

四类全球土地覆盖数据在中国区域的精度评价

吴文斌^{1,2,3}, 杨 鹏^{1,2}, 张 莉^{1,2*}, 唐华俊^{1,2}, 周清波^{1,2}, Shibasaki Ryosuke³

(1. 中国农业科学院农业资源与农业区划研究所, 北京 100081; 2. 农业部资源遥感与数字农业
重点开放实验室, 北京 100081; 3. 日本东京大学空间情报科学研究中心, 东京 153-8505)

摘 要: 该研究以中国耕地类别为研究对象, 选择 2000 年中国土地利用数据 (NLCD-2000) 为参考数据, 利用比较分析法, 从面积数量精度和空间位置精度两方面对目前 4 类全球土地覆盖数据 (UMD、IGBP-DISCover、MODIS 和 GLC2000) 产品进行了精度验证, 并分析研究了 4 类数据精度的异同性。结果表明, 4 类全球数据对中国耕地数量特征和空间位置特征的估测具有明显的区域差异性。MODIS 数据集和 GLC2000 数据集对中国耕地制图的总体精度要高于 UMD 数据集和 IGBP-DISCover 数据集。4 类数据制图精度高的区域主要分布在中国的农业主产区, 而误差大的区域主要分布在中国山区或耕地比例低的区域。低空间分辨率的信息源、基于像元的分类方法, 以及中国复杂地形特征是 4 类全球土地覆盖数据精度差异的主要原因。

关键词: 数据系统, 数据转换, 图像处理, 全球土地覆盖数据, 中国土地利用数据 (NLCD-2000), 耕地, 精度评价
doi: 10.3969/j.issn.1002-6819.2009.12.030

中图分类号: TP753

文献标识码: A

文章编号: 1002-6819(2009)-12-0167-07

吴文斌, 杨 鹏, 张 莉, 等. 四类全球土地覆盖数据在中国区域的精度评价[J]. 农业工程学报, 2009, 25(12): 167—173.

Wu Wenbin, Yang Peng, Zhang Li, et al. Accuracy assessment of four global land cover datasets in China[J]. Transactions of the CSAE, 2009, 25(12): 167—173. (in Chinese with English abstract)

0 引 言

全球尺度的土地覆盖数据是全球陆地表层过程研究的重要基础^[1-2]。20 世纪 90 年代以来, 随着管理决策和科学研究的迫切需求、卫星传感器的不断发展和计算机制图技术的进步, 基于遥感技术的全球土地覆盖数据制图已经取得了长足发展^[3]。目前, 基于遥感影像数据获取的全球土地覆盖数据产品主要包括: 1) 美国马里兰大学的全球土地覆盖数据 (即 UMD 数据集); 2) 国际地圈-生物圈计划的全球土地覆盖数据 (即 IGBP-DISCover 数据集); 3) 美国波士顿大学的全球土地覆盖数据 (即 MODIS 数据集); 4) 欧盟联合研究中心的全球土地覆盖数据 (即 GLC2000 数据集)。以上 4 类全球土地覆盖数据 (以后简称 4 类数据) 产品的空间分辨率都是 1 km, 其建设的共同目标是为提高人类对全球不同土地覆盖类型的数量和空间分布格局的认识和了解。目前, 这 4 类数据产品已经广泛应用于全球或区域尺度的生态学和地理学^[4-6]、气候变化评价和环境模拟等研究^[7-10]。

由于卫星传感器、数据源、分类系统与制图技术的

不同, 这 4 类全球数据集各自具有优缺点, 彼此间也存在明显差异性。因此, 在选择利用这 4 类数据进行全球或区域尺度的应用服务时, 十分有必要对数据产品精度进行验证, 弄清它们之间的异同性。虽然数据生产者对各自的数据产品精度进行了定量评价, 但由于各自使用了不同的验证框架和参考数据, 使得这些独立完成的精度验证结果不具有可比性^[11-12]。此外, 也有部分国外学者在不同区域对这 4 类数据产品精度进行了验证和比较分析研究^[13-17], 但目前在中国区域还没有开展类似的精度验证研究, 限制了这 4 类数据在中国的应用前景。因此, 本研究选择耕地类别为研究对象, 利用统一的参考数据, 在中国区域对以上 4 类数据进行精度验证, 定量评价这 4 类全球数据对中国耕地制图的异同性。

1 数据来源和预处理

1.1 数据来源

表 1 列出了 4 类全球土地覆盖数据集的基本信息。UMD 数据集利用 1992—1993 年的 1 km NOAA/AVHRR 影像数据生成, 其使用了简化的 IGBP 土地覆盖分类系统, 共包括 14 个类别^[18]。IGBP-DISCover 数据集和 UMD 数据集使用了相同的遥感数据源, 但其利用了完整的 IGBP 分类系统, 包括 17 个类别^[19]。MODIS 数据集也遵循了 IGBP 土地覆盖分类系统, 但其使用的是 2000—2001 年 1 km TERRA/MODIS 影像数据^[20-21]。GLC2000 数据是基于 1999—2000 年 1 km SPOT4/VEGETATION 数据生成, 其使用了联合国粮食与农业组织倡导的 LCSS 分类系统^[22], 全球尺度共包括 22 个类别。GLC2000 数据加工采用了“自下而上”的技术流程, 即全球不同区域的数据

收稿日期: 2009-03-29 修订日期: 2009-11-06

基金项目: 国家自然科学基金项目 (40930101 和 40971218); 国家高技术研究发展计划 (863 计划) (2009AA12Z143); 农业部“948”项目 (2009-Z31) 和中央级公益性科研院所专项资金项目 (IARRP-2010-02) 资助

