见，类 EANTLI 夜间灯光亮度均值、标准差、信息熵和平均梯度均大于类 NPP/VIIRS 夜间灯光数据，类 EANTLI 夜间灯光数据各项指标变化趋势较为平滑，表现为稳定上升的趋势，而类 NPP/VIIRS 夜间灯光数据各项指标存在一定波动，变化不稳定。

夜间灯光数据与社会经济数据具有较强的相关性，因此图 7 展示了省级尺度下类 EANTLI 数据和类 NPP/VIIRS 数据与 GDP 和电力消费的线性回归决定系数 R^2 的对比结果。由图 7 (a) 可知，除了 2018–2020 年三年，其余年份类 EANTLI 与 GDP 的 R^2 都优于类 NPP/VIIRS 数据，二者与 GDP 的 R^2 平均值分别为 0.851,6 和 0.787,7，省级尺度下的类 EANTLI 比类 NPP/VIIRS 与 GDP 的相关性显著增加。而在图 7(b)中，类 EANTLI 数据与电力消费的 R^2 数值始终高于类 NPP/VIIRS，二者平均值分别为 0.859,7 和 0.768,9，说明在省级尺度下，类 EANTLI 数据与社会经济参数相关性更高。

5 讨论和总结

基于 DMSP/OLS 年度数据和 NPP/VIIRS 月度数据两种夜间灯光影像数据，通过影像

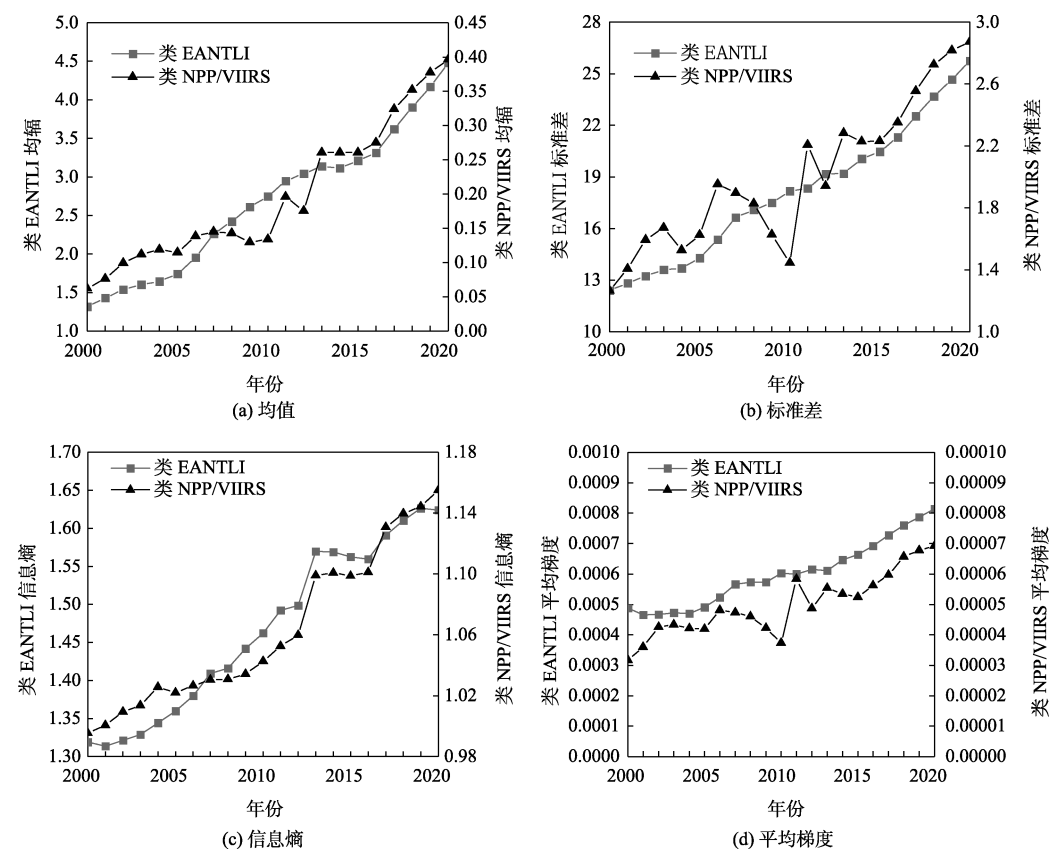


图 6 长时间序列夜间灯光影像质量统计

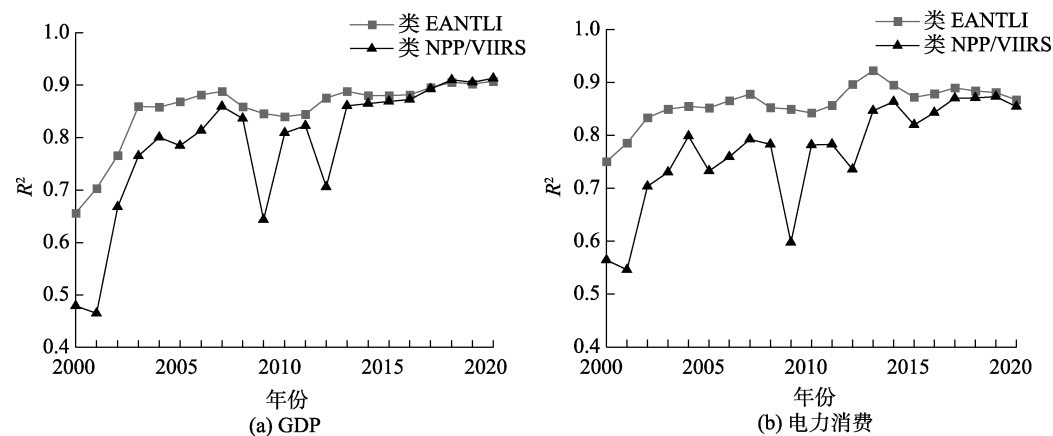


图 7 省级尺度下长时间序列夜间灯光数据与 GDP 和电力消费的线性回归决定系数

校正和数据融合获得了中国长时间序列夜间灯光数据，其中 EANTLI 数据和 NPP/VIIRS 月度数据和年度数据作为中间数据成果提交共享。数据集覆盖了 2000–2020 年的全部年度夜间灯光影像，以及 2012 年 4 月至 2020 年 12 月的全部月度夜间灯光影像。在对 NPP/VIIRS 月度数据进行异常值处理时，提出了基于灵活阈值去除噪声的方法。该方法无需借助其他辅助数据，就能够去除大部分噪声，同时还可以保留正常亮度的暗像元。

已有研究在构建长时间序列夜间灯光数据时,大多基于 DMSP/OLS 和 NPP/VIIRS 数据进行模型构建,获得较低分辨率的类 DMSP/OLS 数据^[13],或结合 EVI 数据使用机器学习获得较高分辨率的类 NPP/VIIRS 数据^[14]。前者没有对 DMSP/OLS 数据的过饱和效应进行处理,而是降低 NPP/VIIRS 数据亮度的值域,难以完全发挥 NPP/VIIRS 数据亮度值高动态范围的优势;后者,使用机器学习的方式能够提升 DMSP/OLS 数据空间分辨率和亮度的值域范围,但没有对时间序列中的不稳定像元进行处理,对数据的时间连续性有一定的影响。本文结合 EVI 数据对 DMSP/OLS 数据进行饱和校正,获得了 EANTLI 数据,减轻了 DMSP/OLS 数据的过饱和效应并提高了 DMSP/OLS 数据亮度的动态范围;在统一 EANTLI 和 NPP/VIIRS 数据的时间和空间分辨率后,通过连续校正保证了像元亮度在时间序列的一致性,去除了不稳定像元。该数据集的长时间序列夜间灯光数据拥有较强的时间连续性,与社会经济数据相关性强,精度验证表明数据集结果可靠。数据集可用于中国城市夜间灯光演变分析,服务于城市活力评价、城市扩张、城市收缩等城市发展研究。

作者分工: 仲晓雅完成了数据集的采集和处理工作,并撰写了数据论文;闫庆武提供了对数据集开发的整体思路和对数据论文的修改审定;李桂娥进行了数据的验证工作。

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] 郑启明. 夜间灯光遥感数据处理关键技术及多中心城市监测研究[D]. 杭州: 浙江大学, 2020.
