

云南沿边地区 1992–2013 年 GDP 空间化数据集

卢秀^{1,2}, 李佳^{1*}, 段平¹, 李晨¹, 程峰¹, 王金亮¹

1. 云南师范大学旅游与地理科学学院, 昆明 650500; 2. 南京师范大学虚拟地理环境教育部
重点实验室, 南京 210023

摘要: 云南沿边地区地处云南省的西南边陲地带, 处于缅、印、中、孟经济走廊, 具有重要的地缘位置。因此, 较高精度地实现云南沿边地区国内生产总值 (Gross Domestic Product, GDP) 空间化拟合显得尤为重要。本研究以云南沿边地区为研究区, 以 DMSP/OLS 夜间灯光数据、土地利用数据、云南省统计数据等为数据源, 实现 1992–2013 年研究区的 GDP 空间化拟合。对夜间灯光数据进行饱和校正、相互校正、年内融合、年际间校正、重投影、重采样及裁剪等处理, 基于土地利用数据实现第一产业的空间化, 基于夜间灯光数据, 采用“分类回归”的方法实现第二、三产业的空间化, 从而得到 GDP 的空间化, 并且对拟合结果进行验证。结果表明: 各期的第一产业拟合值相对误差最大为 1.12%, 各期误差均较小; 第二、三产业经过分类回归后的拟合相对误差均不高于 -6.40%, 最终 GDP 拟合的相对误差每期误差均小于 -4.24%, 精度较高。该数据集存储为.tif 格式, 单个文件的空间分辨率为 1 km, 共 22 组文件, 数据量为 68.6 M (压缩为 1 个文件, 3.43 MB)。

关键词: GDP 空间化; 云南沿边地区; 夜间灯光数据; 土地利用数据

DOI: 10.3974/geodp.2020.02.08

数据可用性声明:

本文关联数据已出版, 可获取: 卢秀, 李佳, 段平等. 云南沿边地区 GDP 公里网格数据集 (1992–2013) [J/DB/OL]. 全球变化数据仓储电子杂志, 2020. DOI: 10.3974/geodb.2020.03.11.V1.

1 前言

云南沿边地区地处于中、印、缅、孟四国的经济走廊, 起到中国与南亚地区、东南亚地区国家进行物质文化交流及经济贸易往来的桥梁和纽带作用, 因此地理位置较为重要。国内生产总值 (Gross Domestic Product, GDP) 是衡量区域经济发展、研究城市发展状况及反映区域规划等的评价指标^[1]。传统的 GDP 统计数据受行政区划单元的限制, 无法获得较为精细的能够显示区划单元内部差异的经济数据, 无法结合生态环境等栅格数据进行综合分析等^[2], 通过对 GDP 数据的空间化拟合, 可以解决以上问题。随着遥感技术的不断进步和发展, 夜间灯光、土地利用等遥感数据被广泛应用在 GDP 数据的空间化拟合。由美国国防军事气象卫星获取的 DMSP/OLS 夜间灯光数据自 1970s 已经被应用^[3],

收稿日期: 2020-05-06; 修订日期: 2020-06-10; 出版日期: 2020-06-25

基金项目: 中华人民共和国科学技术部 (2018YFE0184300); 国家自然科学基金 (41561048)

*通讯作者: 李佳 AAS-6000-2020, 云南师范大学旅游与地理科学学院, keguigiser@163.com

数据引用方式: [1] 卢秀, 李佳, 段平等. 云南沿边地区 1992–2013 年 GDP 空间化数据集[J]. 全球变化数据学报, 2020, 4(2): 155–162. DOI: 10.3974/geodp.2020.02.08.

[2] 卢秀, 李佳, 段平等. 云南沿边地区 GDP 公里网格数据集 (1992–2013) [J/DB/OL]. 全球变化数据仓储电子杂志, 2020. DOI: 10.3974/geodb.2020.03.11.V1.

