

全球工业增加值公里网格数据集

薛倩^{1,2}, 宋伟^{1*}, 朱会议¹

1. 中国科学院地理科学与资源研究所陆地表层格局与模拟院重点实验室, 北京 100101;

2. 重庆交通大学建筑与城市规划学院, 重庆 400074

摘要: 工业增加值空间化数据是评估气候变化对工业经济系统影响的一项重要指标, 特别是对气候变化背景下工业经济的暴露评估具有重要意义。本文利用 2010 年全球植被指数数据 (Enhanced Vegetation Index, EVI, 来自 MODIS 月平均数据) 和 2010 年夜间灯光遥感数据 (DMSP/OLS), 通过构建夜间灯光调整指数 (Night Light Adjustment Index, EANTLI) 处理夜间灯光数据过饱和问题, 得到全球最佳灯光数据; 然后利用最佳灯光数据和来自世界银行的各国工业增加值数据构建回归模型, 进行工业增加值空间化反演, 生成了全球工业增加值公里格网数据集。在全球范围内, 随机选择了 178 个省级 (州) 区域的工业增加值和空间化后的工业增加值进行相关性检验。结果表明, 数据集的工业增加值与统计的工业增加值相关系数为 0.93; 以统计数据作为真值, 178 个区域的工业增加值平均精度为 80.14%。数据集包括全球 1 km 工业增加值数据和验证精度的表格数据。数据集存储为.tif、xlsx 格式, 原始数据文件个数为 6 个, 数据量 1.79 GB, 压缩为 1 个文件, 共 90.8 MB。

关键词: 全球; 工业增加值; 遥感反演; 公里格网

DOI: 10.3974/geodp.2018.01.03

1 前言

全球变化问题自上世纪八十年代以来已成为全球各界学者研究的焦点问题^[1-2], IPCC 第五次评估报告^[3]中认为: 全球地表温度近 120 年上升了 0.85 °C, 化石燃料燃烧等是气候变化的主要因素, 特别是近几十年, 极端气候事件发生频率不断增加^[4]对社会经济发展影响重大, 工业作为美国气象局对气候变化影响部门评估^[5]中敏感性较高的行业, 其受到气候变化影响主要包括气候均值波动变化^[6-7]和极端气候^[8]两个方面, 而辨析气候变化趋势, 分析其可能对工业行业造成的负面影响, 制定相应的适应性措施及政策是目前气候变化对工业经济影响及风险评估研究的重点。

根据 IPCC 发布的有关极端灾害风险特别报告^[9-10], 对于气候变化导致的风险灾害主要跟气候事件本身以及承载体的暴露度和脆弱性有关。在气候风险评估中, 载体在气候事件特别是极端气候事件中的暴露程度往往直接影响到其受灾程度及经济损失情况^[11]。但是

收稿日期: 2018-01-31; 修订日期: 2018-03-19; 出版日期: 2018-03-25

基金项目: 中华人民共和国科学技术部 (2016YFA0602402)

*通讯作者: 宋伟 E-8333-2018, 中国科学院地理科学与资源研究所, songw@igsrr.ac.cn

数据引用方式: [1] 薛倩, 宋伟, 朱会议. 全球工业增加值公里网格数据集[J]. 全球变化数据学报, 2018, 2(1): 9-17. DOI: 10.3974/geodp.2018.01.03.

[2] 薛倩, 宋伟, 朱会议. 全球工业增加值公里网格数据集[DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI: 10.3974/geodb.2018.01.14.V1.

暴露度的研究受人口、政策、环境、行业等众多因素影响，目前大部分有关暴露度的研究多采用构建指标体系^[12]，在一定的行政边界范围内，以行政区均值或总量来表达该区域某种行业的暴露度，在实际脆弱性评估及风险灾害分析中，这种分析存在无法差异性表达内部特征及尺度过大等问题，特别是不能根据气候栅格数据进行叠加评估不同气候情景下的社会经济系统暴露度以及无法对未来暴露度进行模拟分析。所以，目前急需拥有不同尺度大范围的社会经济数据的空间化研究以应对气候变化评估的基础数据需求，特别是不同产业类型的空间化大尺度分辨率数据亦是风险评估中的稀缺资源。