作者简介: 吴文斌 (1976—), 男, 湖北潜江人, 博士, 主要从事农业遥感和土地利用变化等方面的研究。北京 中国农业科学院农业资源与农业区划研究所, 100081。Email: ww@caas.net.cn

*通信作者: 张 莉 (1980—), 女, 主要从事农业及生态环境遥感研究。

北京 中国农业科学院农业资源与农业区划研究所, 100081。

Email: zhangli@caas.net.cn

生产者各自完成本区域的土地覆盖制图，然后将不同区域的制图结果加工综合为全球尺度的土地覆盖数据^[23-24]。

表 1 四类全球土地覆盖数据集基本信息

Table 1 Overview of the four global land cover datasets used in this study

| 数据集名 | 版本 | 卫星和传感器 | 数据源时间 | 数据下载地址 |
|---------------|------|-----------------|-------------|---|
| UMD | | NOAA/AVHRR | 1992—1993 年 | http://www.geog.umd.edu/landcover/1km-map.html |
| IGBP-DISCover | v2.0 | NOAA/AVHRR | 1992—1993 年 | http://edcns17.cr.usgs.gov/glcc/globe_int.html |
| MODIS | v004 | TERRA/MODIS | 2000—2001 年 | http://duckwater.bu.edu/lc/datasets.html |
| GLC2000 | v1.1 | SOPT/VEGETATION | 1999—2000 年 | http://www-gvm.jrc.it/glc2000/products |

本研究使用 2000 年中国土地利用数据(NLCD-2000)作为共同参考数据来评价这 4 类数据产品的精度。NLCD-2000 数据来自中国资源环境数据库的 1:100 000 土地利用/土地覆盖数据集，该套空间数据是由中国科学院组织 8 个研究所 100 多位资源环境遥感专家，利用 30 m 空间分辨率的 Landsat TM/ETM 数据覆盖全国，通过人机交互式图像解译生成。NLCD-2000 数据共包括耕地、林地、草地、水域、建设用地和未利用土地等 6 个一级土地利用类别和 25 个二级类别^[25-26]。

1.2 数据预处理

首先，利用中国国界矢量数据剪切这 4 类全球土地覆盖数据，得到中国区域的 4 类土地覆盖数据。然后，将这 4 类数据统一到相同的坐标系和地理投影下，GLC2000 数据的空间分辨率是 0.008929°，利用最小邻近

法重新采样为 0.008333°，使其和其他 3 类全球数据具有相同的空间分辨率。此外，将比例尺为 1:100 000 万的 NLCD-2000 矢量数据转化为栅格数据，和 4 类全球土地覆盖数据保持相同的栅格单元大小，便于进行空间叠加比较分析。

如前所述，4 类数据产品生产过程中使用了不同的分类系统，其最终数据产品所包含的土地覆盖类别不同。本研究主要对耕地类别的制图精度进行验证。为便于进行精度验证和比较分析，需要对 4 类数据和 NLCD-2000 数据进行二值处理，将 5 类数据产品重新划分为耕地和非耕地，保留耕地类别信息，而剔除其余土地覆盖类别的影响。图 1 是从 4 类全球土地覆盖数据和 NLCD-2000 数据中提取的中国耕地空间分布数据。

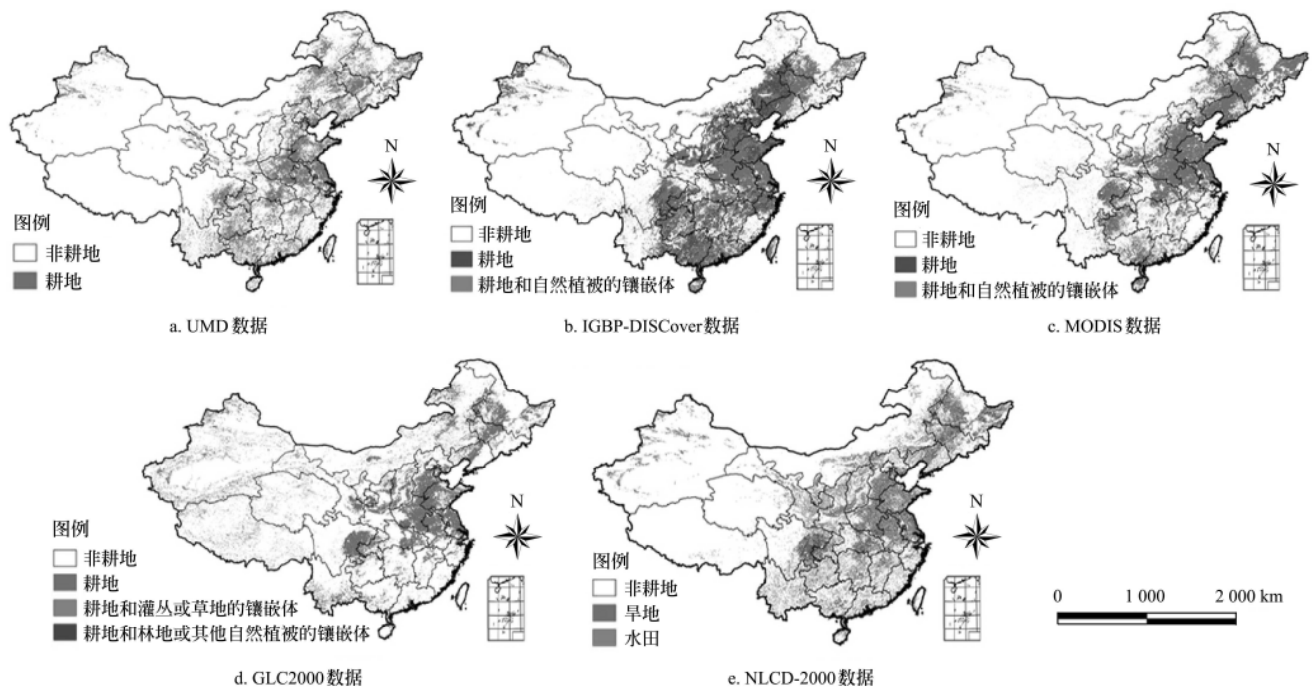


图 1 从 4 类全球土地覆盖数据和 NLCD-2000 数据中提取的中国耕地空间分布

Fig.1 Spatial distribution of cropland in China derived from four global land cover datasets and NLCD-2000 dataset