该数据具有很强的光电放大能力，可直观反映人类活动，且具有时间跨度长、空间覆盖范围广等优点，后来经过不断的完善和更新，被应用在城镇化监测，人口、能量消耗和GDP等的估算等多个方面。因此本研究基于DMSP/OLS夜间灯光数据，结合土地利用数据及统计数据，构建了云南沿边地区1992–2013年的GDP空间化数据集。该数据集能够反映云南沿边地区行政区划单元内部的经济差异性，同时可以与自然生态环境等栅格影像数据进行综合分析。

2 数据集元数据简介

《云南沿边地区GDP公里网格数据集（1992–2013）》^[4]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 《云南沿边地区GDP公里网格数据集（1992–2013）》元数据简表

条目	描述
数据集名称	云南沿边地区GDP公里网格数据集（1992–2013）
数据集短名	YunnanBorderGDP1992-2013
作者信息	卢秀 AAS-6714-2020, 云南师范大学旅游与地理科学学院, lx_rsgis@163.com 李佳 AAS-6000-2020, 云南师范大学旅游与地理科学学院, keguigiser@163.com 段平, 云南师范大学旅游与地理科学学院, duanpingshai@163.com 李晨, 云南师范大学旅游与地理科学学院, lichen924541412@163.com 程峰, 云南师范大学旅游与地理科学学院, chengfeng_rs@163.com 王金亮, 云南师范大学旅游与地理科学学院, jlwang@ynnu.edu.cn
地理区域	云南沿边地区, 包括8个地州市: 保山市、普洱市、临沧市、红河哈尼族彝族自治州、德宏傣族景颇族自治州、怒江傈僳族自治州、文山壮族苗族自治州、西双版纳傣族自治州
数据年代	1992–2013
空间分辨率	1 km
数据量	数据量为68.6 MB（压缩为1个文件, 3.43 MB）
数据集组成	GDP空间化产品数据
基金项目	中华人民共和国科学技术部（SQ2018YFE011725）; 国家自然科学基金（41561048）
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	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 ^[5]
数据和论文检索系统	DOI, DCI, CSCD, WDS/ISC, GEOSS, China GEOSS, Crossref

3 数据研发方法

3.1 原始数据介绍及预处理

DMSP/OLS 夜间灯光数据包括三种图像数据, 分别为非辐射定标夜间平均灯光强度影像数据、平均灯光 XPct 据、全球辐射定标数据。其中非辐射定标夜间平均灯光强度影像数据包括 3 种数据类型, 分别是平均可见灯光影像、稳定灯光影像、无云观测频数影像, 本研究使用的数据是非辐射定标夜间平均灯光强度影像数据中的稳定灯光影像。时相为 1992–2013 年, DN 值范围为 0–63, 像元大小为 0.008,333°。由于该数据产品传感器未做星上定标, 各传感器获得的不同灯光影像之间缺乏连续性和可比性, 且城市中心区域存在像元 DN 值饱和现象^[6–7]。因此研究结合 Elvidge、Wu、Liu 等^[8–10]的方法对夜间灯光数据进行相互校正和饱和校正、年内融合、年际间校正, 重投影为 Albers 投影及重采样为 1,000 m, 裁剪云南沿边地区范围等处理。

校正步骤: (1) 相互校正与饱和校正。将 1992–2013 年待校正影像与辐射定标数据中的 2006 年 F16 传感器的数据进行相互校正与饱和校正。参考区域为黑龙江省鸡西市。校正模型为:

$$DN_{correct} = a \times DN^2 + b \times DN + c \quad (1)$$

式中, DN 表示校正前的亮度值, a 、 b 、 c 为回归系数, $DN_{correct}$ 表示校正后的 DN 值。

(2) 年内融合。不同传感器获取的同一年份的数据不一致, 按照式(2)对步骤(1)校正后的部分影像进行年内融合。

$$DN_{(n,i)} = \begin{cases} 0 & DN_{(n,i)}^a = 0 \mid DN_{(n,i)}^b = 0 \\ (DN_{(n,i)}^a + DN_{(n,i)}^b) / 2 & \text{其他} \end{cases} \quad (2)$$

($n = 1994, 1997, 1998, \dots, 2007$)

式中, $DN_{(n,i)}^a$ 和 $DN_{(n,i)}^b$ 分别表示相互校正和饱和校正后的 n 年的两个不同传感器获取的 i 像元的 DN 值, $DN_{(n,i)}$ 表示影像年内融合校正后的 n 年的 i 像元的 DN 值。