本文基于夜光遥感反演空间化的方法,采用全球 DMSP/OLS 夜光遥感数据、World Bank 中全球各个国家工业增加值数据、MODIS 植被指数产品,对全球工业经济系统进行空间化,构建全球工业增加值公里网格数据集,将有助于理清全球工业增加值的每公里格网分布时空特征,为全球变化中工业产业的暴露度及脆弱性评估提供科学数据,并为推动工业产业规划布局、全球工业行业防灾减灾及应对全球变化提供重要科学依据。

2 数据集元数据简介

全球工业增加值公里网格数据集的名称、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 全球工业增加值公里网格数据集元数据简表

条 目	描 述
数据集名称	全球工业增加值公里网格数据集 ^[13]
数据集短名	GlobeIndusAddV1km
作者信息	薛倩 E-8310-2018, 中国科学院地理科学与资源研究所, 15123085462@163.com 宋伟 E-8333-2018, 中国科学院地理科学与资源研究所, songw@igsnr.ac.cn 朱会义 E-9420-2018, 中国科学院地理科学与资源研究所, zhuhy@igsnr.ac.cn
地理区域	全球
时间分辨率	1 年
数据格式	.xlsx, .tif, .zip
数据量	1.79 GB (压缩前), 90.8 MB (压缩后)
数据集组成	主要由 2 部分组成: IndusAdd_Global.tif、DataValidation.xlsx。其中: IndusAdd_Global.tif 包括全球工业增加值公里格网数据, 数据量为 1.79 GB。 DataValidation.xls 是全球州、省级精度验证统计表, 数据量为 29 KB
基金项目	中华人民共和国科学技术部 (2016YFA0602402)
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)、实体数据 (中英文) 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 ^[14]

3 数据研发方法

3.1 基础数据

本数据集中所采用的源数据包括全球夜间灯光数据、MODIS 数据、相关的统计数据等，具体来源如表 2 所示。

表 2 数据来源统计表

数据名称	数据来源	数据描述	数据用途
夜间灯光数据	https://ngdc.noaa.gov ^[15]	本文所用数据 2010 年全球灯光数据，其空间分辨率为 1,000 m，亮度值范围 0-63	获取灯光亮度值、灯光归一化值，构建 EANTLI 模型以及分配模型
MODIS 数据	https://modis.gsfc.nasa.gov ^[16]	本文所用数据为全球 2010 年 1-12 月的 MOD13A3 产品，其中包含 11 个有效波段	提取增强型植被指数（EVI），用以构建 EANTLI 模型来校正夜间灯光数据的溢出和饱和效应
统计数据	http://data.worldbank.org.cn ^[17]	本文所用数据来自于世界银行中世界发展指标数据集和各国统计局	用以构建分配模型以及精度验证
矢量数据	http://www.gadm.org ^[18]	行政区边界	成图

夜间灯光数据来源于美国气象卫星通过低光源成像技术生产的非辐射定标夜间平均灯光产品^[19]，由于其大尺度及较高分辨率及能够通过灯光亮度值反映地理实体特征的特性所以被广泛应用于经济数据反演^[20]，城镇区域提取^[21]，光污染范围确定^[22]等研究中。夜间灯光数据的灯光强度为 0-63，当灯光亮度值达到 63 时，中心灯光就不再随着灯光强度增加亮度值继续增加，而是产生饱和现象^[23]，导致所示区域实际灯光值远远大于影像所表示的灯光强度值，且城市内部灯光差异较小，无法进行精准的遥感反演。目前对于灯光数据过饱和校正常用的方法包括利用美国国家地球物理数据中心发布的辐射定标数据^[24]进行纠正、电力消耗与灯光强度线性纠正^[25]、不变区域法影像纠正以及植被指数法影像纠正^[26]等方法，由于尺度较大，考虑到数据获取的可实施性，本数据集选取植被指数对夜间灯光数据进行饱和校正。

3.2 数据研发技术路线

本数据集主要技术步骤包括数据预处理、夜间灯光数据过饱和校正、构建分配模型以及精度验证几个步骤。具体技术路线如图 1 所示。

3.3 数据研发算法原理

基于夜间灯光影像与经济数据、增强型植被指数（EVI）之间的关系，首先消除夜间灯光影像的过饱和现象，其次构建经济数据和夜间灯光影像的回归模型，对工业增加值进行空间化处理。具体过程描述如下：

（1）数据预处理。将灯光数据和 MODIS 植被指数数据进行投影、剪裁、镶嵌、二值化、归一化等处理，获取年平均灯光强度值（NTL），计算归一化灯光亮度指数（NTLn）及年