2 研究方法

比较分析法是遥感数据产品精度验证的常用方法，即通过对比分析被验证数据和参考数据（通常假定为真实数据），估算被验证数据和参考数据的一致性 or 相似性。如果两类数据间吻合度高，那么认为被验证数据精度较高^[27]。空间数据产品精度验证通常包括数量精度验

证和空间位置精度验证两部分内容。数量精度验证在于评价被验证数据各类别的面积等数量特征是否与参考数据对应类别的数量特征一致，空间位置精度验证则侧重于评估被验证数据各类别的空间分布是否和参考数据相应类别的空间分布相似。

本研究以 NLCD-2000 数据为参考数据，利用比较分析法，从面积数量和空间位置精度两方面对 4 类数据进

行精度验证。首先，利用中国省级行政边界矢量图分别从 4 类全球数据和 NLCD-2000 数据中提取不同省份的耕地面积，进行汇总得到中国不同区域和全国耕地总面积；然后，在省级、区域和全国 3 个层次上比较全球数据的结果和 NLCD-2000 数据的结果，定量评价 4 类全球数据集的面积精度。

然而，仅评价遥感数据产品的数量精度可能会掩盖该数据产品质量的真实性和可靠性，因为遥感制图可以相对容易描述 1 个区域某一土地覆盖类别的总数量特征，但其空间位置制图可能并不准确^[28]。所以，面积精度验证之外的空间位置精度验证也十分必要。本研究采用逐像元比较法来验证 4 类全球数据的空间位置精度。具体地，首先，分别将 NLCD-2000 栅格数据和每类全球土地覆盖数据进行空间叠加，如果 2 类数据同一像元的类别一致，则保留原有的类别值；反之，如果 2 类数据同一像元的类别不一致，则赋予给像元新的类别值。经过如此处理生成的新分类数据包括 4 个类别，即耕地/耕地、耕地/非耕地、非耕地/耕地和非耕地/非耕地，这些新类别描述了被验证数据和参考数据是否存在空间位置一致性。其次，利用省级行政边界矢量图从空间叠加生成的 4 类数据中统计不同类别的像元数，并利用如下公式计算每类全球数据在省级、区域和全国 3 个层次上和 NLCD-2000 数据的空间相似性

$$O = \left(\frac{A}{A+B+C} \right) \times 100\% \quad (1)$$

式中： O ——空间相似系数； A 、 B 、 C ——空间叠加生成的新分类数据中的耕地类别总像元数、耕地/非耕地类别总像元数和非耕地/耕地类别总像元数。

3 结果与分析

3.1 面积精度

表 2 是从 4 类全球土地覆盖数据集和 NLCD-2000 数据中计算得到的不同区域的耕地面积。可以看出，UMD、IGBP-DISCover、MODIS、GLC2000 和 NLCD-2000 数据估测中国耕地总面积分别为 112×10^6 、 194×10^6 、 145×10^6 、 123×10^6 和 $141 \times 10^6 \text{ hm}^2$ 。MODIS 数据集监测的全国耕地总面积和 NLCD-2000 数据最接近，IGBP-DISCover 和 UMD 数据集的估测结果与 NLCD-2000 数据存在较大偏差。平均而言，UMD、IGBP-DISCover、MODIS 和 GLC2000 数据集和 NLCD-2000 数据的偏差分别为 -20.9%、37.4%、2.9% 和 -12.7%。

分区域看，4 类全球土地覆盖数据和 NLCD-2000 数据的异同特征不尽相同。虽然 UMD 数据集和 IGBP-DISCover 数据集使用了相同的遥感数据源，但在中国二者的耕地数量制图结果却明显不同。UMD 数据集总体上较低估测了中国每个区域的耕地面积，尤其以华北区域和西北区域为突出，和 NLCD-2000 数据的平均误差分别是 -28.3% 和 -58.3%；而 IGBP-DISCover 数据集却较高估测了中国多数区域的耕地面积，在东南区域、华中区域和西南区域最为显著，相对于 NLCD-2000 数据的偏差分别为 75.2%、53.7% 和 59.5%。与其他 3 类全球数据相比，MODIS 数据集对中国不同区域的耕地面积估测最为理想，也有部分区域和 NLCD-2000 数据存在较大差异，如在东北区域（32.8%）和西北区域（-39.2%）。和 NLCD-2000 数据比较，GLC2000 数据集过高估测了西北区域的耕地面积，约为 51.4%，除此之外的其他区域的耕地面积都被明显低估。