(3) 年际间校正。经过年内融合之后, 仍然存在不同年份之间的影像不可比的现象。校正公式如式(3)。

$$DN_{(n,i)} = \begin{cases} 0 & DN_{(n+1,i)} = 0 \\ DN_{(n-1,i)} & DN_{(n+1,i)} > 0 \& DN_{(n-1,i)} > DN_{(n,i)} \\ DN_{(n,i)} & \text{其他} \end{cases} \quad (3)$$

($n = 1992, 1993, 1994, \dots, 2013$)

式中, $DN_{(n-1,i)}$ 、 $DN_{(n,i)}$ 、 $DN_{(n+1,i)}$ 分别表示影像的像元 i 在 $n-1$ 年, n 年, $n+1$ 年的 DN 值。

3.2 算法原理

第一产业与土地利用数据中的耕、林、草和水域地类的相关性较高^[11]。因此将第一产业中的农、林、牧、渔产值数据分别与土地利用数据中对应的地类进行建模, 建模公式如下:

$$\text{GDP}_{ij} = \sum_{k=1}^4 Gk_{ij} \quad (4)$$

式(4)中, GDP_{ij} 表示第 i 个行政区第 j 个栅格的第一产业产值; Gk_{ij} 分别表示这四种产值; $k=1-4$, 分别对应第一产业中农、林、牧、渔产值的序号。

DMSP/OLS 夜间灯光数据的灯光亮度与第二、三产业存在一定的相关性, 可利用夜间灯光不同灯光亮度值分布拟合第二、三产业。拟合模型如下:

$$\text{GDP}_{23} = a \cdot \text{SOL} \quad (5)$$

GDP_{23} 表示第二、三产业拟合值, a 是拟合系数。其中, SOL 表示:

$$\text{SOL} = \sum_{m=1}^{\max} N_m \cdot B_m \quad (6)$$

式中, N_m 代表亮度值为 m 的像元数量, B_m 代表像元自身的亮度值。

基于夜间灯光指数对第二、三产业进行的全局建模的拟合结果与统计数据的误差较大, 为了解决全局建模精度不足问题, 本研究采用“分类回归”的思想, 即对全局建模的拟合模型加入相对误差 δ , δ 的公式如下:

$$\delta = \frac{\text{GDP}_{23} - \text{GDP}_S}{\text{GDP}_S} \times 100\% \quad (7)$$

GDP_S 为研究区第二、三产业统计值。“分类回归”建模的具体流程:

(1) 根据最初始的拟合方程, 筛选出 $|\delta| < 25\%$ 的县域组合成第一个方程, 这些县域采用第一个方程进行拟合;

(2) 将 $|\delta| > 25\%$ 的县域分成两部分, 其中 $\delta > 25\%$ 的部分县域重新建立高估部分的拟合模型, $\delta < -25\%$ 的部分县域重新建立低估部分的拟合模型;

(3) 经过上述两步骤建立模型后, 仍存在 $|\delta| > 25\%$ 的情况, 且数量大于总拟合建模数的 10%, 则继续分区建模, 直到拟合方程的 $|\delta| < 25\%$ 或剩余拟合建模数小于 10%, 终止建模。

将以上得到的第一产业建模拟合结果与第二、三产业建模拟合结果进行叠加, 得到 GDP 拟合结果。得到的拟合结果仍然存在部分县域的拟合误差较大。因此, 对得到的 GDP 拟合结果进行县域范围内的线性纠正, 纠正模型如式(8):

$$\text{GDP}_{\text{correct}} = \text{GDP}_{\text{estimated}} \cdot (\text{GDP}^* / \text{GDP}_{\text{all}}) \quad (8)$$

式中, $\text{GDP}_{\text{correct}}$ 表示纠正后的 GDP 栅格数据, $\text{GDP}_{\text{estimated}}$ 为估算的每个栅格格网的 GDP 数据, GDP^* 表示县域的统计 GDP 数据, GDP_{all} 为对应县域估算的 GDP 数据。