平均增强型植被指数 (EVI), 其中, EVI 数据为剔除值小于 0.01 的区域 (包括水域、裸岩、沙漠等) 后的年平均数据。

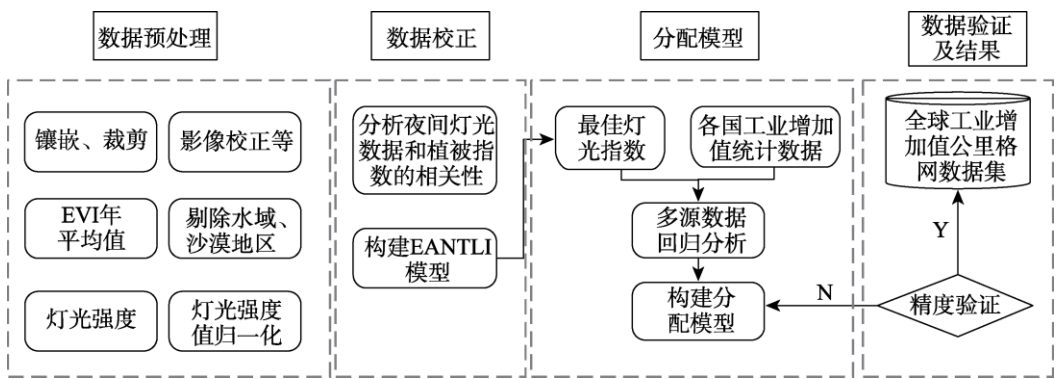


图 1 全球 1 km 工业增加值算法流程

(2) 夜间灯光数据过饱和校正。构建灯光指数修正模型 (EANTIL 模型) 以削减灯光溢出和饱和现象, 利用 NTL、NTLn、EVI 数据构建 EANTIL 模型, 其计算公式^[27]如下:

$$EANTLI = \left(\frac{2}{1 - NTLn + EVI_i} - 1 \right) \times NTL \tag{1}$$

(3) 分配模型构建。根据最佳灯光指数 EANTLI 以及各国统计数据, 构建回归分配模型, 反演每公里工业增加值 I, 计算公式如下:

$$I = \frac{I_i g EANTLI}{EANTLI_i} \quad (i = 1, 2, 3 \cdots n) \tag{2}$$

式中, I_i指各个国家工业增加值, EANTLI_i指各个国家最佳灯光值, 通过 ArcGIS 软件统计分析得来, 通过回归分配模型得到初步全球工业增加值公里格网数据。

(4) 精度验证及结果。随机选取大量州、省级行政单元的工业增加值数据对第三步得到的数据进行精度验证, 若精度符合要求, 则得到最终全球工业增加值公里网格数据集, 否则重新构建分配模型。

4 数据结果与验证

4.1 数据结果组成

全球工业增加值公里网格数据集, 空间分辨率为 1,000 m, 数据年份为 2010 年, 数据集的组成包括实体数据和验证数据两部分组成:

(1) 空间分布数据

实体数据以 IndusAdd_Global.tif 格式存储, 可用 ArcGIS 软件对数据进行读写, 也可用 Python 等主流编程语言调用相关函数库读写数据; 2010 年全球工业增加值格网分布如图 2 所示。全球工业增加值数值分布范围为 0-988, 223×10⁴ \$/km², 其中工业增加值高值区域主要分布在北美洲东海岸, 南美洲巴西沿太平洋海岸、西欧多个国家 (德国、比利时、荷兰、丹麦、瑞士、法国、卢森堡、英国、意大利、奥地利), 中国沿海区域及日本韩国部分地区以及南亚、东南亚区域。