表 2 四类全球土地覆盖数据集计算的中国不同省份和区域的耕地面积以及其与 NLCD-2000 数据的偏差
Table 2 Aggregated cropland areas and their percentage deviations from the NLCD-2000 dataset for the UMD, IGBP-DISCover, MODIS, and GLC2000 datasets in different areas of China

| 区域 | 省、市、 自治区 | UMD | | IGBP-DISCover | | MODIS | | GLC2000 | | NLCD-2000 |
|----------|-------------|------------------------------|----------------------|------------------------------|----------------------|------------------------------|----------------------|------------------------------|----------------------|------------------------------|
| | | 耕地面积/ 10^3 hm^2 | 与 NLCD-2000 的偏差/% | 耕地面积/ 10^3 hm^2 | 与 NLCD-2000 的偏差/% | 耕地面积/ 10^3 hm^2 | 与 NLCD-2000 的偏差/% | 耕地面积/ 10^3 hm^2 | 与 NLCD-2000 的偏差/% | 耕地面积/ 10^3 hm^2 |
| 东北 区域 | 内蒙古自治区 | 12 672 | 33.5 | 15 854 | 67.0 | 11 221 | 18.2 | 12 546 | 32.1 | 9 495 |
| | 黑龙江省 | 10 088 | -30.6 | 12 152 | -16.5 | 18 894 | 29.9 | 9 847 | -32.3 | 14 546 |
| | 吉林省 | 5 962 | -4.6 | 8 295 | 32.7 | 8 375 | 34.0 | 4 974 | -20.4 | 6 251 |
| | 辽宁省 | 4 493 | -15.1 | 7 925 | 49.7 | 8 770 | 65.7 | 3 360 | -36.5 | 5 293 |
| | 汇总 | 33 215 | -6.7 | 44 226 | 24.3 | 47 260 | 32.8 | 30 727 | -13.7 | 35 585 |
| 华北 区域 | 北京市 | 299 | -12.3 | 679 | 99.1 | 389 | 14.1 | 298 | -12.6 | 341 |
| | 天津市 | 512 | 6.4 | 764 | 58.8 | 585 | 21.6 | 596 | 23.9 | 481 |
| | 河北省 | 5 492 | -22.5 | 10 740 | 51.5 | 8 105 | 14.3 | 6 208 | -12.4 | 7 089 |
| | 山西省 | 1 280 | -74.1 | 7 700 | 55.6 | 3 850 | -22.2 | 4 548 | -8.1 | 4 949 |
| | 山东省 | 6 573 | -20.9 | 10 358 | 24.6 | 9 617 | 15.7 | 7 484 | -10.0 | 8 312 |
| | 河南省 | 7 064 | -16.3 | 9 681 | 14.6 | 9 843 | 16.6 | 8 229 | -2.5 | 8 444 |
| | 汇总 | 21 220 | -28.3 | 39 922 | 34.8 | 32 389 | 9.4 | 27 363 | -7.6 | 29 616 |
| 东南 区域 | 上海市 | 230 | -29.7 | 375 | 14.7 | 188 | -42.5 | 256 | -21.7 | 327 |
| | 江苏省 | 4 963 | -2.3 | 6 079 | 19.7 | 5 355 | 5.4 | 4 935 | -2.9 | 5 080 |
| | 浙江省 | 1 626 | -24.8 | 5 141 | 137.9 | 1 205 | -44.2 | 1 463 | -32.3 | 2 161 |
| | 福建省 | 1 202 | -26.2 | 2 089 | 28.2 | 823 | -49.5 | 313 | -80.8 | 1 629 |
| | 广东省 | 3 824 | 16.1 | 7 976 | 142.1 | 4 121 | 25.1 | 922 | -72.0 | 3 295 |
| | 海南省 | 427 | -37.8 | 1 189 | 73.1 | 1 023 | 48.9 | 177 | -74.2 | 687 |

接上页

| 区域 | 省、市、 自治区 | UMD | | IGBP-DISCover | | MODIS | | GLC2000 | | NLCD-2000 |
|----------|-------------|--|----------------------|--|----------------------|--|----------------------|--|----------------------|--|
| | | 耕地面积/ 10 ³ hm ² | 与 NLCD-2000 的偏差/% | 耕地面积/ 10 ³ hm ² | 与 NLCD-2000 的偏差/% | 耕地面积/ 10 ³ hm ² | 与 NLCD-2000 的偏差/% | 耕地面积/ 10 ³ hm ² | 与 NLCD-2000 的偏差/% | 耕地面积/ 10 ³ hm ² |
| 东南 区域 | 台湾省 | 589 | -15.1 | 1 460 | 110.4 | 354 | -49.0 | 279 | -59.8 | 694 |
| | 汇总 | 12 861 | -7.3 | 24 309 | 75.2 | 13 069 | -5.8 | 8 345 | -39.8 | 13 873 |
| 华中 区域 | 安徽省 | 4 890 | -19.8 | 8 063 | 32.2 | 6 494 | 6.5 | 5 444 | -10.7 | 6 097 |
| | 江西省 | 2 635 | -21.1 | 6 683 | 100.1 | 2 871 | -14.0 | 1 520 | -54.5 | 3 339 |
| | 湖北省 | 4 739 | -8.7 | 6 214 | 19.8 | 6 064 | 16.9 | 4 278 | -17.5 | 5 188 |
| | 湖南省 | 3 699 | -18.7 | 8 512 | 87.0 | 3 061 | -32.7 | 1 498 | -67.1 | 4 551 |
| | 汇总 | 15 963 | -16.8 | 29 472 | 53.7 | 18 490 | -3.6 | 12 740 | -33.6 | 19 175 |
| 西南 区域 | 广西壮族自治区 | 5 741 | 50.1 | 11 446 | 199.2 | 5 229 | 36.7 | 1 440 | -62.4 | 3 826 |
| | 重庆市 | 2 116 | -25.6 | 3 950 | 38.8 | 2 461 | -13.5 | 2 392 | -15.9 | 2 845 |
| | 四川省 | 4 635 | -47.2 | 9 805 | 11.6 | 7 181 | -18.2 | 7 189 | -18.1 | 8 782 |
| | 贵州省 | 4 696 | -7.5 | 8 385 | 65.2 | 5 571 | 9.7 | 1 355 | -73.3 | 5 077 |
| | 云南省 | 3 483 | -11.9 | 5 456 | 38.1 | 2 390 | -39.5 | 3 759 | -4.9 | 3 952 |
| | 汇总 | 20 671 | -15.6 | 39 042 | 59.5 | 22 832 | -6.7 | 16 135 | -34.1 | 24 482 |
| 西北 区域 | 西藏自治区 | 27 | -94.2 | 785 | 67.4 | 482 | 2.8 | 5 764 | 1 129.0 | 469 |
| | 陕西省 | 1 308 | -76.8 | 3 980 | -29.6 | 4 473 | -20.8 | 3 484 | -38.3 | 5 650 |
| | 甘肃省 | 1 535 | -72.0 | 3 078 | -43.9 | 1 816 | -66.9 | 4 122 | -24.8 | 5 482 |
| | 青海省 | 1 333 | 105.7 | 272 | -58.0 | 677 | 4.5 | 2 235 | 244.9 | 648 |
| | 宁夏回族自治区 | 240 | -83.4 | 447 | -69.0 | 202 | -86.0 | 492 | -65.9 | 1 443 |
| | 新疆维吾尔自治区 | 3 235 | -31.5 | 8 337 | 76.6 | 3 543 | -25.0 | 11 771 | 149.3 | 4 721 |
| | 汇总 | 7 678 | -58.3 | 16 899 | -8.2 | 11 193 | -39.2 | 27 868 | 51.3 | 18 413 |
| 全国 | | 111 608 | -20.9 | 193 870 | 37.4 | 145 233 | 2.9 | 123 178 | -12.7 | 141 144 |