- [2] 陈颖彪, 郑子豪, 吴志峰等. 夜间灯光遥感数据应用综述和展望[J]. 地理科学进展, 2019, 38(2): 205–223.
- [3] 董鹤松, 李仁杰, 李建明等. 基于 DMSP-OLS 与 NPP-VIIRS 整合数据的中国三大城市群城市空间扩展时空格局[J]. 地球信息科学学报, 2020, 22(5): 1161–1174.
- [4] Zhao, X., Xia, X., Liu, X., *et al.* Combining DMSP/OLS nighttime light with echo state network for prediction of daily PM_{2.5} average concentrations in Shanghai, China [J]. *Atmosphere*, 2015, 6(10): 1507–1520.
- [5] Elvidge, C. D. A fifteen year record of global natural gas flaring derived from satellite data [J]. *Energies*, 2009, 3(2): 595–622.
- [6] Levin, N., Kyba, C., Zhang, Q., *et al.* Remote sensing of night lights: a review and an outlook for the future [J]. *Remote Sensing of Environment*, 2020, 237(C): 111443.
- [7] 仲晓雅, 闫庆武, 李桂娥. 中国长时间序列夜间灯光数据集 (2000–2020) [J/DB/OL]. 全球变化数据仓储电子杂志, 2022. <https://doi.org/10.3974/geodb.2022.06.01.V1>. <https://cstr.science.org.cn/CSTR:20146.11.2022.06.01.V1>.
- [8] 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).
- [9] 曹子阳, 吴志峰, 米素娟等. DMSP/OLS 稳定夜间灯光影像在中国的分类校正方法[J]. 地球信息科学学报, 2020, 22(2): 246–257.
- [10] 仲晓雅, 闫庆武, 厉飞等. 一种面向城市收缩地区的 DMSP/OLS 夜间灯光影像多年连续校正方法[J]. 地理与地理信息科学, 2021, 37(6): 46–51.
- [11] 卓莉, 张晓帆, 郑璟等. 基于 EVI 指数的 DMSP/OLS 夜间灯光数据去饱和方法[J]. 地理学报, 2015, 70(8): 1339–1350.
- [12] 李明峰, 蔡炜珩. NPP/VIIRS 多时相夜光遥感影像校正方法[J]. 测绘通报, 2019(7): 122–126.
- [13] Zhao M., Zhou Y., Li X., *et al.* Building a series of consistent night-time light data (1992–2018) in Southeast Asia by integrating DMSP-OLS and NPP-VIIRS [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 99: 1–14.
- [14] Chen, Z., Yu, B., Yang, C., *et al.* An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration [J]. *Earth System Science Data*, 2021, 13(3): 889–906.