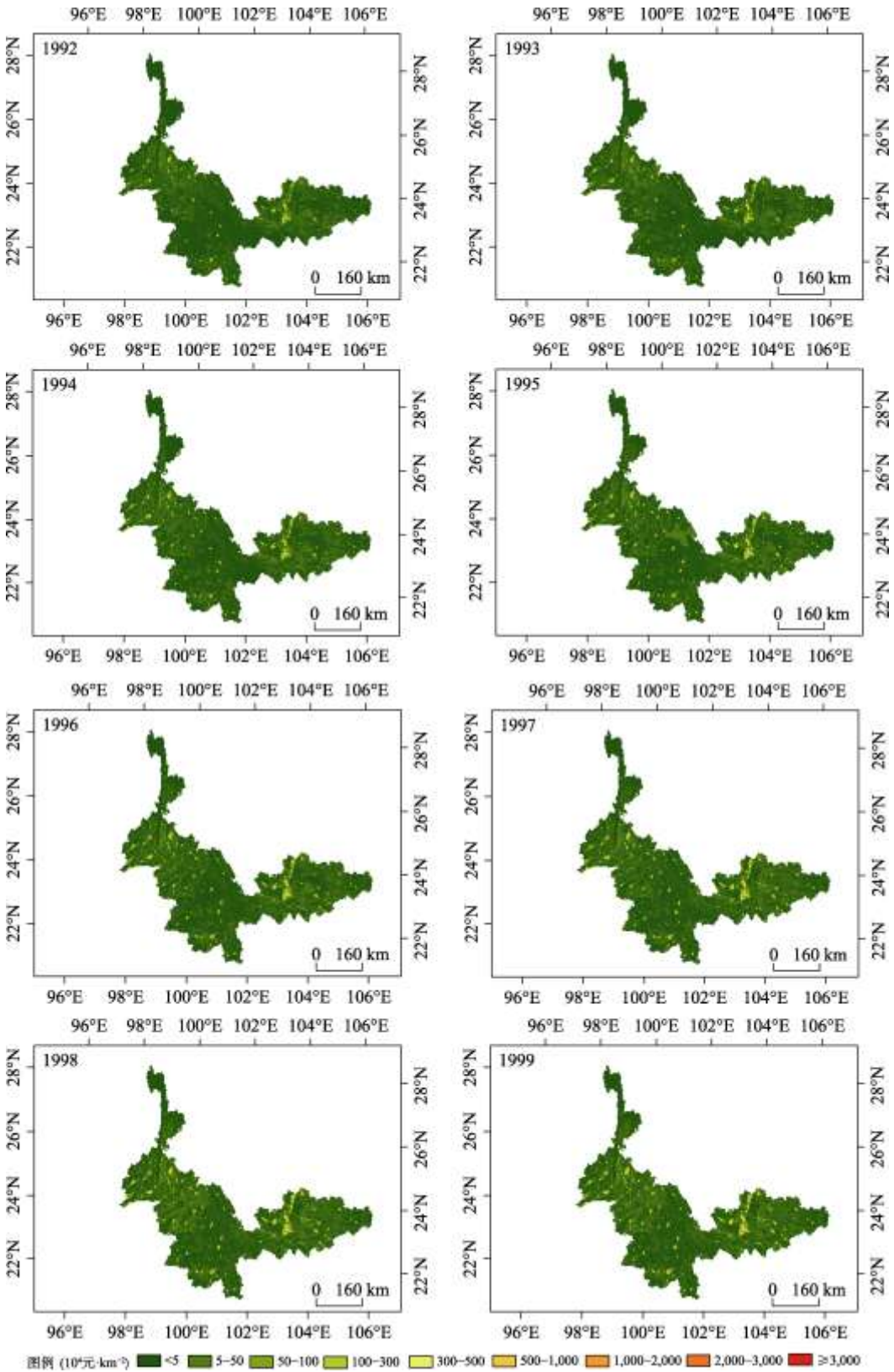
4 数据结果与验证

4.1 数据集组成

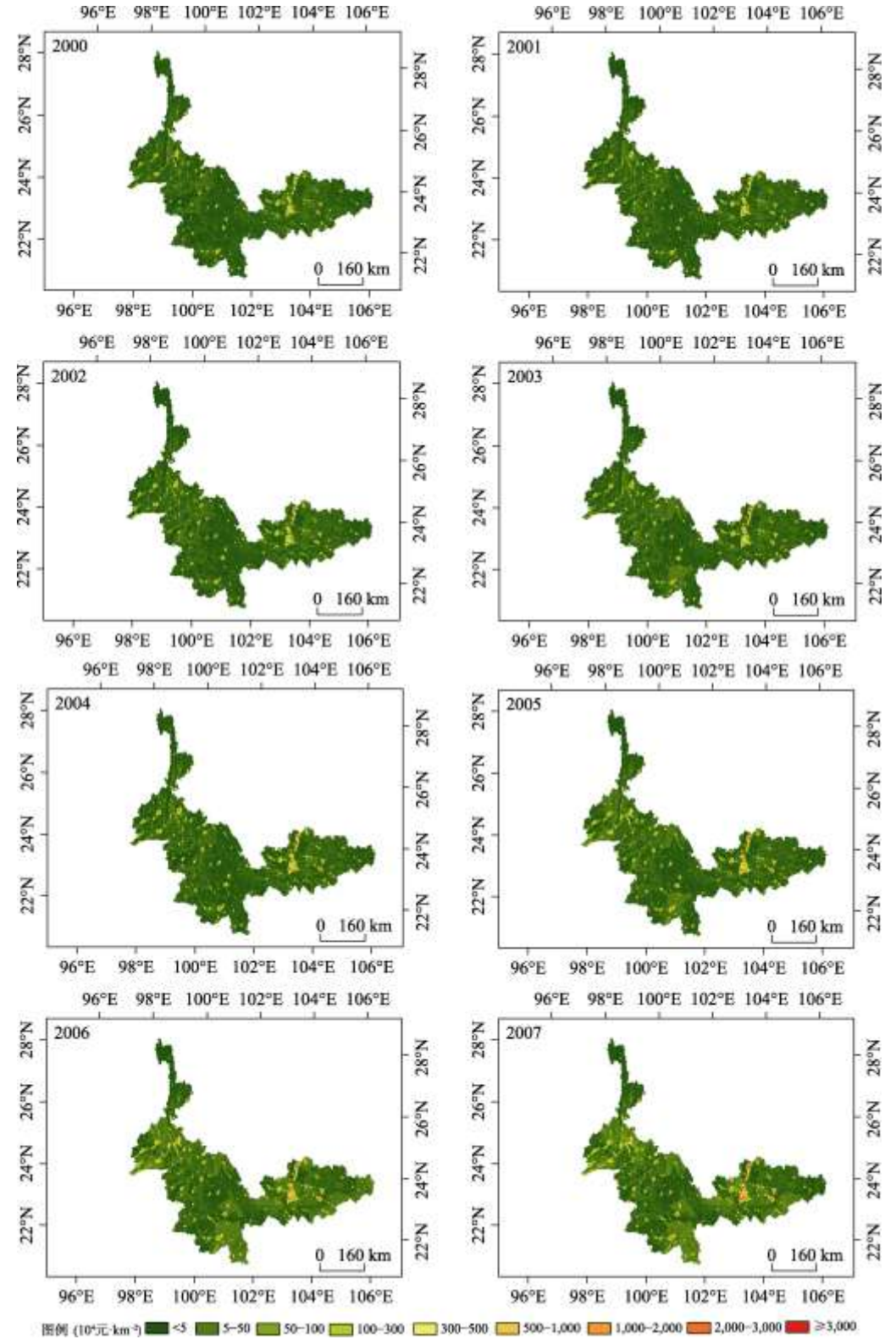
云南沿边地区 GDP 数据集包括 1992–2013 年间各年度的 GDP 空间分布 (.tif 格式), 其空间分布数据可视化图^[12–13]如图 1 所示。

4.2 数据结果

1992–2004 年间, GDP 较高即经济发展较好的地区主要分布在云南沿边地区的东部地区-红河哈尼族彝族自治州区域内以及西部地区的零散分布。从 2005 年开始, 保山市、