图 2 描述了 4 类全球土地覆盖数据集与 NLCD-2000 数据在省级耕地面积估测的相关性。UMD、IGBP-DISCover、MODIS 和 GLC2000 数据集与 NLCD-2000 数据的相关系数分别为 0.735、0.661、0.876 和 0.707。可以看出，在 4 类全球土地覆盖数据集中，

MODIS 数据对中国不同省份耕地面积的监测效果最好，UMD 和 GLC2000 数据集总体上较低估测了中国省级的耕地面积（约 20%），而 IGBP-DISCover 数据集则过高估测了耕地面积（约 26%）。此外，从表 2 中可以发现，耕地面积占全国耕地总面积比例高于 5% 的省份主要包括

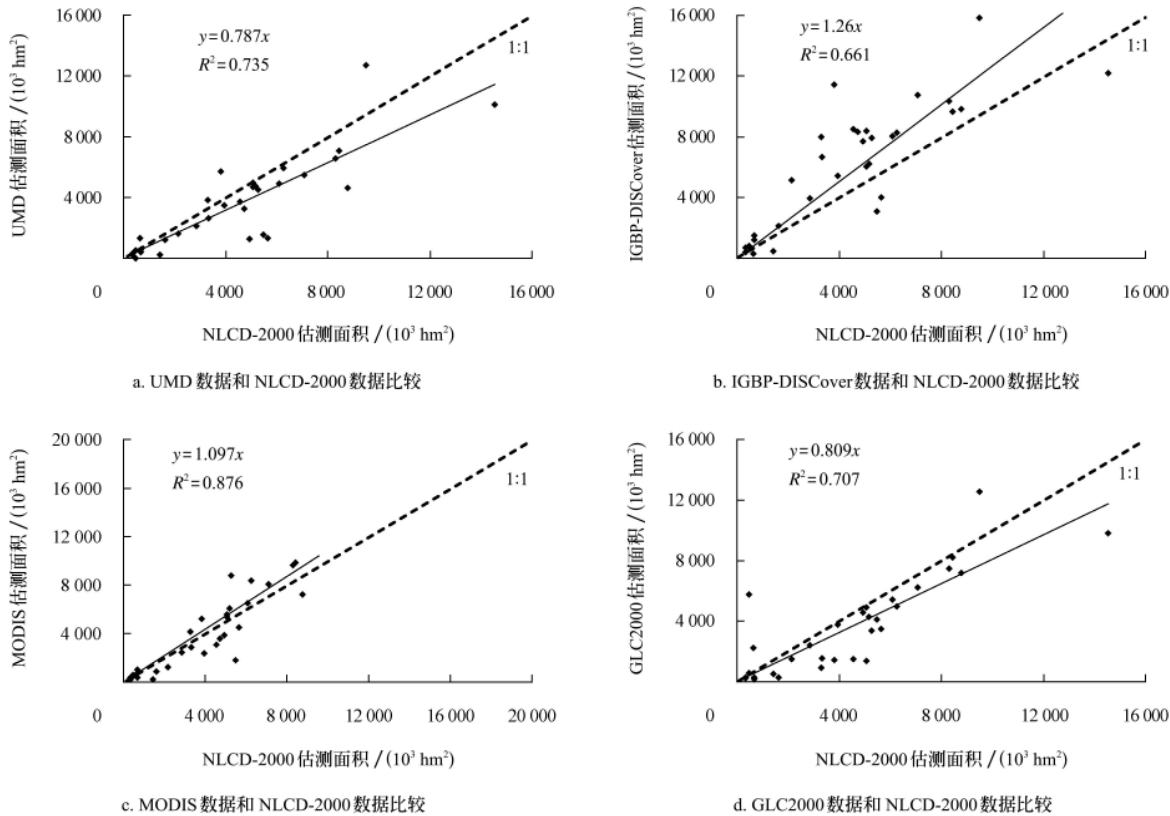


图 2 四类全球土地覆盖数据估测的省级耕地面积和 NLCD-2000 估测的省级耕地面积的比较
Fig.2 Comparison of aggregated cropland areas estimated from individual global land cover datasets against the NLCD-2000 dataset at provincial level