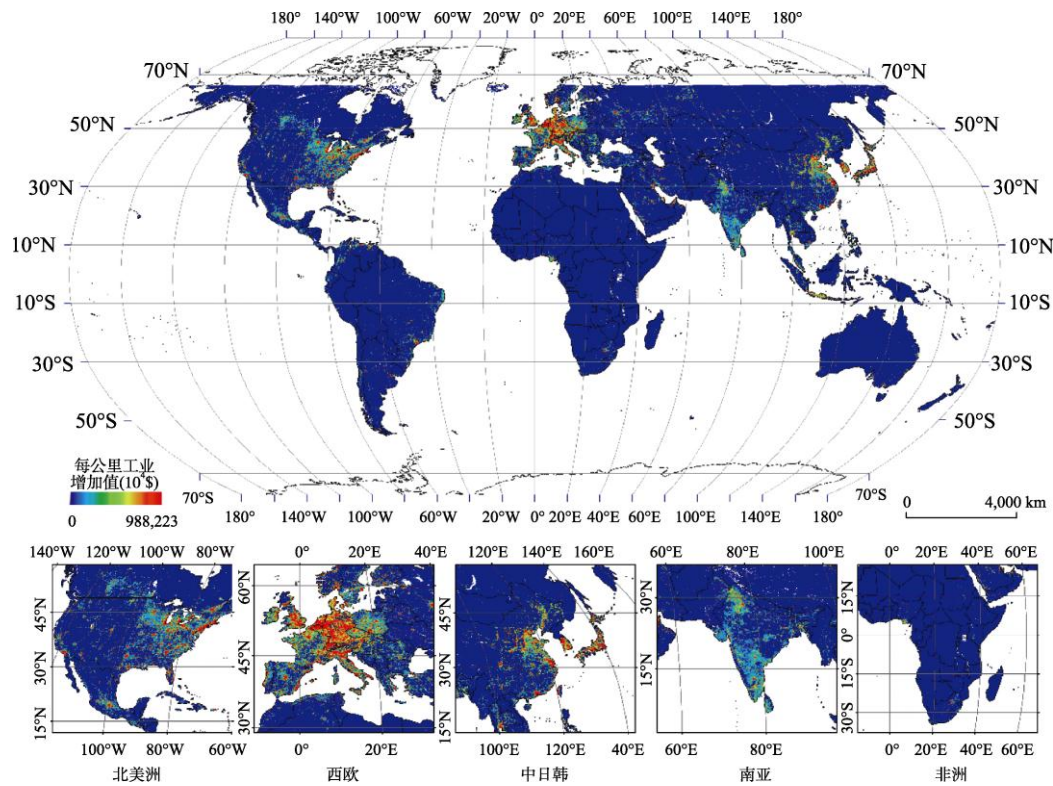


图 2 全球 1 km 工业增加值空间分布图

(2) 验证数据统计表

DataValidation.xlsx 数据为本数据集选取的 178 个区域验证精度数据。验证内容及结果列于表 3。

本数据集随机选取了 178 个区域进行精度验证，如表 3 所示进行部分数据说明，区域名称为所验证区域的中英文名称，统计数据为所选验证区域的统计局所公布的工业增加值数据，GlobeIndusAddV1km 为所选验证区域空间化的结果，精度为两者之间的相对精度，统计数据来源为所选验证区域的统计局网站，例如北阿古桑省（Agusan del Norte）2010 年工业增加值为 1.30×10^8 \$/km²，统计数据来源于 <http://psa.gov.ph/>，本数据集经过空间化后得到北阿古桑省 2010 年工业增加值为 1.64×10^8 \$/km²，精度为 73.71%。

4.2 数据结果验证

4.2.1 夜光遥感数据去饱和程度检验

夜光遥感数据由于其过饱和现象造成经济数据空间化精度较低，所以本数据集利用植被指数构造 EANTLI 模型对夜光遥感数据进行去饱和处理，得到结果如图 3 所示，可见未处理前灯光强度 NTL 值在城市核心区域及城郊结合区域差异较小，无法区分城市内部灯光差异，致使空间化时，精度较差，而相比之下，EANTLI 模型处理后的灯光值能够很好地

区分城市内部及周边灯光变化。

表 3 全球 1 km 工业增加值数据集在 178 个区域的验证情况(部分)

工业增加值 ($\times 10^8$ \$/km ²)					
区域名称 (英文)	区域名称 (中文)	统计数据	GlobeIn- dusAddV1km	精度	统计数据来源
Agusan del Norte	北阿古桑省	1.30	1.64	73.71	http://psa.gov.ph/
Anhui Province	安徽省	942.75	669.17	70.98	http://www.stats.gov.cn/
Ehime Prefecture	爱媛县	110.50	116.87	94.24	http://www.stat.go.jp/english/index.htm
Alaska	阿拉斯加州	213.89	213.45	99.80	https://www.bea.gov/
Ohio	俄亥俄州	1,043.95	1,229.37	82.24	https://www.bea.gov/
Amazonas	亚马逊	50.66	53.32	94.77	https://ww2.ibge.gov.br/english/default.php/
Antioquia	安蒂奥基亚	122.30	112.08	91.65	http://www.dane.gov.co/index.php/en/
Rostov region	罗斯托夫州	61.88	61.41	99.24	http://www.gks.ru/