续图



续图

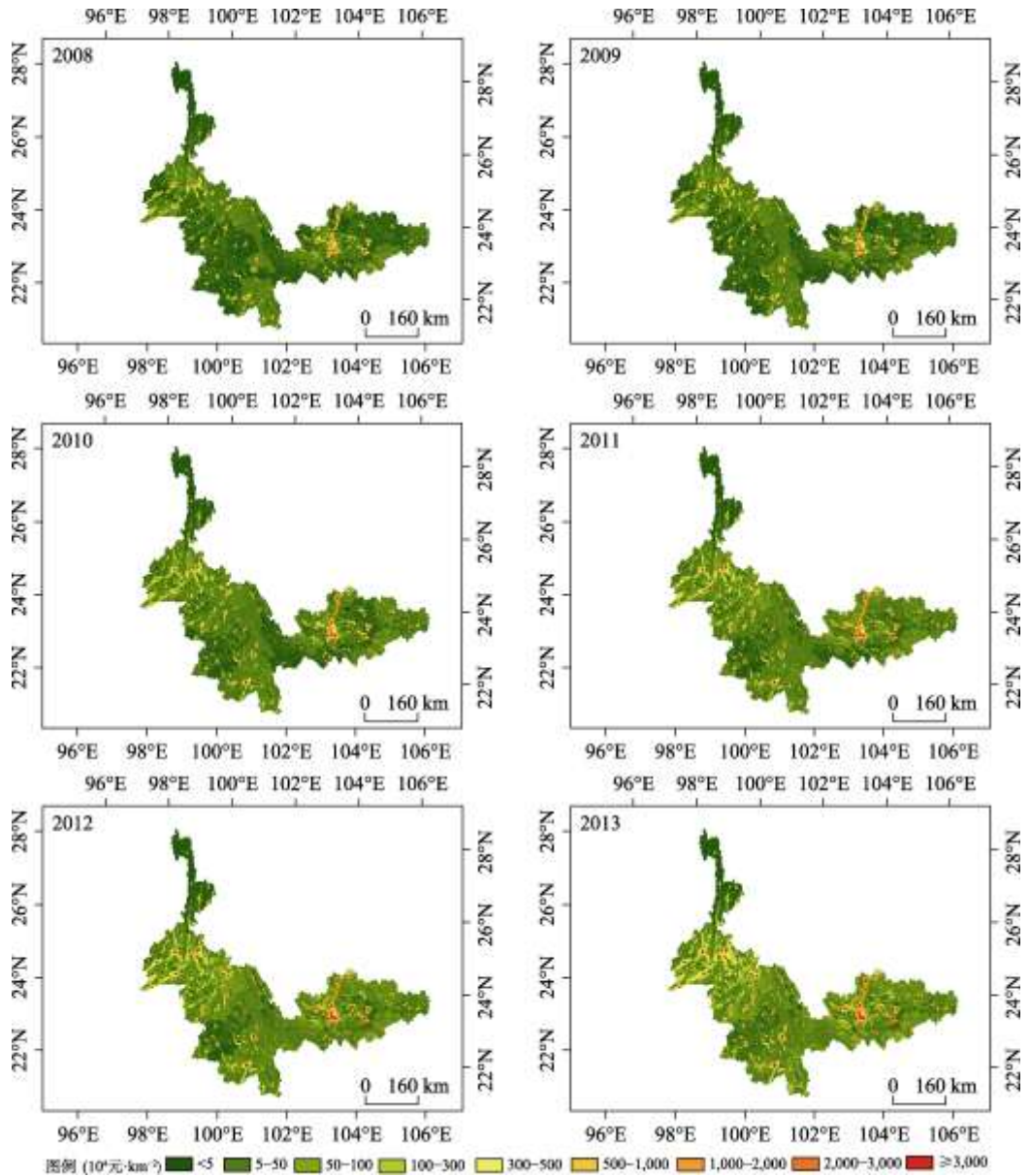


图 1 1992–2013 年历年云南沿边地区 GDP 空间分布图

德宏傣族景颇族自治州、西双版纳傣族自治州等区域的经济逐渐提升，而中部地区的临海市、普洱市的经济的发展落后于其他地区。2010 年以后，GDP 高值像元区域逐渐散布整个沿边地区，各个地州的经济的发展水平逐渐增强。从县域来看，GDP 高值像元区域如个旧市、开远市、弥勒县等县域一直处于沿边地区的发展前列，尤其是弥勒县，从 2001 年，一直位列沿边地区 GDP 首位。2001 年以后，保山市、景洪市、腾冲市、蒙自市和建水县等逐渐发展，GDP 像元值逐年增多，高值像元区域逐渐从城中心区域往外扩散。而 GDP 像元低值主要集中分布在贡山、福贡、西盟、梁河等县域，经济发展缓慢。