内蒙古自治区、黑龙江省、河北省、山东省、河南省和四川省等 6 个省区，这些省份是中国的农业主产区，其耕地面积之和约占全国总耕地面积的 40%。如果考虑这些农业主产省份的耕地面积，UMD、IGBP-DISCover、MODIS 和 GLC2000 数据与 NLCD-2000 数据的差异并不大，分别为 17.9%、21%、14.5% 和 10%。4 类全球土地覆盖数据和 NLCD-2000 数据差异大的区域主要分布在山西、浙江、福建、广东、海南、台湾、江西、湖南、广西、贵州、西藏、青海和新疆等省和自治区。

3.2 空间位置精度

图 3 是 4 类全球土地覆盖数据和 NLCD-2000 数据的逐像元比较结果。图 4 是利用式 (1) 计算的 4 类数据和 NLCD-2000 数据在全国和区域尺度的空间相似系数。从全国看，每类数据和 NLCD-2000 数据的空间相似性总体一致，约为 30%~41%。从不同区域看，在华北区域，UMD、IGBP-DISCover、MODIS 和 GLC2000 数据与 NLCD-2000 数据的空间一致性较相似，分别为 43.2%、

58.8%、55.8% 和 54.6%。华北地区是中国的主要产粮基地，地形平坦，土地覆盖类型相对单一，遥感制图相对容易，因此，4 类数据的耕地制图结果较为接近。在东南区域和华中区域，4 类全球土地覆盖数据和 NLCD-2000 数据的空间一致性也相似，约为 40%。在东北区域，UMD 和 GLC2000 数据与 NLCD-2000 数据的一致性分别是 27.7% 和 32.9%，而 IGBP-DISCover 和 MODIS 数据分别为 41.0% 和 43.8%，较前两者高约 10%。从图 4 也可以看出，4 类全球数据在西南区域和西北区域耕地的空间位置精度要比其他区域低，尤其在西北区域，4 类数据对耕地空间分布提取误差最大，UMD、IGBP-DISCover、MODIS 和 GLC2000 与 NLCD-2000 数据的空间相似系数分别为 8.3%、22.5%、18.1% 和 13.2%。此外，4 类数据估测的西南区域耕地空间分布和 NLCD-2000 数据相比，其相似系数也仅为 23%~30%。这或许是由于西南和西北两个区域以丘陵和山地分布为主，地形状况复杂，耕地地块破碎，混合像元现象普遍，在很大程度上影响了遥感制图精度。

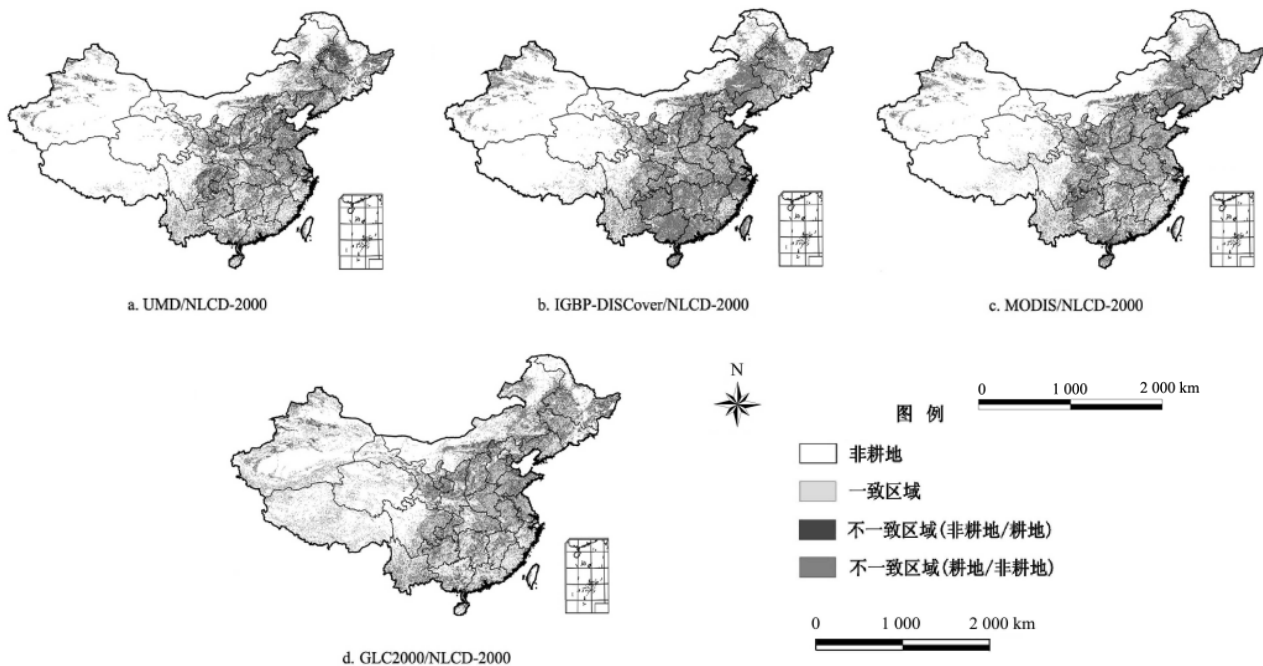


图 3 四类全球土地覆盖数据和 NLCD-2000 数据的逐像元空间比较

Fig.3 Uniformity of cropland areas for the UMD, IGBP-DISCover, MODIS, and GLC-2000 datasets versus the NLCD-2000 dataset