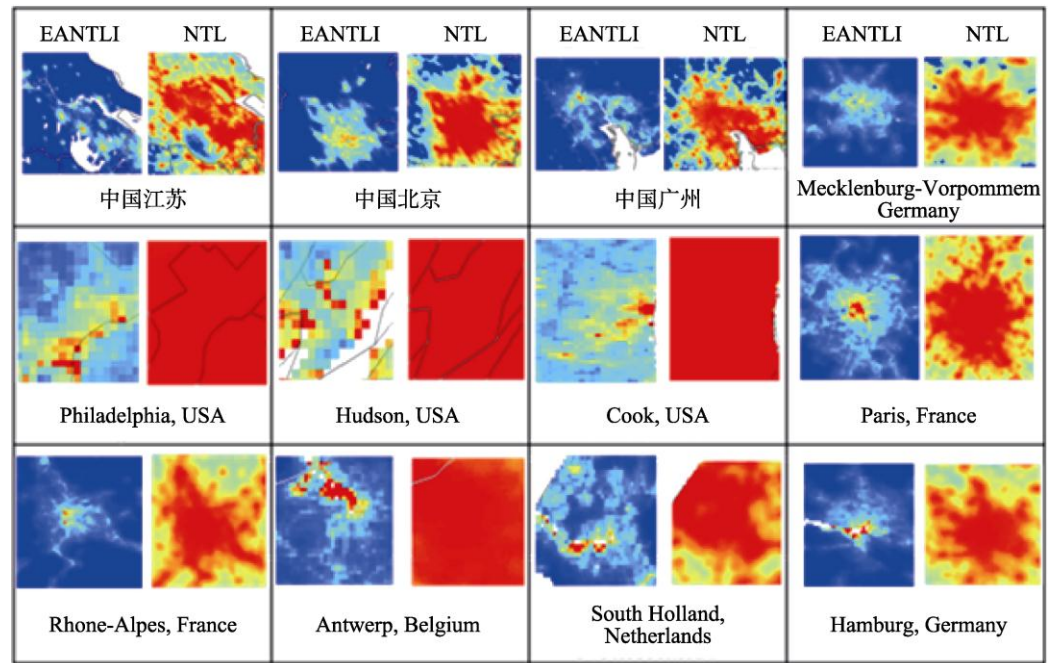


图 3 NTL 和 EANTLI 对比图

4.2.2 州、省级尺度工业增加值空间化精度检验及误差分析

为验证工业增加值栅格分布数据集的精度，本数据集选择 Alaska 等 178 个省、州（图 4）级别的各国家及地区统计年鉴中工业增加值统计数据，以各个省、州行政区矢量图为统

计图层，以全球工业增加值栅格分布数据为目标图层利用 ArcGIS 进行分区统计分析，得到省、州级工业增加值，并与各统计年鉴中数据进行对比分析，以统计数据为横坐标，空间化得到的工业增加值数据为纵坐标做散点图（图 5），从图 5 可以看出，空间化得到的工业增加值栅格数据集与统计数据呈现明显的线性相关关系，在 95% 的置信度水平下相关系数为 0.93，呈现显著相关，且该数据集的平均精度达到 80.14%，可见空间化的工业增加值数据集能够表达工业增加值的每公里分布规律。