4.3 数据结果验证

对第一产业拟合值,第二、三产业拟合值,GDP 拟合值和对应的 GDP 统计数据进行误差比较。误差分析结果表明:各期拟合数据的第一产业拟合值相对误差最小可达到0.09%,最大不高于1.12%,误差较小,拟合效果好,说明第一产业通过土地利用数据拟合是可行的并且精度较高;第二、三产业未经分类回归拟合的最大误差为-29.09%,经过分类回归后的拟合相对误差最大仅为-6.40%,最小可达到-0.40%,说明经过分类回归拟合的方法,很大程度地提高了第二、三产业的拟合精度。最终拟合的 GDP 相对误差最大仅为-4.24%,精度较高,拟合效果较好,该拟合结果可用于后续的数据计算、分析等应用。

5 讨论和结论

本研究基于土地利用数据对 GDP 第一产业进行拟合,基于 DMSP/OLS 夜间灯光数据采用分类回归的拟合方法对第二、三产业拟合,拟合时间段为 1992–2013 年。通过对拟合结果的精度验证,结果表明各产业的拟合精度均较高,第一产业拟合相对误差不高于1.12%,第二、三产业拟合的相对误差最大不高于6.40%,最终拟合的 GDP 相对误差最大仅为-4.24%,拟合效果较好,反演精度高。该数据可进一步用于行政区划内部的经济差异分析、结合其他栅格影像数据进行空间计算、分析等。

参考文献

- [1] 王金南, 逯元堂, 周劲松等. 基于 GDP 的中国资源环境基尼系数分析[J]. 中国环境科学, 2006, 26(1): 111–115.
- [2] 韩向娣, 周艺, 王世新等. 夜间灯光遥感数据的 GDP 空间化处理方法[J]. 地球信息科学学报, 2012, 14(1): 128–136.
- [3] Croft, T. A. Burning waste gas in oil fields [J]. *Nature*, 1973, 245(5425): 375–376.
- [4] 卢秀, 李佳, 段平等. 云南沿边地区 GDP 公里网格数据集 (1992–2013) [J/DB/OL]. 全球变化数据仓储电子杂志, 2020. DOI: 10.3974/geodb.2020.03.11.V1.
- [5] 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).
- [6] Abrahams, A., Oram, C., Lozano-Gracia, N. Deblurring DMSP nighttime lights: a new method using Gaussian filters and frequencies of illumination [J]. *Remote Sensing of Environment*, 2018, 210: 242–258.
- [7] 倪愿, 周小成, 江威. 结合 Landsat 数据的 DMSP/OLS 夜间灯光影像去饱和方法研究[J]. 遥感技术与应用, 2017, 32(4): 721–727.
- [8] Elvidge, C. D., Ziskin, D., Baugh, K. E., et al. A fifteen year record of global natural gas flaring derived from Sstellite data [J]. *Energies*, 2009, 2(3): 595–622.
- [9] Wu, J. S., He, S. B., Peng, J., et al. Intercalibration of DMSP-OLS night-time light data by the invariant region method [J]. *International Journal of Remote Sensing*, 2013, 34(20): 7356–7368.
- [10] Liu, Z. F., He, C. Y., Zhang, Q. F., et al. Extracting the dynamics of urban expansion in China using DMSP-OLS nighttime light data from 1992 to 2008 [J]. *Landscape and Urban Planning*, 2012, 106(1): 62–72.
- [11] 钟凯文, 黎景良, 张晓东. 土地可持续利用评价中 GDP 数据空间化方法的研究[J]. 测绘信息与工程, 2007, 32(3): 10–12.
- [12] 卢秀, 李佳, 段平等. 基于夜间灯光和土地利用数据的云南沿边地区 GDP 空间差异性分析[J]. 地球信息科学学报, 2019, 21(3): 455–466.
- [13] 卢秀, 李佳, 段平等. 基于夜间灯光和土地利用数据的 GDP 空间化及预测——以云南沿边地区为例[J]. 地域研究与开发, 2020, 39(2): 36–39, 81.