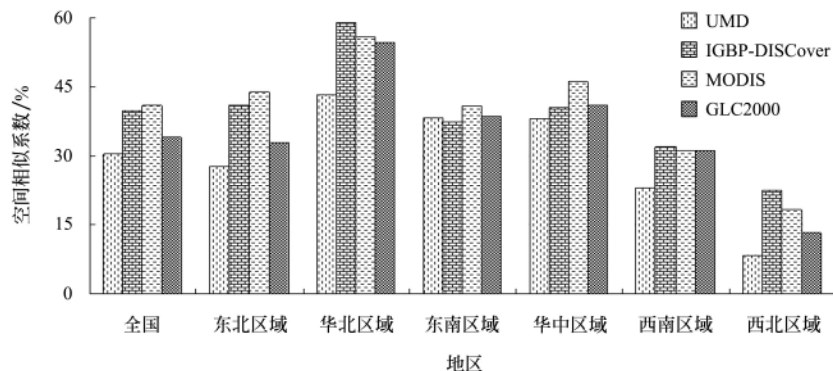


图 4 四类全球土地覆盖数据和 NLCD-2000 数据在全国和区域尺度的空间相似系数

Fig.4 Spatial similarity coefficients between individual global land cover datasets and the NLCD-2000 dataset for cropland spatial distribution at national and regional scales

4 结 论

本研究以中国耕地类别为研究对象, 利用 NLCD-2000 数据作为参考数据, 从面积数量精度和空间位置精度两方面对 4 类全球土地覆盖数据产品进行了精度验证, 并比较分析了它们之间的异同点。研究结论如下:

1) MODIS 和 GLC2000 数据集对中国耕地制图的整体精度要高于 UMD 和 IGBP-DISCover 数据集。MODIS 和 GLC2000 数据产品的生产年份比 UMD 和 IGBP-DISCover 要晚近 10 a, 其制图整体精度的差异反映了全球土地覆盖制图在数据源和技术方法等方面的全面进步。

2) 四类全球数据对中国耕地的数量特征和空间特征的估测精度具有显著的区域差异性。制图精度高的区域主要分布在中国农业主产区, 如华北区域和东部区域, 而制图误差大的区域主要发生在中国山区或耕地比例低的区域, 如西北区域和东南区域。这种空间差异性主要与全球土地覆盖制图使用的低空间分辨率遥感数据和基于像元的分类方法有关。在中国农业主产省份, 耕地大面积成片分布, 地表覆盖类型单一, 遥感数据的光谱分辨能力高, 进行土地覆盖类型识别也就相对容易。然而, 在中国非农业主产区或山地丘陵区, 地形起伏大, 耕地零星分布, 而且与其他植被类型高度混合, 使得地表复杂程度高于卫星传感器的识别能力。利用低空间分辨率遥感数据在这些区域进行土地覆盖制图时, 显然会有更多误差和不确定性。可见, 如何解决“混合像元”问题是目前利用低空间分辨率数据进行全球土地覆盖制图所面临的最大挑战之一。

本研究的精度评价结果是通过比较高空间分辨率的参考数据和低空间分辨率数据的被验证数据得到的。需要说明的是, 高空间分辨率的 NLCD-2000 数据本身也存在一些不足。如 NLCD-2000 数据也是利用基于像元的分类方法, 30 m 大小像元中可能存在的非耕地信息未予以剔除, 使得 NLCD-2000 数据对中国耕地的监测结果总体偏高。即便如此, 本研究结果仍可以帮助数据用户根据研究目的和区域选择适宜的土地覆盖数据, 提升区域尺度的生物地球化学循环、土地利用和土地覆盖变化等研究。此外, 研究结果还可以为数据生产者提供反馈信息, 以促进数据处理算法和分类技术方法的改进, 更好地服务于未来的全球土地覆盖制图活动。

[参 考 文 献]

- [1] Cihlar J. Land cover mapping of large areas from satellites: status and research priorities[J]. *International Journal of Remote Sensing*, 2000, 21(6/7): 1093—1114.
- [2] 李晓兵, 陈云浩, 喻锋. 基于遥感数据的全球及区域土地覆盖制图: 现状、战略和趋势[J]. *地球科学进展*, 2004, 19(1): 71—80.
Li Xiaobing, Chen Yunhao, Yu Fent. Global and regional cover mapping from remote sensing data: Status quo, strategies and trends[J]. *Advance in Earth Sciences*, 2004, 19(1): 71—80. (in Chinese with English abstract)
- [3] 杨立民, 朱智良. 全球及区域的土地覆盖土地利用遥感研究的现状和展望[J]. *自然资源学报*, 1999, 4(4): 340—344.
Yang Liming, Zhu Zhiliang. The status quo and expectation of global and local land cover and land use RS research[J]. *Journal of Natural Resources*, 1999, 4(4): 340—344. (in Chinese with English abstract)
- [4] Belward A S, Estes J E, Kline K D. The IGBP-DIS global 1 km land-cover dataset DISCover: a project overview[J]. *Photogrammetric Engineering and Remote Sensing*, 1999, 65(9): 1013—1020.
- [5] Defries R S, Hansen M C, Townshend J R, et al. A new global 1-km dataset of percentage of tree cover derived from remote sensing[J]. *Global Change Biology*, 2000, 6(2): 247—254.
- [6] 吴文斌, 杨鹏, 周清波, 等. 2005~2035年全球农作物播种面积变化情景模拟研究[J]. *农业工程学报*, 2007, 23(10): 378—390.
Wu Wenbin, Yang Peng, Zhou Qingbo, et al. Modeling sown area change for major crops during 2005-2035 at a global scale[J]. *Transactions of the CSAE*, 2007, 23(10): 378—390. (in Chinese with English abstract)
- [7] Hese S, Lucht W, Schimmlus C, et al. Global biomass mapping for an improved understanding of the CO₂ balance—the Earth observation mission Carbon-3D[J]. *Remote Sensing of Environment*, 2005, 94(1): 94—104.
- [8] Ramankutty N, Foley J A, Norman J, et al. The global distribution of cultivable lands: current patterns and sensitivity to possible climate change[J]. *Global Ecology & Biogeography*, 2002, 11(5): 377—392.
- [9] Giri C, Defourny P, Shrestha S. Land cover characterization and mapping of continental Southeast Asia using multi-resolution satellite sensor data[J]. *International Journal of Remote Sensing*, 2003, 24(2): 418—4196.
- [10] Mucher C A, Steinocher K T, Kressler F P, et al. Land cover characterization and change detection for environmental monitoring of PAN-Europe[J]. *International Journal of Remote Sensing*, 2000, 21(6/7): 1159—1181.
- [11] Jung M, Henkel K, Herold M, et al. Exploiting synergies of global land cover products for carbon cycle modeling[J]. *Remote Sensing of Environment*, 2006, 101(4): 534—553.
- [12] Herold M, Mayaux P, Woodcock C, et al. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets[J]. *Remote Sensing of Environment*, 2008, 112(5): 2538—2556.
- [13] Hansen M C, Reed B. A comparison of the IGBP DISCover and University of Maryland 1km global land cover products[J]. *International Journal of Remote Sensing*, 2000, 21(6/7): 1365—1373.
- [14] Giri C, Zhu Z, Reed B. A comparative analysis of the Global Land Cover 2000 and MODIS land cover datasets[J]. *Remote Sensing of Environment*, 2005, 94(1): 123—132.
- [15] See L M, Fritz S. A method to compare and improve land cover datasets: application to the GLC2000 and MODIS land cover products[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(7): 1740—1746.
- [16] Herold M, Woodcock C, Di Gregorio, et al. A joint initiative