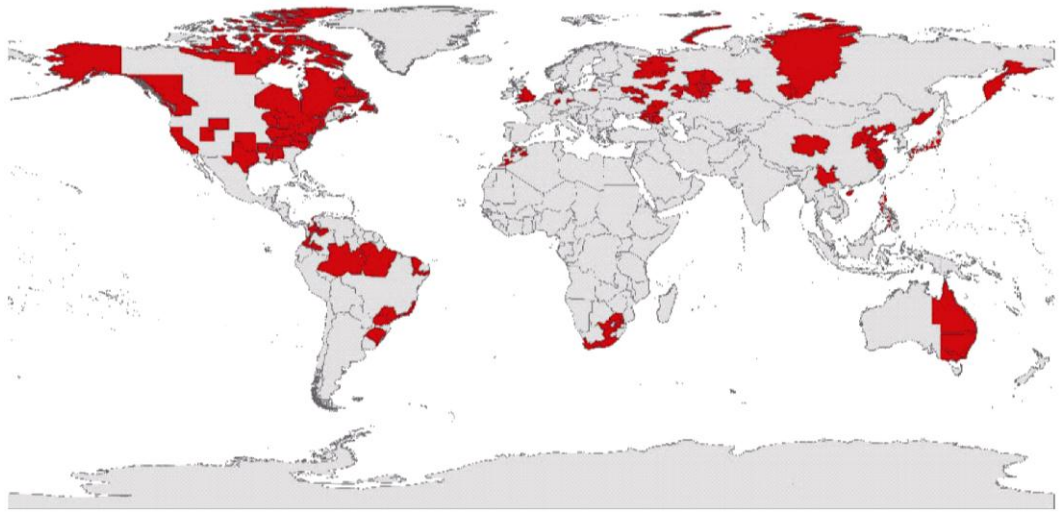


图 4 Alaska 等 178 个省、州验证区域位置示意图

5 讨论和总结

全球工业增加值公里网格数据集，打破了行政边界对统计数据限制，实现了统计数据高时空分辨率大尺度的反演，且精度验证结果较好。该数据集可用于辨析全球历史时期 2010 年工业增加值时空分布特征，判别全球工业产业集群及工业产值区域划分；同时，因为数据反演方法可适用于大尺度、长时间序列的工业增加值反演，所以，本数据集可以进行长时间序列的工业增加值数据研发，从而可以服务于全球变化研究中有关工业风险评估以及工业灾害适应性评估及措施制定、工业产业选址及结构升级等研究，为其提供数据支撑。此外，全球尺度的空间化产品大多数为 GDP、人口的空间化产品，特别是现有 GDP 空间化数据集大多是对二三产业合并处理，鲜有单独的二产或三产空间化数据集产品。本数据集依据全球各国工业增加值及夜间灯光数据进行工业增加值反演，丰富了社会经济数据，且本数据集的平均精度达到了 80.14%，保障了数据的使用精度。

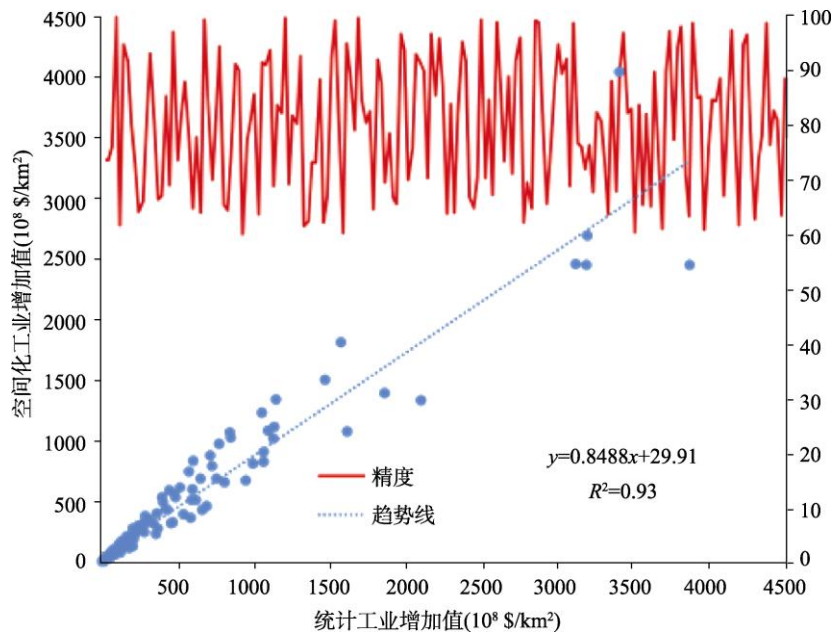


图 5 全球工业增加值公里网格数据集精度验证

作者分工：宋伟和朱会义对数据集的开发做了总体设计；薛倩采集和处理了全球工业增加值数据并设计了模型和算法、做了数据验证、撰写了数据论文等。

参考文献

[1] Bryson, R. A. A perspective on climatic change [J]. *Science*, 1974, 184(4138): 753.

[2] Zeng, N., Ding, Y., Pan, J., *et al.* Climate change: the Chinese challenging [J]. *Science*, 2008, 319(5864): 730-731.

[3] IPCC, WGII. Climate change 2013: the physical science basis [J]. *Contribution of Working*, 2013, 43(22): 866-871.