- for harmonization and validation of land cover datasets[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(7): 1719—1727.
- [17] Mccallum I, Obersteiner M, Nilsson S, et al. A spatial comparison of four satellite derived 1 km global land cover datasets[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2006, 8(4): 246—255.
- [18] Hansen M C, Defries R S, Townshend J R G, et al. Global land cover classification at 1 km spatial resolution using a classification tree approach[J]. *International Journal of Remote Sensing*, 2000, 21(6/7): 1331—1364.
- [19] Loveland T R, Reed B, Brown J F, et al. Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data[J]. *International Journal of Remote Sensing*, 2000, 21(6/7): 1303—1330.
- [20] Friedl M A, Mciver D K, Hodges J C F, et al. Global land cover mapping from MODIS: algorithms and early results[J]. *Remote Sensing of Environment*, 2002, 83(1/2): 287—302.
- [21] Justice C O, Townshend J R G, Vermote E F, et al. An overview of MODIS land data processing and product status[J]. *Remote Sensing of Environment*, 2002, 83(1/2): 3—15.
- [22] Di Gregorio A, Jansen L J M. *Land Cover Classification System (LCCS): Classification Concepts and User Manual* [M]. Rome: Food and Agriculture Organization of the United Nations Publishing Service, 2000.
- [23] Bartholome E, Belward A S. GLC2000: a new approach to global land cover mapping from Earth observation data[J]. *International Journal of Remote Sensing*, 2005, 26(9): 1959—1977.
- [24] 徐文婷, 吴炳方, 颜长珍, 等. 用SPOT-VGT数据制作中国2000年度土地覆盖数据[J]. *遥感学报*, 2005, 9(2): 204—214.
Xu Wenting, Wu Bingfang, Yan Changzhen et al. China land cover 2000 using SPOT-VGT S10 data[J]. *Journal of Remote Sensing*, 2005, 9(2): 204—214. (in Chinese with English abstract)
- [25] 王思远, 刘纪远, 张增祥, 等. 近10年中国土地利用格局以及演变[J]. *地理学报*, 2002, 57(5): 523—530.
Wang Siyuan, Liu Jiyan, Zhang Zengxiang, et al. Spatial pattern change of land use in China in recent 10 years[J]. *Acta Geographica Sinica*, 2002, 57(5): 523—530. (in Chinese with English abstract)
- [26] Liu J, Liu M, Tian H, et al. Spatial and temporal patterns of China's cropland during 1990—2000: An analysis based on Landsat TM data[J]. *Remote Sensing of Environment*, 2005, 98(4): 442—456.
- [27] Lambin E F, Geist H. *Land-use and land-cover change: Local processes and global impacts*[M]. Germany: Springer, 2006.
- [28] Foody G M. Status of land cover classification accuracy assessment[J]. *Remote Sensing of Environment*, 2002, 80(1): 185—201.

Accuracy assessment of four global land cover datasets in China

Wu Wenbin^{1,2,3}, Yang Peng^{1,2}, Zhang Li^{1,2*}, Tang Huajun^{1,2}, Zhou Qingbo^{1,2}, Shibasaki Ryosuke³

(1. *Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China;*

2. *Key Laboratory of Resources Remote Sensing and Digital Agriculture, Ministry of Agriculture, Beijing 100081, China;*

3. *Center for Spatial Information Science, University of Tokyo, Tokyo 153-8505, Japan)*

Abstract: This study aims to examine the suitability of four global land cover datasets (UMD, IGBP-DISCover, MODIS and GLC2000) for their accuracies in mapping and monitoring cropland across China. For that, four global land cover products were firstly compared with the national land cover dataset 2000 (NLCD-2000) at provincial, regional and national scales to evaluate the accuracies of estimation of aggregated cropland area in China. This was followed by a spatial comparison to assess their accuracies in estimating the spatial distribution of cropland across China. The results showed that there were varying levels of apparent discrepancies in estimating China's cropland among these four global datasets, and that both aggregated areas and spatial agreement between them varied from region to region. MODIS and GLC2000 datasets had a relatively higher accuracy in depicting China's cropland than UMD and IGBP-DISCover datasets. The coarse spatial resolution and per pixel classification approach, as well as landscape heterogeneity, are the main reasons for large discrepancies between these four global land cover datasets and NLCD-2000 dataset.

Key words: database systems, data transfer, image processing, global land cover datasets, national land cover dataset 2000(NLCD-2000), cropland, accuracy assessment