[4] Orłowsky, B., Seneviratne, S. I. Global changes in extreme events: regional and seasonal dimension [J]. *Climatic Change*, 2012, 110(3-4): 669-696.

[5] Jager, J. 气候与能源系统[M]. 北京: 气象出版社, 1988.

[6] 王守荣. 气候变化对中国经济社会可持续发展的影响与应对[M]. 北京: 科学出版社, 2011.

[7] Li, D. H. W., Liu, Y., Lam, J. C. Impact of climate change on energy use in the built environment in different climate zones—a review [J]. *Energy*, 2012, 42(1): 103-112.

[8] Rosenberg, E. A., Keys, P. W., Booth, D. B., *et al.* Precipitation extremes and the impacts of climate change on stormwater infrastructure in Washington State [J]. *Climatic Change*, 2010, 102(1-2): 319-349.

[9] Aven, T., Renn, O. An evaluation of the treatment of risk and uncertainties in the IPCC reports on climate change [J]. *Risk Analysis*, 2015, 35(4): 701-712.

[10] Mastrandrea, M. D., Mach, K. J. Treatment of uncertainties in IPCC assessment reports: past approaches and considerations for the Fifth Assessment Report [J]. *Climatic Change*, 2011, 108(4): 659.

[11] Fussler, H. M. Development and climate change: review and quantitative analysis of indices of climate change exposure, adaptive capacity, sensitivity, and impacts [R]. 2009.

[12] Benzie, M., Hedlund, J., Carlsen, H. Introducing the transnational climate impacts index—indicators of

- country-level exposure [R]. *Stockholm Environment Institute*, 2016. DOI: 10.13140/RG.2.1.2839.7044.
- [13] 薛倩, 宋伟, 朱会义. 全球工业增加值公里网格数据集[DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI: 10.3974/geodb.2018.01.14.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]. DOI: 10.3974/dp.policy. 2014.05 (2017 年更新).
- [15] NOAA. Defense meteorological satellite program/operational linescan system (DMSP/OLS) [EB/OL]. (2013-12-30) <https://ngdc.noaa.gov>.
- [16] DAACNL. Vegetation indices monthly L3 global 1 km [DB]. Version 5. <https://modis.gsfc.nasa.gov/>.
- [17] The World Bank. World development indicators [EB/OL]. <https://data.worldbank.org/>.
- [18] Areas, G. A. GADM database of global administrative areas [EB/OL]. <http://www.gadm.org/>.
- [19] Cho, K., Ito, R., Shimoda, H., *et al.* Fishing fleet lights and sea surface temperature distribution observed by DMSP/OLS sensor [J]. *International Journal of Remote Sensing*, 2010, 20(1): 3–9.
- [20] Wu, J., Wang, Z., Li, W., *et al.* Exploring factors affecting the relationship between light consumption and GDP based on DMSP/OLS nighttime satellite imagery [J]. *Remote Sensing of Environment*, 2013, 134(7): 111–119.
- [21] Amaral, S., Cmara, G., Miguel, A., *et al.* Assessing nighttime DMSP/OLS data for detection of human settlements in the Brazilian Amazon [J]. *Nature*, 2001, (4): 895–903.
- [22] Yang, M., Wang, S., Zhou, Y., *et al.* Review on applications of DMSP/OLS night-time emissions data [J]. *Remote Sensing Technology & Application*, 2011, 26(1): 45–51.
- [23] Ma, L., Wu, J., Li, W., *et al.* Evaluating saturation correction methods for DMSP/OLS nighttime light data: a case study from China's cities [J]. *Remote Sensing*, 2014, 6(10): 9853–9872.
- [24] Jing, X., Shao, X., Cao, C., *et al.* Comparison between the Suomi-NPP day-night band and DMSP-OLS for correlating socio-economic variables at the provincial level in China [J]. *Remote Sensing*, 2015, 8(1): 17.
- [25] Letu, H., Hara, M., Tana, G., *et al.* A saturated light correction method for DMSP/OLS nighttime satellite imagery [J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2012, 50(2): 389–396.
- [26] Zhang, Q., Seto, K. C. Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data [J]. *Remote Sensing of Environment*, 2011, 115(9): 2320–2329.
- [27] Li, Z., Zhang, X., Jing, Z., *et al.* An EVI-based method to reduce saturation of DMSP/OLS nighttime light data [J]. *Acta Geographica Sinica*, 2015, 70(8): 1339–1350.